THE COMPLETE
TECHNICAL PAPER PROCEEDINGS
FROM:

Published by: ncta SCTE · ISBE

Compiled by:
Mark Bell, VP, Industry and Association Affairs
Katie Mercier, Director, Programs & Events

Current and past editions of the *Technical Forum Proceedings* and *NCTA & SCTE·ISBE Technical Papers* are available online at www.nctatechnicalpapers.com.

ISBN Number: 0-940272-60-1

5G Fronthaul Over DOCSIS

Transporting O-RAN's Split 7-2x over DOCSIS

A Technical Paper

Mark Grayson
Distinguished Engineer
Cisco Systems
10 New Square Park, Feltham, Middlesex, TW14 8HA, United Kingdom
+44 20 882 43403
mgrayson@cisco.com

John Chapman
Fellow, CTO Broadband Technologies
Cisco Systems
408 526-7651
jchapman@cisco.com

1. Introduction

The transition to open RAN (Radio Access Network) based on interoperable splits is gaining significant momentum across the mobile industry. Conventional composed base stations are being decomposed into separate radio units (RU), distributed units (DU) and centralized units (CU), as illustrated in Figure 1. However, where best to split the open RAN is a complex compromise between RU simplification, DU functionality, CU functionality, support of advanced co-ordinated multipoint RF capabilities, consequential limitations on transport delay budgets as well as interface bandwidth expansion.

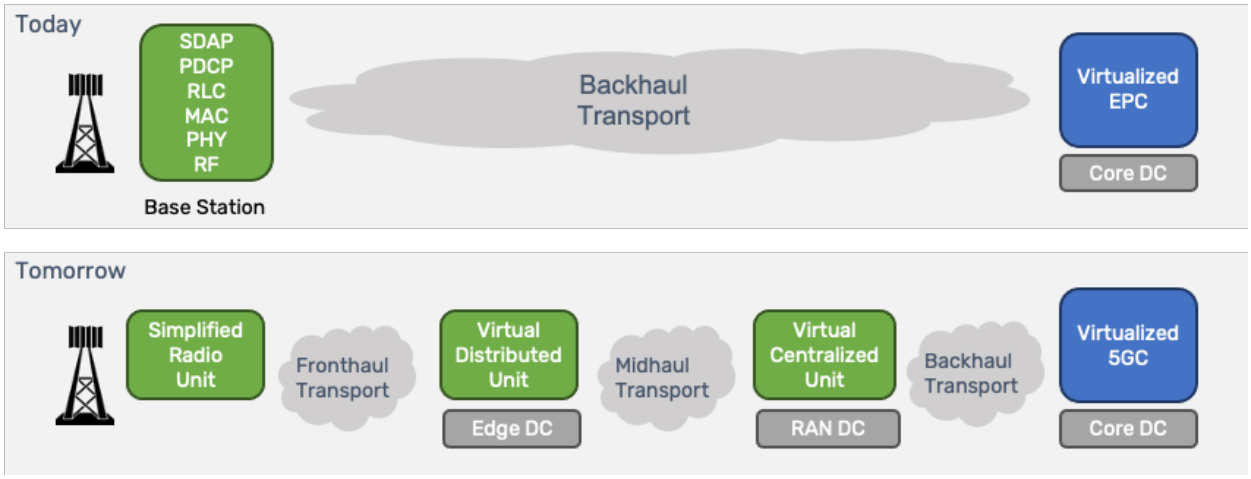


Figure 1 - The Transition to a Decomposed Radio Access Network

During the study into 5G's Radio Architecture, several alternative splits were analysed [1]. These were broadly categorized into "higher layer splits" (HLS) and "lower layer splits" (LLS). The demarcation between these two categories is the location of the scheduler, with the term LLS being applicable to deployments with lower latency transport where it is possible to realize enhanced performance through centralized scheduling, and HLS being applicable to deployment with higher transport latencies that operate with a distributed scheduler.

To help in comparing alternative options, different splits have been assigned numbers with higher numbers representing splits "lower down" in the protocol stack, meaning less functionality being deployed "below" the split in the RU. Lower layer splits occur below the medium access control (MAC) layer in the protocol stack which contains the scheduler functionality. Several lower later splits are possible, including Split 6 - between the MAC and physical layers, Split 7 - within the physical layer, and Split 8 - between the physical layer and the RF functionality. Interestingly, Split 6 is analogous to the Remote PHY interface as defined at CableLabs [2].

Before 5G's analysis into splits, the de facto approach to split the RAN was to use an interface based on the Common Public Radio Interface (CPRI). Back in 2003, the CPRI industry co-operation (www.cpri.info) had defined an interface between a Radio Equipment Control (REC) element implementing all the RAN baseband functions and a Radio Equipment (RE) element implementing the RF functions, to enable the RE to be located at the top of a cell tower and the

REC to be located at the base of the cell tower. This interface was subsequently repurposed to support relocation of the REC to a centralized location that could serve multiple cell towers via a fronthaul transport network. Using the split numerology introduced in 3GPP 38.801, with the RE implementing the RF and REC implementing the physical layer and above, the CPRI-based split is identified as a split 8 approach.

Split 8 has been characterized as requiring up to a 30-fold bandwidth expansion compared to HLS approaches [2]. Furthermore, in LTE the transport latency budget of all LLS approaches is constrained by the operation of hybrid automatic repeat request (HARQ) functionality in the MAC layer. This results in the oft-quoted delay requirement of 250 micro-seconds for one way transport delay budget between the radio and the element implementing the MAC layer's uplink HARQ functionality. Finally, with 5G's increasing focus on multi-antenna systems, the CPRI split 8 approach is hampered by the need to scale linearly with the number of RF elements.

With such extreme requirements, it is evident why the focus of DOCSIS based transport has been on HLS approaches to RAN decomposition, where there is nominal bandwidth expansion compared to conventional RAN backhaul systems and delay budgets can be measured in milli-seconds instead of micro-seconds. This paper takes an alternative view and compares the requirements for supporting two different lower layer splits, namely the network functional application platform interface (nFAPI) Split 6 as defined by the Small Cell Forum (www.smallcellforum.org) and the Split 7-2x as defined by the O-RAN Alliance (www.o-ran.org). Whereas the pre-conception is that lower layer splits are incompatible with fronthaul being transported using DOCSIS, this paper examines the LLS requirements associated with these splits and demonstrates how, given correct configuration, fronthaul deployments can be compatible with DOCSIS based transport.

2. Small Cell Splits

The Small Cell Forum (SCF) took the initial lead in defining a multivendor lower layer split, taking its FAPI platform application programming interface (API) that had been used as an informative split of functionality between small cell silicon providers and the small cell RAN protocol stack providers, and enabling this to be "networked" over an IP transport.

This "networked" FAPI, or nFAPI, enables the Physical Network Function (PNF) implementing the small cell RF and physical layer to be remotely located from the Virtual Network Function (VNF) implementing the small cell MAC layer and upper layer RAN protocols. First published by the SCF in 2016, the specification of the MAC/PHY split has since been labelled as "Split 6" by 3GPP TR38.801 that studied 5G's New Radio access technology and architectures. The nFAPI architecture is shown in Figure 2.

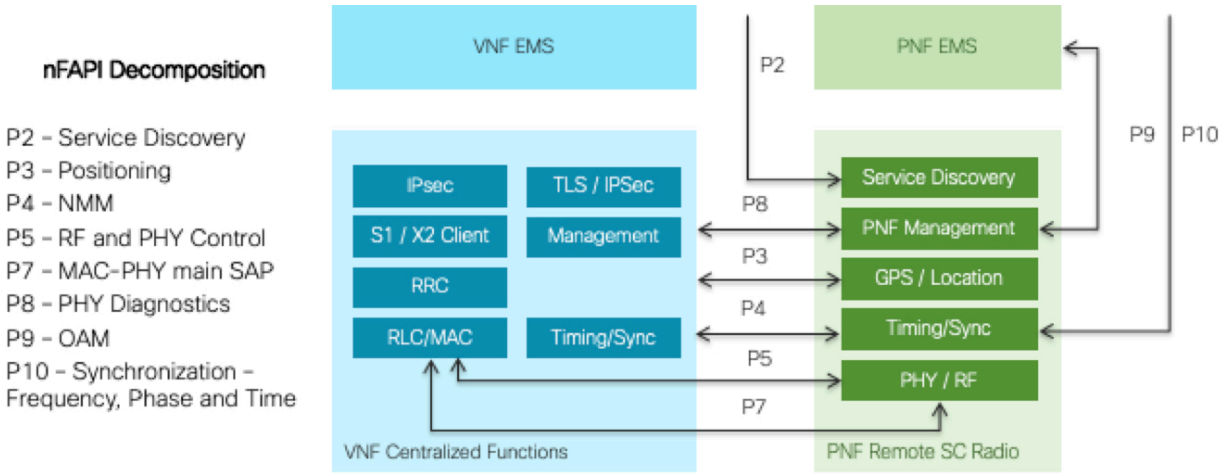


Figure 2 - Small Cell Forum's nFAPI MAC/PHY Split

The initial SCF nFAPI program delivered important capabilities that enabled small cells to be virtualized. Importantly, when comparing the transport bandwidth requirements for the fronthaul interface, nFAPI/Split 6 does not significantly expand the bandwidth required compared to more conventional small cell backhaul deployments. Moreover, just like the backhaul traffic, the nFAPI transport bandwidth varies according to served traffic, enabling statistical multiplexing to be used over the fronthaul IP network. This can be contrasted with the alternative CPRI/Split 8 that requires bandwidth expansion up to 30-fold and a constant bit rate connection, even if there is no traffic being served in a cell.

3. HARQ Latency Constraints

LTE contains a retransmit mechanism called the hybrid automatic repeat request (HARQ). DOCSIS does not have an equivalent mechanism. There is an uplink synchronous HARQ which has tight timing constraints and a downlink asynchronous HARQ that does not. The timing constraints discussed here are unique to LTE and do not apply to 5G. This is important as in the Cable industry, some operators may only consider deploying 5G on their DOCSIS network.

Whereas nFAPI/Split 6 offers significant benefits over CPRI/Split 8 in terms of bandwidth expansion, both splits are below the hybrid automatic repeat request (HARQ) functionality in the MAC layer that is responsible for constraining the transport delay budget for LTE fronthaul solutions. LTE-based Split 6, Split 7 and Split 8 all have a common delay constraint equivalent to 3 milliseconds between when uplink data is received by the RF to the time when the corresponding downlink ACK/NAK needs to be ready to be transmitted by the RF, as illustrated in Figure 3.

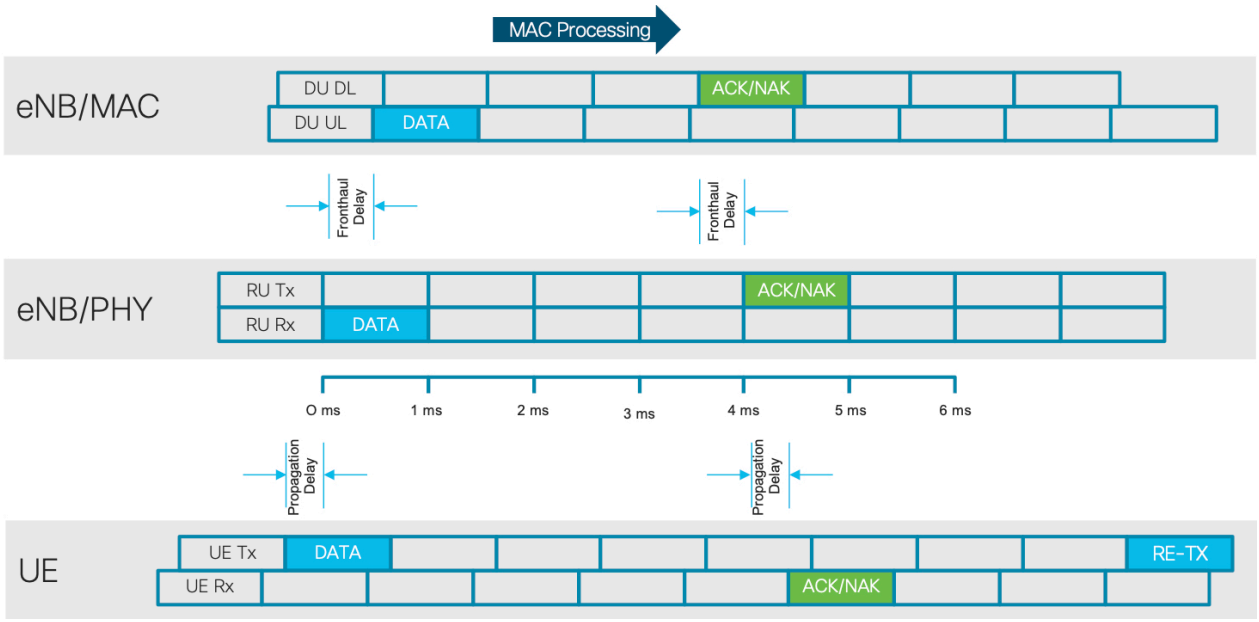


Figure 3 - HARQ Latency Constraints for LTE

These 3 milliseconds need to be allocated to HARQ processing by the MAC layer and fronthaul transport delay, with a common assumption being that 2.5 milliseconds are allocated to processing, leaving 0.5 milliseconds allocated to round trip transport, or 250 micro-seconds for one way transport delay budget between the element implementing the RF and the element implementing the MAC layer's uplink HARQ functionality.

The Small Cell Forum acknowledges such limitations when using its nFAPI split. Because the 250 micro-seconds one way transport budget severely constrains nFAPI deployments, SCF defines the use of HARQ interleaving that leverages standardized signalling to defer HARQ buffer emptying, enabling higher latency fronthaul links to be accommodated. Although HARQ interleaving buys additional transport delay budget, the operation has a severe impact on single user equipment (UE) throughput; as soon as the delay budget exceeds the constraint described above, the per UE maximum throughput is immediately decreased by 50%, with further decreases as delays in the transport network increase, as illustrated in Figure 4.

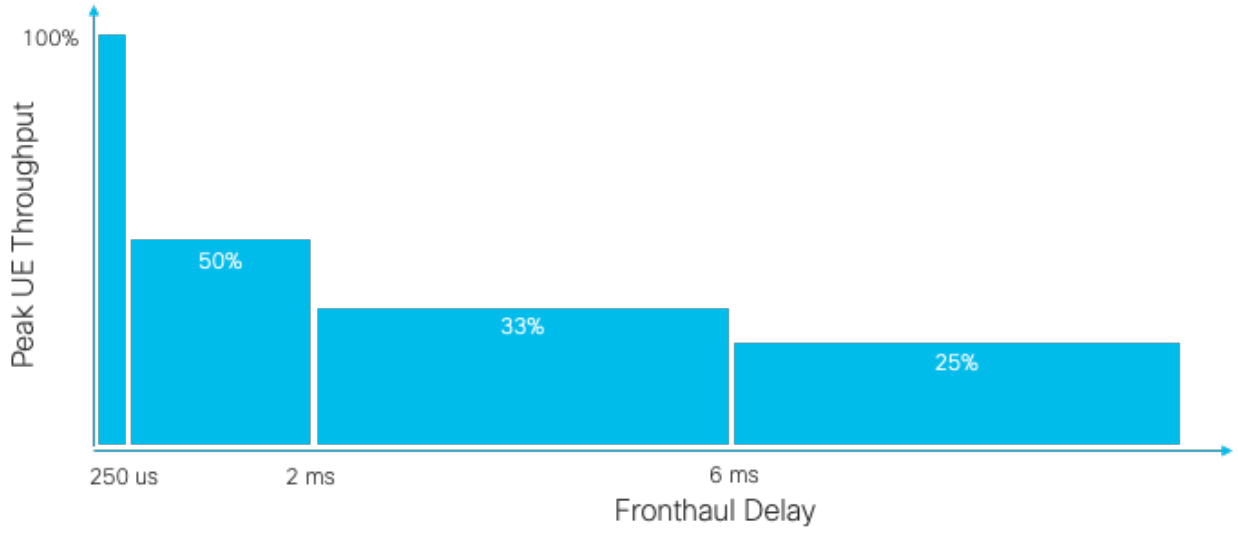


Figure 4 - Impact of HARQ interleaving on peak UE throughput

Some proponents have advocated that in certain LTE deployment scenarios, the operation of HARQ can be disabled to avoid the associated 250 micro-second delay constraint. However, analysis indicates that HARQ operation plays an important role in improving performance as signal to node ratio (SNR) falls below 8 dB [4], meaning that a significant number of cell edge users will likely be impacted if HARQ is disabled.

The restriction associated with the operation of LTE uplink HARQ is due to the definition of HARQ operation as being synchronous, whereby the identity of a HARQ process is derived from its transmission timing. This can be contrasted to the operation of the LTE downlink HARQ that defines the signalling of a HARQ process identity along with the data which means that there are no equivalent downlink timing constraints.

Importantly, 5G new radio (NR) does not implement the same synchronous uplink HARQ procedures, instead defining the use of HARQ process identities in both the downlink and the uplink. Consequently, 5G fronthaul systems do not suffer the same HARQ-based transport delay constraints as LTE. Instead, the limiting factor constraining the transport budget in 5G fronthaul systems is the operation of the windowing during the random-access procedure.

When a UE wants to establish a connection, it first recovers information on the system information broadcast type 2 (SIB2) message broadcast by the cell. SIB2 includes information about the random-access channel (RACH) configuration. The UE uses the RACH configuration to determine the time, frequency, preamble identity and repetition information to use when sending the physical random-access channel (PRACH) preamble, sometimes referred to as “message 1” (MSG1). If received correctly by the network, the base station will transmit a random-access response (RAR) message to the UE. The UE will monitor the physical downlink control channel (PDCCH) for reception of the RAR message sent by the network, sometimes referred to as “message 2” (MSG2). Importantly, the monitoring period is controlled by a

parameter termed *raResponseWindow* which has a maximum value of 10 milliseconds [5], as illustrated in Figure 5.

This maximum value of 10 milliseconds needs to be partitioned between the PRACH preamble processing by the MAC layer, round-trip transport delays and over the air delays. Allocating 2.5 milliseconds to PRACH preamble processing leaves 7.5 milliseconds to be allocated between the over the air transmissions and round-trip transport delay. This effectively means that fronthaul round-trip transport delays of up to 5 milliseconds can be accommodated without impacting the currently defined RACH processing.

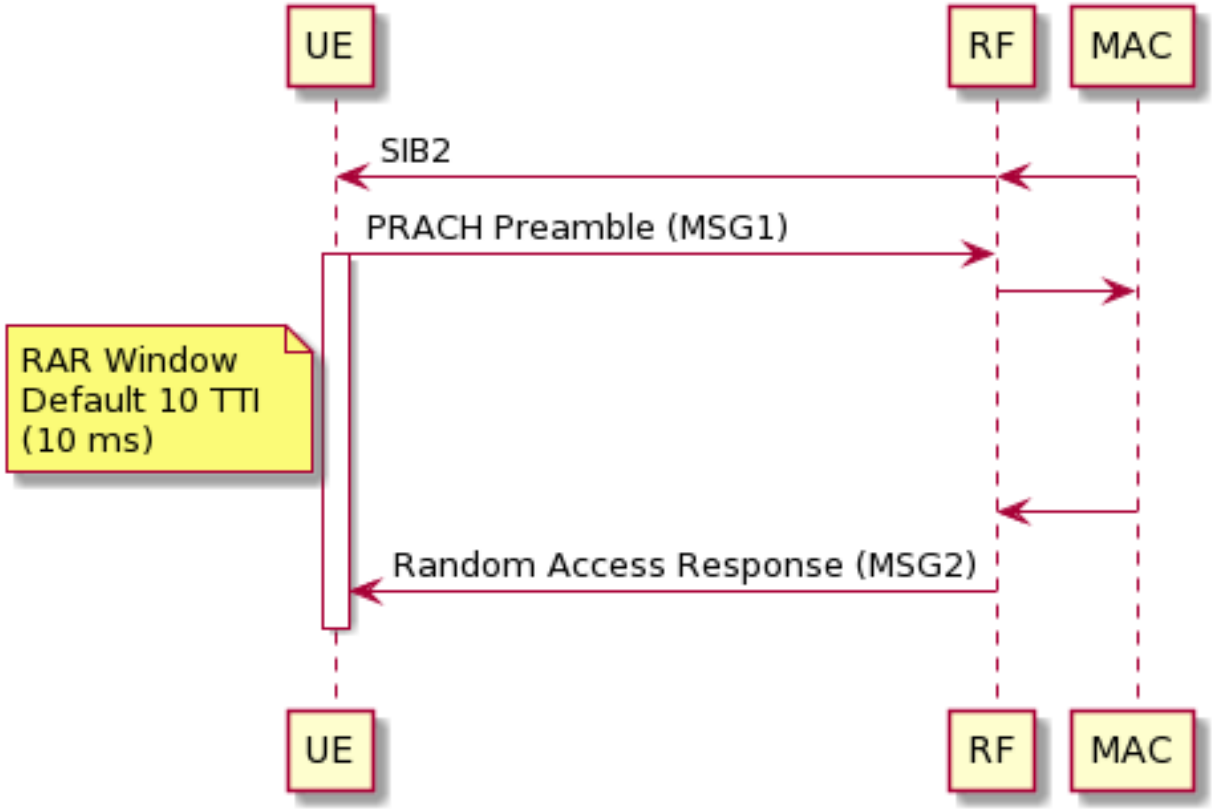


Figure 5 - Random Access Timing

Importantly, whereas the 500 microsecond round-trip transport delay requirement necessitated by LTE's synchronous uplink HARQ cannot be met by the DOCSIS system, the ~5 millisecond round-trip transport delay for 5G might be achieved with an optimized DOCSIS configuration, and thus splits 6 through 8, will work for DOCSIS, at least from the HARQ viewpoint. The Low Latency Xhaul (LLX) feature is targeted at getting the majority mobile traffic to be within 5 milliseconds round trip.

4. Split PHY Alternatives

Unlike in nFAPI/split 6 where there is a clear demarcation between MAC and PHY layers, the newer split 7 intra PHY approach can have multiple realizations. 3GPP 38.801 describes three alternative realizations:

- Split 7-1: whereby the FFT, Cyclic Prefix handling and uplink PRACH processing are distributed into the RU
- Split 7-2: In addition to those functions defined in Split 7-1, the RU additionally includes layer mapping/de-mapping and optionally the precoding functionality
- Split 7-3: In addition to those functions defined in Split 7-2, in the downlink, the RU additionally includes the modulator

These split PHY alternatives, together with the nFAPI/split 6 and the CPRI/split 8 are illustrated in Figure 6.

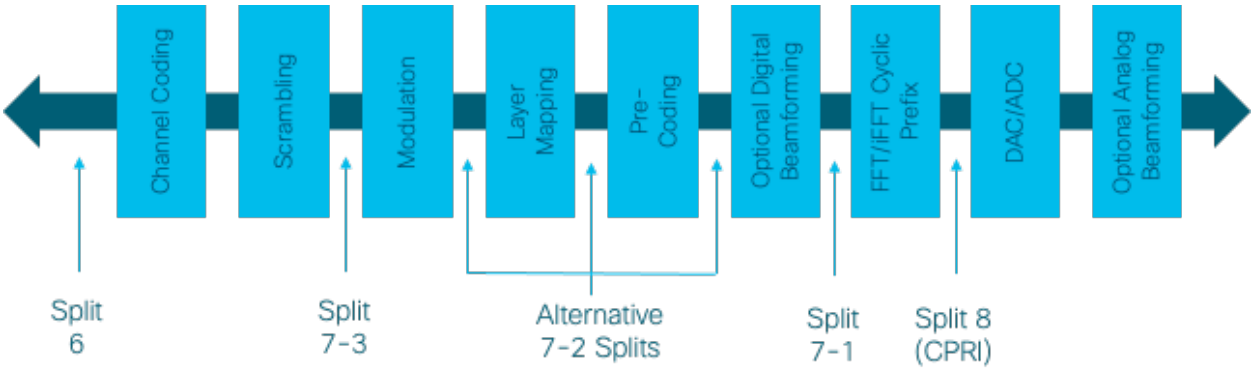


Figure 6 - Alternative Lower Layer Splits

As reported by the Small Cell Forum [2], these alternative split PHY approaches offer benefits in terms of advanced RF combining capabilities. Table 1 is taken from the Small Cell Forum's study into virtualization which highlights that "the lower down in the protocol stack the decomposition occurs, the greater the ability to benefit from the enhanced co-ordination techniques".

However, the alternative approaches to split PHY realizations risk fragmenting the industry in its effort to define a multi-vendor interoperable split PHY approach. This issue was taken up in 2016 by a group of operators and vendors in the xRAN Forum. The xRAN Forum worked on comparing alternative vendor views on split PHY realization and managed to coalesce these into a single approach. Then in 2018, the xRAN Forum announced its merger with the C-RAN Alliance to form a world-wide, carrier-led effort to drive new levels of openness in the radio access network of next-generation wireless systems named the O-RAN Alliance.

Table 1 - Comparing Advanced RF Combining Capabilities of Lower Layer Splits

Advanced RF Combining Capability	PDCP/ RLC	Split MAC	MAC/ PHY	Split PHY
Carrier Aggregation		x	x	x
Cross Carrier Scheduling		x	x	x
Higher order MIMO				x
Downlink Joint Processing- Joint Transmission			x	x
Uplink Joint Reception independent PHY decoding			x	x
Uplink Joint Reception joint equalization PHY decoding				x
Joint Processing – Dynamic point Selection		x	x	x
Co-ordinated Scheduling/Beamforming up and downlink	x	x	x	x

5. O-RAN Alliance's Lower Layer Split

Taking its lead from earlier xRAN work, the O-RAN Alliance published its "7-2x" Split PHY specification in February 2019 [6]. All Split 7 alternatives offer significant benefits over the legacy CPRI/Split 8, which includes avoiding split 8 requirements to scale fronthaul bandwidth on a per antenna basis, as well as introducing transport bandwidth requirements that vary with served traffic in the cell, compared to Split 8 which has a near constant network data rate even when there is no cell traffic. Moreover, when compared to Split 6, the O-RAN lower layer Split 7-2x supports all advanced RF combining techniques, including the higher order multiple-input, multiple-output (MIMO) capability that is viewed as a key enabling technology for 5G deployments.

However, instead of supporting individual transport channels over the nFAPI interface, Split 7-2x defines the transport of frequency domain IQ defined spatial streams or MIMO layers across the lower layer fronthaul interface. The use of frequency domain IQ symbols can lead to a significant increase in fronthaul bandwidth when compared to the original transport channels.

Figure 7 illustrates the bandwidth expansion due to split 7-2 occurring "below" the modulation function, where the original 4 bits "1110" to be transmitted are expanded to over 18 bits after 16-QAM modulation is applied, even when using the block floating point compression scheme defined by O-RAN Alliance.

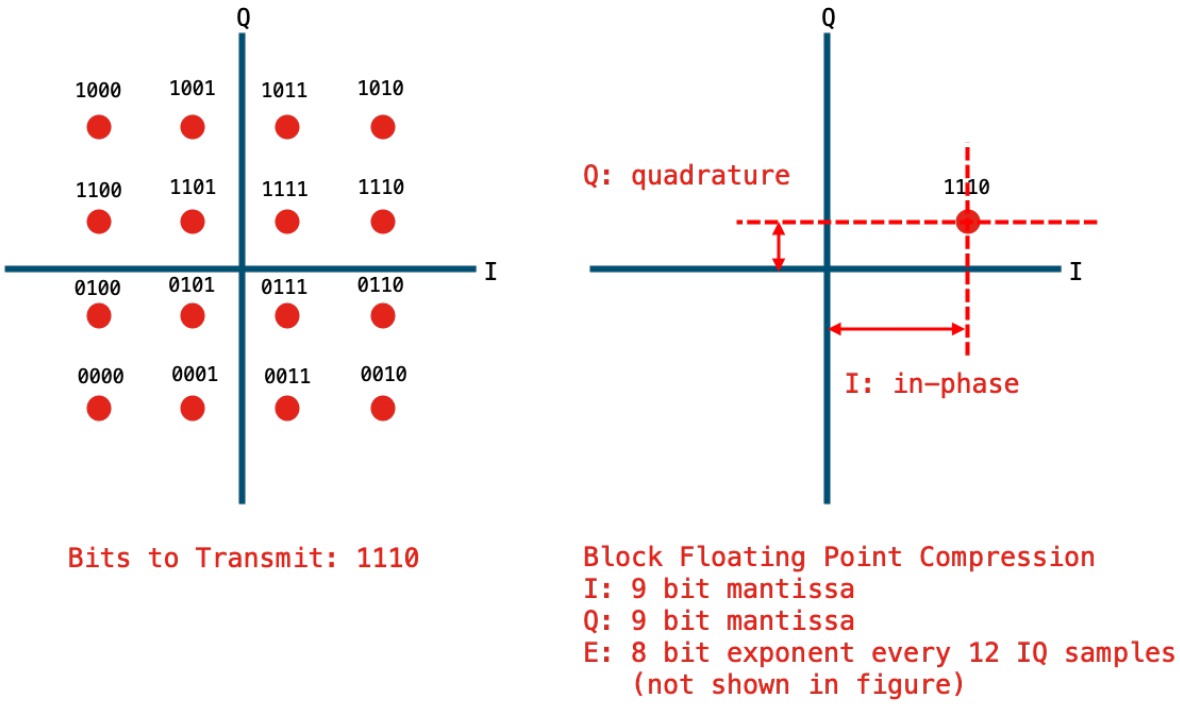


Figure 7 - Bandwidth Expansion with Block Floating Point Compressed Split 7-2x

The bandwidth expansion is a function of the modulation scheme, with higher expansion required for lower order modulation, as shown in Table 2.

Table 2 - Bandwidth Expansion for Split 7-2x with Block Floating Point Compression compared to Split 7-3

Modulation Scheme	Bits to modulate	Block floating point bits	Bandwidth expansion ratio
16 QAM	4	18.67	4.67
256 QAM	8	18.67	2.33

Such a bandwidth expansion was one of the reasons that proponents of the so called Split 7-3 advocated a split that occurred “above” the modulation/demodulation function. To address such issues, and the possible fragmentation of different Split 7 solutions, the O-RAN Alliance lower layer split includes the definition of a technique termed *modulation compression*.

The operation of modulation compression of a 16-QAM modulated waveform is illustrated in Figure 8. The conventional Split 7-2 modulated constellation diagram is shifted to enable the modulation points to lie on a grid that then allows the I and Q components to be represented in binary instead of floating-point numbers. Additional scaling information is required to be signalled across the fronthaul interface to be able to recover the original modulated constellation points in the RU, but this only needs to be sent once per data section.

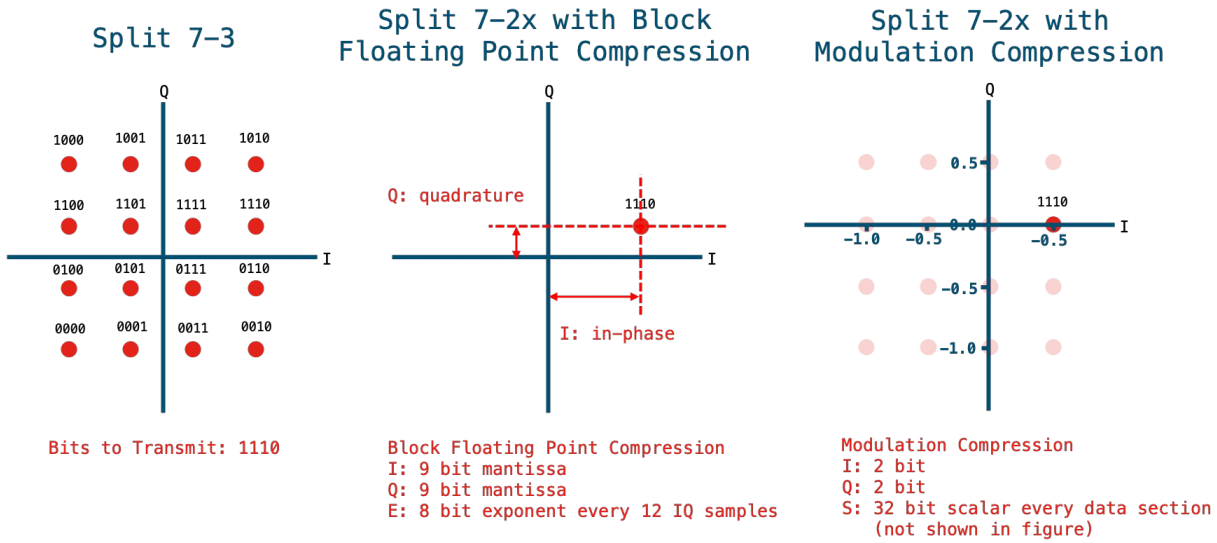


Figure 8 - User Plane Bandwidth Reduction Using Modulation Compression with Split 7-2

Because modulation compression requires the in-phase and quadrature points to be perfectly aligned with the constellation grid, it can only be used in the downlink. However, when used, it decreases the bandwidth expansion ratio of Split 7-2x, where the expansion compared to Split 7-3 is now only due to the additional scaling and constellation shift information. This information is encoded as 4 octets and sent every data section, meaning the bandwidth expansion ratio will vary according to how many Physical Resource Blocks (PRBs) are included in each data section. This value can range from a single PRB up to 255 PRBs, with Table 3 showing the corresponding Split 7-2x bandwidth expansion ratio over Split 7-3 is effectively unity when operating using large data sections.

Table 3 - Bandwidth Expansion for Split 7-2x with Modulation Compression compared to Split 7-3

Modulation Scheme	Bits to modulate	7-2x Block FP Compression BW Expansion Ratio	PRBs per data section	Modulation Compression (ModComp) bits	7-2x ModComp BW Expansion Ratio
16 QAM	4	4.67	1	6.67	1.67
			10	4.27	1.07
			255	4.01	1.00
256 QAM	8	2.33	1	10.67	1.33
			10	8.27	1.03
			255	8.01	1.00

Note, even though modulation compression is only applicable to the downlink (DL), the shift of new frequency allocations to Time Division Duplex (TDD) enables a balancing of effective fronthaul throughput between uplink (UL) and downlink. For example, in LTE, 4 of the 7 possible TDD configurations have more slots allocated to downlink traffic, compared to 2

possible configuration that have more slots allocated in the uplink. Using a typical 12-to-6 DL/UL configuration, with 256-QAM and 10 PRBs per data section, the overall balance of bitrates for modulation compression in the downlink and block floating point compression in the uplink will be (1.03 x 12) to (2.33 x 6), or 12.40:13.98, i.e., resulting in a relatively balanced link as it relates to overall bandwidth.

A more comprehensive analysis by the O-RAN Alliance has examined control and user-plane scaling requirements for Split 7-2x with modulation compression and compared the figures with those for Split 7-3. When taking into account other overheads, this analysis indicated that the difference in downlink bandwidth between Split 7-3 and Split 7-2x with Modulation Compression was estimated to be around 7%. Using such analysis, it is evident why the O-RAN Alliance chose not to define a Split 7-3, instead advocating a converged approach based on Split 7-2x that can be configured to address a variety of lower layer split deployment scenarios.

6. Comparing Split 7-2x and nFAPI

Material from the SCF clearly demonstrates that, in contrast to Split 7, their nFAPI/Split 6 approach is challenged in supporting massive MIMO functionality that is viewed as a key enabling technology for 5G deployments. However, massive MIMO is more applicable to outdoor macro-cellular coverage, where it can be used to handle high mobility and suppress cell-edge interference use cases. Hence, there may be a subset of 5G deployments where massive MIMO support is not required, such as 5G fronthaul over DOCSIS, so let's compare the other attributes.

With both O-RAN's Split 7-2x and SCF's nFAPI lower layer split occurring below the HARQ processing in the MAC layer, both are constrained by exactly the same delay requirements as it relates to LTE HARQ processing and fronthaul transport budgets. Both O-RAN's Split 7-2x and SCF's nFAPI lower layer split permit the fronthaul traffic load to match the served cell traffic, enabling statistical multiplexing of traffic to be used within the fronthaul network. Both O-RAN's Split 7-2x and SCF's nFAPI split support transport using a packet transport network between the RU and the DU.

The managed object for the SCF's PNF includes the ability for a single PNF to support multiple PNF Services. A PNF service can correspond to a cell, meaning that a PNF can be shared between multiple operators, whereby the PNF operator is responsible for provisioning the individual cells. This provides a foundation for implementing Neutral Host. More recently, the O-RAN Alliance's Fronthaul Working Group has approved a work item to enhance the O-RAN lower layer split to support a "shared O-RAN Radio Unit" that can be parented to DUs from different operators, thus facilitating multi-operator deployment.

Both SCF and O-RAN Split 7-2x solutions have been influenced by the Distributed Antenna System (DAS) architectures that are the primary solution for bringing the RAN to indoor locations. The SCF leveraged the approach to DAS management when defining its approach to shared PNF operation. In contrast, O-RAN's Split 7-2x has standardized enhanced "shared cell" functionality where multiple RUs are used in creating a single cell. This effectively uses the

eCPRI based fronthaul to replicate functionality normally associated with digital DAS deployments.

Comparing fronthaul bandwidth requirements, it's evident that the 30-fold bandwidth expansion of CPRI was one of the main reasons for SCF to embark on its nFAPI specification program. However, the above analysis highlights how O-RAN has delivered important capabilities in its Split 7-2x to limit the necessary bandwidth expansion and avoid fragmentation of the lower layer split market between alternative split PHY approaches.

The final aspect when comparing these alternatives is how much the bandwidth is expanded when going from Split 6 to Split 7-2x. Figure 6 illustrates that the bandwidth expansion between Split 6 and Split 7-3 is due to the operation of channel coding. With O-RAN having already estimated that Split 7-3 offers a 7% bandwidth savings compared to Split 7-2x with Modulation Compression, we can use typical channel coding rates to estimate the bandwidth expansion between Split 6 and Split 7-2x.

Table 4 uses typical LTE coding rates for 64-QAM modulation to calculate the bandwidth expansion due to channel coding rate, where the coding rate is the ratio of the useful data transmitted in a subframe to the total amount of data transmitted. This is combined with the additional 7% expansion due to Modulation Compression to estimate the differences in required bandwidth. This table shows that the difference in bandwidth between nFAPI/Split 6 and Split 7-2x is a function of channel coding rate and can be as high as 93% for 64QAM with $^{1}/_{2}$ rate code, and as low as 16% for 64 QAM with an $^{11}/_{12}$ rate code.

Table 4 - Example LTE 64QAM Channel Coding Bandwidth Expansion

Name	Effective Code Rate	Channel Coding BW Expansion	Channel Coding Expansion plus 7%
64-QAM $^{1}/_{2}$	0.554	1.81	1.93
64-QAM $^{3}/_{5}$	0.650	1.54	1.64
64-QAM $^{3}/_{4}$	0.754	1.33	1.42
64-QAM $^{5}/_{6}$	0.852	1.17	1.26
64-QAM $^{11}/_{12}$	0.926	1.08	1.16

Whereas the above analysis indicates that the cost of implementing the Channel Coding above the RU in Split 7-2x is a nominal increase in bandwidth, the benefit to such an approach is the significant simplification of the RU by removing the need to perform channel decoding. Critically, the channel decoder requires highly complex arithmetic and can become the bottleneck in physical layer processing. Often, this results in the use of dedicated hardware accelerators that can add significant complexity and cost to the Split 6 Radio Unit. In contrast, O-RAN's split 7-2x allows the decoding functionality to be centralized, where it is expected that it can benefit from increased utilization and associated efficiencies, while simplifying the design of the O-RAN Radio Unit. A summary of these comparisons is illustrated in Table 5.

Table 5 – Summarizing Differences Between nFAPI and Split 7-2x

Characteristic	nFAPI	Split 7-2x	Comment
Advanced RF Techniques	Supports 6 out of 8 RF techniques	Supports 8 out of 8 RF techniques	Split 7-2x supports higher order MIMO
Round-trip Transport Latency for LTE	Hard limit of 0.5 milliseconds	Hard limit of 0.5 milliseconds	Identical delay constraints as both splits are below HARQ
Round-trip Transport Latency for NR	Soft limit of ~5 milliseconds	Soft limit of ~5 milliseconds	Identical delay constraint
Bandwidth Expansion compared with HLS	Limited bandwidth expansion	~16-93% bandwidth expansion for 64 QAM	Split 7-2x has lower bandwidth expansion for higher modulation rates
MIMO Layer Bandwidth Expansion	None	Bandwidth scales with MIMO layers (Cat-B)	Key delta in bandwidth is due to expansion due to MIMO layers
Statistical Multiplexing in Transport	Yes	Yes	Both splits enable statistical multiplexing
RU Complexity	Similar to composed base station	Removes requirement for channel decoder in RU	Split 7-2x enables RU simplification

7. MIMO Layer Optimization

Both LTE and 5G define the use of MIMO that use multiple transmitting and receiving antennas and exploit multi-path to enable multiple MIMO layers to be supported. In cable, an analogy would be that each port on a fully segmented fiber node is a MIMO layer. The comparison above indicates that there may be limited bandwidth expansion possible with Split 7-2x compared with Split 6 when a single MIMO Layer is being sent over the fronthaul interface. However, where the RAN is configured to support multiple MIMO layers, then this will be the primary parameter that governs the effective bandwidth expansion of Split 6 versus Split 7-2x deployments.

The separate MIMO layers can be used in various configuration, including spatial multiplexing where different layers are used to transmit separate information in order to increase the capacity of the channel, and transmit diversity where different layers are used to transmit the same information in order to enhance the quality of the received signal. Earlier analysis of LTE field trials has compared the transmission modes used in a congested multi-cell environment, contrasting spatial multiplexing, transmit diversity and massive MIMO techniques [7]. These results indicate that the most common multi-antenna technique operated in the network is transmit diversity. Significantly, the O-RAN Split 7-2x supports 3 different transmission schemes:

- Spatial Multiplexing with Cyclic Delay Diversity
- Spatial Multiplexing without Cyclic Delay Diversity
- Transmit Diversity

Specifically, with transmit diversity, the two or four transmit diversity MIMO layers can be packed into a single physical resource block sent over the Split 7-2x interface, meaning that the fronthaul bandwidth is not expanded compared with Split 6, even when transmitting multiple MIMO layers. This is illustrated below for the downlink direction, showing a single set of resource elements (shown in yellow) being used to drive multiple antenna streams (green for antenna 0, blue for antenna 1) and where the individual antenna reference symbols necessary for the operation of MIMO are time multiplexed across the fronthaul interface and then unpacked into the separate streams used to transmit over the respective antennas.

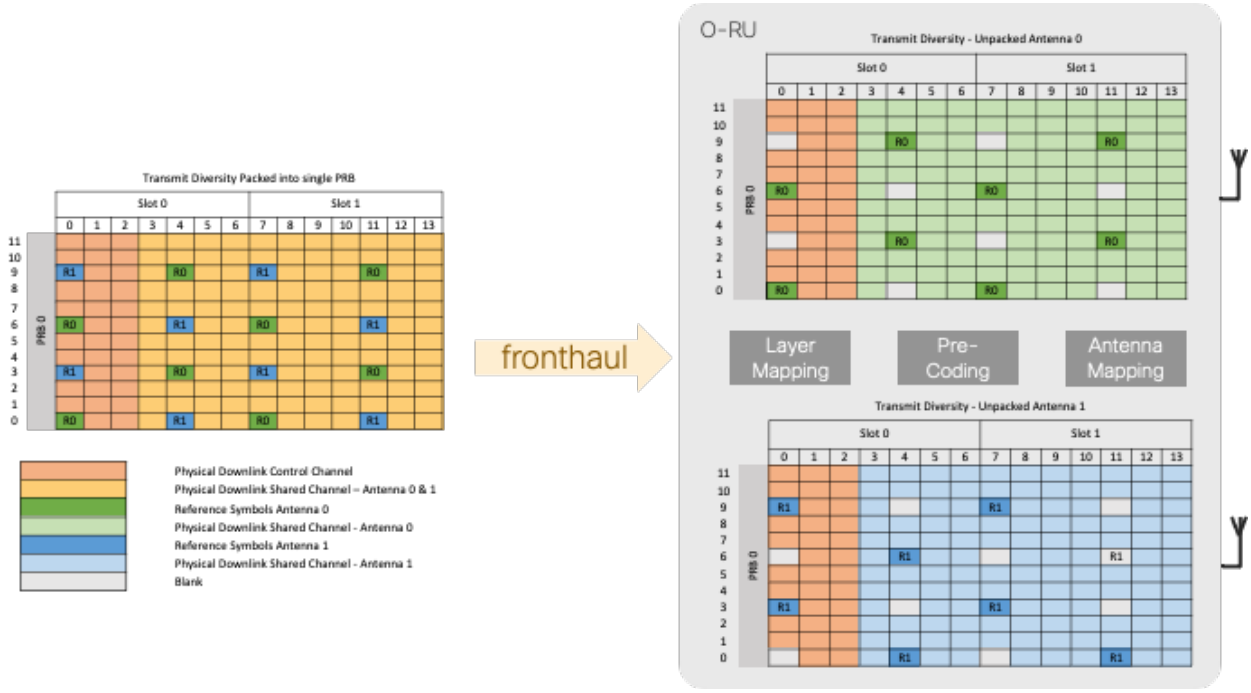


Figure 9 - Packing Multiple Transmit Diversity Layers into a Single Physical Resource Block

Hence, in scenarios where transmit diversity is the most common multi-antenna technique, the options available in O-RAN's split 7-2x avoids any additional fronthaul bandwidth expansion compared to Split 6.

8. Minimizing Transport Delays with Co-operative Scheduling

The above analysis has shown the variety of techniques embedded within the O-RAN fronthaul specification that are aimed at reducing the bandwidth requirements when transporting the multi-vendor interoperable Split 7-2x interface. The final aspect covered in this section deals with the

O-RAN Alliance specification for optimizing fronthaul deployments transported using resource allocation-based transport networks such as DOCSIS and passive optical networking (PON).

Heavily influenced by earlier CableLabs low latency mobile xhaul (LLX) technology [8][9] , O-RAN has specified a “co-operative transport interface” (CTI) targeted at minimizing transport delay requirements through the coordination of resource scheduling between the RAN elements and the transport network. Leveraging the same concepts in CableLabs’ bandwidth report (BWR) concept, the mobile scheduler in O-RAN’s O-DU can signal in real-time information such as fronthaul bandwidth and associated latency or QoS requirements to the transport scheduling entity in the transport node (TN). Here the reported information can be used to expedite resource allocation for the uplink RAN traffic, including adapting the transport control traffic sent to the distributed transport unit (TU) used to support the packetized xhaul traffic.

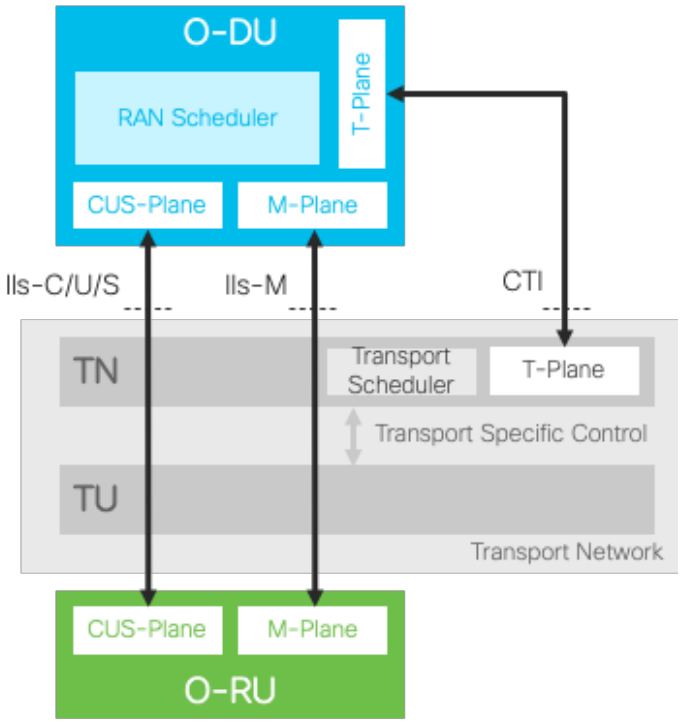


Figure 10 - O-RAN’s Co-operative Transport Interface

Earlier trials of the BWR system have demonstrated the benefits of scheduler co-ordination [10]. Trials have shown that even when channel utilization is high and many users are trying to access the channel, BWR ensures a 1-2 millisecond latency with a higher DOCSIS traffic priority applied to the BWR flow. Significantly, at the 95th percentile, BWR has been demonstrated to reduce DOCSIS upstream latency by almost an order of magnitude, from 22 milliseconds to 2.5 milliseconds.

With 5G new radio (NR) avoiding the strict sub-millisecond latency constraints associated with LTE’s synchronous uplink HARQ procedures, the specification of CTI by the O-RAN alliance,

coupled with the above delay measurement results, indicate that DOCSIS will be able to support the transport delay requirements associated with O-RAN's 5G fronthaul interface.

9. Scaling For Fronthaul Bandwidth

Network densification is a key driver for enabling 5G. This will see conventional cell tower sites be upgraded with 5G capability, but also a raft of new sites being deployed. These new sites will likely be targeted at delivering "hotspot" capacity to meet consumers' ever-increasing demand for more mobile broadband, or targeted coverage for supporting new vertical value chains which 5G aims to support. The 5G radio market can then be viewed as split into conventional "on-tower" based and newer "off-tower" based deployments.

The on-tower radios will be attempting to deliver increased capacity across the tower's coverage area through the use of MIMO. These MIMO systems could range from 8x8 MIMO up to 64x64 massive MIMO systems. Due to their larger radius, there will also be more UEs per radio which means the average traffic rate will be higher. The 5G on-tower market will almost always use a fiber-based fronthaul transport, although some instances do have radio backhaul.

The off-tower market is the emerging small cell market. The small cell will have a smaller radius with a less dense MIMO (2x2 or 4x4) and with fewer UEs, which means the average traffic rate will be lower. These small cells may be pole/building mounted, strand mounted, or indoor mounted. This is the market for HFC connectivity using the fiber or coax side of the hybrid fiber-coax (HFC) plant. The coax side of the HFC plant would use the DOCSIS protocol. This market segment is somewhat analogous to Wi-Fi with a potentially larger radius.

In order to understand the impact of fronthaul on the off-tower market, we first define the expected bandwidth model of the RAN system that will be backhauled over DOCSIS. Then we compare the bandwidth load to the service level of a cable modem.

With so many variations in how to configure an LTE or 5G system, it is further complicated by different approaches of how to configure the fronthaul. We will focus our attention on the split 7-2x Modulation Compression profiles approach that leverage the bandwidth saving capabilities described above.

Moreover, we will re-use a TDD test profile that looks to be applicable to the majority of new spectrum being allocated to 5G, using the so called "DDDSU" frame structure configuration. In this TDD frame configuration, D represents a slot configured for downlink operation, U represents a slot configured for uplink operation, and S represents a special slot that includes a number of symbols for downlink, a number of symbols for uplink and a number of symbols for a guard period between the downlink and uplink symbols. In an example the 14 symbol special slot can be configured as 10:2:2 (D:G:U) where G is guard time. So, this is equivalent to a time multiplexing ratio of approximately 4:1 for DL:UL.

An example LTE profile would be for 2x20MHz radio that uses 2x2 MIMO, for example corresponding to a CBRS deployment that uses 4 x 10 MHz licenses. Using Modulation Compression in the downlink and block floating point with 9-bit mantissa in the uplink, our calculations are that when considering control plane and transport overheads, a peak of 330

Mbit/s of throughput will be required on the downlink and a peak of 320 Mbit/s of throughput will be required on the uplink. Note, because of the unequal fronthaul compression techniques applied to the UL and DL, the 4:1 time multiplex results in a roughly symmetrical fronthaul bandwidth being transported over Ethernet.

An example 5G NR profile using the same 40 MHz of spectrum would be for a 40 MHz radio that uses 4x4 MIMO. Using Modulation Compression in the downlink and block floating point with 9-bit mantissa in the uplink, our calculations are that when considering control plane and transport overheads, a peak of 950 Mbit/s of throughput will be required on the downlink and a peak of 690 Mbit/s of throughput will be required on the uplink.

Whereas these rates represent the peak uplink and downlink speeds, it is recognized that real-world deployments operate with a non-uniform spatial distribution of traffic such that not all radios will be simultaneously operating at their peak capacity. The SCF analyzed such a phenomenon as part of their nFAPI virtualization deliverables [12]. While the exact peak-to-mean spatial distribution across a RAN will be a function of the deployment, e.g., including the use case being address (e.g., residential/enterprise/urban), the SCF report that a value of a peak-to-mean ratio of 3.5-to-4.0 across a 200 radio node "off-tower" network can be used to help dimension the virtualized RAN.

The bandwidth of the DOCSIS plant is well documented in [13] and results are show in Figure 11 with 96 channels (6 MHz per channel) of MPEG video, and in Figure 12 with no MPEG video. Figure 11 represents a typical case today on the HFC plant while Figure 12 represents a point in 3 to 5 years when MPEG video has been retired from the HFC plant and video services are all video over IP over DOCSIS.

The bandwidth numbers in the table represent the peak bandwidth of the DOCSIS spectrum. Each cable modem (CM) will be provisioned with some value less than this. The Distributed Access Architecture (DAA) nodes are described by the number of unique DOCSIS ports in the downstream (DS) and upstream (US) direction. A 2x4 DAA node has two unique DOCSIS DS (equivalent to a downlink) and four unique DOCSIS US (equivalent to an uplink) ports.

2 x 4 Node Capacity	User	DOCSIS 3.1 with video					DOCSIS 4.0 ESD with video					
Scenario	Calc	1	2	3	4	5	6	7	8	9	10	11
DS End MHz	1002	1002	1218	1002	1218	1218	1794	1794	1794	1794	1794	1794
DS Start MHz	54	54	54	108	108	258	108	258	372	492	606	834
US Rtn Path MHz	42	42	42	85	85	204	85	204	300	396	492	684
VOD/SDV MPEG-TS	0	32	32	32	32	32	32	32	32	32	32	32
Linear Video MPEG-TS	0	64	64	64	64	64	64	64	64	64	64	64
DOCSIS DS port Gbps	8.7	2.9	5.1	2.6	4.8	3.3	10.5	9.0	7.9	6.7	5.6	3.3
DOCSIS US port Gbps	0.10	0.10	0.10	0.47	0.47	1.48	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	17.3	10.7	15.0	10.1	14.4	11.4	25.8	22.8	20.6	18.2	15.9	11.4
Ethernet US Gbps	0.4	0.4	0.4	1.9	1.9	5.9	1.9	5.9	8.4	11.7	15	22

2	DS ports per Node	4096	OFDM Modulation	24	DS MHz skipped below 108 MHz	32	VOD/SDV MPEG-TS
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band	64	Linear Video MPEG-TS
24	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1794	D4.0 Stop Frequency (MHz)	ESD	D4.0 FDX or ESD
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)	120	FDX Trans Band (MHz)

Figure 11 - DOCSIS Bandwith for a 2x4 DAA Node with Video

2 x 4 Node Capacity	User	DOCSIS 3.1 with no video					DOCSIS 4.0 ESD with no video					
Scenario	Calc	1	2	3	4	5	6	7	8	9	10	11
DS End MHz	1002	1002	1218	1002	1218	1218	1794	1794	1794	1794	1794	1794
DS Start MHz	54	474	54	108	108	258	108	258	372	492	606	834
US Rtn Path MHz	42	42	42	85	85	204	85	204	300	396	492	684
VOD/SDV MPEG-TS	0	0	0	0	0	0	0	0	0	0	0	0
Linear Video MPEG-TS	0	0	0	0	0	0	0	0	0	0	0	0
DOCSIS DS port Gbps	8.7	4.7	10.8	8.4	10.5	9.0	16.2	14.8	13.6	12.4	11.3	9.0
DOCSIS US port Gbps	0.10	0.10	0.10	0.47	0.47	1.48	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	17.3	9.4	21.6	16.7	21.0	18.0	32.5	29.5	27.2	24.8	22.6	18.0
Ethernet US Gbps	0.4	0.4	0.4	1.9	1.9	5.9	1.9	5.9	8.4	11.7	15	22

2	DS ports per Node	4096	OFDM Modulation	24	DS MHz skipped below 108 MHz	0	VOD/SDV MPEG-TS
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band	0	Linear Video MPEG-TS
24	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1794	D4.0 Stop Frequency (MHz)	ESD	D4.0 FDX or ESD
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)	120	FDX Trans Band (MHz)

Figure 12 - DOCSIS Bandwidth for a 2x4 DAA Node

The DOCSIS plant, as are all networks, is built with over-subscription in mind. In Figure 11, scenario 1, the DOCSIS plant supports 3 Gbps DS and 100 Mbps US. This is a typical deployment in 2021. As a typical example, this bandwidth may support 200 CMs with highest provisioned rate of 1 Gbps DS and 20 Mbps US for about 10% to 20% of the CMs, while the majority of the CMs are 200 Mbps x 10 Mbps and 100 Mbps x 5 Mbps. This defines an over-subscription case that recognizes that not all CMs will transmit or receive at the same time.

As discussed before, the small cell will not use all four of its MIMO channels simultaneously all the time. Nor will it use its full spectral capacity all the time. In fact, it depends upon the reach of the small cell and how many subscribers it connects to. Since the CM is part of a large service group (SG) of say 200 homes, the reach of the small cell and any other small cells on that node, will have the same geographical footprint. An analysis of the number of small cells per fiber node and small cells per macro-cell can be found in [14].

So, at a first pass, if a DOCSIS SG shares the same region as a set of small cells, and hence the same customers and same traffic load, then their traffic patterns will be similar. In this scenario, think of Wi-Fi connected laptops to a DOCSIS network, mixed with cell phones that are either Wi-Fi to DOCSIS connected or mobile connected.

At a second pass, the small cell fronthaul downstream bandwidth (950 Mbps) matches the max CM bandwidth (1 Gbps) and thus fits. The small cell fronthaul upstream (650 Mbps) does not match the upstream bandwidth of scenario 1 (100 Mbps). Instead, the DOCSIS upstream spectrum will have to increase from 42 MHz return (100 Mbps) to 204 MHz return (1.48 Gbps). This assumes a common spectrum starting point of 16.4 MHz.

The tables show that the DOCSIS DS bandwidth can be increased dramatically to about 10 Gbps with DOCSIS 3.1 by removing the legacy MPEG video. DOCSIS 4.0 with extended spectrum DOCSIS (ESD) can take the DS limit further as well as increase the upstream bandwidth.

Another issue that arises in support small cells, especially if they are in the home, is the support of IEEE 1588 timing. The DOCSIS Time Protocol (DTP) provides this service and is defined in [15] with further support in [16][17][18][19][20].

10. Conclusion

Fronthaul is perceived as being synonymous with fiber-based transport systems; extreme bandwidth requirements and sub-millisecond latency requirements may cause many to reach the conclusion that DOCSIS is only suitable for transporting higher-layer splits. However, in this paper we have described the enhanced capabilities in O-RAN's Split 7-2x lower-layer split architecture that are targeted at minimizing the impact on transport networking requirements.

Standardized techniques have been described that enable the bandwidth expansion of a single fronthaul stream to be reduced to low percentages when compared to alternative higher-layer split alternatives. Moreover, spatial stream optimization techniques have been described that enable a single transport stream to drive certain MIMO antenna configurations. The hard delay requirements for LTE fronthaul are deprecated in favour of soft delay requirements for 5G New Radio, with delay constraints now measured in milliseconds instead of microseconds.

Taking the lead from earlier CableLabs BWR concepts, O-RAN has fully specified a co-operative transport interface to link the RAN and transport schedulers, an approach that has already demonstrated low millisecond transport latencies and, at the 95th percentile, a reduction in DOCSIS upstream latency of almost an order of magnitude, from 22 milliseconds to 2.5 milliseconds.

Finally, we use example LTE and 5G profiles to set the parameters used in fronthaul bandwidth calculation and compare them with existing and evolving DOCSIS bandwidths. It was clear that the DOCSIS downstream had ample bandwidth, but the HFC plant needs to be upgraded to 204 MHz to allow for small cell fronthaul.

Abbreviations

ACK	acknowledgement
ADC	analog to digital conversion
BWR	bandwidth report
CB	coordinated beamforming
CM	cable modem
CMTS	cable modem termination system
CPRI	Common Public Radio Interface
CS	coordinated scheduling
CTI	co-operative transport interface
DAC	digital to analog conversion
DAS	distributed antenna system
DL	downlink
DOCSIS	data over cable service interface specifications
DPS	dynamic point selection
DS	downstream
DTP	DOCSIS Time Protocol
FAPI	functional application platform interface
FFT	fast Fourier transform
HARQ	hybrid automatic repeat request
HLS	higher layer split
iFFT	inverse fast Fourier transform
IQ	in-phase and quadrature
JR	joint reception
JT	joint transmission
LLS	lower layer split
LTE	long term evolution
MAC	medium access control
MIMO	multiple-input multiple-output
MPEG	Moving Picture Experts Group
NAK	negative acknowledgement
nFAPI	networked functional application platform interface
NMM	network monitor mode
NR	new radio
OAM	operations and maintenance
PDCCH	physical downlink control channel
PDCP	packet data convergence protocol
PNF	physical network function
PON	passive optical networking
PRACH	physical random-access channel
PRB	physical resource block
QAM	quadrature amplitude modulation
RACH	random-access channel
RAN	radio access network
RAR	random-access response
RE	radio equipment

REC	radio equipment control
RF	radio frequency
RLC	radio link control
RRC	radio resource control
RU	radio unit
SCF	Small Cell Forum
SG	service group
TDD	time division duplex
UE	user equipment
UL	uplink
US	upstream
VNF	virtual network function

Bibliography & References

[1] *Study on new radio access technology: Radio access architecture and interfaces*, 3GPP 38.801, https://www.3gpp.org/DynaReport/38801.htm

[2] *Virtualization for small cells: Overview*, Small Cable Television Laboratories, Inc., "Remote PHY Specifications for DAA", CM-SP-R-PHY, CableLabs. [link]

[3] Cell Forum, SCF106, https://scf.io/en/documents/106__Virtualization_for_small_cells_Overview.php

[4] *LTE physical layer: Performance analysis and evaluation*, Applied Computing and Informatics Volume 15, Issue 1, January 2019, Pages 34-44, https://www.sciencedirect.com/science/article/pii/S2210832717301990

[5] *Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification*, 3GPP 36.321, https://www.3gpp.org/DynaReport/36321.htm

[6] *O-RAN Fronthaul Control, User and Synchronization Plane Specification*, O-RAN Alliance, https://www.o-ran.org/specifications

[7] *Performance Evaluation of a Live Multi-Site LTE Network*, IEEE Access, Volume 6, 2018, https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8453797

[8] *Low Latency Mobile Xhaul over DOCSIS® Technology*, Data Over Cable Service Interface Specification Mobile Applications, CM-SP-LLX-I01-190628 [link]

[9] John T. Chapman, Jennifer Andreoli-Fang, *Low Latency Techniques for Mobile Backhaul over DOCSIS*, SCTE Cable-Tec Expo Fall Technical Forum, Denver, October, 2017. [link]

[10] Jennifer Andreoli-Fang, John T Chapman, Tong Liu, Damian Poltz, *Blueprint for Mobile Xhaul over DOCSIS*, SCTE Cable-Tec Expo Fall Technical Forum, Sep, 2019. [link]

[11] *O-RAN Fronthaul Interoperability Test Specification (IOT)*, O-RAN Alliance, https://www.o-ran.org/specifications

[12] *Network Aspects of Virtualized Small*, Small Cell Forum, SCF-161, https://scf.io

[13] John T. Chapman, "The Path to 100 Gbps DAA Nodes", *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, Oct, 2021. [link]

[14] John T. Chapman, "Small Cell Traffic Engineering," *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, Oct, 2020. [link]

[15] Cable Television Laboratories, Inc., "Synchronization Techniques for DOCSIS Technology Specification," CM-SP-SYNC, CableLabs. [link]

[16] Roy Sun, John T. Chapman, et. al., "DOCSIS Time Protocol Proof of Concept", *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, Oct, 2021. [link]

[17] Roy Sun, John T. Chapman, Rahil Gandotra, "Designing a Cloud-based DOCSIS Time Protocol Calibration Database", *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, Oct, 2021. [link]

[18] Elias Chavarria Reyes, John T. Chapman, "How the DOCSIS Time Protocol makes the SYNC Specification Tick," *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, Oct, 2020. [link]

[19] Jennifer Andreoli-Fang, John T. Chapman, "Mobile Backhaul Synchronization Architecture," *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, October, 2017. [link]

[20] John T. Chapman, Rakesh Chopra, Laurent Montini., "The DOCSIS Timing Protocol (DTP), Generating precision timing services from a DOCSIS system," *INTX/SCTE Spring Technical Forum*, 2011. [link]

5G Security & Protection Framework

Vasu Dalal
Director, Product Management
NOKIA
vasu.dalal@nokia.com

Patrick Nta
Chief Security Architect (Consulting)
NOKIA
patrick.nta@nokia.com

1 Abstract

Cable companies offer multiple services – TV, broadband (cable, fiber, ethernet), voice, business services, home security and many others. And Cable companies now offer mobile service. They have purchased CBRS spectrum and are developing 5G service offerings for a Quad-Service play.

Mobile is an entirely new technology area for the industry and the service must be brought online in double-quick time to meet market demand, monetize CBRS spectrum purchases, and meet business commitments. At the same time, they must invest in the mobile network which is evolving to 5G. Operators must also be in position to offer these three (3) major 5G communications use cases: Ultra Reliable and Low Latency (URLLC), Massive Machine Type (mMTC), and Enhanced Mobile Broadband (eMBB).

5G is complex:

- New User Endpoints (UEs in 5G speak)
- New radios (CBRS), DOCSIS/PON backhaul and/or hybrid RAN (own & partner)
- An SDN/NFV-based Core, both centralized & distributed and/or hybrid Core
- Edge clouds and cloud-based environments (mix of bare-metal, VNF, & CNF deployments)
- A diverse network with a multitude of vendors & customers
- New, unique use cases including but not limited to Fixed Wireless Access (FWA), network "slicing", autonomous vehicles, IoT etc.
- Exposure to developer APIs (for even more app use cases) and many others
- Subscriber needs are changing, and new experiences are being created at a rapid pace

Security is critical in such a diverse, evolving, and complex 5G network while at the same time these services must be bought online faster to market with limited budgets and strained resources.

Traditional Enterprise-based security solutions will not be adequate based on the scope of the challenge, the size, diversity and scale of the network and the numerous new, unique & evolving use cases.

A 5G-based security solution requires not only adherence to traditional IT concepts of availability, integrity & confidentiality but also provide:

- A centralized, multi-vendor, end-to-end network control & management ("single pane of glass")
- Be adaptive, self-learning (AI/ML) with real-time threat updates
- Highly-scalable to support millions of diverse (UE & network) elements
- Customizable and automated (auto-discovery, audit & auto-remediation) to handle the volume of threats & scale and extent of the 5G network

2 5G: Key Security aspects

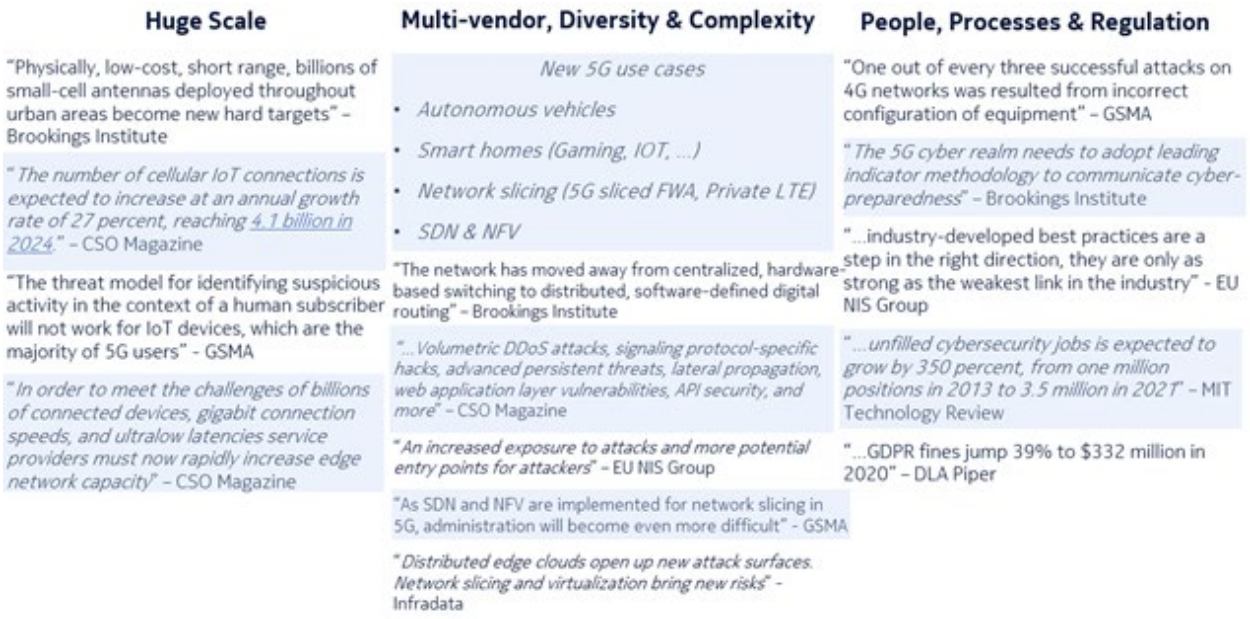


Figure 1 – 5G: Key security aspects

The key 5G security aspects can be summarized in three broad categories

1. Stringent requirements (latency, reliability & security) at high scale
2. Multi-vendor, Diversity & Complexity
3. People, Processes & Regulations

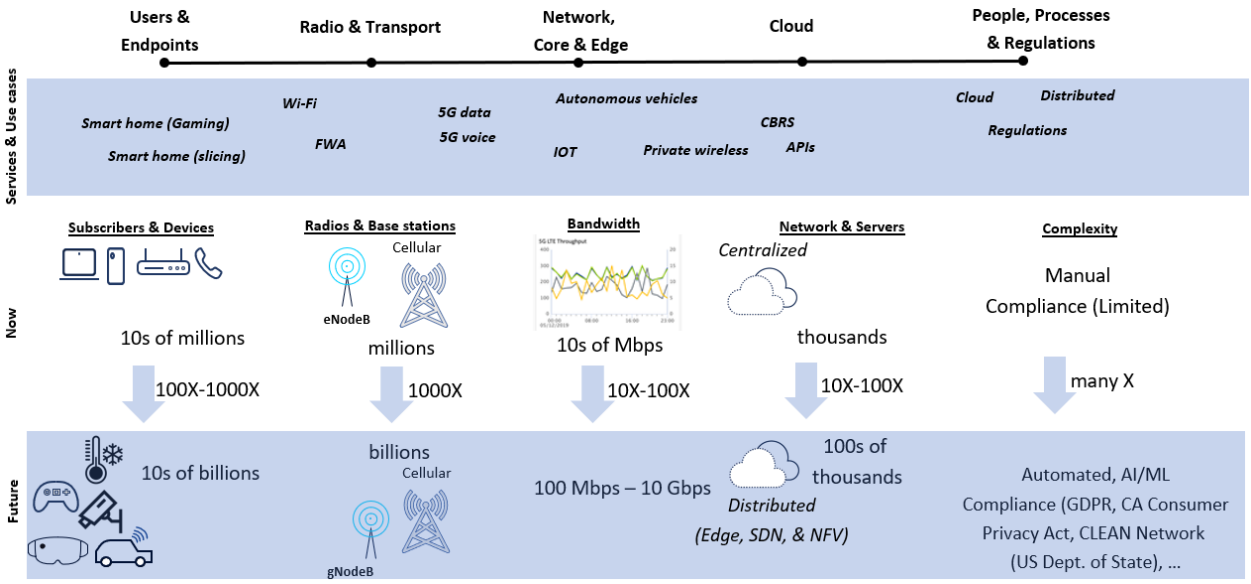


Figure 2 – Scale challenges at a glance

2.1 Stringent requirements (latency, reliability & security) at high scale

Table 1 - Speeds, Latency, Reliability requirements & Security implications

	Use-Case	DL	UL	Network Latency	Reliability	Cost Sensitivity	Security
Consumers	Mobile Broadband	100-300M	10-50M	15-25ms	Medium	Medium	Medium
	Fixed Wireless Access	1-5G	100-200M	1-20ms	High	High	Medium
	Event experience	1-100M	1-5G	1-5ms	Medium	Medium	Medium
	In-vehicle Infotainment	5-100M	1k-1M	1-20ms	Medium	Medium	Medium
Industries	Critical automation	1M	1-10M	1-5ms	Very high	Low	Very High
	Tele-operation	1M	1-10M	1-25ms	Very high	Low	Very-High
	Highly interactive AR	5-100M	1-100M	1-10ms	High	Medium	High
	Mass sensor arrays	1k-1M	1k-1M	200-500ms	Low	Very High	Medium-High

CBRS & 5G network requirements are in order(s) of magnitude higher in terms of # of UEs, # of radios, a software-based Edge, Core & Cloud network compared to prior networks (fixed or mobile). This compounded with the scale and new use cases leads to stringent security implications.

2.2 Multi-vendor, Diversity & Complexity

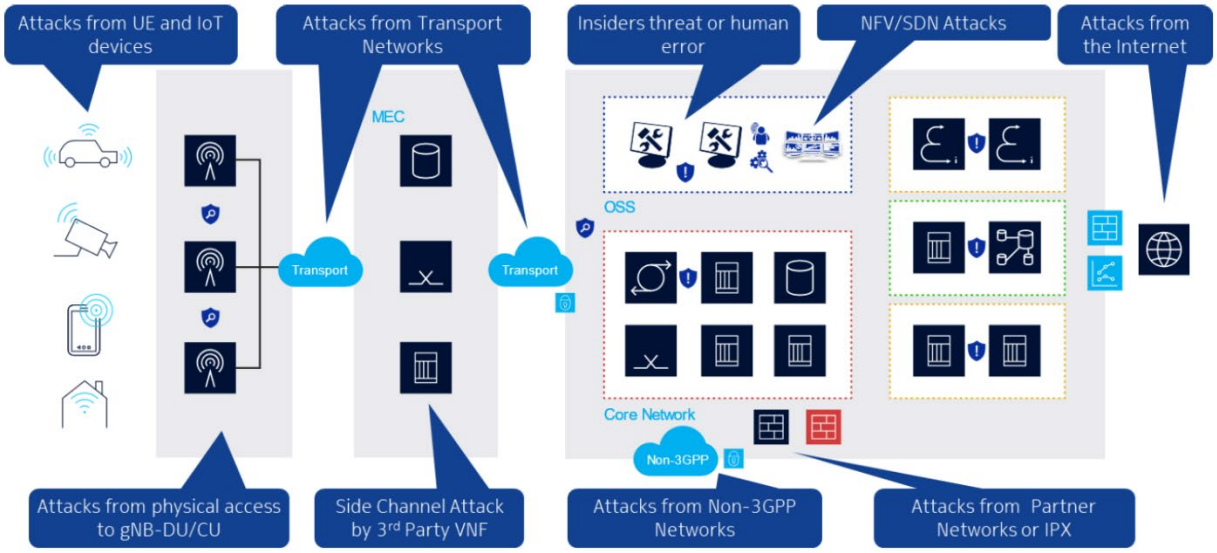


Figure 3 – Multi-vendor, diversity & complexity

Rather than having single monolithic providers controlling everything from the infrastructure up to the service layer, there are multiple stakeholders involved. The cloud, mixed edge/cloud and hybrid cloud environments have changed the notion of a perimeter. With 5G, there are many more players involved in the delivery of a service, with a much more diverse set of roles, and different understanding of risks.

Disaggregated network with lots of software solution components, new, & complex use cases (network "slicing", augmented reality, V2X, IoT & APIs).

Specific attack threat vectors for 5G networks include, but are not limited to:

- Users, Devices & Endpoints
    - Protection against eavesdropping, DOS, traffic injection, & rogue gNB attacks
    - Many times, attacks may occur without the owner of the device even being aware of it. It could be triggered by malware that has infected the device. Botnets are among the biggest threats. For example, large sets of infected devices that are controlled by an attacker and used to carry out large scale attacks, such as distributed denial of service attacks (DDoS). Such attacks can happen in 4G, too. But in 5G, we assume a different level of magnitude with higher speeds and larger device numbers. Many of them will be cheap and poorly managed IoT devices, which may easily become part of a botnet, possibly due to missing security patches, for example

- UE interaction is complex with DSDS-controlled handoff between CBRS <-> 5G (MNO) <-> 5G (MNO roaming)

- CBRS/5G Radio, Network and Transport
  - Network distribution & 5G services increasing the overall attack surface (DU, CU, (v)CMTS, OLT, IWF, MEC, Edge, Core, Cloud, IoT)
  - Transport technology is a mix of DOCSIS & PON backhaul for the CRBS radio sites where traffic will be backhauled over broadband (BB) networks to the Hub (or Headend). And the distributed user plane with IWF function at the interconnect point to the 5G Core network
  - Attacks may also derive from transport networks, as in 4G. Base stations are physically exposed, and therefore particularly endangered. That the base station may be split into a central unit (CU) and several distributed units (DU) is an infrastructure feature that is specific to 5G. In this case, it is mostly the DUs that will be physically exposed, but the network interconnecting CU and DUs is also at risk
- Edge, Cloud & Infrastructure Security

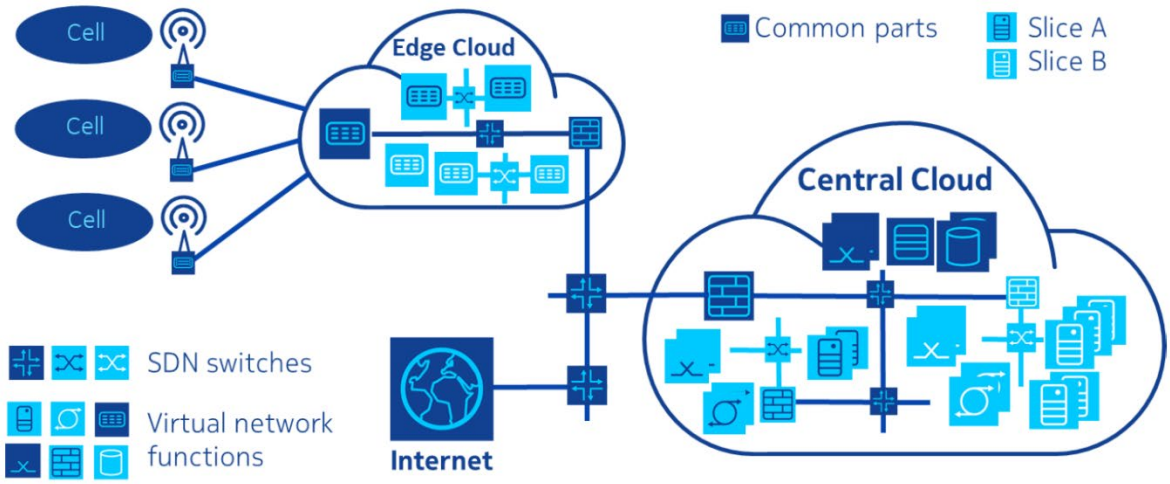


Figure 4 – Edge, cloud & infrastructure

  - 5G networks will adopt new networking paradigms. Network Function Virtualization (NFV) with both CNFs & VNFs and Software Defined Networking (SDN) will make networks much more dynamic. Cloud-centric networking is characterized by massive-scale, software-driven infrastructure and continuously shifting traffic flows, bandwidth demands, and network topologies. New attack vectors come up due to the use of Network Function Virtualization and Software Defined Networking. In particular, the sharing of infrastructure may allow so-called ‘side channel attack’. When two different applications share common hardware, for example a CPU, there is always a risk that information may leak from one application to another. Other side channels may be opened by flaws in the virtualization layer. For example, a hypervisor may have a flaw that allows one virtual machine running on said hypervisor to access the memory of another virtual machine

running on the same hypervisor. Such side channel attacks have the potential to break the isolation between different applications or slices, and they can be very subtle and hard to detect.

If only trusted software under the control of one organization (e.g. the network operator) runs in a cloud, the risk of side channel attacks may be low. However, in 5G low-latency scenarios, third party software, such as AR/VR applications, may need to be deployed on the same edge computing infrastructure as the operator software. In this case, the side channel attack threat becomes a very real one.

- Scale
  - Discrete physical devices replaced by multiple VNFs & CNFs
  - Number of network entities under configuration management increasing dramatically
  - Expansion in the # of VNFs, CNFs & NEs requiring authenticated communications
  - Dramatic increase in volume of security information & alerts generated across the 5G network
  - Dramatic increase in the number & type of user accesses to the 5G virtualized infrastructure
  - Large increase in the amount & type of logs & data generated across the distributed network

- New, complex use cases (network “slicing”, IoT & APIs)
  - Given that more and more parts of the overall solution are accomplished in software and the "web speed" need for CI/CD, means even more attack vectors. Security must be designed in.

2.3 People, Processes & Regulations

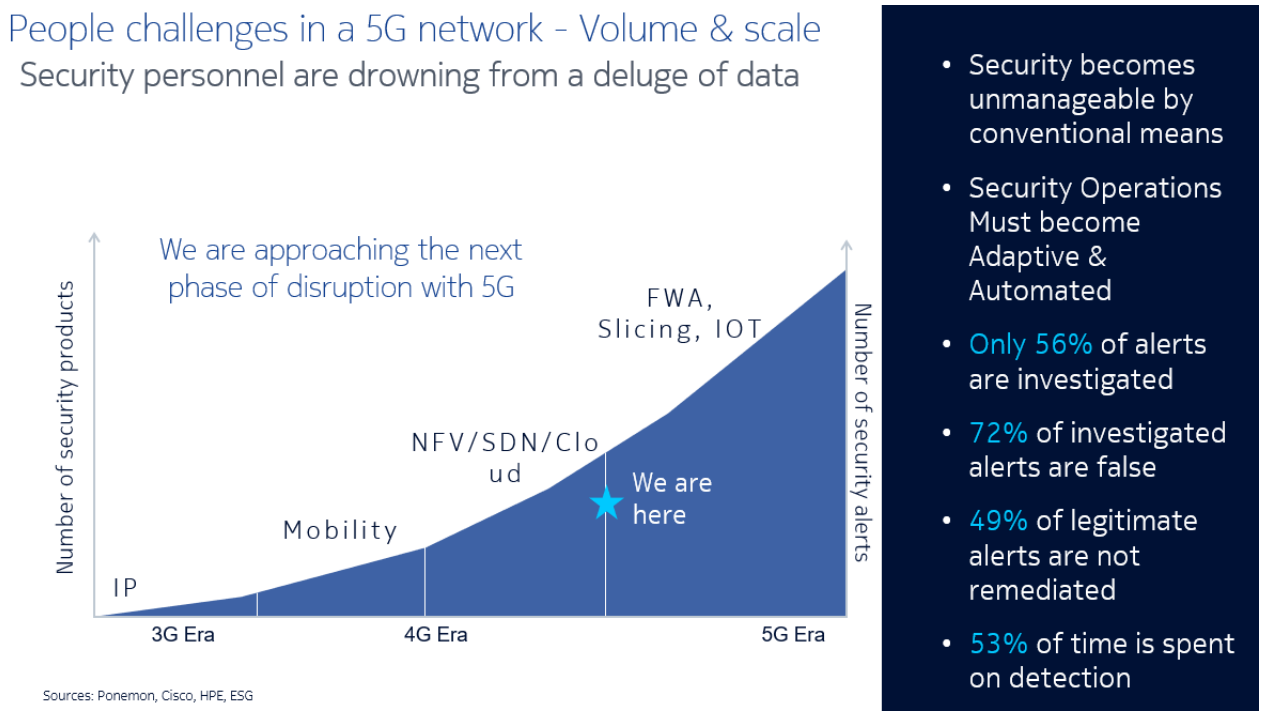


Figure 5 – People, processes & regulations

There is always the threat of security breaches by human errors or by malicious insiders. This is so in all networks, but considering mission critical 5G services, the impact of insider attacks may be even more devastating than in earlier mobile network generations.

- 65% of cyberattacks exploit configuration-related vulnerabilities
- 62% of causes of downtime are configuration errors (user error)

As the security threat landscape of mobile network is evolving very fast, it creates a lot of concerns on industries and governments and drives need to impose stricter security regulation on critical information infrastructures (CII) including mobile networks.

Many countries have passed cybersecurity or privacy laws which have important implications on the design, implementation, and operation of mobile networks. Examples include:

- EU issued a report on 5G risk assessment in Q3 2019 which identifies the main threats and threat actors, the most sensitive assets, the main vulnerabilities, and strategic risks Mitigating measures will be announced end 2019 to address the identified cybersecurity risks at national and union level
- General Data Protection Regulation (GDPR) in EU is passed in April 2016 and implemented in May 2018. GDPR has global significance as it governs not only EU companies but also all companies that process EU resident's data

- In France, law has been issued in August 2019 to preserve the interests of defense and national security in connection with the operation of mobile radio networks (5G and further) It modifies the Posts and Electronic Communications Code to introduce an authorization request for the operation of radio network equipment (BTS and core).
- China Internet Security Law enacted in November 2016 and in force in June 2017.
- Canada's Personal Information and Electronic Documents Act (PIPEDA)
- In USA, if government project is involved, contractors need to comply with Federal Information Security Management Act (FISMA) and National Institute of Standards and Technology (NIST) requirement FIPS 200 and SP 800-53
- California Consumer Privacy Act (CCPA)
- Clean Network (US Department of State) and many others

Finally, there is a geo-political and nationalistic aspect to 5G, and supply-chain provenance must be evaluated in depth.

Before 5G, mobile network was primarily focusing on voice and internet services. Moving toward 5G, mobile network operators will unavoidably get into other business segments such as banking, energy, healthcare, public safety or even military. Network operators must be ready to comply with all industry specific and government regulatory security requirements.

3 5G Security Framework

3.1 Vision

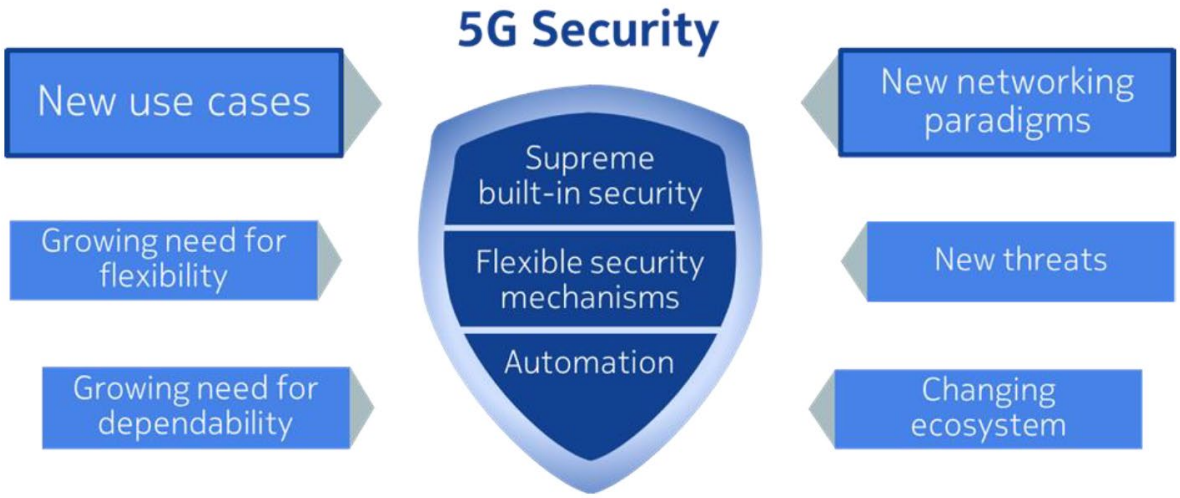


Figure 6 – 5G Security framework

Mobile networks are moving into a post-perimeter world where the network boundaries have disappeared, and the difference between insiders and outsiders has been eradicated. Given this situation there is pressing need to gain total visibility and intelligence on what is happening within service provider infrastructures, services, applications, data and people, to detect security breaches and respond to them.

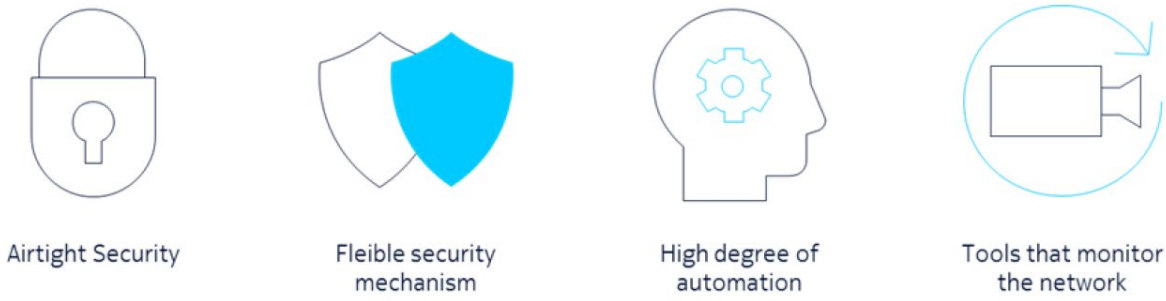


Figure 7 – 5G Security vision

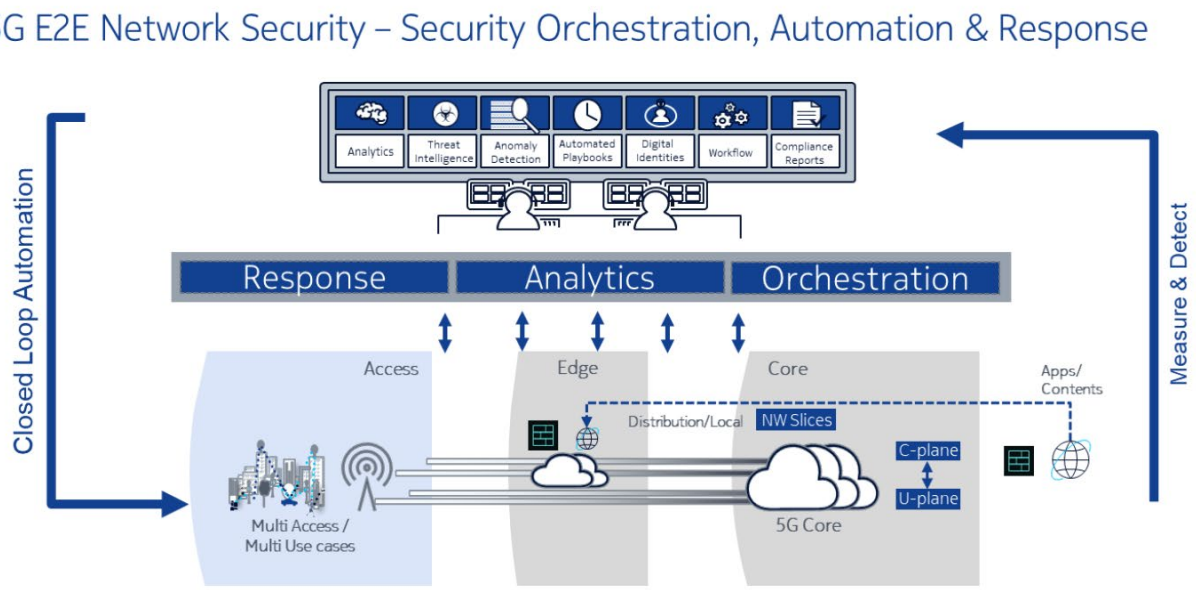


Figure 8 - SOAR

A Security Orchestration, Automation & Response (SOAR) strategy provides a centralized security command-and-control structure to protect the network from the most advanced persistent attack with the fusion of threat intelligence, analytics, machine learning & automated response.

Security must limit attack vectors, detect all threats when they do occur and respond faster to eliminate time between detection and mitigation.

Some of the key themes include:

- Constantly measure your security posture and risk levels. This is more than compliance auditing, but rather an ongoing near real time assessment of your network security posture. To do this effectively requires automated software security systems
- Control and limit access to key operational systems and assets. In addition to perimeter security defenses, this includes access governance and management
- Detecting threats earlier in the kill chain requires an ability to perform multi-dimensional analytics across a variety of systems and resources in order to identify threats that may be otherwise missed. The goal is to identify anomalies from normal behavior, and this is where data analytics and machine learning (ML) for security are emerging. Analytics and machine learning (ML) are needed to spot indicators of compromise, proactively identify harmful actors, help security analysts prioritize risk and initiates the appropriate rapid response

Rapid response is key to minimize the impact of cyber-attacks. The time between detection and mitigation needs to be eliminated. One of the big challenges security teams face is the inability to keep pace with the diversity and velocity of threats.

Combined with a global cyber skillset shortage, traditional incident response strategies rely on too many manual processes performed by limited security expert resources. This is where security process automation or orchestration plays a key role.

Adaptive security is about transforming security operations to be predictive & automated by using machine learning and multi-dimensional analytics and threat intelligence in order to drive rapid, automated, and predictive responses to threats.

3.2 Security Orchestration Analytics and Response (SOAR)

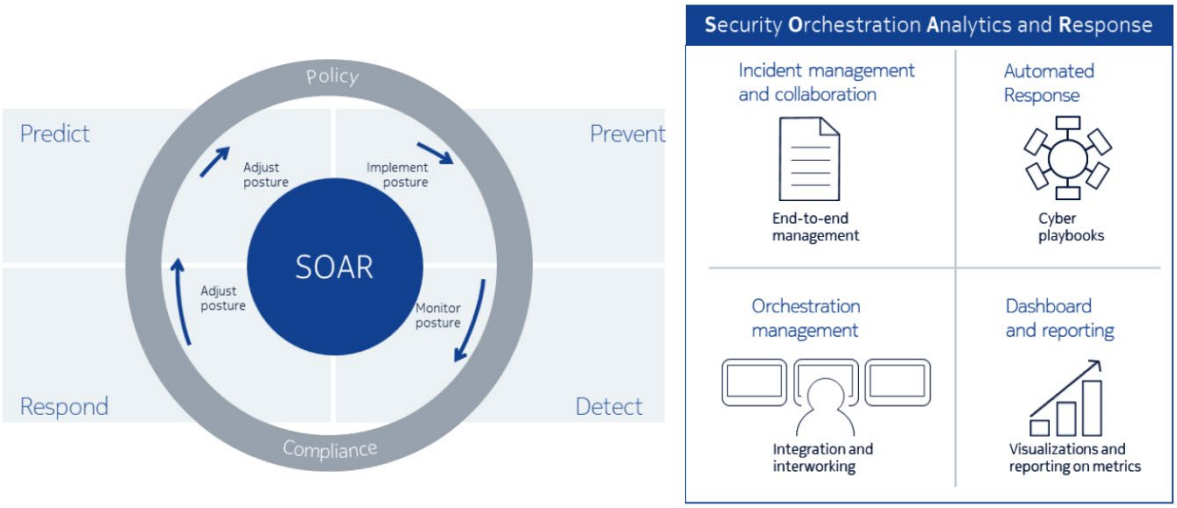


Figure 9 – SOAR in action

An "optimized" state cannot be achieved using conventional approaches to security operations. The priority for any digital strategy is to build an adaptive security architecture that automates security driven by intelligence and analytics.

These are the basic principles of Security Orchestration, Analytics and Response (SOAR): using security analytics in order to drive an orchestrated automated response.

SOAR systems aggregate, correlate, and analyze data from disparate point tools into cohesive and enriched security intelligence with business-specific context.

By analyzing user behavior to identify bad actors and providing threat indicators to potential insider threats. These capabilities help security professionals prioritize risks and automate security operations activities in the context of the attack surface and business and improves alert management by correlating and consolidating alerts from existing systems.

Security operations workflow automation and orchestration are at the heart of the transition from static defense to agile and adaptive response. Security automation involves more than just operations; it must be aware of and encode business processes, regulations, and customer-specific policies. Automation is the

process executing repeatable actions without human intervention while orchestration is the concept chaining these automated tasks into executed playbooks to perform workflows to accelerate both investigation and mitigation.

3.3 5G Security Architecture

3.3.1 Defense in Depth

Below is how the 5G architecture is depicted by 3GPP in its Technical Specification TS 23.501.

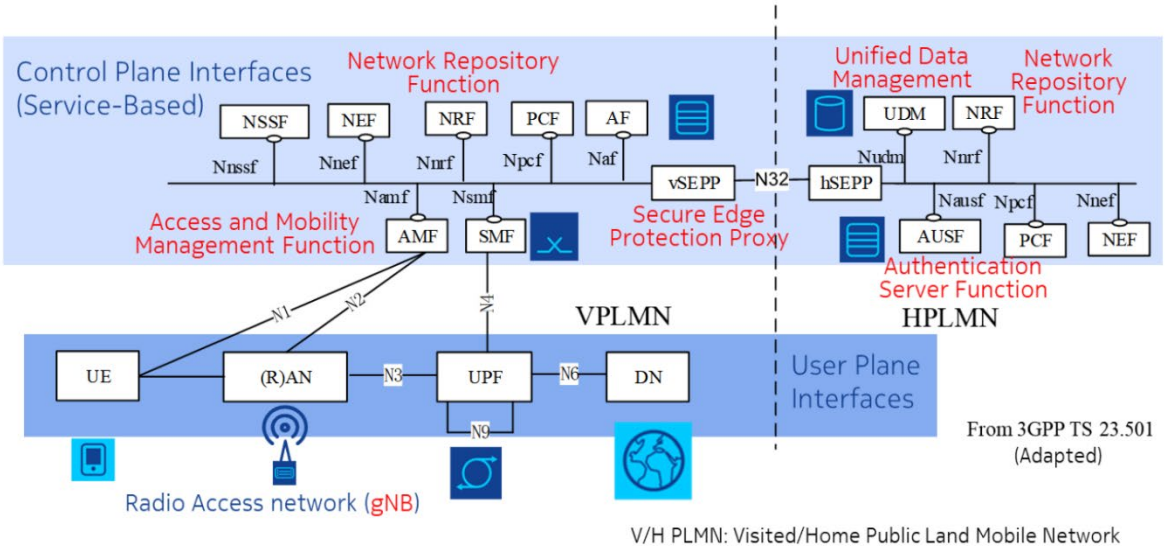


Figure 10 – 3GPP specifications

The basic concept of 5G Mobile Network Security is known as “Defense in Depth”. This describes multiple layers of overlapping security measures protecting the valuable assets, preventing from potential attack, and causing impact to the important asset of the network.

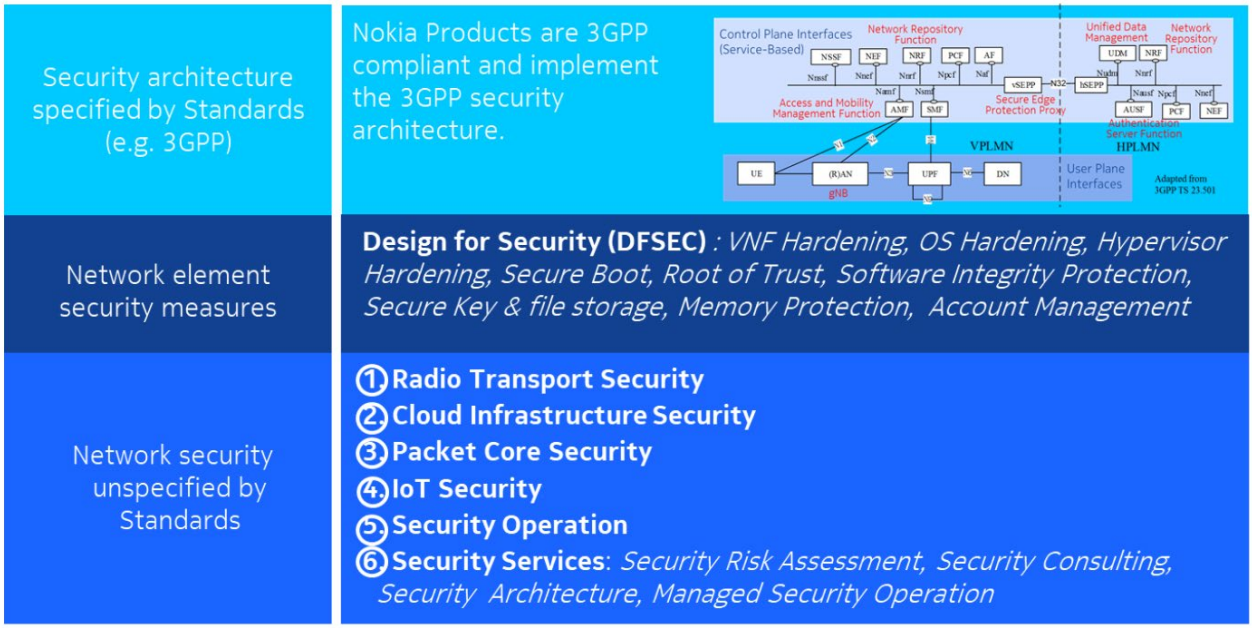


Figure 11 – Defense in Depth

The 1st layer is the 3GPP specified security architecture. Security features are implemented in most 5G solutions and products, supporting 3GPP-recommended security architecture.

The 2nd layer is the hardening of the Virtual Network Function & cloud infrastructure and is vendor and network dependent. Vendors require well-defined security management processes built into their Design & Development. This ensures all products in the portfolio are implementing a baseline set of security features and hardening according to best practice within the industry. Examples include secure key and file storage, secure boot, root-of-trust, account management, and Software Integrity Protection.

The 3rd layer is also vendor and operator dependent. Vendors must identify gaps that are not covered by 3GPP and standard VNF & CNF hardening steps and fill the gaps by offering comprehensive security solutions and services.

The diagram below illustrates the essential elements of a security architecture for a 5G network implemented on distributed service provider networks:

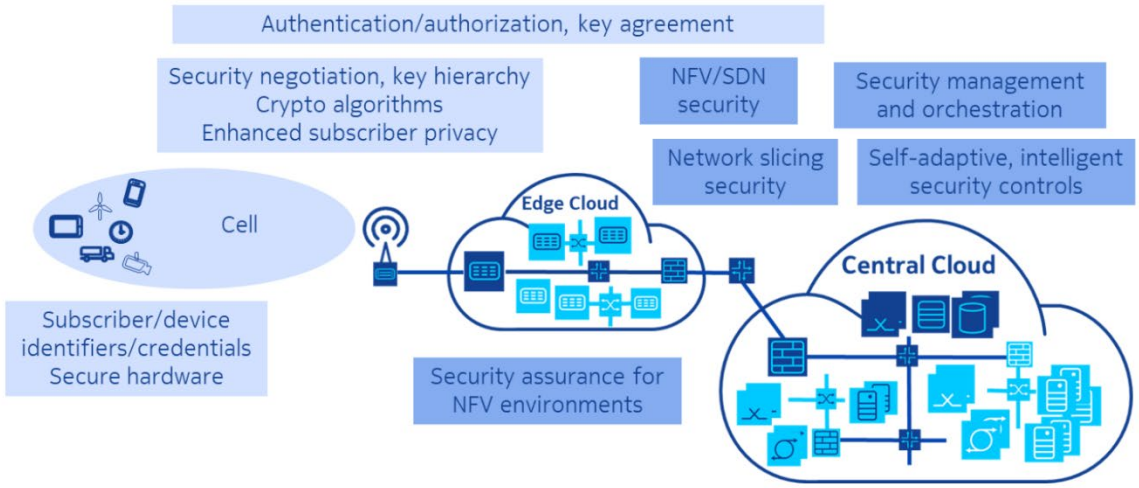


Figure 12 – Edge, infrastructure & cloud security

The key security areas are split into Radio Transport Security, Network & Packet Core Security, Cloud Infrastructure Security, Network Slicing Security, Security Operations, Design for Security (DFSEC) and last but not the least are the security professional services which help operators to understand their existing security risks, to design, implement and operate their network the most secure way.

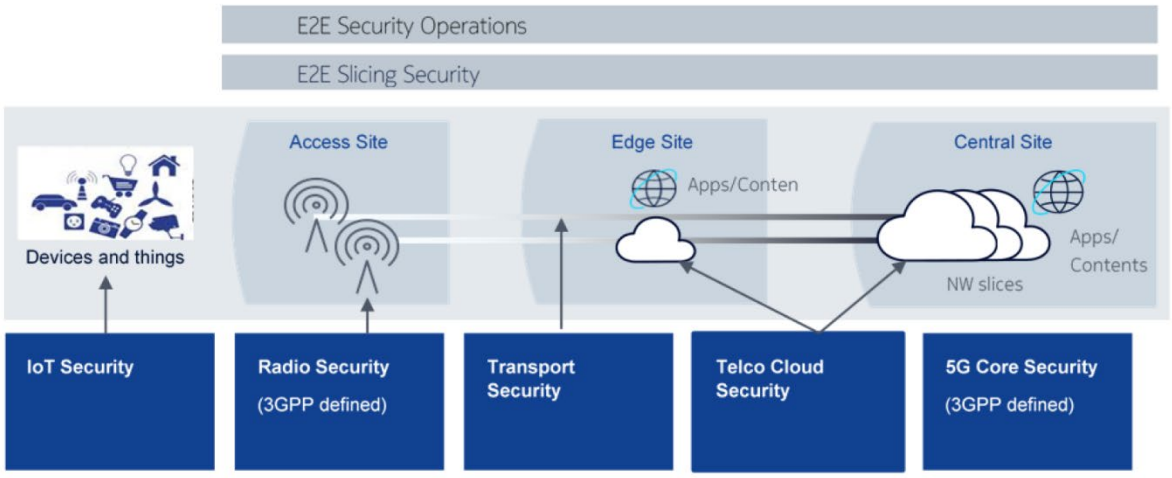


Figure 13 – 5G security overview

3.4 Key defense building blocks

Defense is the deployment of security controls in layers to eliminate or mitigate against threats. Defense includes:

- Layered protections – security layers complement one another, such that what one layer misses, another layer will catch
- Defense in multiple places – security defenses are pervasively located in different places within the network
- Defense through diversification – when possible, using different security controls will limit the effect that a fault or a vulnerability in one part of the network will have on the rest of network

3.4.1 Identity

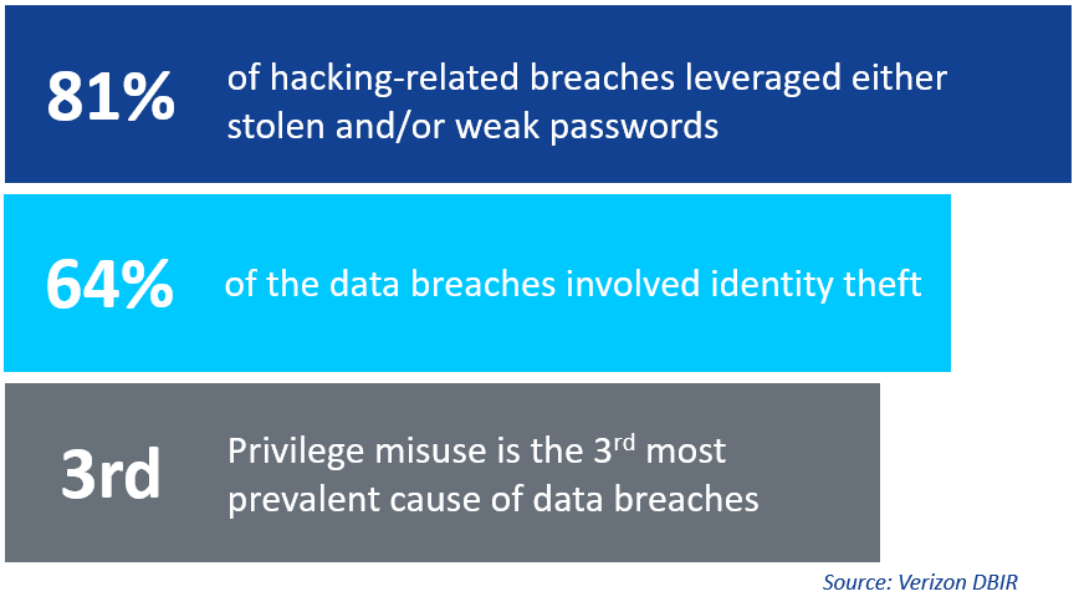


Figure 14 - Identity

Identity is the basis of any sound defensive posture. One must know who a threat is (and who isn't) with certainty, quickly and robustly at scale.

Two primary functions required to maintain identity in the end-to-end network:

- Certificate issuance
- Certificate deployment & Lifecycle Management

These tasks are performed by a Certificate Authority (CA) which must be:

- Centralized

- Issues certificates to NFs for NF-NF mutual authentication
- Certificates enable TLS and IPSec for secure NF-NF communications
- Highly scalable – support millions of certificates

3.4.2 Abstraction (or Zoning)

Abstraction is the classification of objects into security classes or groups and assigning security controls, rights, permissions and privileges to these classes or groups. An object in a 5G network could be any network element, application, device or asset which is part of the infrastructure providing the mobile service.

The classification of objects in mobile operator networks should be made based on their criticality to the business and their level of exposure to external threats.

Examples of abstraction in a mobile operator network include division of the network into security zones and the assigning of security controls to mitigate threats specific to each zone. Abstraction also involves the methods used to control how access to different domains of the network is controlled for administrators and network operations personnel, along with controls which are applied to ensure security of the network element configurations.

3.4.3 Zero Trust

As 5G networks are being exposed to a large array of threats, classification of objects based on a "Zero Trust" concept is required.

In response to the NIST RFI for Developing a Framework to Improve Critical Infrastructure Cybersecurity (RFI # 130208119-3119-01), Forrester Research introduced the security concept of "Zero Trust"

For years in information security most security concepts were designed based on a model of a "hard shell and soft core". Alternatively called the "castles-and-moat" framework. This model is based on threats always coming from outside the network, requiring a "hard shell", while permitting loose controls on the inside of the network "soft core", where free access to systems is generally employed. This concept has become outdated for many reasons, including:

- The distributed & hybrid nature of the network with edge, core & cloud locations
- The model does not consider or expect internal threats, where "trust" is implicitly granted. However, the reality of the last few years is that internal threats are a significant attack vector into networks, including not only malicious employees, but also human error and identity spoofing. Once in the internal network, the attacker has a free reign to penetrate other systems with almost no risk of detection. And the attacker can stay in the network for extended periods of time, sometime even years
- The model does not account for machine-to-machine/IoT communication, as well as complex automation processes, which cannot be trusted by default just because they are "internal"

The Zero Trust model is straightforward:

- Zero Trust implies that there is no such thing as a pre-granted trust status – nothing and no one is trusted – not even inside your network. This applies to everything in the network including

traffic, machines, devices, and people. Zero Trust mandates that since nothing is trusted, controls must be implemented on a per service per request level to protect the business. This is unlike a host or IP-based security mechanism where trust is wholesaled to anyone with a hostname/IP and thus to any application running on that device.

Zero Trust requires that security professionals protect internal data from insider abuse in the same manner they protect external data on the public Internet, following three main principles:

1. All resources must be accessed securely, regardless of their location (logical or physical) in the network
2. Access control and least privileges are key security mechanisms to employ across the entire network
3. The security framework for the network is built based on all available "if-then" scenarios. Logs are gathered, making sure all traffic is inspected, to cover future "what-if" scenarios

Zero Trust also means that security controls are not built on top of each other, but rather controls are built from the inside out, starting with each system and network element, and using that as the basis for building security across the complete network.

3.4.4 Data Hiding

Data hiding is the concept of revealing to any subject only the minimum level of information the subject requires to perform their task. System hardening standards must also be employed to ensure that only the required minimal number of services for a specific implementation are running, thereby reducing the attack vector open to malicious traffic, and minimizing the potential for widespread disruption of service.

3.4.5 Encryption

Information in a mobile operator's network is a combination of data in transit, as well as data at rest. Some of this data is sensitive subscriber data, while other data is critical information needed to configure and manage the network. To protect against attack vectors attempting to exploit this data, encryption is used to control sensitive and critical information while the information is traversing the network (in-transit) or is stored (at rest).

3.5 UEs, Radio & Transport Security

In every mobile network, there is a radio interface. This interface is inherently exposed to attacks and must therefore be secured carefully. Traffic over the radio interface is already encrypted as publicly known, but radio air interface security must go beyond encryption.

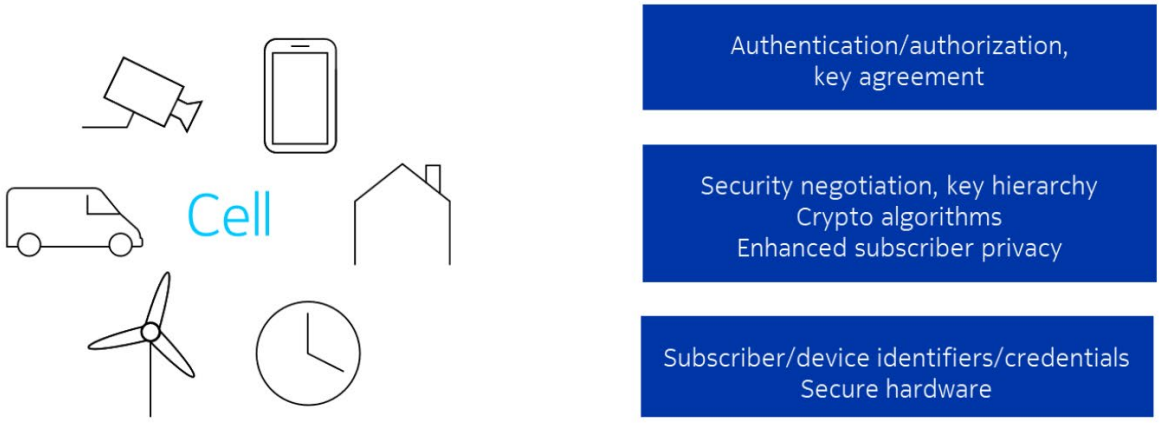


Figure 15 – UE security & privacy protection

First, UEs (mobiles) must be authenticated and authorized to use the network or specific services. The authentication relies on means to identify subscribers or devices, and on credentials that should be stored on the devices securely, making use of specific secure hardware, such as the SIM card. The SIM card is technically called UICC, or Universal Integrated Circuit Card, on which a USIM (Universal Subscriber Identity Module) is implemented. During authentication, a key is agreed upon, from which a hierarchy of session keys is derived to secure the subsequent communication. How security is applied must be negotiated, for example, determining which type of traffic will be secured and by which crypto algorithms this will be achieved.

“Enhanced subscriber privacy” refers to the fact that in earlier network generations, an attacker can trick mobiles into revealing the true identity of the subscription, a practice known as “IMSI catching”, one that is applied not only by attackers but also by law enforcement. In this mechanism, the subscription identifier is never passed in the clear over the air but encrypted. Protection against this kind of attack is considered a requirement for 5G. All falls under the scope of 3GPP.

The attacks on the confidentiality and integrity of the traffic can be mitigated by state-of-the-art cryptography. This has been standardized by 3GPP for many interfaces.

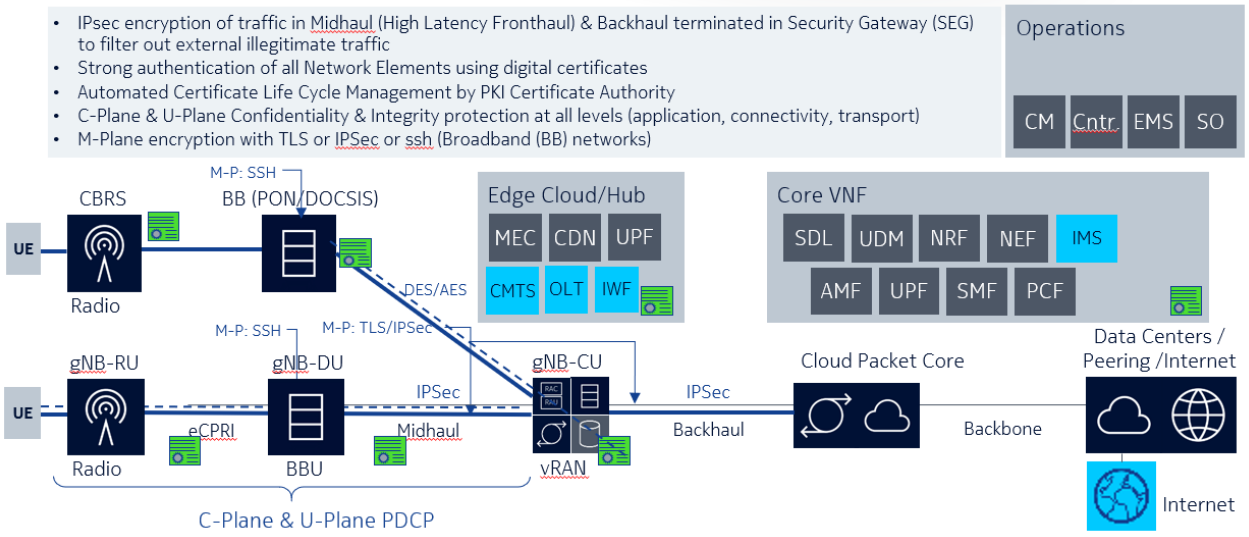


Figure 16 – UE, radio & transport security

Next-generation Node B (gNB), especially CBRS radios, gNB-CU & gNB-DU are located in unsecured locations. Hackers can easily attack the core network from any unsecured DU or CU through the transport network interface. In order to protect the edge cloud and cloud core data centers from illegitimate traffic, the best practice, as recommended by 3GPP, is to encrypt the traffic between gNBs and the core network using IPSec. The concept to encrypt all traffic from DU or CU is not only to ensure confidentiality and privacy of user traffic but more important to ensure all traffic entering the core network is not tainted by hackers in any way. The 3GPP recommendation to filter out unwanted traffic is to encrypt the legitimate traffic with IPSec and authenticate with asymmetrical keys using with Public Key Infrastructure (PKI).

Key solution components for this include:

- Security Gateway (SeGW) devices with GTP firewall and IPsec capabilities
- the centralized security gateway management system and
- PKI, or Public Key Infrastructure. Public Key Infrastructure consists of the Certificate Manager, which is served as the PKI certificate authority. A strong authentication of gNBs is provided with PKI certificates
- Automated Certificate Life Cycle Management is also required to be provided by the Certificate Manager/Authority

3.6 Network & Packet Core Security

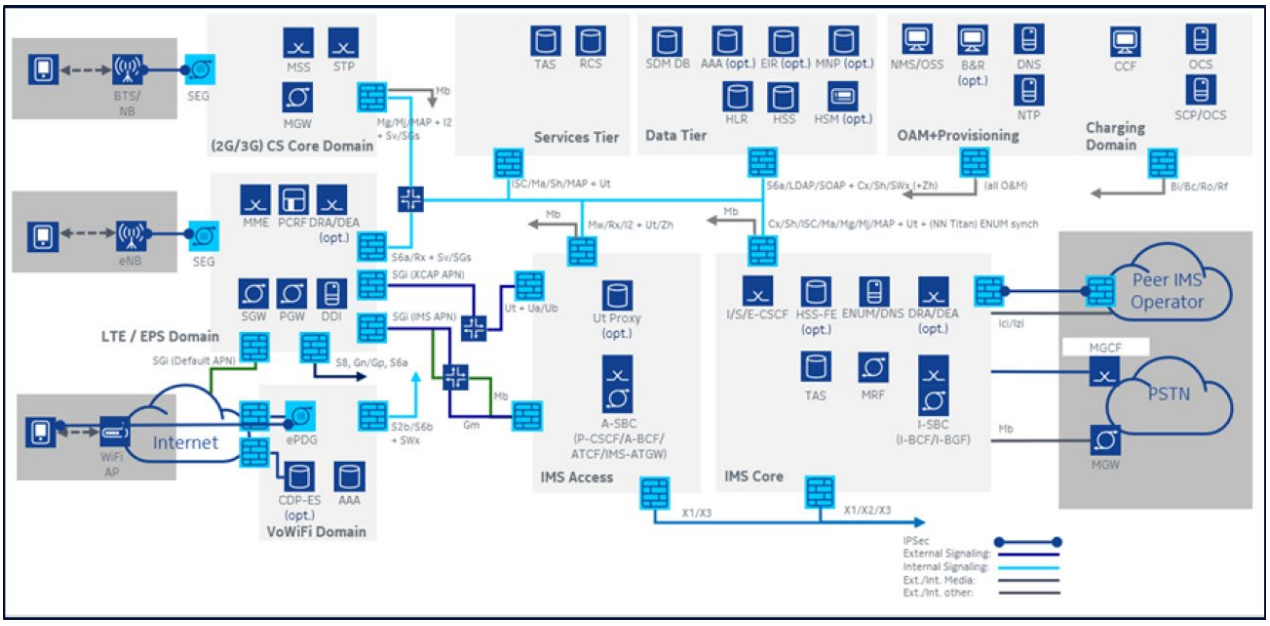


Figure 17 – Network & Packet Core security

A comprehensive security architecture meeting stringent data and roaming security standards includes:

• E2E traffic separation and zoning are mandatory to isolate high exposure equipment from high value assets. In case one security zone is compromised, the exposure can be easily contained before the high value asset is also compromised. Security Zone must be implemented from the beginning or retrofit will be difficult and costly

• Virtualized security appliances provide isolations between security zones or domains

• GTP, SCTP firewall to protect the eNB or gNB interface from radio access network

• Diameter firewall is required to protect the DIAMETER roaming interface

• Physical or virtualized firewall with Intrusion Detection System and Intrusion Protection System (IDS/IPS) is required at DN or SGi interface to internet

• Secure DNS to protect against infiltration of network via DNS

• Protect network from DDoS volumetric attacks

3.7 Cloud Infrastructure, NFV & SDN Security

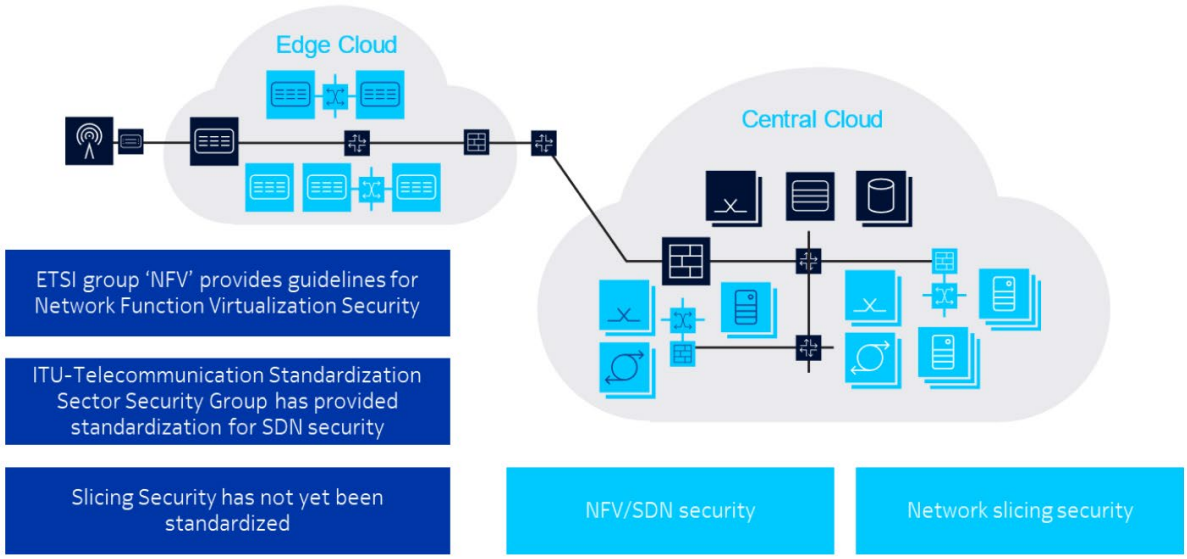


Figure 18 – Cloud, NFV & SDN security

Cloud infrastructure using NFV and SDN technology is crucial in 5G as it enables the flexibility and the elasticity needed for the diverse use cases.

All 5G networks adopt the new networking paradigms Network Function Virtualization and Software Defined Networking, as well as supporting slicing. In this section, we do not cover specific NFV and SDN security measures but only the threat of side channel attacks in NFV environments.

One can conclude that it is important to design and implement the shared cloud platform with a high degree of care, in a way that keeps the residual, exploitable vulnerabilities to a minimum, whilst still being prepared to patch the system quickly, in case one of these residual vulnerabilities are detected. SDN and NFV security are not specified by 3GPP. It is the security group of ETSI that provides guidelines and recommendations for Network Function Virtualization security.

To secure a network implemented in an NFV environment, the following security measures are recommended:

• Secure implementations of the virtualization layer and the overall cloud platform software. The focus here is about implementing a robust hypervisor and good isolation of traffic data. Examples include:

- Root of Trust: software integrity protection, secure boot, & vendor certificate
- Security hardening including zoning, segmentation & traffic filtering
- Security management: Certificate management, malware protection, identity management, & image signing
- Security orchestration: automation of security policies, breach remediation, monitoring & API security
- Robust security implementation of the VNFs & CNFs

- Good logical separation of VNFs provided by the virtualization layer. It is possible to have a physical separation of VNFs, but this comes at a cost and less flexibility
- Traffic separation by dedicated virtual switches, VLANs and wide-area VPNs
- Perimeter security and network internal traffic filtering by virtual firewalls
- Logically or even physically separated security zones
- Secure operation and maintenance, secure operation of IP services (e.g. DNS)
- Cryptographic protection of traffic and of data on storage

To ensure the SDN is secure, the SDN controller must be secured. Implement important measures, such as:

- Cryptographic protection
- Authentication and authorization
- Robust implementation of overload control

Microservices architecture introduces different requirements around how applications are developed, deployed, and managed across their lifecycle. It presents new attack vectors – need to be concerned with the fact that containers share a common kernel. Keeping malicious container applications from exploiting kernel and container security holes is a top concern.

All CNF (Container Network Functions) owners must run security audits and security scans on all container images which are produced. The audits must include:

Container provided default security mechanisms e.g. process restrictions, file & device restrictions, sandboxing using Linux namespaces, & Linux kernel hardening

Image provenance: Check if images and packages inside images are up-to-date and are free of security vulnerabilities

Namespace quotas

Application separation by namespaces or clustering or zoning (Kubernetes provided)

- Audit automatization, we must be able to automatize all checks. That will save precious time and one can run it as often as one requires. Manual audit is not an option unless one is just testing or learning
- Container links and volumes. If you use read-only filesystem in your running container “docker diff” can help you to find issues
- The bigger an image is, the harder the audit will be. Reduce the size of your images as much as you can
- The host kernel is the shared point between all containers in the same server, keep that kernel up-to-date

3.8 Self-adaptive security management and orchestration

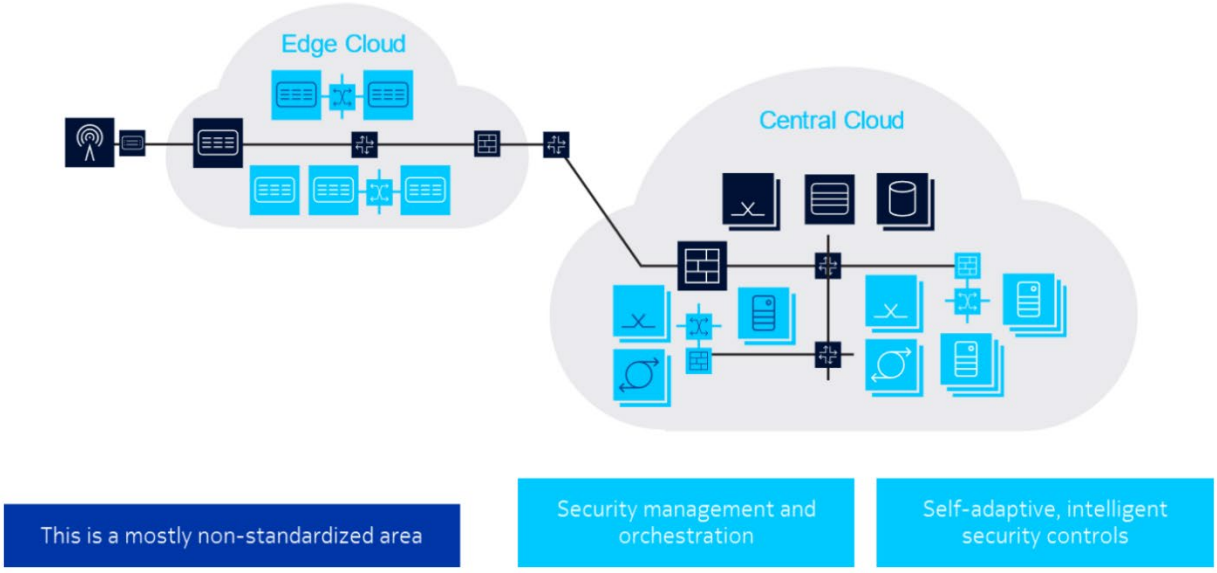


Figure 19 – Security management & Orchestration

There are two key parts of the 5G security architecture Security Management and Orchestration, and what is known as “self-adaptive, intelligent security controls”, which describes tools that monitor the network pervasively, analyze the information gained to detect anomalies and attacks, and trigger suitable countermeasures, with as little need for human interaction as possible. Again, this is an area that is mostly non-standard.

3.9 Network Slicing Security

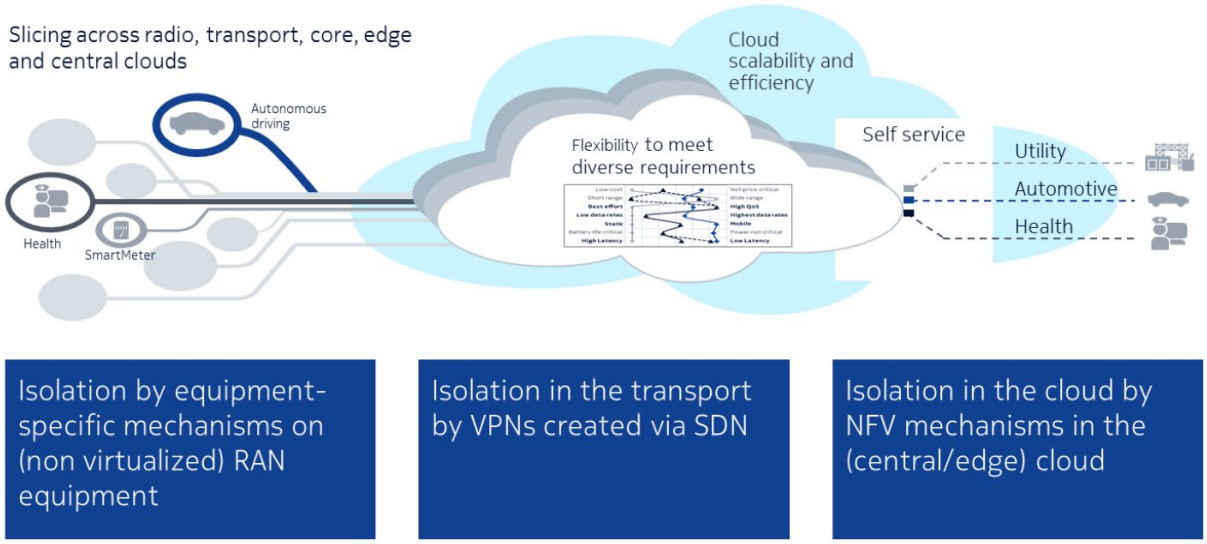


Figure 20 – Network slicing security

The crucial aspect in slicing security is slice isolation. Isolation has two angles:

- Availability: resources dedicated to one slice cannot be consumed by another slice
- Confidentiality: data/traffic cannot be intercepted/faked by entities of another slice

Network Slice Isolation = Resource Isolation + Security Isolation

Isolation will confine any effects of a potential cyber-attack to a single network slice. It is obvious that perfect isolation is required in a multi-tenant setup, where tenants may be competing organizations, such as different manufacturers running each running its own industrial automation slice.

As mentioned earlier, 5G security must be flexible. Instead of a one-size-fits-all approach, the security setup must optimally support each application. In a sliced network, this can be achieved by customizing the security setup per slice. Security features subject to this flexibility may comprise the mechanisms for identifying and authenticating mobile devices and/or their subscriptions, or for determining the way that user traffic is protected. For example, some applications may rely on security mechanisms offered by the network. These applications may require not only encryption, as in LTE, but also user plane integrity protection. However, other applications may use end-to-end security on the application layer. They may opt out of network-terminated, user-plane security because it does not provide additional security in this case (but rather increases the energy consumption of mobile devices).

Below are recommendations for network slicing security:

1. Better isolation if less components are shared

No side channel attacks if computer hardware and hypervisor are not shared

Tradeoff between resource usage efficiency and degree of isolation

Sharing increases the resource usage efficiency but potentially lowers the isolation

In a big central data center, there may be abundant physical resources, so some physical resources may be dedicated exclusively to one application, e.g. the UDM or a network management system. In small (far) edge cloud deployments, it may not be possible to set aside part of the physical resources for a single application only

2. Cloud Infrastructure security is mandatory. Cloud infrastructure must be carefully designed, implemented, and hardened to minimize vulnerability and side channel attack

3. Secure, trusted parties operating the shared parts are required

In most cases, a mobile network operator MNO may be considered a trusted party by its customers. There can also be use cases where the tenant cannot afford to rely on the security provided by an MNO, but needs to establish its own security mechanisms, including the use of tenant-owned infrastructure (for sensitive data, e.g. subscription data), where the MNO has no access at all

If untrusted parties need to deploy their application in a shared infrastructure, for example, online game vendor may want to install their gaming software in the Mobile Edge Compute (MEC) platform in order to reduce network latency, it is suggested operators must have a well-defined onboarding process to ensure the 3rd party software is fully tested and validated in a sandbox environment before deployment

4. Security automation & orchestration is needed to cope with dynamic nature of slicing. Security management and orchestration must be aware of slicing, same for the reactive security controls. Some security tools may only run within one slice, not aware of other slices, but there must be others that have the complete network view

3.10 Design for Security (DFSEC)

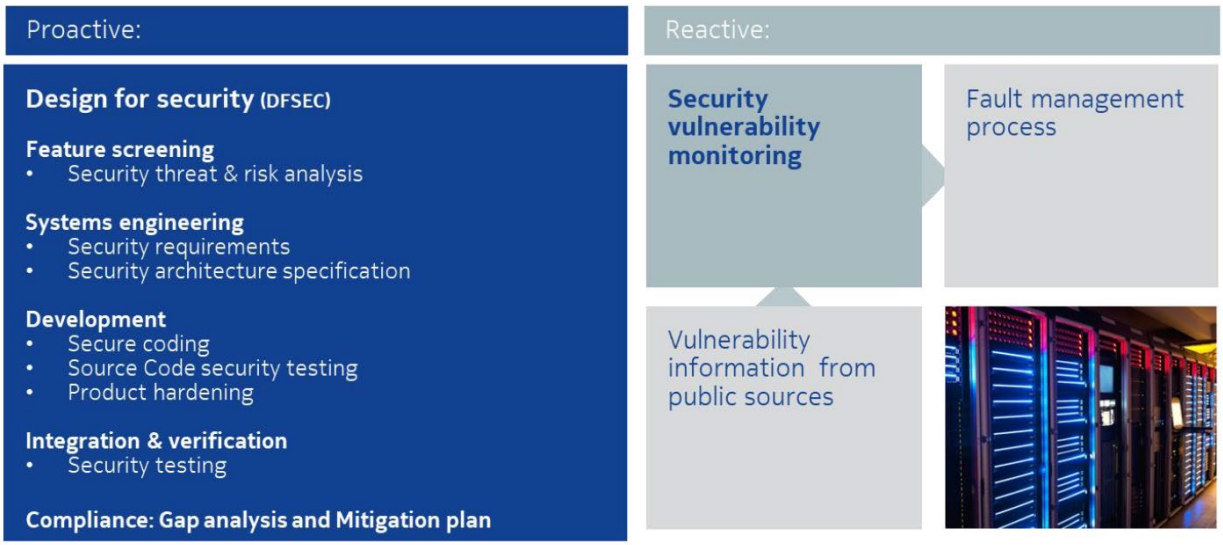


Figure 21 – Design for Security (DFSEC)

The vendor community recognizes that product security is not limited to security functions and protocols implemented in the product itself. Product security is strengthened and augmented by information security, incident response/vulnerability management and independent checks and audits.

Vendors must be capable to comply with legal and regulatory requirements around the world. And implement a security management processes and contractual security requirements for their internal teams and suppliers. Privacy by design is a part of Design for Security (DFSEC). Data privacy modelling and security features must be inbuilt to protect sensitive and private customer information in their products and ensure that product design complies with applicable regulatory requirements such as GDPR.

DFSEC is based on standards, industry best practices and customer requirements. Vendors' processes must be aligned with global security assurance frameworks from 3GPP (GSMA NESAS), TL9000 etc. Products must undergo customer acceptance tests, and security is a part of this testing. Below is a list of industry standards that must be employed in different phases of the Design & Development process.

Table 2 – DFSEC requirements

DFSEC Phase	Standard Compliancy
DFSEC	• SSE-CMM (SSE-CMM)
Threat / Risk Analysis	• ISO/IEC 27001 • 3GPP TS 21.133 Security Threats and Requirements • ITU-T X.805 (threat categories) • Guidance on the Privacy and Electronic Communications (EC Directive) Regulations 2003 Part 2
Security Requirements	• ISO/IEC 27001 • ISO/IEC 17799
Security Architecture	• ITU-T X.805
Secure Coding	• MISRA C
Security Testing	• NIST-1 (2003). NIST Guideline for Network Security Testing.
Security Auditing	• Common Criteria, Common evaluation methodology
	• ISO/IEC 19011 Guidelines for quality and/or environmental management system auditing.

Vendors must have a Design for Security (DFSEC) process embedded in the product development lifecycle and applies security requirements and security architecture at the beginning of the lifecycle, shifting the product security process from being reactive to proactive. DFSEC covers all the product development phases from feature screening, systems design, software development, integration, to verification processes.

Security is never a one-time effort. Every modification makes it possible for software bugs and security vulnerabilities to emerge. Therefore, security development must be process-oriented. Every release will undergo same threat & risk assessment, security checks, tests etc.

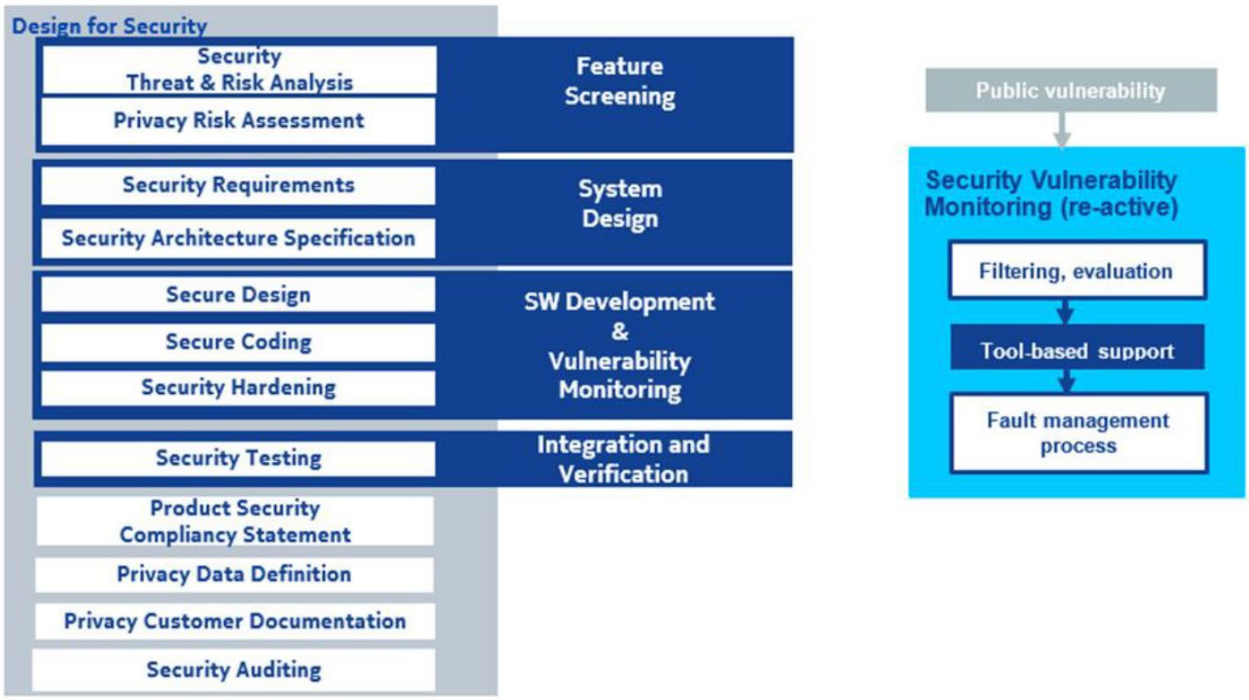


Figure 22 – DFSEC in action

Independent audits give customers further confidence in security promises and helps vendors identify areas where they can improve to meet the latest threats. Vendor must conduct third-party vulnerability scanning and penetration testing for their product samples. And ensure that operations support centers are certified to ISO 27001 standards.

3.11 Security Operations

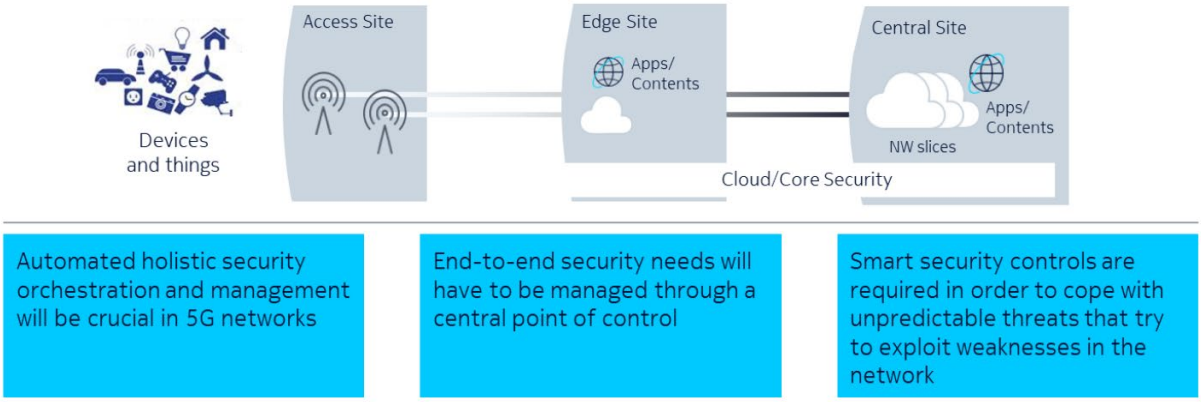


Figure 23 – Security Operations

Today, security professionals monitoring service provider and critical infrastructure networks often get thousands of cyber security alerts each day. Many are false alerts and duplicates. Yet, the sheer number of alerts can overwhelm a company's security team, resulting in incidents that are not investigated. For example, the "2018 Ponemon Security study" found that on average, 44 percent of alerts are not investigated, and of those investigated and deemed legitimate, nearly half (49 percent) go un-remediated. Teams need better ways to automatically prioritize alerts that allows them to focus on the most severe ones first. It is not an option to stay with the manually-intensive approach in the 5G era and must be migrated to an automated approach supported by artificial intelligence, data analytics and machine learning.

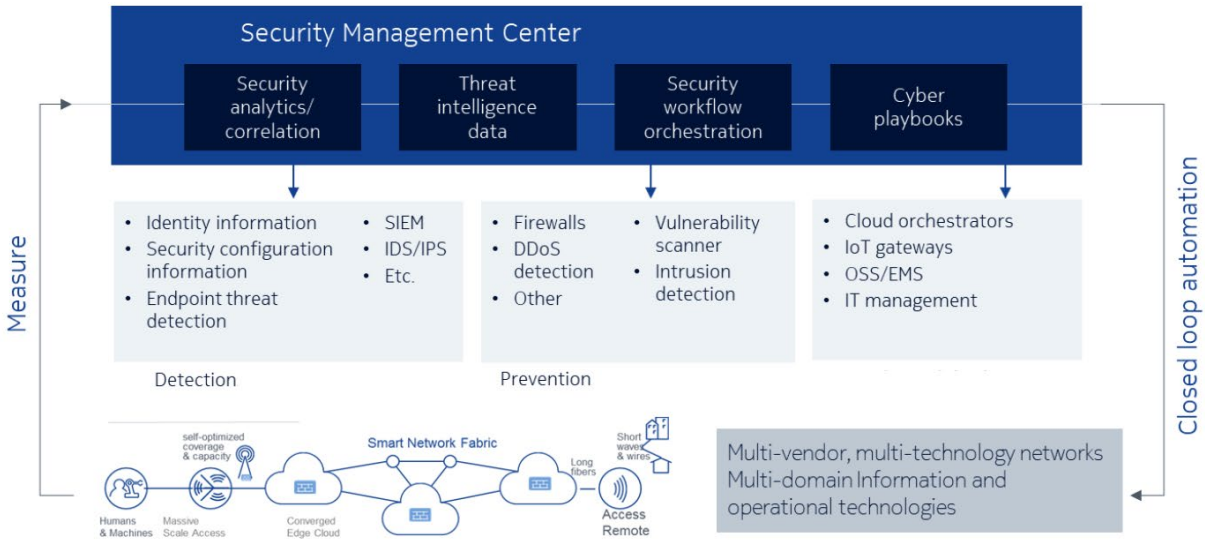


Figure 24 – Security Management & XDR

SOAR solution has been developed to replace today's manually-intensive approaches with security systems built on three pillars – intelligence gathering and analysis with machine learning and automation.

Intelligence gathering and analysis correlates data from across the network, devices, and cloud layers to spot suspicious anomalies, and provide insight into the nature of the threat, the associated business risk, and the recommended response.

Security Orchestration, Analytics & Response (SOAR) must be supported by an umbrella of security applications covering all aspects covered earlier in the paper:

1. Identity Access Manager

implements access controls to ensure least privileged access to key operational systems. It provides jump-host for CLI and GUI access to all (physical or virtual) network elements. i.e. no direct access from operator's consoles to management interfaces of the network elements. Human credentials are segregated from network elements' credentials. It also assures traceability of operator actions

2. Audit Compliance Manager

automatically and continuously audits the configurations of network entities for compliance with golden configuration

3. Certificate Lifecycle Manger

automates certificate enrollment and deployment, discovery, and audit

4. Endpoint Security

- analyzes network traffic in order to detect malware and anomalies of all end-user devices

5. Security Management Center

- aggregates logs and data for real-time monitoring & management. It provides a consolidated view for efficient reporting and simplifies incident management and forensics

3.12 Extended Detection & Response (XDR)

A fragmented security posture must move towards integration across Security Operations, Security Tools and Threat Intelligence. Extended Detection and Response (XDR) enables all three areas of integration. XDR takes security orchestration, automation, and response (SOAR) to even greater effectiveness through a cloud-native architecture built to accommodate the ever-growing and increasingly complex volumes of data coursing through 5G networks. XDR-based security operations are anchored by a robust data pipeline. That makes it possible to collect more data from more sources, all processed and analyzed through one cohesive security management system — so threats can be acted on faster and more effectively than ever before.

With machine learning (ML), the effectiveness of intelligence gathering, and analysis will improve continuously. Having access to a massive amount of high-quality data is the basis for training an AI/ML system. When using a security product that includes ML, you will want to augment the things you have

done in the past, like signature collection and automated malware analysis, and combine them with the machine's capability to determine new, malicious content.

A data pipeline also allows XDR to provide overarching security lifecycle management, orchestrating and automating all aspects of risk and threat prediction, detection, and response. Security teams can more easily integrate disparate threat intelligence data that is tailored to their unique requirement, model specific threats and attacks to their networks, and automatically apply the most appropriate preventative controls.

XDR solutions should provide the following features:

- Integrated security operations: End-to-end visibility across networks, clouds, and endpoints through a "single pane of glass" management interface — allowing security teams to quickly pinpoint the exact source of any potential breach
- Integrated security tools: Manage and administer disparate point products in a coherent and consistent way, providing a library of interfaces and connectors that bring a range of end-to-end infrastructure components and multi-vendor security tools under a single security management platform
- Integrated threat intelligence: Cognitive threat detection analyzes network sessions for malware or anomalous device behavior, and interprets the global threat landscape in a consistent, actionable way. Automated alert prioritization and classification eliminate the need for security teams to investigate redundant or lower-priority notifications so they can focus on blocking legitimate attacks

1. Summary

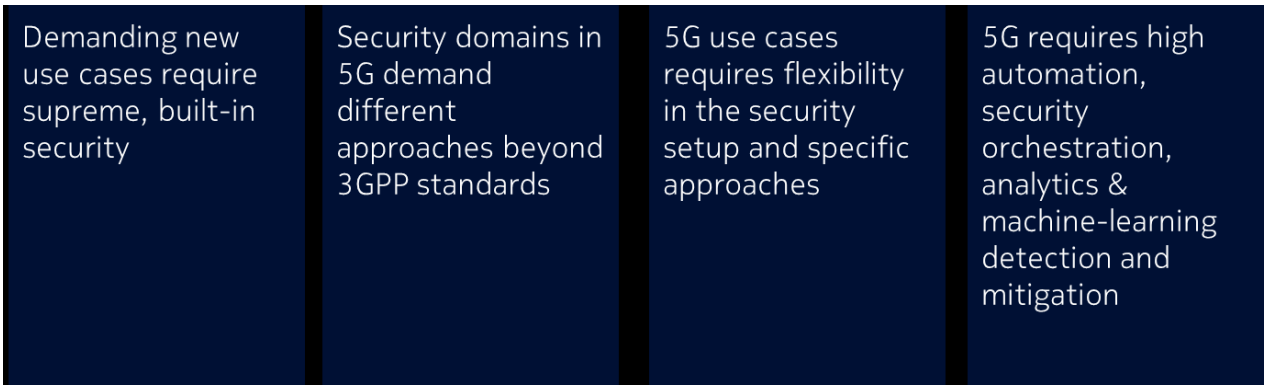


Figure 25 – End-to-end 5G security

End-to-end holistic security is mandatory in 5G network deployment and it is clear that:

1. 5G has a lot of mission critical use cases requiring supreme, built-in security. All network functions in the system must be hardened to avoid known vulnerabilities. Retrofit is always challenging and costly and, in some cases, service impacting
2. Number of network functions in a typical 5G network will be an order of magnitude more than 4G or fixed (cable or fiber) BB networks. Number of incidents and security logs will increase in the same order of magnitude. Adding the dynamic nature of the network configuration, 5G requires automation, security orchestration and machine-learning to identify and mitigate security threats
3. Different use cases may have different security requirements in different domains. Additional or overlapped security measures need to be implemented on top of recommendations from 3GPP standard
4. Because of dynamic nature of 5G network slicing, 5G use cases require flexibility in the security configuration

By introducing the security solution early on, operators can leverage security from the very beginning of their 5G rollout, rather than shelling out for an expensive retrofit further down the line. Operators can start to enrich security workflows, analytics, and training prior to the mass-deployment.

Security is a process (not a destination). Operators require a partner who is experienced in both 5G and security realms who has the skills, deployment experience in numerous 5G networks to be on top of any evolving threats and strong engineering and financial capabilities to continue investing in the 5G.

And most of all, a partner who is aligned in the overall goal to make 5G a secure, safe environment for everyone.

4 Acronyms

3GPP	3rd-Generation Partnership Project
AI	Artificial Intelligence
AKA	Authentication & Key Agreement
AMF	Access & Mobility Management Function
AR	Augmented Reality
AUSF	Authentication Server Function
BB	Broadband
CA	Certificate Authority
CMTS	Cable Modem Termination System
CNF	Container Network Function
C-Plane	Control Plane
CU	(gNode B) Central Unit
DFSEC	Design for Security
DN	Data Network
DNS	Domain Name Server
DOS	Denial of Service
DDOS	Distributed Denial of Service
EAP	Extensible Authentication Protocol
gNB	5G gNodeB base station
GTP	GPRS Tunneling Protocol
IDS	Intrusion Detection System
IMSI	International Mobile Subscriber Identity
IoT	Internet of Things
IPS	Intrusion Protection System

IPX	IP Exchange
MEC	Multi-access Edge Computing
ML	Machine Learning
MNO	Mobile Network Operator
N3IWF	Non-3GPP Interworking Function
NEF	Network Exposure Function
NFV	Network Function Virtualization
NGFW	Next Generation Firewall
NRD	Network Resource Directory
NRF	Network Repository Function
NSSF	Network Slice Selection Function
OSS	Operation Support System
OLT	Optical Line Terminal
PCF	Policy Control Function
PKI	Public Key Infrastructure
RAN	Radio Access Network
RU	(gNode B) Remote Unit
SBA	Service Based Architecture
SDN	Software Defined Networking
SeGW	Security Gateway
SEPP	Security Edge Protection Proxy
SIM	Subscriber Identification Module
SMF	Session Management Function
SOAR	Security Orchestration Analytics Response
SUCI	Subscription Concealed Identifier
SUPI	Subscription Permanent Identifier

TLS	Transport Layer Security
UDM	Unified Data Management
UE	User Equipment
UPF	User Plane Function
U-Plane	User Plane
V2X	Vehicle to Anything
VNF	Virtual Network Function
VR	Virtual Reality
VSR	Virtual Service Router
XDR	eXtended Detection & Response

10G And FTTP:
Drivers, Considerations And Strategies

A Technical Paper prepared for SCTE by

Michael T Scardina
Director, Network Strategies and Technologies
Armstrong
One Armstrong Place, Butler, PA 16001
(724) 283-0925 ext 50288
mscardina@agoc.com

Jess Beihoffer
Director, Sales Engineering
ADTRAN
901 Explorer Boulevard
Huntsville, AL 35806
(678) 772-0559
jess.beihoffer@adtran.com

Introduction

Armstrong is well recognized as America's 11th largest multiple system operator (MSO), serving thousands of homes and businesses across Pennsylvania, Ohio, West Virginia, Kentucky and Maryland. Their first broadband Internet customer was connected in 1997 and over the years they've earned a reputation as a fast-to-market and efficient disruptor.

As other MSOs debate fiber-to-the-premise (FTTP), Armstrong actually made the decision to build FTTP – 15 years ago … using radio frequency over glass (RFoG) technology to reach down to ten homes per mile in rural areas. With over 27,000 RFoG FTTP customers, four years ago Armstrong made the switch to light up the same 1x32 distributed split passive optical network (PON) infrastructure with gigabit passive optical networking (GPON). In fact, in dense areas, Armstrong activates GPON on the same glass as RFoG allowing one customer at a time to be moved from RFoG to GPON.

Why did Armstrong pursue FTTP in the first place and what drove the later transformation with XGS-PON? In this paper, we will explore the business, network, operations and financial reasons as to why Armstrong pursued this path to 10 gigabits per second (10G) and bypassed other options. We also outline how these network advancements and market factors have caused consumers to enhance their home wireless fidelity (Wi-Fi).

The Original FTTP Decision by Armstrong

Armstrong has experienced an impressive compound annual growth rate (CAGR) of combined bandwidth consumption for all Armstrong customers across the entire network. Historically, it had been 42%. More recently, the average usage CAGR has been 55%; daily average usage jumped from 290 Gbps (gigabits per second) in Feb 2020 to 450 Gbps a year later, (Gbps) on 2/1/2020 to 450 (Gbps on 1/31/2021 as shown in Figure 1.

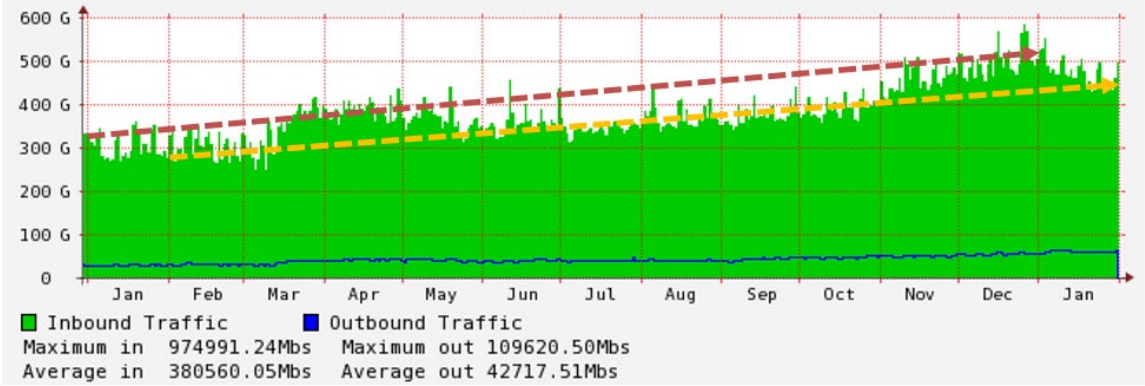


Figure 1 - Cumulative Usage for (Wide Area Network) WAN Plus (Content Delivery Network) CDN and Peering Cumulative Ports (Jan'20-Jan'21). (Source: Armstrong)

To achieve and sustain the stated goal of "growth in perpetuity," the company pivoted to a fiber-fueled growth strategy. Armstrong realized that each Data Over Cable Service Interface Specification (DOCSIS®) upgrade would be costly and, at the current growth rate, could likely be over subscribed upon being fully rolled out.

Even though FTTP was not the lower cost solution Day 1, it would be Day 2. To reach a fact-based decision, Armstrong created a spreadsheet containing all available historical data. “All the numbers proved out the fiber decision,” according to Michael Scardina, Director of Network Strategies and Technologies at Armstrong. Mid-term and long-term evaluations showed FTTP to be both the lower cost and longer life solution – enabling Armstrong to remain highly competitive long into the future.

Power Savings

A significant HFC and DOCSIS operating expense is electricity (e.g., powering the nodes and the downstream amplifiers.) Switching to fiber and PON would yield a substantial energy savings in millions of dollars annually. Then, there’s the labor savings resulting from a major decrease in the need for hybrid fiber-coaxial (HFC) plant maintenance (e.g., avoid ingress/egress leakage repair, eliminate battery upkeep).

Customer Premises Equipment (CPE) Costs

Another financial factor in Armstrong’s decision process involved CPE costs as they form a significant part of any network upgrade. Upgrading to DOCSIS 3.1 (then 4.0) or FTTP requires replacing CPE. For Armstrong, a GPON ecosystem was more mature and costs were more predictable. And a GPON 1Gbps system is easily upgraded to a XGS-PON 10Gbps system, with no change to outside plant necessary. XGS-PON also provides higher capacity and, like GPON, offers a very simple and scalable method to extend high throughput layer 2 connectivityto enterprise customers on the same fiber route. Armstrong uses these layer 2 connections to extend a full suite of enterprise services by virtually and securely connecting the customer to the hub distribution router. There are substantial financial and operational benefits to this as it typically removes the need for dedicated direct fiber with MPLS CPE routers. According to Scardina, “The cost of XGS-PON CPE is $1/10^{th}$ that of MPLS CPE used for enterprise customers.” For Armstrong, it was a more economical use of last mile fiber.

Upstream Limitations

Consumers and their behavior also played a key role in Armstrong’s decision. The need for speed applied not only to the downstream, but also the upstream. Inherent upstream limitations made DOCSIS technology a less viable path on which to continue. Upgrading to DOCSIS 3.1 technology would be “a stopgap solution at best” says Scardina. “Moving to orthogonal frequency-division multiplexing / orthogonal frequency-division multiple access (OFDM/OFDMA) is one thing, but moving the forward/reverse split to add upstream spectrum is a much more complex problem with compounding effects.”

Scardina continues, “For example, while Armstrong’s HFC plant architecture from the mid-1990s would allow amplifier module-only change to move from 5-42 megahertz (MHz) sub-split to 5-85MHz, because of limits in the legacy set-top box command and control downstream channel, upstream frequencies above 85MHz would require new diplex filter technology to notch the upstream to support a small amount of downstream. 5-85MHz isn’t enough upstream spectrum to support growth demands and 5-204MHz isn’t either in the long term. For the time being, Armstrong uses node/distribution leg splits to reduce the number of modems sharing the 5-42MHz upstream as FTTP replacement of HFC fully ramps.”

The HFC plant change from sub-split to mid-split, as shown in Figure 2, enables up to 300Mbps in the upstream. However, for the reasons articulated by Scardina above, the shift to a high-split to support symmetrical 1Gbps speeds is more complex and costly.

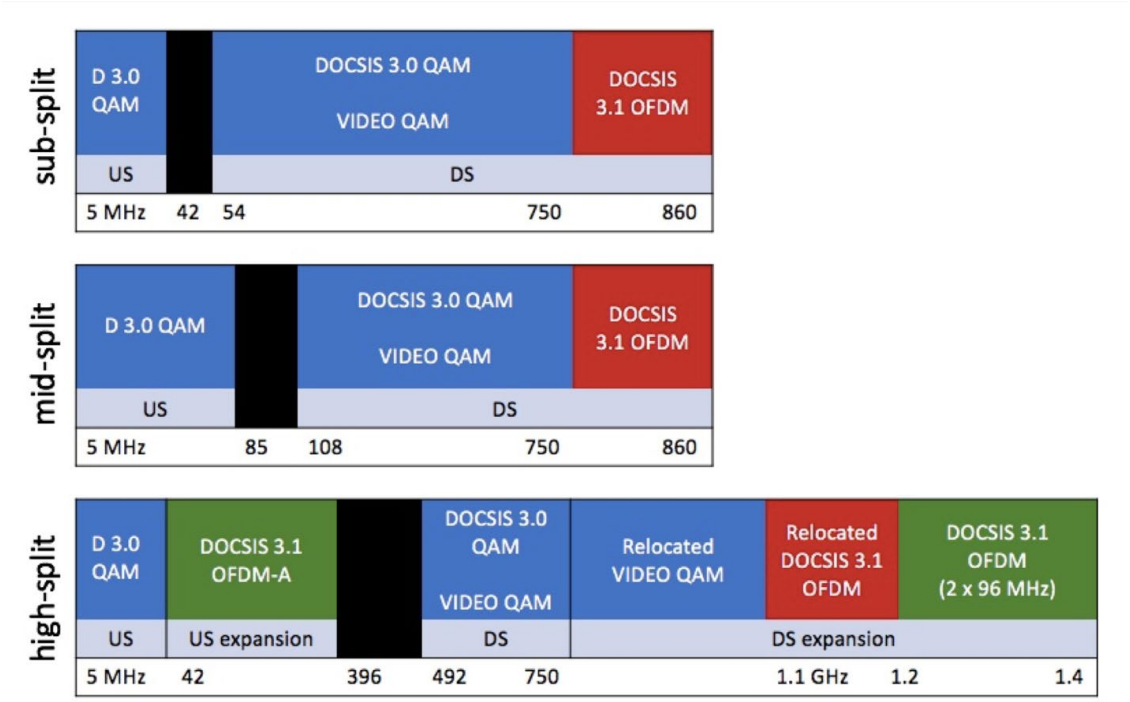


Figure 2 - HFC Spectrum Impact of Support for Higher Speed Services

The Consumer Electronics Show (CES) Horizon

The annual CES show in Las Vegas also provided ample proof that this was the right decision. Viewing the event as a "weather forecast" of what is on the horizon, Scardina took note of what the Armstrong network would need to support with the emergence of augmented reality (AR) and virtual reality (VR) centric gaming applications. Beyond software used in the home, there's the proliferation of devices. "On average, four new devices entered each home in one month earlier this year," according to Scardina. Washers, dryers, dishwashers, exercise equipment have all evolved into IoT devices – connected to the network. "There are so many new bandwidth uses that we simply didn't have in the past."

Field Tech Training & Equipment

In deciding upon fiber, Armstrong had an important operational consideration: the degree of difficulty to transition from DOCSIS technology to FTTP. What impact would this decision have upon the people, the processes and the equipment? Armstrong discovered that the technical skill sets honed through years of DOCSIS deployments could fairly easily be transferred to the deployment of fiber. Many of the same basic concepts applied. The move to fiber yielded a simpler process in the field as handheld splicing tools took the place of complicated connectorized cables. As for the use of spectrum analyzers, the technique that puts radio frequency (RF) signals in a radio spectrum for DOCSIS signals applies in a very similar way to optical signals in a fiber. And PON meters are simpler to operate than RF signal level meters.

The rationale for Armstrong's FTTP decision has been validated time and time again. An additional lesson that has been learned is that the company could have deployed FTTP even faster. With Rural Digital Opportunity Fund (RDOF) and other broadband initiatives driving increased fiber rollouts, Armstrong now competes with new and existing fiber entrants for the same labor and materials,

challenging the rate that fiber expansions can be built. However, broader operator adoption of fiber is a double-edged sword and Armstrong benefits from lower technology prices as a result. In fact, cost declines contributed to Armstrong's move from GPON to XGS-PON.

Provisioning: DOCSIS to GPON

"We initially started with Ethernet passive optical network (EPON) because of the DOCSIS Provisioning of EPON (DPOE) standard", says Scardina. "However, EPON didn't work out. DPOE and its virtual cable modem approach performed as expected, but the available CPE and services delivery did not. So after investing many months and the need to pacify angry customers, we switched to GPON due to its proven track record. A GPON system utilizes its own set of service templates for provisioning customers and a DOCSIS Provisioning of GPON (DPoG) standard never came to fruition. Nonetheless, by layering in commercial middleware, we were able to convert DOCSIS service codes to GPON templates; an outcome that surprised us in its simplicity."

Scardina continues, "Manual provisioning just wasn't an option, so we initially purchased middleware to marry our back office DOCSIS provisioning system to the GPON Network Management System (NMS). That worked very well and remained in place for more than a year. We then explored replacing the commercial middleware with internal resources for OpEx savings. In less than a week, an in-house developer was able to replace the commercial middleware with our own adapter. Additionally, designing our own middleware to interface between our traditional billing and CPE provisioning systems created an agnostic approach to flow-through provisioning. This allows field premise techs to provision a cable modem, GPON ONT, Wi-Fi, Telephony multimedia terminal adaptor (MTA)/number porting, and video set-top boxes all from one mobile application on their smart phones. Change can be hard, but it doesn't have to be as hard as you think."

The 10G Platform with XGS-PON and Combo PON

Armstrong's innovation with FTTP did not end with their original decision. For some Armstrong customers, GPON fits their needs perfectly. But, for the high-bandwidth consuming customers, there was a need for a smooth shift to accessing higher speeds. Service providers can co-mingle GPON and XGS-PON on the same fiber strand to provide more network flexibility and better address customer demands. For example, you can maintain customers on GPON while utilizing XGS-PON as a higher-end option for commercial customers (Figure 3). According to Scardina, "there's no need for a dedicated fiber using different technology to a single customer. Instead, we use an XGS-PON wavelength to deliver prioritized services."

"Armstrong was able to simplify its deployment process of XGS-PON over its existing GPON network through the use of Combo PON technology, which allows providers to combine these disparate technologies over the same infrastructure – from a single OLT port," according to Jess Beihoffer, Director of Sales Engineering at ADTRAN.

The need for separate overlay networks to serve homes and businesses can now be avoided. The technologies can be converged on the same platform. A single fiber. A single PON (up to 128 customers). All yielding economies of scale. Armstrong recently announced the launch of 10Gbps services starting in the town of Wexford, Pennsylvania using this architecture on a 32 PON split.

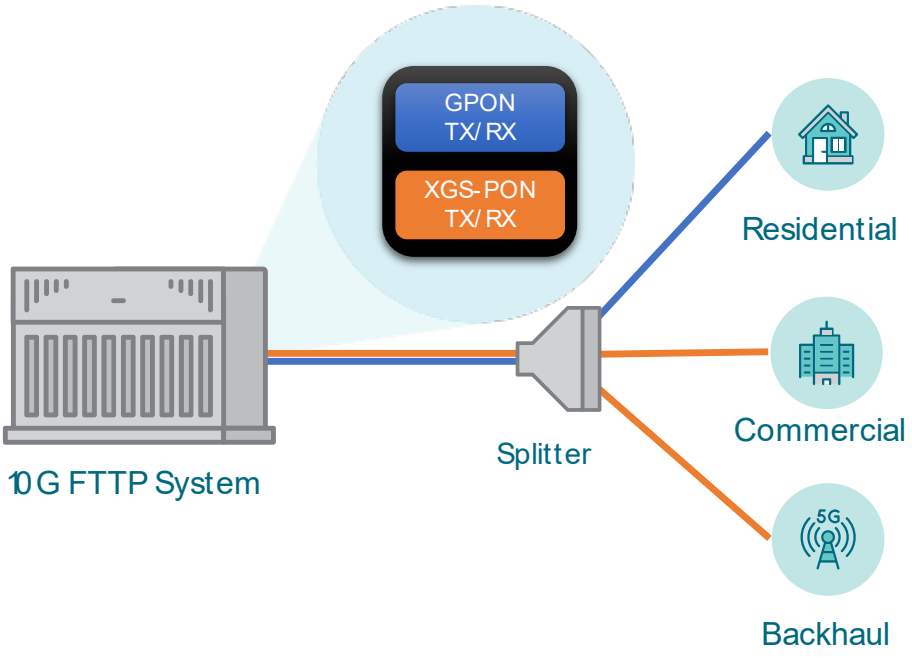


Figure 3 - Combined XGS-PON and GPON technologies on a common optical distribution network (ODN) and optical line terminal (OLT) system. (Source: ADTRAN)

The co-existence of PON technologies enabling backward compatibility has been a point of discussion since the inception of 10G PON standards in 2009. By combining both wavelengths within the OLT optics, Armstrong not only simplified the optical distribution network (ODN) when introducing higher capacity XGS-PON, but also took advantage of added port density as a single port can now simultaneously support both GPON and XGS-PON. Armstrong was able to gain greater network flexibility. The growing popularity of this technology has led industry and analysts such as Omdia to converge around the term Combo PON. Refer to Figure 4.

This technology is seen as key to simplifying the path to 10Gigabit networks, as well as providing improved economics for operators building new FTTP networks or modernizing existing GPON fiber networks to support the economic and social development of the communities they serve.

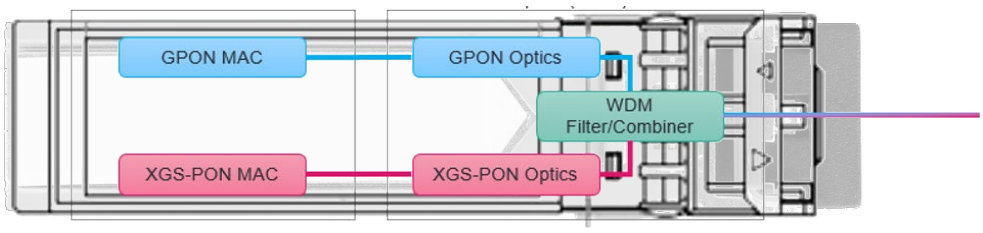


Figure 4 - Combining XGS-PON and GPON technologies within a common OLT module. (Source: ADTRAN)

Some of the immediate economic benefits that providers are seeing is a reduction in power consumption (50+%) as well as rackspace (up to 50% lower), mitigating two of the major concerns that providers face in their hubsites.

A more flexible network contributes to higher customer satisfaction. Moving high-usage GPON customers to the XGS-PON overlay alleviates congestion for customers staying on the GPON network and results in improved performance for those who remain and those who migrate. As Scardina explained, “If you have one customer on a street that is a very high bandwidth consumer, you can easily shift them over to XGS-PON to better fit their broadband needs without causing any disruption to your other customers in the area.”

Next generation FTTP technology like Combo PON provides all of the XGS-PON capacity and operational benefits, affords the opportunity to leverage the full value of mass market GPON technology, while further simplifying fiber network modernization processes. The promise of such a network should serve to “demystify the perceived level of effort and lessen intimidation,” according to Beihoffer of ADTRAN. “It’s not a forklift upgrade. Folks shouldn’t be scared to touch PON.”

With 10G networks historical bottlenecks no longer exist in the access network. The bottlenecks have migrated into the home and the backbone. Regarding the home, a stellar Wi-Fi experience is all-important – especially since most consumers view Wi-Fi as the Internet.

Delivering the In-Home Gigabit Experience

The in-home experience is critical. “If 10G is at the side of the house but can’t be accessed by the client device, what’s the point?” Scardina noted. The importance of excellent home Wi-Fi was recently confirmed with executives at Tier 2-3 service providers. They cited that the primary benefits of offering a residential managed Wi-Fi solution are Improved Customer Experience and Reduced Churn – two sides of the same coin. Not surprisingly, every executive said “Wi-Fi coverage throughout the house” is a very important to their residential customers.

Managed Wi-Fi

As MSOs increasingly focus on being broadband-first providers, they are focusing on providing a high-quality managed Wi-Fi service within the home. They also know that their customers view Wi-Fi as synonymous with the Internet and a Wi-Fi issue often triggers a customer service call. By providing a managed Wi-Fi solution, the service provider can ensure that their customers are leveraging the latest technology (currently Wi-Fi 6, soon to be Wi-Fi 6E), and placing advanced Wi-Fi system components throughout the house to ensure maximum coverage.

Analytics and Insight

Leveraging cloud-based data and analytics packages integrated into the Wi-Fi system solution, operators are given a window into the customer’s network, behavior, and experience. They are then equipped to efficiently provide enhanced customer service and present new offerings tailored to the subscriber.

The Core Network: Scaling to 100G

The backbone network between the OLTs and the core is another area that requires focus. As the last mile access network scales to 10G, Armstrong also is further strengthening and scaling core network and

transport systems to support the growing needs of the last mile. “Don’t just look at your last mile, it all comes back to your core. Build your core and make sure it’s highly scalable and redundant”, said Scardina. Several years ago, Armstrong began upgrading its core network backbone links from multiple (N) x 10 gigabit Ethernet (Nx10GE) ports to (N) x 100 gigabit Ethernet (Nx100GE). As XGS-PON deployments increase, Armstrong is preparing for similar upgrades to the uplinks from their 10G fiber access, aggregation and transport platform. Refer to Figure 5.

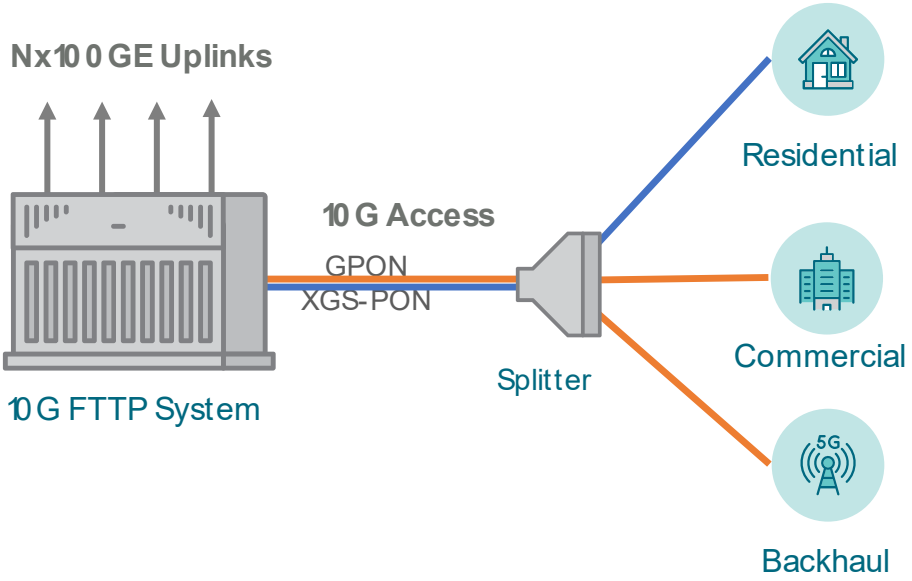


Figure 5 - Strengthening the Network Core for 10G Access. (Source: ADTRAN)

Why Not DOCSIS 3.1 / 4.0 (with fiber deep)

If service groups are kept small enough, DOCSIS 3.1 technology can reliably support the speeds that GPON enables (at least in the downstream direction), but not those delivered with XGS-PON. To achieve speeds approaching 10G, DOCSIS 4.0 Extended Spectrum or Full Duplex DOCSIS techniques are required. These are certainly valid choices for some, but many will find that the plant upgrades required to extend the spectrum to 1.8 GHz and beyond, and to modify the split and/or change the architecture to Node + 0 can add these complications that are not found with an FTTP overlay, or with an upgrade from GPON to Combo PON:

- Service disruption during upgrade activities;
- No operating cost reduction;
- Disruption of legacy services due to interference from new upstream channels (requiring installation of filters or change of CPE), and
- Multiple upgrade steps may be required on the way to 10G.

Conclusion

The Armstrong FTTP story is a compelling one. The business, network, operations and finance reasons as to why Armstrong pivoted to FTTP long ago are ones which other MSOs could well use today to make a strong argument in favor of this transformative move. These are just some of the reasons which are as relevant as ever:

- Performance of bandwidth-hungry (gaming) applications and escalating volume of connected internet of things (IoT) devices;
- Long-term, competitive, lower cost solution;
- Lower labor expense with reduction in plant maintenance;
- Significant energy savings, and
- Maturity of fiber and PON ecosystem due to greater FTTP adoption by more operators.

That argument is only bolstered by the advantages realized with XGS-PON together with Combo PON technology. Service providers can co-mingle GPON and XGS-PON on the same fiber strand to provide more network flexibility and better address customer demands. Combo PON technology yields a significant reduction in space use, power consumption and capital expense which enabling a longer GPON life span and an increase in gigabit service coverage.

As Scardina states, “The time for XGS-PON is now. Find a way to do it that makes financial sense. Spend the capital expense (Capex) once and be prepared for when customer demands increase. Be proactive, not reactive.”

Abbreviations

100G	100 gigabits per second
10G	10 gigabits per second
AR	augmented reality
CAGR	compound annual growth rate
CapEx	capital expense
CDN	content delivery network
CES	Consumer Electronic Show
CPE	customer premises equipment
DOCSIS (3.0, 3.1, 4.0)	data over cable service interface specification (version)
EPON	Ethernet passive optical network
FTTP	fiber to the premises
Gbps	gigabit per second
GPON	gigabit passive optical networking
HFC	hybrid fiber-coaxial
MHz	megahertz
MSO	multiple system operator
MTA	multimedia terminal adaptor
Nx100GE	number (N) x 100 gigabit Ethernet
Nx10GE	number (N) x 10 gigabit Ethernet
ODN	optical distribution network
OFDM	orthogonal frequency-division multiplexing
OFDMA	orthogonal frequency-division multiple access
OLT	optical line terminal
ONT	optical network termination
PON	passive optical network
QAM	quadrature amplitude
RDOF	Rural Digital Opportunity Fund
RF	radio frequency
RFoG	radio frequency over glass
SLA	service level agreement
Virtual Reality	virtual reality
WAN	wide area network
Wi-Fi (6, 6E)	wireless fidelity (version)
XGS	10 gigabit symmetrical

10G Full Duplex DOCSIS Implementation Exceeds Expectations

A Technical Paper prepared for SCTE by

Richard S Prodan, Ph.D.
Engineering Fellow
Comcast Cable
1401 Wynkoop Street #300
720-512-3742
rich_prodan@comcast.com

1. Introduction

The advent of DOCSIS 4.0 Full Duplex (FDX) technology has arrived after several years of industry collaboration and the publication the DOCSIS 4.0 FDX specifications. Implementation of an FDX node reference design incorporating the remote PHY ASIC SoC has been initially evaluated in the Comcast labs. The key technology to enable simultaneous transmission and reception in the same spectrum requires real-time removal of the high-power cable plant reflections of the transmitted downstream signal from the low-power upstream signal received using “echo cancellation”.

The successful performance of this enabling technology demonstrates the potential for greatly increasing upstream throughput in an additional nearly 600 MHz of spectrum shared concurrently with the transmission of the downstream signal. Evaluations indicate increased spectral efficiency beyond the minimum performance requirements in the FDX specification. This paper will analyze the implied capacity limitations in the specs. A comparison to the measured performance is made which demonstrates exceeding these limits. This comparison will conclude that the performance limits in the FDX specification should be reexamined to update the expected efficiency gains enabled by full duplex echo cancellation technology as currently realized.

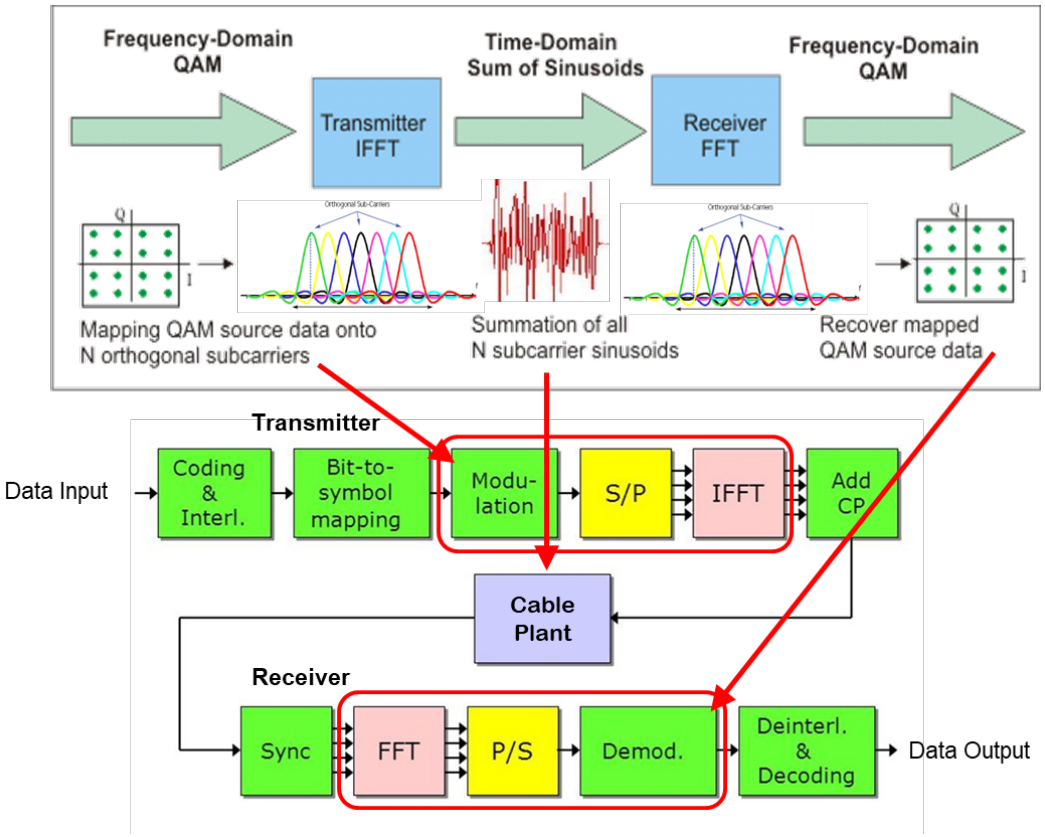
2. DOCSIS 3.1 – The Basis of DOCSIS 4.0 Full Duplex Modulation

DOCSIS 3.1 provides higher throughput and greater flexibility for cable data systems. The PHY layer based on Orthogonal Frequency Division Multiplexing (OFDM) differs significantly from that in previous generations based on Single Carrier Quadrature Amplitude Modulation (SC-QAM).

OFDM is a modulation scheme where many closely spaced, complex valued, harmonically related (orthogonal) QAM data subcarriers of various modulation orders are transformed into a time domain waveform or OFDM symbol. The resulting OFDM symbol in the time domain can be reversibly transformed back into the original QAM data subcarriers in the frequency domain. This reversible process uses the mathematically efficient implementation of the Discrete Fourier Transform (DFT) called the Fast Fourier Transform (FFT) to accomplish this reversible transformation. The N-point FFT transforms a group of N time samples into an equal number of N frequency domain complex valued subcarriers. The inverse FFT reverses this transformation from the frequency domain into the time domain.

As shown in Figure 1, forward error correction (FEC) coding and interleaving is applied to the input data bits. The error protected data bits are mapped into N frequency domain QAM symbols modulated onto N orthogonal subcarriers. These QAM subcarriers are serial-to-parallel converted and input into an inverse Fast Fourier Transform (iFFT). This transformation produces a single time domain OFDM symbol comprised of the summation of all N subcarriers. A portion of the end of each OFDM symbol known as a cyclic prefix (CP) is prepended to the beginning of the same symbol to prevent inter-symbol interference (ISI) due to signal micro-reflections or “echoes”.

Any contiguous group of samples within each OFDM symbol can be used to recover the subcarriers using the periodicity property of the DFT. If the latter part of the symbol past the prepended cyclic prefix is used, and the micro-reflections are confined in time to the duration of the cyclic prefix, then the remaining samples in the OFDM symbol are free of inter-symbol interference. Consecutive groups of QAM subcarriers are thus transformed into OFDM symbols and transmitted successively over the cable plant.

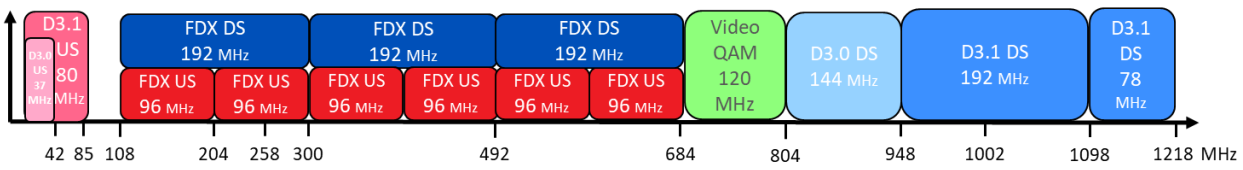


The reverse process at the cable modem receiver synchronizes the ISI free portion of each OFDM symbol, performs an FFT and parallel-to-serial conversion to recover the QAM modulated data subcarriers which are then demodulated, deinterleaved, and FEC decoded to recover the error-corrected data bits. Details of the DOCSIS 3.1 OFDM/OFDMA transmission and reception are described in a prior paper [1].

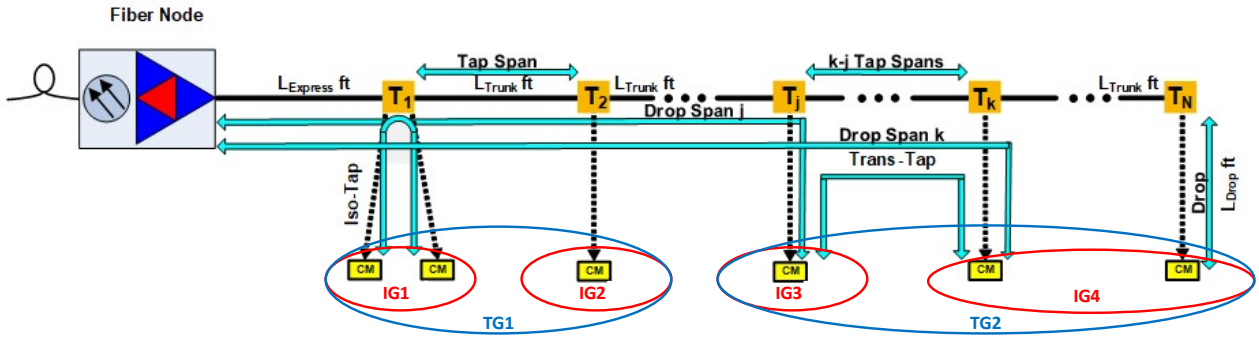
3. DOCSIS 4.0 - Full Duplex Spectrum Sharing

Full Duplex DOCSIS 4.0 uses the same OFDM/OFDMA modulation as DOCSIS 3.1 but overlaps downstream transmissions from node to modem and upstream transmissions from modem to node within the same frequency spectrum in the 108 MHz to 684 MHz FDX band. This greatly increases the upstream bandwidth yielding a total channel data rate up to 5.7 Gbps with 1024 QAM subcarriers and over 4 Gbps total data payload rate without overhead.

Legacy upstream DOCSIS 3.0 and 3.1 signals remain in the 5 to 85 MHz (mid-split) band. Legacy downstream DOCSIS 3.0 and 3.1 signals occupy the 804 MHz to 1002 MHz band, and only DOCSIS 3.1 OFDM signals occupy 1002 MHz to 1218 MHz. An example of such spectrum allocation is shown in Figure 2 with downstream (DS) and upstream (US) frequency bands.



The FDX band is divided into three 192 MHz sub-bands in the FDX band containing three downstream OFDM channels overlapped with six upstream OFDMA channels. The FDX node simultaneously transmits and receives these channels within the FDX band. Cable modems operate in a dynamic FDD mode with the direction of each sub-band upstream transmission or downstream reception dynamically assigned with a Resource Block Allocation (RBA) message allocating the upstream or downstream capacity within the FDX band in a flexible manner as capacity demands require.



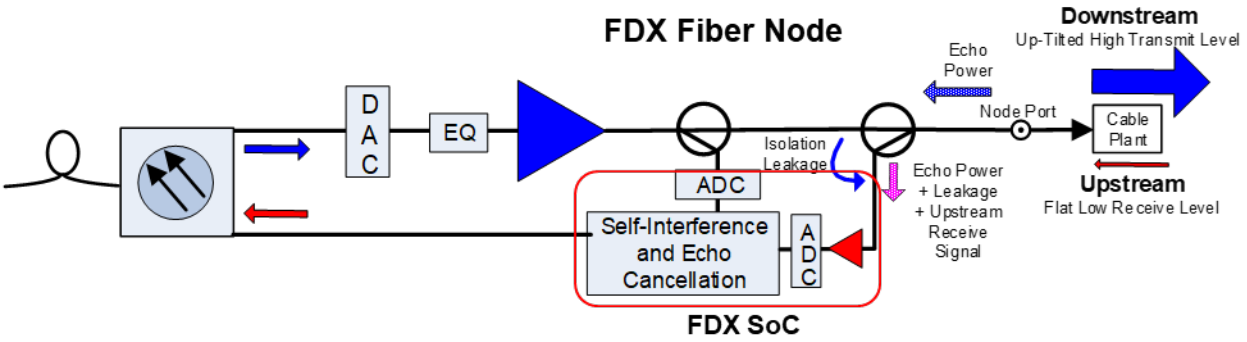
Tap Position (Highest to Lowest):	Tap 1	Tap 2	Tap 3	Tap 4	Tap 5	Tap 6	Tap 7
Tap Type (Number of Ports):	4	4	4	4	4	4	4
Tap Value (dB):	29	29	29	26	20	17	8
Tap Conditioner Type (CS or EQ):	CS	CS	CS	EQ	EQ	EQ	EQ
Tap Conditioner Value (dB):	12	6	6	4	10	10	22

Full Duplex DOCSIS 4.0 was originally specified for a Node + 0 plant without amplifiers beyond the node. The Node + 0 cable plant architecture is shown in Figure 3. When a modem transmits in an FDX sub-band, modems on the same tap (Iso-Tap) cannot receive on the downstream in the same sub-band due to the high upstream interference introduced across tap ports into the low downstream receive level of the other modems on the same tap. These modems are said to belong to the same interference group (IG). However, modems separated sufficiently apart across different taps can have sufficient isolation to interference. Such separated modems can belong to different IGs where the transmission of a modem in a sub-band does not significantly interfere with the reception of other modems in other IGs within the same sub-band. Identification of each modem IG is performed in the “sounding” process in which each modem in turn transmits a test upstream transmission while all other modems measure the received level relative to the downstream received level in a modulation error ratio (MER) triggered measurement window. Low received MER modems are placed in a common IG while higher MER modems remain in other IGs that are isolated from the test modem interference.

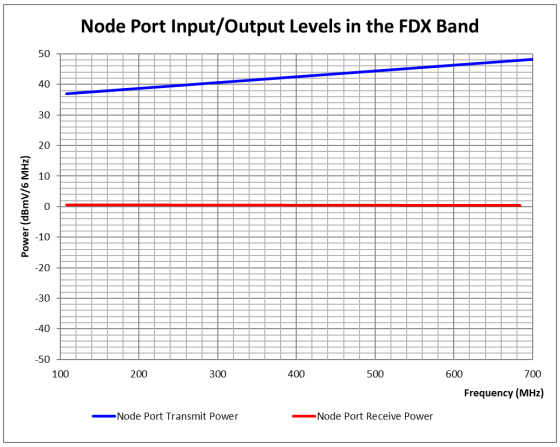
Modems in different IGs can be grouped into a single Transmission Group (TG) with common sub-band RBA transmit/receive assignments. Other TGs can operate with complementary RBA assignments providing full duplex simultaneous transmission and reception across all TG sub-bands at the node. A detailed treatment of FDX PHY layer operation of interference management is described in [2].

4. FDX Node Evaluation in the Comcast Lab Cable Plant

The operation of the simultaneous transmission and reception in the FDX node reference design that was evaluated in our lab. The 7 tap Node + 0 cable plant design used in this evaluation is shown in Figure 3. The FDX node reference design is depicted in Figure 4. The Remote PHY device (RPD) containing the FDX system-on-chip (SoC) in the FDX node receives legacy upstream below 85 MHz and legacy downstream from 804 MHz to 1218 MHz (not shown). These legacy bands are needed to initialize the node as well as cable modems in DOCSIS 3.1 mode prior to adding DOCSIS 4.0 FDX functionality in the FDX band.

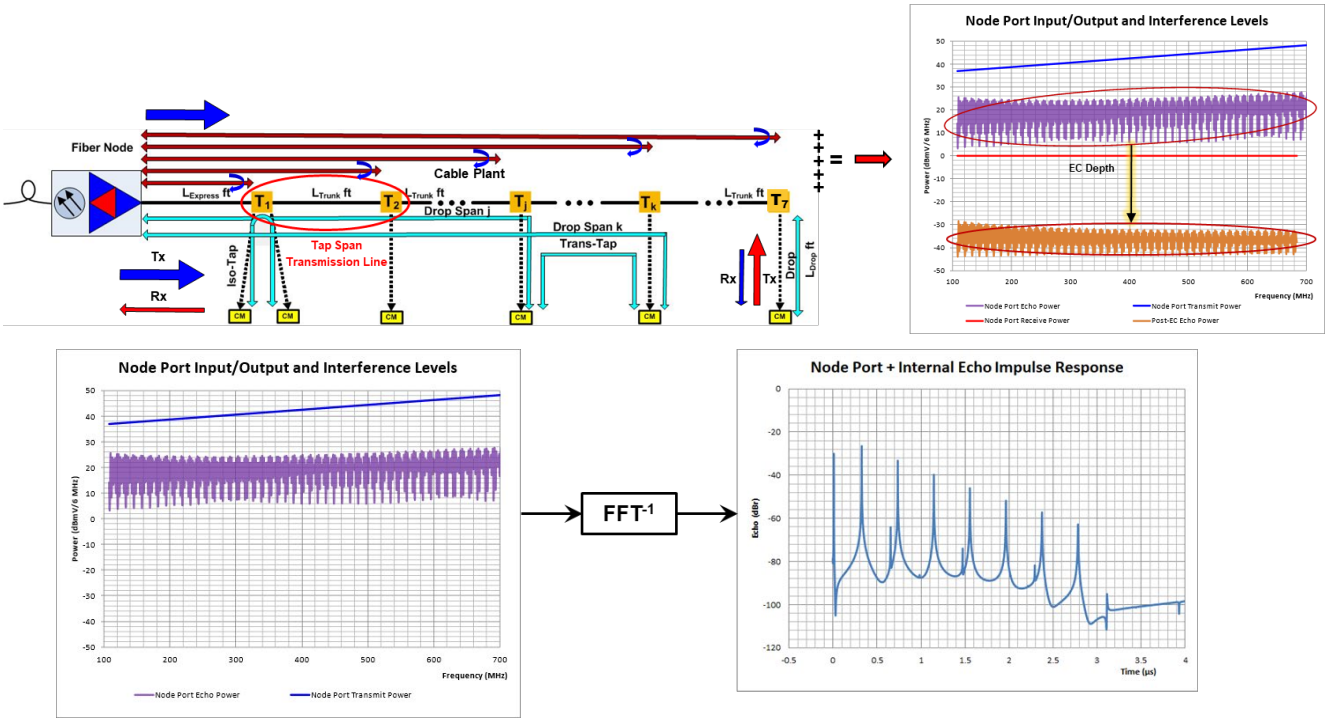


The downstream continuous OFDM signal is launched from the node port at a level from 37 to 48 dBmV/6 MHz with a 10 dB uptilt in the FDX band (see Figure 5). The downstream level determines the maximum reach (number of equalized taps and feeder plus drop cable lengths) for a modem receive level of 0 dBmV/6 MHz across the entire FDX and legacy bands.



The combined upstream burst OFDMA signals of all granted modem minislots arrives at the node port without significant tilt at a low level near 0 dBmV/6.4 MHz. The upstream receive level is limited by the 65 dBmV total composite power (TCP) of the cable modems in the highest loss path from the modem to the node. Thus, the upstream path being essentially equal path loss will transmit with the same power spectral density resulting in the same 0 dBmV/6.4 MHz across the FDX band at the node port.

The high-level downstream signal launched into the cable plant will result in reflected signal energy back toward the node due to tap return losses causing impedance mismatches with the cable characteristic impedance in the cascade of connected tap-to-tap transmission lines. The node echo paths and the resulting reflected signals or echoes are shown in Figure 6.



Note that the downstream echo power is below the downstream launch power at the node port but the echo is higher than the received upstream signal at the node port. Taking the inverse FFT of the reflected signal power shows the distribution of the relative echo level and delay of the echo path components from the node port and each tap relative to the 0 dB node launch signal reference. This echo level being higher than the received upstream signal at the node port results in a negative signal-to-noise ratio (SNR). The upstream signal would not be recoverable with a negative SNR. Hence echo cancelation technology of sufficient cancelation depth to suppress this interference and increase upstream received SNR is required.

5. Echo Cancellation Performance Exceeded

The RPD SoC in the FDX node reference design of Figure 4 reduces the echo level below the upstream receive level using self-interference and echo cancelation of the downstream echo. The node downstream signal is sampled and digitized as a reference for echo cancelation. The upstream signal with the

downstream echo through and the leakage across the FDX node port directional coupler of Figure 4 (replacing the frequency division duplex diplexer) is also digitized. Using the downstream reference signal, the echo cancelation operation within the FDX SoC subtracts the echo plus leakage signals from the upstream input plus interference.

The evaluated performance of the echo cancelation is shown in Figure 7. The resultant bit-loading achieved was 1024-QAM subcarriers in the received OFDMA symbols with zero codeword errors after LDPC decoding. The echo is substantially cancelled. The residual post EC signal to echo ratio is increased to 35 dB to support 1024-QAM an shown in Table 2, Upstream CNR Performance for D3.1 – a significantly positive upstream receiver SNR!

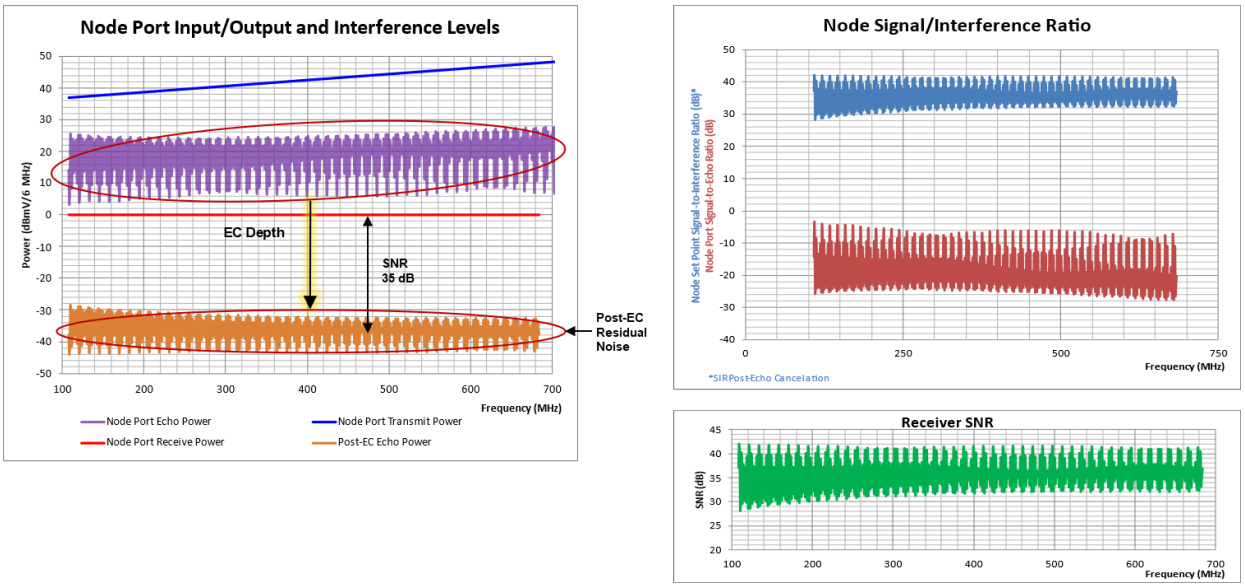


Figure 7 – Node Signal to Echo Interference (Pre and Post Echo Cancelation)

The upstream and downstream data rates per channel type and the total aggregate data rates for the channel plan of Figure 2 are calculated in Table 1.

Table 1 – FDX Cable System Data Rates

FDX OFDMA Upstream per 96 MHz Channel

Modulation	SNR (dB) @ BER=1E-8	Bits/Symbol	Spectral Efficiency (b/s/Hz)	Spectral Efficiency w/ Total Overhead (b/s/Hz)	Loss w/ Total Overhead (b/s/Hz)	Modulated Bandwidth (MHz)	Non-Excluded Subcarriers	Minislot Subcarriers	Minislot Symbols	Pilots (#4 :2^=16)	Low Density Pilots	Cyclic Prefix (us)	FEC Rate	FEC Rate w/ Minislot Overhead	FEC Rate w/ Total Overhead	Bit Rate (Mbps)
1024-QAM	35.5	10	8.888888889	6.772486772	2.116402116	94.8	1896	8	14	16	2	2.5	0.888888889	0.761904762	0.677248677	642

Non-FDX SC-QAM Upstream per 6.4 MHz Channel

Modulation	SNR (dB) Low Pkt Loss	Bits/Symbol	Bandwidth (MHz)	Symbol Rate (Msymbol/s)	Efficiency	Bit Rate (Mbps)
64-QAM	22.8	6	6.4	5.12	0.834	25.6

FDX OFDM Downstream per 192 MHz Channel

Modulation	SNR (dB) @ BER=1E-8	Bits/Symbol	Spectral Efficiency (b/s/Hz)	Spectral Efficiency w/ Total Overhead (b/s/Hz)	Loss w/ Total Overhead (b/s/Hz)	Modulated Bandwidth (MHz)	Non-Excluded Subcarriers	PLC	Continuous Pilots	Scattered Pilots	NCP	Cyclic Prefix (us)	FEC Rate	FEC Rate w/ CW Hdr (2 B PDU Ptr)	FEC Rate w/ Total Overhead	Bit Rate (Mbps)
1024-QAM	27.2	10	8.785185185	7.527264746	1.257920439	190	3800	8	48	29	48	2.5	0.8785185	0.8775309	0.7527265	1430

Non-FDX OFDM Downstream per 192 MHz Channel

Modulation	SNR (dB) @ BER=1E-8	Bits/Symbol	Spectral Efficiency (b/s/Hz)	Spectral Efficiency w/ Total Overhead (b/s/Hz)	Loss w/ Total Overhead (b/s/Hz)	Modulated Bandwidth (MHz)	Non-Excluded Subcarriers	PLC	Continuous Pilots	Scattered Pilots	NCP	Cyclic Prefix (us)	FEC Rate	FEC Rate w/ CW Hdr (2 B PDU Ptr)	FEC Rate w/ Total Overhead	Bit Rate (Mbps)
4096-QAM	27.2	12	10.54222222	9.032717695	1.509504527	190	3800	8	48	29	48	2.5	0.8785185	0.8775309	0.7527265	1716

Non-FDX SC-QAM Downstream per 6.4 MHz Channel

Modulation	SNR (dB) @ BER=1E-8	Bits/Symbol	Spectral Efficiency (b/s/Hz)	Spectral Efficiency w/ Total Overhead (b/s/Hz)	Loss w/ Total Overhead (b/s/Hz)	Bandwidth (MHz)	Symbol Rate (Msymbol/s)	FEC Rate RS Code	FEC Rate Trellis Code + Framing	FEC Rate w/ Total Overhead	Bit Rate (Mbps)
256-QAM	30	8	7.625	7.2400768	0.3849232	6	5.360537	0.953125	0.953125	0.9050096	38.8

Frequency Band	DOCSIS Bit Rate (Mbps)
FDX US	3852
Non-FDX US	394
Total US	4246
FDX DS	4291
Non-FDX DS	3343
Total DS	7634

Table 2 – DOCSIS Bit-Loading Specifications

Upstream CNR Performance for D4.0 FDX			
QAM Order	Modulation Efficiency (bits/subcarrier)	CNR Threshold (dB)*	Minimum Power (dBmV/6 MHz)
QPSK	2.0	12.5	0.0
8-QAM	3.0	15.5	0.0
16-QAM	4.0	18.5	0.0
32-QAM	5.0	22.0	0.0
64-QAM	6.0	25.5	0.0
128-QAM	7.0	29.0	1.0
256-QAM	8.0	32.0	3.0
512-QAM	9.0	36.0	5.0
1024-QAM	10.0	44.0	7.0

Upstream CNR Performance for D3.1			
QAM Order	Modulation Efficiency (bits/subcarrier)	CNR Threshold (dB)*	Minimum Power (dBmV/6 MHz)
QPSK	2.0	11.0	-4.0
8-QAM	3.0	14.0	-4.0
16-QAM	4.0	17.0	-4.0
32-QAM	5.0	20.0	-4.0
64-QAM	6.0	23.0	-4.0
128-QAM	7.0	26.0	0.0
256-QAM	8.0	29.0	0.0
512-QAM	9.0	32.5	0.0
1024-QAM	10.0	35.5	0.0
2048-QAM	11.0	39.0	7.0
4096-QAM	12.0	43.0	10.0

The upstream bit-loading is a function of the echo level present relative to the minimum upstream receive level. This relationship is illustrated in Figure 8 where lower echo levels yield higher bit-loading post echo cancelation and vice versa. The assessment of this trade-off for a given minimum upstream receive level is ongoing for a more complete characterization of the echo cancelation performance as implemented.

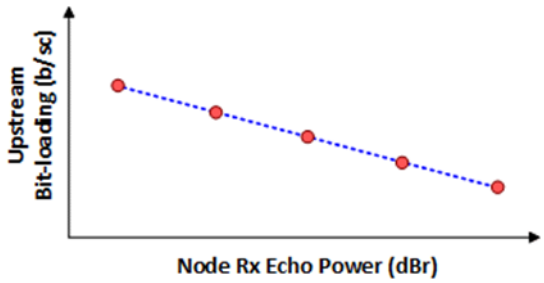


Figure 8 – Upstream bit-loading vs. relative echo power

6. The Future Evolution of FDX Technology

The demonstrated results of this technology in our labs indicates a path forward to extend the reach of echo cancelation beyond all-passive Node + 0 cable systems. The depth of echo cancelation with the system designed transmit and receive signal levels and resulting echo levels shows 1024-QAM OFDMA modulation is achievable in a Node + 0 cable system. The use of this technology in FDX amplifiers can extend the reach into Node + N system designs.

Managing downstream output and upstream input levels in amplifier cascades such that the upstream signal to echo ratio is reduced below that of the node can increase the achievable SNR at the upstream amplifier output for the same echo cancelation depth in the node. Longer amplifier cascades will accumulate the reduced echo cancelation residual interference lowering the received SNR and achievable bit-loading at the node. But as shown in the Table 2 SNR vs. bit-loading, a 6 dB decrease in SNR to 29 dB lowers the modulation efficiency from 10 to 8 bits per subcarrier. This compromise could enable the use of FDX amplifiers with echo cancelation in limited cascade depths for a 20 percent decrease in overall upstream capacity over the 576 MHz wide FDX band. Such a compromise may be entirely acceptable in exchange for the still very significant increase in wideband upstream capacity with amplifier cascades extending the reach of FDX technology to Node + N cable systems.

7. Conclusion

The initial evaluation of echo cancelation technology of an ASIC SoC enabled RPD in a DOCSIS 4.0 FDX node reference design was described. The performance in a representative Node + 0 cable system design in our labs demonstrated the successful implementation of FDX echo cancelation in a representative but challenging Comcast Node + 0 cable system design. The measured 1024-QAM OFDMA error-free performance exceeded the upstream capacity of 64-QAM OFDMA in the DOCSIS 4.0 FDX PHY specification – a 67 percent capacity increase.

The echo cancelation performance utilized in a future FDX amplifier can extend the reach of FDX beyond all-passive Node + 0 cable systems. A trade-off of amplifier cascade depth vs reduced upstream spectral efficiency could provide significant increases in total upstream capacity in existing 1.2 GHz Node + N cable systems.

Abbreviations

CP	cyclic prefix
DFT	Discrete Fourier Transform
DS	downstream
FDX	Full Duplex
FEC	forward error correction
FFT	Fast Fourier Transform
iFFT	inverse Fast Fourier Transform
IG	interference group
ISI	inter-symbol interference
MER	modulation error ratio
OFDM	Orthogonal Frequency Division Multiplexing
RBA	Resource Block Allocation
RPD	Remote PHY device
SC-QAM	Single Carrier Quadrature Amplitude Modulation
SNR	signal-to-noise ratio
SoC	system-on-chip
TCP	total composite power
TG	transmission group
US	upstream

Bibliography & References

[1] *Demystifying the DOCSIS 3.1 PHY*, R. Prodan, L. Montreuil, A. Kliger, BZ Shen, SCTE Cable-Tec Expo 2014

[2] *Full Duplex DOCSIS PHY Layer Design and Analysis for the Fiber Deep Architecture*, R. Prodan, SCTE Cable-Tec Expo 2017

A Common Remote PHY Software Stack for all RPDs?

A Technical Paper prepared for SCTE by

Michael Robinson
Remote PHY Software Architect
Comcast Cable Communications
1701 JFK Blvd – Philadelphia, PA 19103
+1 (404) 723-7847
michael_robinson2@comcast.com

Jorge Salinger
VP, Access Architecture
Comcast Cable Communications
1701 JFK Blvd – Philadelphia, PA 19103
+1 (215) 439-1721
jorge_salinger@cable.comcast.com

1. Introduction

We have all heard the term “Software Fragmentation”, especially in the context of mobile device applications. Software fragmentation occurs because of the variety of hardware and software platforms. As a consequence, applications may not be compatible with that new hardware, forcing the need for application changes or the need for another application altogether.

The above also happens with Remote PHY Devices (RPDs). A single cable MSO will typically deploy RPDs from multiple suppliers, each with a different application software, and even multiple RPDs from the same supplier that have different hardware components and require different application software. Although RPD software design is based on common specifications, the behavior of the software differs between RPDs implemented with different hardware components and/or by different suppliers. As these differences accumulate, managing the system and maintaining interoperability becomes more complex.

How can this be mitigated?

A common software base can be shared between RPDs, even when they come from different suppliers. This paper will explain how sharing a common software base is implemented, and how it is applied to the RPD platform and supplier ecosystem.

2. Typical Cable Access Network

Most MSOs’ hybrid fiber-coax (HFC) networks have been designed with an upper spectral boundary of 750 or 860 MHz, while some are designed to support 1 GHz and other newer networks designed to support 1.2 GHz. For the more abundant 750 or 860 MHz networks, if not already fully utilized, it is expected that the use of their capacity will soon be increased to the point of exhaustion.

The increased utilization of this access network capacity has been driven by the success of cable MSOs’ service offerings.

As it is well known and understood, for many years the growth in, and demand for, more video programming resulted in the need to allocate large numbers of EIA (Electronic Industries Association) channels for video services, mostly known for their fixed, 6 MHz size, and used both for BC (Broadcast) and NC (Narrowcast). These 6 MHz EIA channels have filled every available portion of the spectrum.

Additionally, the success of, and growth in, HSD / broadband services continues. Most, if not all, operators offer increased service tiers and observed a growth in the use of HSD service capacity for well over a decade now, which amounts to a significant year-over-year compounded growth. This phenomenal success drives the need for increased network capacity, which is implemented by either expanding the spectrum allocation for HSD, or continuous service group segmentation, to reduce the number of users per service group. Or both.

In a separate but parallel trend, operators have been actively converging video and data services into a common Converged Cable Access Platform (CCAP) platform. This evolution, which has been underway for many years now, is intended to reduce the environmental requirements in HEs (headends) that result from service group segmentation. This is because as more service groups are added through segmentation, more HE equipment is needed, which drives the need for space, power and cooling. These trends imply an evolution towards newer, more modern, and denser equipment.

However, as the success of high-speed data and on-demand services continued, the evolution of the access network progresses towards further expanded capacity and ever-smaller service groups. For the

former, the spectrum allocated to narrowcast services increases, driving operators to deploy Data Over Cable Service Interface Specification (DOCSIS©) services including more SC-QAM (Single Carrier Quadrature Amplitude Modulation) channels, and more recently, with DOCSIS 3.1, the allocation of network spectrum for more or wider OFDM (Orthogonal Frequency Division Multiplexing) channels, as well as more narrowcast video services. For the later, more CCAP ports are needed, which drives the deployment of more line cards and eventually more chassis. These expansion trends result in a continuous growth of headend equipment, which is already starting to exceed the capacity that headend facilities can support.

3. Benefits of DAA

The above trends are now intractably linked to two additional evolutions: distribution of components of the access network, implementing a Distributed Access Architecture (DAA), and virtualization of the core network functions.

There are many benefits from the implementation of DAA. One key benefit is the improvement on performance, which is achieved in multiple ways, including: improved SNR characteristics, enable longer link distances between the headend and the nodes, provide higher service reliability, better use of capacity, etc. Beyond the performance improvements, a key benefit of DAA is the increased headend capacity. The implementation of DAA makes it possible to improve the density of CCAP devices in several ways, including the implementation of denser equipment, the use of Ethernet technology which is simpler and smaller in footprint, which more than doubles the capacity as compared to traditional CCAP chassis.

DAA can be implemented in many ways. As the optical link from the headend to the node is converted from analog modulated forward to digital, using Ethernet as the transport, several approaches can be taken for the implementation of the remaining headend components. One approach, for which a key goal is to convert all required components into functionally individual software pieces implemented independently, is to implement DAA in various discrete SW components as shown below.

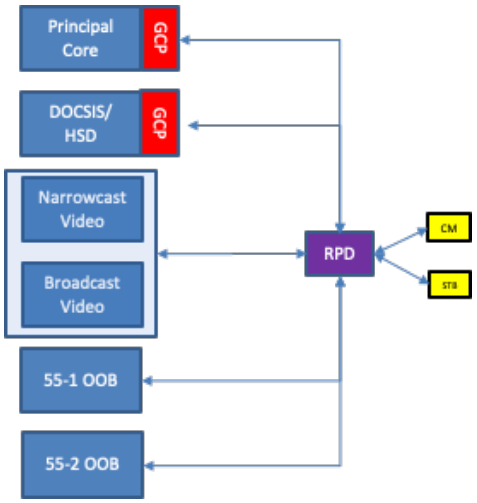


Figure 1 – DAA Implementation Components

The key components depicted in the DAA implementation above include:

- The Principal GCP Core, or GCPP, is the first component that the RPD will contact after receiving an IP address. GCPP is implemented such that it will configure all the RPD functions except DOCSIS channels and behavior, which will be implemented by the DOCSIS CMTS. As the name implies, GCPP communicates with the RPD using the GCP protocol. Included in the GCP Principal Core are all the non-DOCSIS command and control functions for the RPD, including configuration, management and reporting.

- The DOCSIS Core, which is the second component that the RPD will contact in the network. The DOCSIS Core also communicates with the RPD using GCP, and provides all configuration, command and control for DOCSIS channels, both downstream and upstream.

- Narrowcast and Broadcast Video engines, which can be implemented as separate components or combined into a single device, provide all the video content services for the various RPDs in the network. It is important to note that neither the Narrowcast nor the Broadcast Video engines communicate with, nor have knowledge of, the RPDs. Instead, services are configured statically in the Narrowcast and Broadcast Video engines upon their bring-up, and are multicast to all RPDs, which listen for these services as configured by the GCP Principal.

- Out-of-Band engines or cores are implemented separately from the video engines. Given that video systems are implemented using a single encryption and command/control technology, only one (i.e., either SCTE 55-1 or SCTE 55-2) of them is deployed in any one system. The OOB function may or may not implement GCP for communicating with the RPD. When GCP is implemented the OOB server is a Core, and it will configure the OOB downstream and upstream OOB channels in the RPD. However, when GCP is not implemented the OOB server is an engine, and the GCP Principal will configure the downstream and upstream OOB channels.

- Finally, not depicted in Figure 1, is a very important component: the Timing Server. Also known as the Grand Master, the Timing server provides the critical timing synchronization for all the DAA components. Each of the DAA components will include a Timing Client, which will communicate with the Timing Server to maintain timing synchronization. While timing synchronization is not absolutely critical for video services, it is imperative for DOCSIS service to operate. Therefore, video services may be initiated before timing synchronization is achieved, but DOCSIS services will not.

The key advantages for the above architecture are: implementation consisting of a multi-supplier platform where each component can be developed independently, smaller functional components with simpler implementation, and generally reduced time-to-market. However, implementation of smaller discrete components has its downsides, such as: the implied requirement to more tightly specify the behavior of each component to ensure that the overall system will operate as intended, management of the various components including their configuration and upgrade, and the need for more elaborate orchestration.

4. Common Issues Faced with RPD Deployments

The base implementation of a DAA system is generally simple. However, significant complexity is introduced when interoperability with different suppliers' components is introduced.

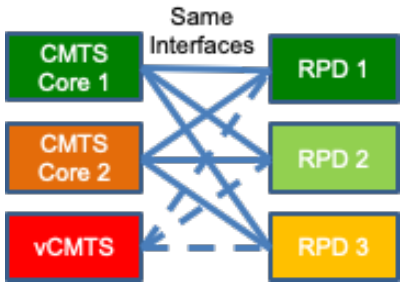


Figure 2 – Functional CMTS-RPD Interoperability Matrix

When considering the overall CCAP system, the complexity to achieve multi-supplier interoperability is even larger. Over the years, larger MSOs have deployed multiple CMTSs, and eventually multiple CCAPs, experiencing the complexities associated with such equipment implementation interoperability.

As depicted in Figure 2, the number of combinations of interoperable components increases geometrically as additional components are added on either side of the interoperability matrix. Having a single CMTS to interoperate with multiple suppliers' RPDs is complex and requires a lot of careful planning and implementation. But if the number of CMTS implementations is increased to 2 or 3, the interoperability complexity doubles and triples respectively. Furthermore, in the case of DAA, if multiple GCP Principals and/or multiple DOCSIS cores, and/or multiple video engines, and/or multiple OOB engines/cores are introduced into the mix, the amount of complexity and work required for lab and field testing to maintain interoperability increases by orders of magnitude.

Therefore, a multi-supplier RPD deployment coupled with a single headend implementation is a sensible approach to an interoperable DAA ecosystem. In this way, the HE components of the DAA, which are deployed in the hundreds, are kept the same for all MSO sites, but the component that is deployed in the thousands, tens of thousands, or even hundreds of thousands, are obtained from multiple sources, ensuring an innovative and competitive ecosystem.

Finally, it is important to recognize that there are numerous types of HFC nodes and network use cases, which will drive the need for an even larger variety of RPD types.

In the same way as there are different kinds of nodes for different HFC network applications, there are also Remote PHY devices with different characteristics that are best suited for each of the specific HFC network use cases. For example, nodes that are best suited for N+x network architectures may have different RF characteristics, and their corresponding RPDs may be implemented to support more service groups. By contrast, N+0 network architectures have different node RF characteristics and require RPDs that are intended for fewer subscribers.

Similarly, while there are use cases for RPDs in the outside plant, there are also applications for RPDs in headends, or "inside plant" as it is frequently called, which will have different implementation characteristics. This use case diversity has driven suppliers to offer RPDs packaged for nodes, where there is a single RPD, or in some cases 2 RPDs, and the only opportunity for cooling is through convection by contact with the outside node enclosure, while other RPDs are packaged in shelves, where the number of RPDs is much larger, even requiring the need to support line-cards, and it is possible to implement cooling through forced air movement.

Therefore, while it is possible to maintain the HE equipment constant, the variety of MSO's HFC networks and application use cases, plus the desire to maintain a multi-supplier RPD ecosystem, drives the need for a large number of RPD types (e.g. upstream/downstream port counts and frequency splits), models and suppliers.

5. Solutions Attempted To Date

Over the last few years, as the development of RPDs proceeded through more use cases and newer models, suppliers have made every effort to reduce the number of implementations, and to the extent possible maintain a single RPD design. While the RPD hardware implementation is different between RPDs intended for different use cases, it is especially beneficial to maintain the same software base. To that end, most suppliers have tried to implement a single software base for their RPDs.

This is possible while the key hardware components used in the implementation of RPDs are the same or are of the same hardware generation. However, when the hardware implementation is different, it is no longer possible to maintain the same software base. Such is the case for RPDs that were implemented initially with discrete hardware components, and eventually were migrated to functionality integrated transceiver-type hardware devices, it is especially the case when migrating to generations of RPD hardware components that implement newer versions of the DOCSIS specifications, as is the case now with the advent of DOCSIS 4. In such cases, a change in the software base becomes inevitable.

To mitigate the need for software diversity, an approach was launched within the industry, led by CableLabs® and supported by a multitude of suppliers and operators, known as OpenRPD. The approach, implemented as an open-source project, consisted of developing an application layer software, which interfaced with the RPD hardware via a hardware abstraction layer and device drivers, and operated above a lower-level kernel software.

While all the above efforts were intended to maintain and/or arrive to a single software base, various ecosystem situations and changes in the hardware approach made it difficult to succeed.

6. Collaborating on a Single Software Solution

Given our industrial experience throughout the evolution described above, it became clear that many of the issues stemmed from the differences in the software implementation of the standard. The variability of behavior increases the complexity of managing and debugging the devices. So how could this be mitigated? The answer is, once again, to move to a single RPD software base. To accomplish this, a common RPD application software program was created to include all suppliers developing RPDs for the operator.

The idea of RPD suppliers sharing a common software base is not new. As described above, CableLabs hosted a similar program known as OpenRPD. The OpenRPD program allowed participating suppliers to access a common software base as well as submit changes for features and fixes. These suppliers showed early successes during the Remote PHY interoperability activities at CableLabs. Issues were certainly found during these activities, as expected. But when solutions to these issues were found, all participants reaped the benefits.

6.1. Chosing the Target Hardware

Before beginning the architectural design of the software, it is important to decide on the initial hardware that will be supported. This decision will have a large impact on the software development scope of work.

The primary question here is, should the currently deployed devices be considered in the software implementation? Or should the focus of the software be on only on new hardware development? Targeting both existing and new hardware would certainly reduce the mix of implementations to support, which is the goal of the project. But at the same time, it would greatly increase the scope of work necessary to deploy the software.

On the other hand, supporting only new hardware development could simplify the software architecture by leveraging the design of the new generation of integrated silicon solutions. Additionally, the new devices would soon be replacing the previous generation in the field. And because of these reasons, it was decided that focusing on the next generation of RPDs was the way to go.

6.2. RPD Software Architecture

Proper architecture of the software is critical when supporting multiple hardware targets. It is important to separate the main components are shared and those that will differ between the various platforms. Figure 3 below shows the basic architecture of the RPD software:

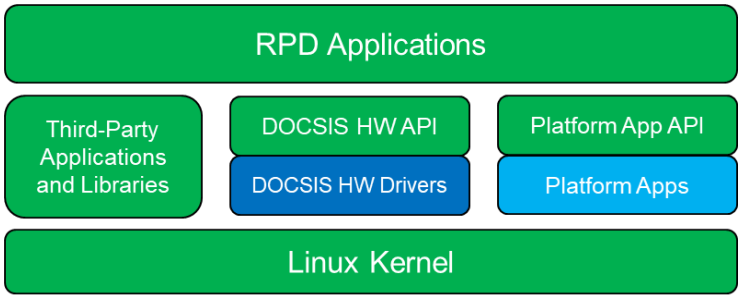


Figure 3 – RPD Software Architecture

The above blocks highlighted in green represent the functional components that are shared across all target platforms.

- **RPD Applications**: This is the collection of applications that provide the primary control interface and orchestrate control operations such as device configuration and monitoring.
- **DOCSIS HW API**: This specifies the programming interface for the DOCSIS-specific hardware. This includes operations such as DOCSIS channel and data-plane configuration.
- **Platform API**: This specifies the programming interface for the platform-specific hardware. This includes operations such as amplifier and attenuation configuration, LED control, timing configuration and monitoring, etc.
- **Third-Party Applications and Libraries**: This includes dependencies for features such as DHCP, SSH, 802.1x, etc.
- **Linux Kernel**: The distribution provides the supported Linux kernel version.

The blocks highlighted in shades of blue are specific to the hardware on the target platform.

- **DOCSIS Drivers**: These drivers conform to the interface defined by the DOCSIS API. They perform the low-level operations necessary to program the DOCSIS-specific hardware. Platforms that use the same DOCSIS hardware solution share the common drivers from the hardware provider.
- **Platform Applications**: These applications conform to the interface defined by the Platform API. These applications are specific to the each of the suppliers' hardware designs.

7. Managing the Software

The common RPD software will be used by multiple RPD suppliers targeting multiple platforms. This introduces some complexity when considering how the software base will be managed. Some of the questions raised are:

- Which build framework should be used to create the OS distribution?
- How should the software repositories be organized?
- How can each supplier separate their IP from the common software?

7.1. Deciding on the Build Framework

There are many choices when it comes to managing the project's build framework and operating system distribution. Some of the more popular frameworks for embedded systems are:

- Buildroot
- Yocto
- OpenWRT
- "Roll your own"

As expected, each of the choices come with its own advantages and disadvantages, briefly summarized below.

Buildroot is known for its simplicity. This means developers new to the framework can get started relatively quickly. However, precisely because of its simplicity, a significant amount of customization may be required to support many platforms within a single project.

The Yocto project, on the other hand, was designed with flexibility in mind. Yocto comes ready with a vast library supporting a large number of platforms. But with this flexibility comes complexity. Yocto is known for its steep learning curve.

The OpenWRT Project's primary focus is building firmware for commercial devices such as routers. It can be used for other types of embedded devices as well, but stepping away from its design for routers means additional customization.

Creating a custom distribution without a third-party framework is another option. This is something to consider if a project has very specific requirements that existing solutions do not provide. But this can introduce more overhead in maintenance when it comes to tasks such as keeping security patches up to date and resolving dependencies for upgrades.

Considering the choices above, the Yocto project was selected for managing the RPD program's OS distribution. In the end, it was decided that the time spent on learning the tools would be well worth the flexibility and hardware support required by the program.

7.2. The Yocto Project: Layers and Recipes

So, what makes the Yocto project flexible?

Let's start with what Yocto refers to as "recipes." A recipe is basically a script that provides instructions on how to build a particular entity (e.g. application, library, file system image, etc.). In its simplest form, a recipe for an application can just provide the location of the source software. Yocto can determine how to build the application and install it as long as the source is using one of the common automated build systems (e.g. autotools, cmake, meson, etc.).

Recipes are located in directory trees where the top directory is referred to as a "layer." A layer is a collection of recipes and configuration files. A single layer typically encompasses a particular category of recipes. For example, one layer might include recipes for shells, or even common RPD software applications. A typical project will have multiple layers, but will not necessarily use every recipe in every layer.

7.3. Organization of Layers in the RPD Project

The RPD project contains four categories of layers:

- **Third-Party Applications Layer**: This is actually several layers. The RPD project pulls in the existing layers from Yocto that provide the third-party open-source applications and libraries such as DHCP, SSH, and 802.1x. These layers are shared among all RPD projects.
- **Common RPD Applications Layer**: This layer contains the recipes for all of the common configuration and monitoring applications for the RPD. In addition, recipes for the DOCSIS HW and Platform APIs are provided. This layer is shared among all RPD projects.
- **SoC Applications Layer**: These layers contain the recipes for the various system on a chip (SoC) solutions available. They provide recipes for the required OS kernel and the DOCSIS HW drivers. Each RPD project selects the SoC layer associated with the hardware chosen for the target platform. Access control is necessary to restrict its use to suppliers that have the associated license agreements with the SoC provider.
- **OEM Applications Layer:** These layers contain recipes for applications, libraries, and drivers specific to the target device. Access control ensures that this layer is restricted to the associated supplier.

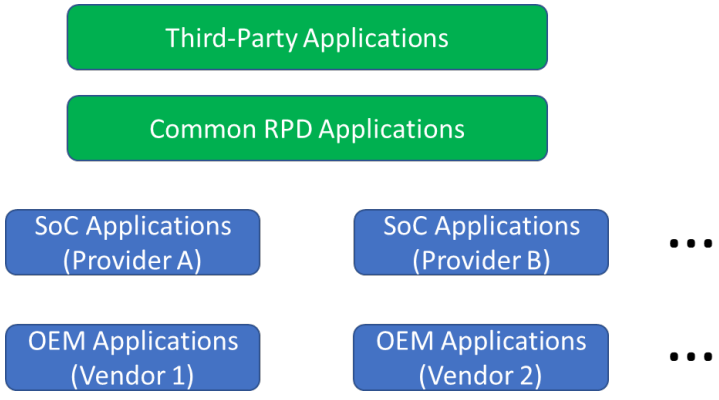


Figure 4 – Yocto Layers for the RPD Project

As described above, some of the layers are shared while others are selected based on the needs of the target device. For example, two different suppliers could use two different SoC solutions in their respective RPDs. Figure 5 below illustrates the combinations of layers required for an RPD developed by supplier "1" using SoC "B":

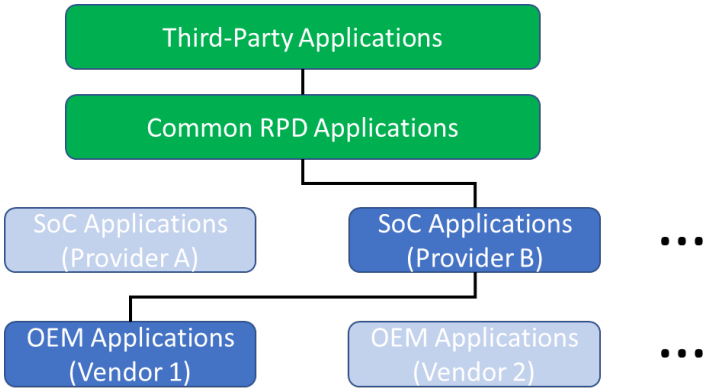


Figure 5 – Example Selection of Yocto Layers

7.4. Managing the Layers

As explained in the previous section, the Yocto layers are useful for organizing the build recipes into logical categories. These categories make sense from a software architectural standpoint. But they are also helpful when it comes to managing software changes and access control.

The layers in the RPD project have different access requirements depending on which category they fall into. For example, the common RPD applications must be accessible to all developers, but a specific supplier's OEM layer must not be available to other suppliers. To handle this requirement, each layer is placed in a separate repository. The proper access permissions can now be set uniquely for each repository.

One concern with this setup may be that developers now have to deal with a large number of repositories just for a single target. It is certainly true that the list of required layers can grow, especially given that just the third-party applications alone can be spread over many layers. But this is easily handled by using an application such as Google's "repo" command. Tools like this can take a single manifest that points to the collection of repositories for a project and manage any bulk operations such as checking out the software, branching, committing, etc. Operations on individual repositories can be performed as usual.

With the access controls in place for each of the repositories, this naturally decides which developers can see and make software changes as well as participate in software reviews for each layer. For example, only the suppliers with licensing agreements for a specific SoC can participate in the development for the associated layer. Likewise, all parties can participate in the development of the common RPD software layer.

8. What Comes Next?

Features continue to be added to the Remote PHY specifications that will need to be supported. The next big feature on the roadmap is streaming telemetry. This is a paradigm change from the pull model

supported today. Adding the vast number of statistics required for monitoring field deployments is no small task. But the results of this effort can be shared among those involved with the project.

Another task on the horizon is virtualization of the RPD control plane. Running the RPD software without the underlying hardware can have multiple uses. A virtual version of the RPD can be used for automated integration testing in the build environment. Additionally, initial development testing for some features can be started with the virtual RPD, reducing the time spent on lab resources.

Overall, the roadmap may be similar to other RPD software programs. The big benefit with this program is that the fruits are shared by all.

9. Conclusion

Managing a large deployment of Remote PHY devices is a challenging task. Problems will arise in any large network. Troubleshooting these issues become more complex when the devices differ in behavior, debug capabilities, and supported features. Moving to a common software base certainly will not solve all issues typically seen in deployments, but standardizing the implementation of the management interface will simplify the configuration, monitoring and troubleshooting of RPDs by reducing the supplier-specific behavior. Accomplishing this goal is possible with a software architecture and development environment that supports multi-supplier collaboration.

Abbreviations

RPD	remote PHY device
RPHY	remote PHY
DOCSIS	data-over-cable service interface specifications
API	application programming interface
PHY	physical layer
SoC	system on a chip

A Comparison of the Energy Consumption Properties of Wi-Fi Backscatter and Bluetooth Devices as it Relates to Sensor and Asset Tracking Solutions

A Technical Paper prepared for SCTE by

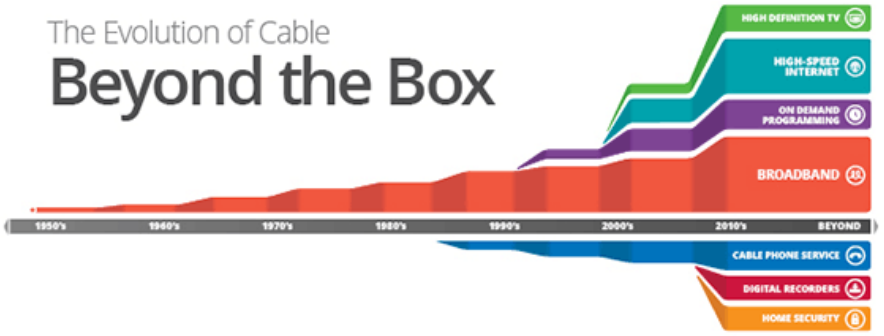
Ty Pearman
Principal Engineer
Comcast Labs, Comcast Cable
4100 East Dry Creek Rd
Centennial CO 80122
ty_pearman@cable.comcast.com

Arun Ravisankar
Principal Engineer
Comcast Labs, Comcast Cable
1800 Arch St, Philadelphia PA 19103
215-286-7558
Arun_Ravisankar@comcast.com

1. Introduction

The airwaves around us are continually filled with the invisible signals emanating from and between the Wi-Fi gateways and devices that use them for broadband connectivity, as well as from and between devices that use Bluetooth for connectivity. Naturally, techniques like Radio Frequency Identification (RF ID), which date back to the 1970s, were developed to use those radio waves for object identification. Commonly, you will see RF ID tags affixed as stickers or built into consumer devices, for the purposes of being read at a distance and without line of sight, as with highway toll readers, employee ID badges, etc. The ambient nature of these signals provides the opportunity to "piggy back" on the existing networks, which can reduce the overall power consumption of RF backscatter solutions. This paper examines the performance tradeoffs between different communication protocols used in both sensor and tracking solutions, including RF backscatter, Wi-Fi, Bluetooth, and LoRa (Long Range). The main area of focus will be around power consumption and data throughput for IoT type solutions. Readers will learn the basics of energy consumption and data transmission rate capabilities from these technologies, along with the power consumption baselines of mainstream IoT devices and how transmission protocols can impact the powering and life of the IoT devices.

Cable network operators are always looking for ways to add services for their customers, especially so since the 1990s. A couple of examples include: Data over cable, voice, and DVR(Digital Video Recorder) evolving to nDVR, were added to the service bundle. More recently, home monitoring and security services have also been offered by many operators. [1]



(Calcable.org, n.d.)

Figure 1 - Evolution of Cable

The Internet of Things is loosely defined as Internet-connected sensors in homes, businesses, and public spaces, as well as the data analytics monitoring of those sensors back in the data center. With the Internet of Things, there is an opportunity to rapidly open up entirely new service opportunities that can differentiate cable network operators from their competition. However, the primary challenge will be to smoothly install, operate and integrate these new services with the operator's existing service bundle.

Cable network operators are uniquely positioned to offer IoT services to new and existing customers. They have four characteristics that industry start-ups and OTT service providers covet:

- Existing service location in millions of homes, businesses, and public spaces

- High speed and reliable network connectivity
- Power for sensors and gateways
- An existing and localized/in-market fleet of fulfillment technicians

Cable network operators have a well-established presence in the home including cable modems, home gateways, set top boxes, Wi-Fi extenders and home security hubs, however the evolution to new services - such as connected healthcare, and smart homes – will require new devices and sensors, as well as increased care to ensure the highest network performance while preventing security breaches.

Where is the world of IoT going? According to a market report from Research and Markets "The total number of IoT connection worldwide will increase from 1.5 billion at the end of 2019 to 5.8 billion in 2029". Using that as a data point one can assume these devices will continue to show up in our daily lives in various locations, in a multitude of forms. Many of these devices are part of larger systems or add-ons to existing systems. In those cases, the powering of the devices come from wired sources or systems recharged on a regular basis. Additional implementations consist of stand-alone devices used for monitoring or tracking. These devices tend to be small and powered by batteries (think of door sensors, window sensors, outdoor sensors for temperature, humidity, noise…etc.). One of the biggest issues with the stand-alone IoT devices is maintaining power to the devices so they can do their job.

For systems or environments with large amounts of sensors, replacing batteries can be an endless task, making the labor cost to replace batteries or devices a non-starter. Just think about how often you swap out your remote batteries and multiply that by 5, 10 or even more for your home IoT devices. Now take that equation and move it to an outdoor or warehouse situation and the problem explodes one or two orders of magnitude. With that said, many IoT vendors target their designs to a single battery that will last the lifetime of the device, which can be anywhere from 3 years all the way out to 10-15 years in a low power wide area network solution (LPWAN).

So how can we extend the life of these devices? Doing so could be a game changer, depending on the application and the amount of data that is being sent out. There are a couple of ways this can be done: use larger batteries, leverage energy harvesting technologies to replace the batteries or extend battery life tend to be the most obvious. However, another interesting option is changing the data transmission protocol used to transmit the data to the access point/base station.

2. Communication Protocols used in IoT(Internet of Things)

IoT devices are ubiquitous and are part of our everyday life, from the smart home devices we use to the health and wellness devices that keep track of our overall health. Have you ever wondered what makes these devices smart? The simple answer is that these devices are now capable of communicating with a computer/machine/internet and provide critical data for algorithms and applications.

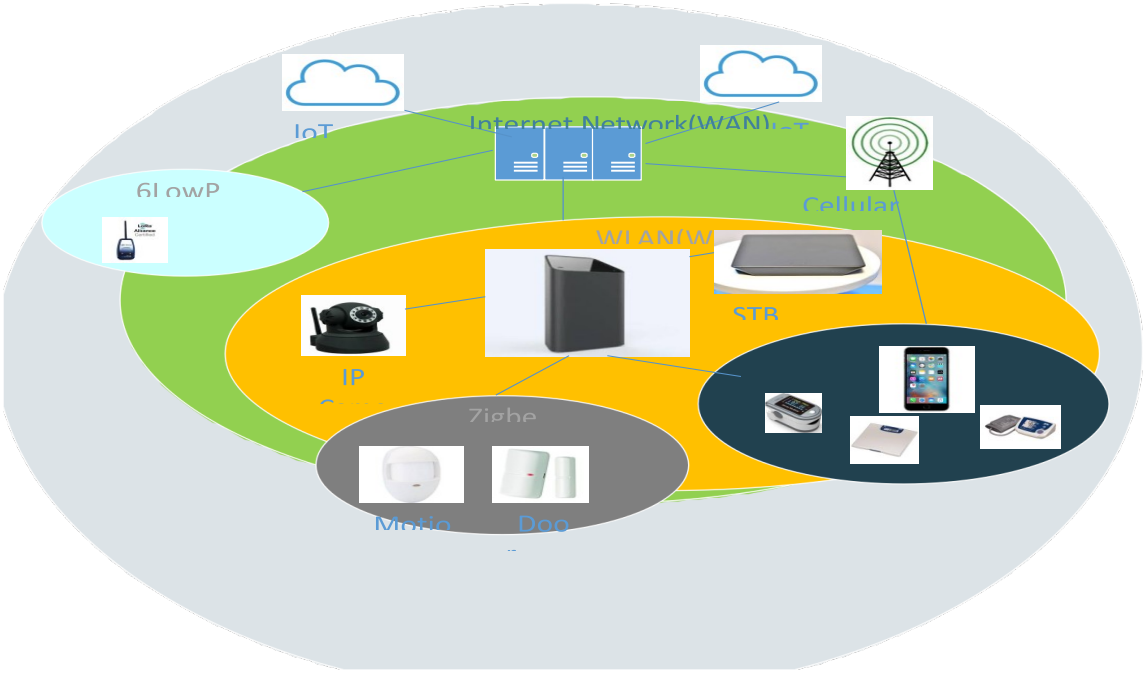


Figure 2 - Devices and Networks used on IoT Applications

Figure 2 shows multiple devices using different communication networks and protocols. The choice of the protocols and networks depends on the type of data involved, range of the sensors, and the applications these devices support. Some networks and protocols support data transfer at longer distances, while other protocols support larger data transfer rates at low range(distance). Depending on the range, these networks could be classified as WAN(Wide Area Network) or PAN(Personal Area Network). PAN includes networks like Bluetooth, Zigbee, Wi-Fi, which support very high data transfer rates, but their range is limited to probably a few meters. On the other hand, WAN include networks and protocols like LoRaWAN, NB-IoT, Mobile(LTE/5G), Cable/Fiber broadband, and can support connectivity at larger distances. Table 1 provides some key details of each of these network protocols. The sensor that uses these networks need to meet the power requirements of the network/protocol, as the sensor might need to engage the antenna and provide enough power to transmit and receive the data. Often, the power characteristics would depend on the data being handled and the distance the data is being transmitted.

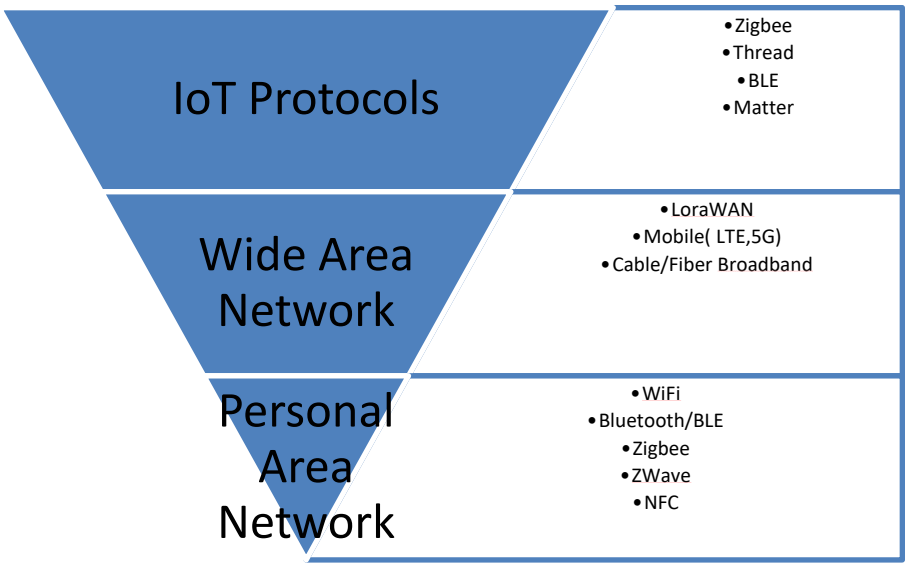


Figure 3 - Components of a typical IoT Application, Networks and Protocols used

As the penetration of these sensors/devices increase, it is critical to note the power consumed by these sensors, both from an efficiency and a sustainability perspective. Most of these sensors are often powered by batteries and the users would ultimately need to replace batteries on these sensors. This is a significant challenge in terms of a sustainable ecosystem. This paper compares the power characteristics of the different protocols and would seek methods to make the IoT ecosystem more sustainable. By understanding the power characteristics and requirements for these networks and protocols, we can explore augmenting the power from sustainable sources, such as ambient light. This would improve the battery life of the sensor, making it more sustainable and provide better customer experience.

Table 1 - Table showing operating conditions of IoT networks

	Personal Area Networks						*Wide Area Networks*	
	NFC	**BLE**	**Wi-Fi**	**Zigbee**	**Zwave**	**WiFi HaLow**	**Mobile(LTE, 5G)**	**LoRa**
Operating Freq	13.56 MHz	2.4 GHz	2.4, 5 and 6 GHz	2.4 GHz	908.42 MHz	< 1GHz	450 MHz to 6GHz and 24.25GHz to 52.6 GHz	902-928 MHz
Max Power	Low power needed	1mW – 100 mW	100mW – 4000mW	10mW – 100mW	10mW – 100mW	10mW – 100mW	Varies as per endpoint	
Battery Operated?	Yes	Yes	Possible	Yes	Yes	Yes	Yes	Yes
Range	10 cm	<100m	Around 50m	<100m	100m	Approx 1Km	Range varies. Up to a few Km. 5G bands offer very low range compared to LTE	Between 15 to 20 Km
Throughput	424 Kbps	2Mbps	WiFi ax max – 3.5Gbps	250kbps	40kbps to 100kbps	150kbps to 4Mbps	50Mbps to 10Gbps	300 bps to 37.5 kbp
Ease of Adaptation	Wide Adoption	Wide Adoption	Wide Adoption	Wide Adoption	Medium Adoption	Low Adoption	Growing Adoption	Low to Medium Adoption

3. Power Characterisctics

Table 1 provides us with some details on the metrics on the networks and protocols used. In this section, we will examine how some of these protocols are efficient with regard to power consumption.

As the world around us gets more complicated, the flow of information grows in importance as we become more connected. Someone is always checking their Facebook feed, e-mail, doggie cam, receiving alerts from their home systems, finding their car, checking the air quality, traffic speeds…etc. As the flow of information becomes increasingly more important for our daily lives, the need for these systems to be connected and always available becomes critical. The devices that make it all possible need to rely on some type of power source to keep us “in the know”.

The easiest power solution for all these are the devices already connected into the regular power grid. With wired solutions, the power usage isn't a big concern, other than a power outage, which can be handled by some type of battery backup, if necessary. Other solutions that tend to be more mobile (the main one being a smartphone) have rechargeable batteries that give the device portability but offer a very limited life before the user must recharge the system. The final case, which we are most interested in, is the IoT case where devices are deployed in a stand-alone fashion and expected to function for an extended period normally measured in years as opposed to days. This is where changes in power consumption can have enormous impacts on the life of the device.

There's a couple of ways to break down the power characteristics of the different devices and protocols. In general, the farther you transmit a signal, the more power you will need; the more data you transfer, the more power you will need; and the longer the device is "awake" and transmitting data, the more power you will need. This addresses the technologies employed with a very broad stroke, but with that in mind, we can break the different solutions down based on their reach, reliability, throughput, and power consumption. Each of these different parameters really speaks to the notion of there not being a "one size fits all" IoT solution.

In general, IoT devices have two main purposes: the first is to collect data and the second is to forward that data to a centralized system where the data can be stored, analyzed, and/or acted upon. Looking at the power budget for the endpoint, a large portion is expended in the transmission of the data back to the gateway. These networks consist of many endpoints to very few gateways. In some cases, a gateway can support up to and potentially over 1 million IoT devices. Here's some of the different networks these IoT devices can be deployed in:

3.1. Personal Area Networks(PAN)

PAN is a network that interconnects computers within a limited area such a residence or a small business facility. Primarily these networks were intended to be the data pipe between computers and modems and hence the focus was more on data throughput. As spaces like residences and businesses grew, the need to provide reliable data connectivity at larger distances became important. This resulted in increased power output on the radios. With the advent of IoT devices and networks, devices that are powered by batteries grew and this put a natural limit on the power we could transmit. There have been significant advancements on the power efficiency of these networks that consume less power, but still manage to support better data rates.

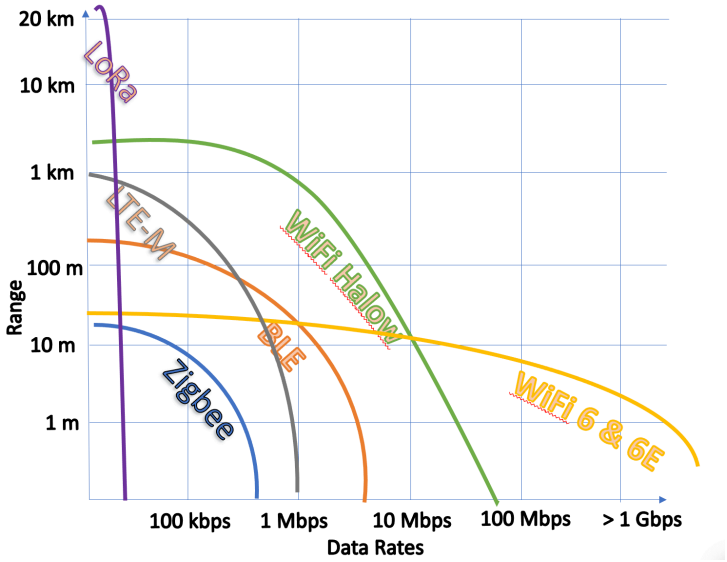


Figure 3 - Range vs Data rates of Wireless PAN

As shown in Figure 3, we see that the data throughput of the networks tends to drop down as the distance between two endpoints increase. The data rates supported at different distances depends on the protocol used and on the specific implementation of the standard. As protocols are pushed to the edge, the need to provide reliable connectivity often results in increasing the output power, which in turn puts a lot of stress on the endpoints, especially if they are battery-powered. Even within range, the pain point in managing battery life of multiple sensors and devices leads to an unpleasant customer experience. Hence, this paper looks at the power characteristics of these network protocols and where power performance can be optimized if we can augment the battery power with energy harvested from ambient power sources, like light, temperature differential, and power over RF. We will look at the power characteristics of some protocols in this section.

- ***BLE (Bluetooth Low Energy)***
  Compared to Classic Bluetooth, Bluetooth Low Energy is intended to provide considerably reduced power consumption and cost while maintaining a similar communication range. Mobile operating systems including iOS, Android, Windows-Phone and BlackBerry, as well as macOS, Linux, Windows 8, and Windows 10, natively support Bluetooth Low Energy.

  A typical BLE SoC (i.e. an all-in-one Application + Radio chip) typically consumes:

  - A few hundreds nA(Nano Amp) while in deep sleep,
  - 2 to 10 µA while a RTC tracks time (needed between radio events while advertising or connected),
  - 10 to 30 mA while CPU or Radio runs (computing data, TX, RX). RX and TX power consumption is roughly the same.

  The life of a BLE peripheral basically consists of 3 main states:

- Be idle (not advertising, not connected). The device would be in a sleep/standby state and consumes just over a few hundred nanoamps though.

- Advertise (before a connection takes place). Peripheral needs to be running approximatively 5ms every 50ms. This is the time when your device actually uses the most power because advertising requires sending many packets, frequently. Average power consumption is in the 1-10 mA range.

- Be connected. Here, consumption is application-dependent. If application is mostly idle, a peripheral is required to wake up periodically and must send a packet each time in order to keep the connection alive. Even if the peripheral has nothing useful to send, an *empty packet* is still sent. Side effect: that means low duty cycle applications basically transmit packets for free.

We could estimate the battery life of a BLE device using a typical BLE example:

- Radio power consumption = (5.4uA * 3V) * 1/2s = 8.1uW
- System power consumption = (2uA * 3V) = 6uW
- Total power consumption = 14.1uW
- Battery life = ((300mAh * 3V) / 14.1uW) * 0.7 = 44681 hours = 1860 days

If we can augment the battery with an energy harvesting source that can supply enough power for the SOC, we could either make the sensor battery free or extend the battery life.

- ***ZigBee***

  ZigBee is a mesh network protocol designed to carry small amounts of data across medium distances. It runs on a mesh topology network, meaning information from a single sensor node, travels across a group (or "mesh") of nodes until the transmission reaches the gateway.

  ZigBee is a local area network (LAN), so unlike BLE, it is not intended to connect to devices directly around a user. Instead, it connects to devices that need a wider range. Because of this, it's an ideal protocol for home automation[1] and smart lighting[2].

  As an example, battery life of a motion sensor with a ZigBee radio could be around 5-7 years, which is comparable to that of BLE

[1] https://www.link-labs.com/applications-of-home-automation/
[2] https://www.link-labs.com/smart-lighting/

- ***WiFi HaLow***
  WiFi HaLow is IEEE 802.11ah standard that uses the WiFi protocol characteristics, using a sub-GHz band. WiFi HaLow promises
    - Lower Power Consumption
    - Longer Distances
    - Better Penetration
    - Lower data rates (150 Kbps at longer range, 10+Mbps at closer range)

  Figure 4 shows the frequency bands used by WiFi Halow in the WiFi standards.

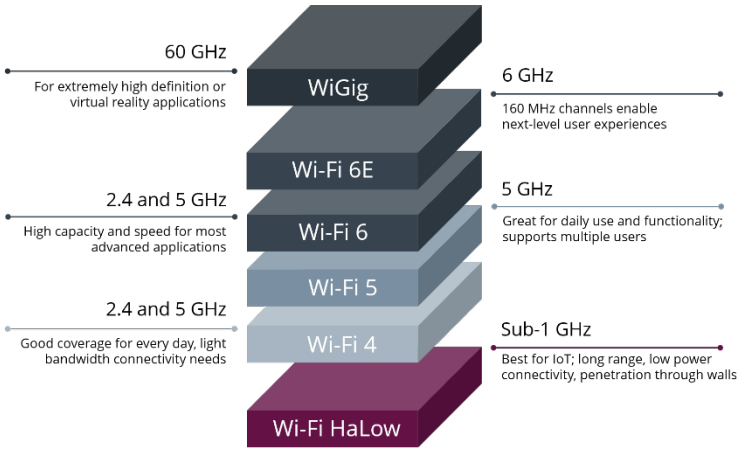


Figure 4 - Wi-Fi Frequency Bands

3.2. Wide Area Network (WAN)

(In the context of this paper, when talking about WAN, we are referencing wireless WAN solutions.)

Some of the most compelling deployment use cases for IoT revolve around the deployment of sensors and asset tags spanning large geographic reach (when the term WAN is used, a range of more than 2km is a good starting point). There are a multitude of use cases for these devices, but in general they can be categorized into a handful different areas…1) highly reliable data with large data throughput requirements, 2) small amounts of data that are not time sensitive and 3) small amounts of data requiring a high level of reliability. Additionally, from a technology perspective, there are two different solutions that are usually discussed for these transmission scenarios; cellular and non-cellular (sometimes referred to as licensed and unlicensed spectrum). Below is a chart that outlines some of the different characteristics of these systems

Table 2 - Characteristics of Cellular and Non-Cellular LPWAN Networks

Parameter	CAT-1	CAT-M1	NB-IoT	LoRa	SigFox
Bandwidth	1.4 - 20 MHz	1.4 MHz	180 kHz	125 kHz or 500 kHz	100 Hz
Data Rate	DL: 10 Mbps UL:5 Mbps	DL: 1 Mbps UL: 1 Mbps	DL: ~20 kbps UL: ~60 kbps	.3 – 50 kbps	100 bps
Latency	50-100 ms	10-15 ms	1.6 – 10 s	Topology dependent	Topology dependent
Spectrum	Licensed	Licensed	Licensed	Unlicensed	Unlicensed
Peak Transmit Power	23 dBm	23 dBm	23 dBm	20 dBm	22 dBm

- Cellular (Licensed spectrum)[3]
  Cellular WAN solutions have evolved with the changing cellular technologies, but the most common ones are CAT-1, CAT-M1 and NB-IoT. All three solutions leverage licensed cellular bands which provide a more controlled environment for the transmissions which give the data transmissions a higher reliability than networks using unlicensed spectrum.

  - CAT-1

  In the world of LPWAN CAT-1 would be considered a power hog but has some of the highest throughputs and lower latency.

    - ◊ Latency typically 50-100 ms
    - ◊ Throughput of 5 Mbps uploads and 10 Mbps downloads
    - ◊ Spectrum bandwidth of up to 20 MHz

[3] https://en.wikipedia.org/wiki/Narrowband_IoT

- ◊ Typical use cases include video surveillance, ATM communication and vehicle telemetry

- CAT-M1(LTE-M)

CAT-M1 or LTE-M uses existing LTE networks but consumes far less battery power than CAT-1 devices. The solution has the ability to handle moderate bandwidth applications with minimal latency requirements.

- ◊ Latency typically 10-15 ms
- ◊ Throughput of 500 kbps uploads and 1 Mbps downloads
- ◊ Spectrum bandwidth of up to 1.4 MHz
- ◊ Ability to handle cell tower handoffs
- ◊ Typical use cases include wearables, high value asset tracking and health monitors

- NB-IoT[4]

This protocol is designed for devices that transmit small amounts of data with fairly loose latency requirements. The solution still leverages the LTE network, but operates in the roll off of the licensed spectrum.

- ◊ Latency typically 1.6 – 10 seconds
- ◊ Throughput of 120 kbps uploads and 160 kbps downloads
- ◊ Spectrum bandwidth of approximately 180 kHz
- ◊ Typical use cases include smart gas, water and electric meters along with smart city applications such as street lighting and parking sensors
- ◊ Stationary application since it cannot handle cell handoffs

Table 3 - LPWAN Network Characteristics

	Data reliability	Data throughput	Power consumption	Endpoint mobility
Grid Powered Cellular	High	High	High	stationary
Battery Powered Cellular	High	Low	Medium	
LPWAN/NB-IoT	Low	Low	Low	broad

- Non-Cellular (Unlicensed spectrum)

In the unlicensed LPWAN space the main technologies used are LoRa and Sigfox. Both solutions are focused on data uplink, where a large number of sensors transmit data to a small number of gateways in a star type topology. With these solutions using unlicensed spectrum, interference in the form of noise and other signals is a much larger concern.

[4] https://www.siretta.com/2021/02/lte-cat-1-vs-lte-cat-4/

Data collected and used in these systems in most cases is slow changing and not latency dependent. Occasional missed data packets shouldn't impact the overall system. Applications like temperature monitoring, air quality and agricultural data are good candidates for both solutions.

- LoRaWAN or LoRa[5]

LoRa networks can be deployed in a very low-cost manner with a minimal number of gateways to support the network. Unlike licensed LPWAN solutions, network gateways are easily deployed and managed. With a focus on low power LoRa endpoints can see device lifetimes from 10-15 years depending on the data being transmitting, frequency of data updates and distance the sensor is from the gateway.

  - ◊ Latency of system is dependent on the number of packets a gateway can process in a day related to the number of endpoints being supported
  - ◊ Throughput of .3 kbps to 50 kbps
  - ◊ Easily transmits through physical barriers
  - ◊ Spectrum bandwidth of 125 kHz or 500 kHz uplink and 500 kHz downlink
  - ◊ Star on Star network topology

- SigFox[6]

SigFox uses a proprietary technology which is optimized for extended distance over LoRaWAN, where up to 10 km is urban and 40km is rural areas, where LoRaWAN would only be capable of roughly 5 km in urban and 20 km in rural areas.

  - ◊ Latency of system is dependent on the number of packets a gateway can process in a day related to the number of endpoints being supported
  - ◊ Throughput max of 100 bps
  - ◊ Easily transmits through physical barriers
  - ◊ Spectrum bandwidth of 100 Hz for the uplink and 600 Hz on the downlink
  - ◊ Star network topology
    - ◊ Limited to 140 messages a day for the uplink and 4 messages for the downlink

[5] https://lora-alliance.org/about-lorawan/

[6] https://www.survivingwithandroid.com/sigfox-protocol-network-architecture-iot-protocol-stack/

4. Energy Harvesting

Harvesting energy for IoT applications can done using a variety of sources. While some include using the ambient environmental conditions to harvest energy, other methods include transmitting energy using mediums like WiFi or RF. Candidates include sensors used in IoT applications such as temperature, humidity sensors, to motion sensors and cameras. The key here is to find the right energy source that can either power the sensor in a battery-free mode, or trickle-charge the batteries to increase the overall life of the sensor.

Figure 5 shows an example of a system that is capable of harvesting energy to augment the battery of a sensor using the BLE communication system.

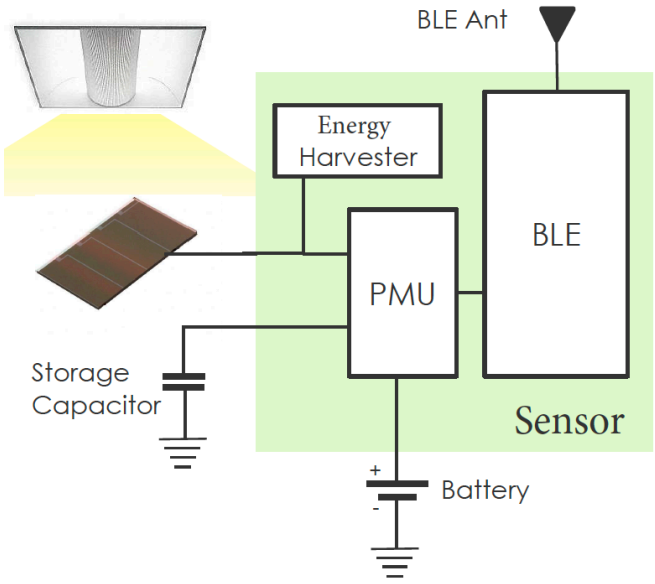


Figure 5 - Typical Energy Harvesting system

There are a variety of energy sources from which we could harvest energy. We will examine a few of the most application or common energy harvesting technologies next. Figure 6 illustrates how different energy harvesting technologies could support applications such as sensors, asset tracking application and many others. Table 4 shows the power density of the each of the energy harvesting technologies.

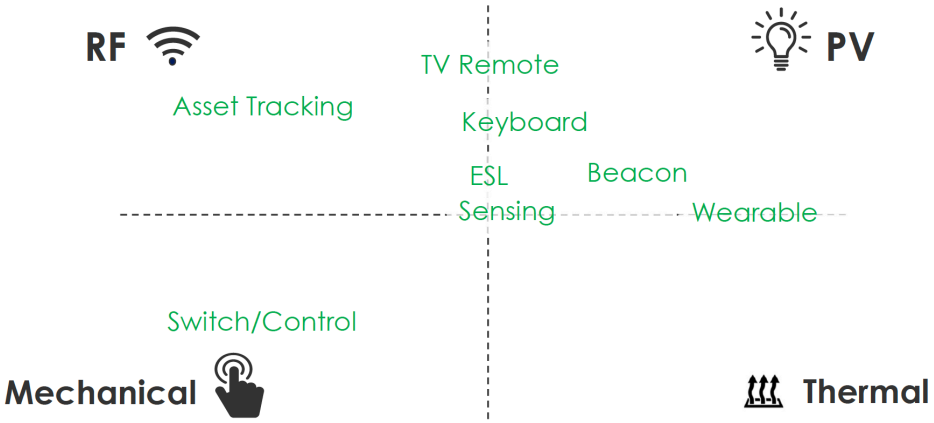


Figure 6 - Applications Supported by Energy Harvesting Technologies

Table 4 - Power density of various Energy harvesting systems

Source	Power Density
Mechanical/Piezoelectric	[0.11-7.31] mW g2/cm3
Radiofrequency	$1.2*10^5$–15 mW/cm2
Solar	[0.006–15] mW/cm2
Thermoelectrical	[15–60] _W/cm3
Wind	[0.065–28.5] mW/cm2

4.1. Photo Voltaic Energy Harvesting

The technology of using photons (sunlight) for generating power/energy has been around for millions of years. The ambient light available around us, even indoors, is sufficient to generate energy that can trickle charge a battery or power low-power sensors. Significant progress has been made in the form of organic photo voltaic cells that can generate sufficient power from indoor lighting conditions.

Figure 7 shows the power needed for various wireless protocols and how they could be powered by a 10 cm^2 IPV (Indoor Photo Voltaic) panel.

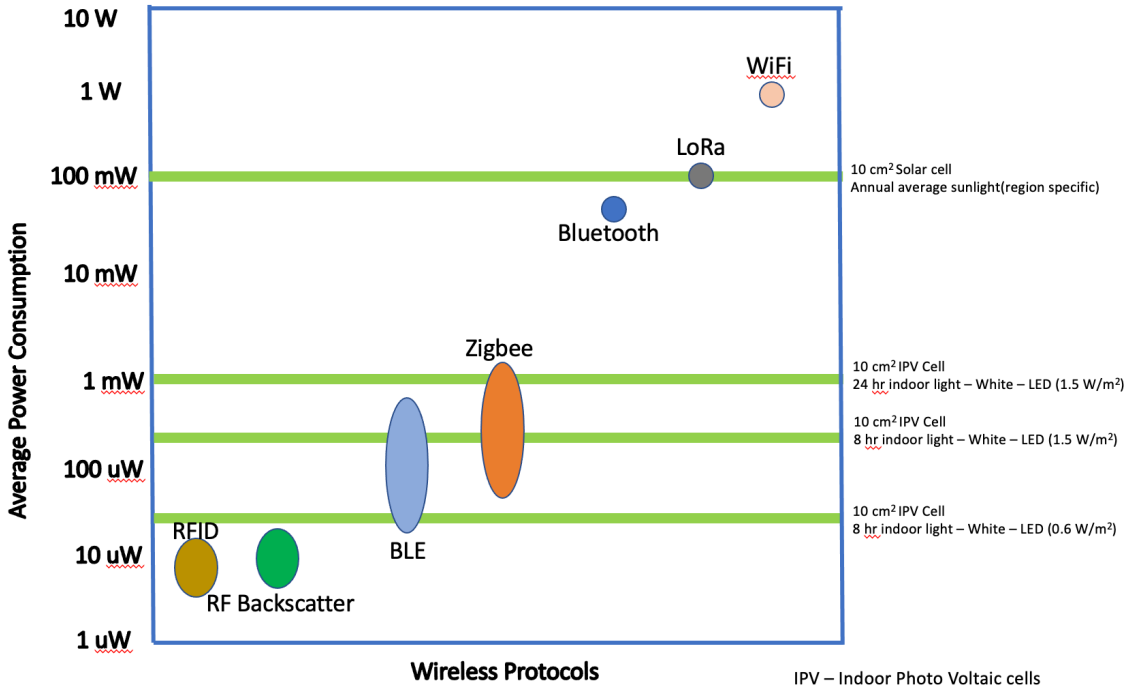


Figure 7 - Mapping wireless protocols to indoor photovoltaic cells based on power needs

4.2. Radio Frequency based Energy Harvesting

RF-based energy harvesting provides a controlled environment for charging/powering IoT devices/sensors over RF/Electromagnetic waves that propagate around us in the form of WiFi/Bluetooth or other wireless signals. RF-based energy harvested signals can carry the information along with the production of energy and can also process that information simultaneously. [7][8]The amount of RF energy available for harvesting at the RF harvester's antenna input will depend on the source's transmitter strength, RF frequency, duty cycle, and range to the receiver.

Figure 8 shows a typical RF harvester that receives energy on a particular frequency. There can be 2 types of RF harvester.

- Ambient RF harvesting: In this case, the receiver is capable of harvesting energy from ambient RF like Bluetooth and WiFi signals
- Active RF harvesting: In this case, a transmitter would send energy on a specific frequency and the receiver would tune to get power. This method offers a more controlled environment.

[7] https://ieeexplore.ieee.org/document/6552840 - **Relaying Protocols for Wireless Energy Harvesting and Information Processing**

[8] https://ieeexplore.ieee.org/document/6623062 - **Wireless Information and Power Transfer**

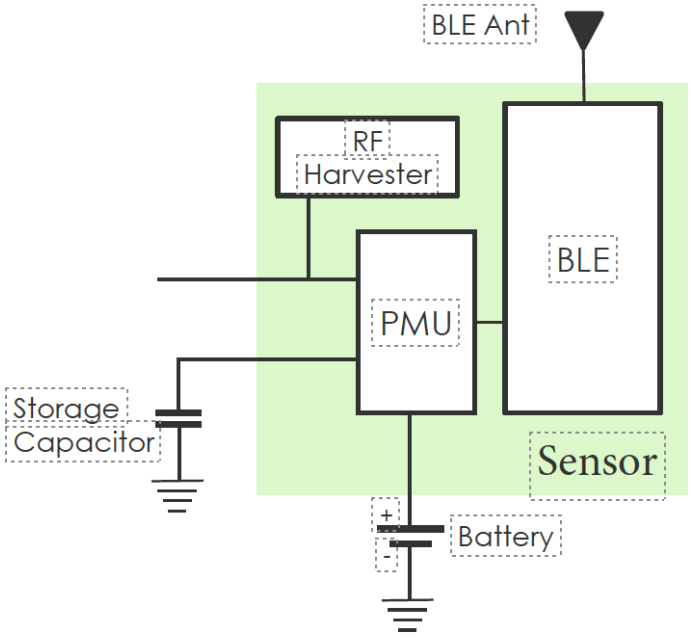


Figure 8 - RF based Energy Harvesting

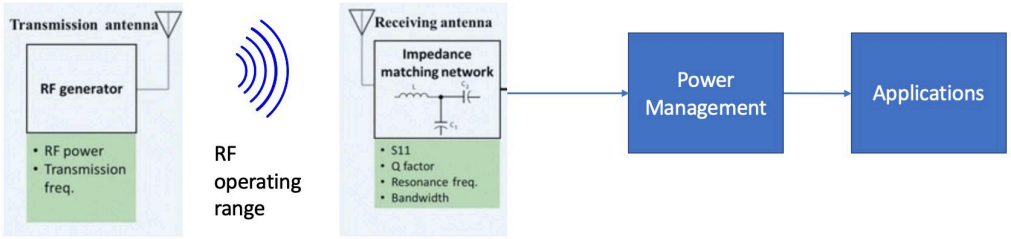


Figure 9 - Active RF based Energy Harvesting

4.3. Mechanical based energy harvesters

Devices called Nanogenerators can convert small amounts of mechanical energy into electric current. The very first nanogenerators were based on the triboelectrification and piezoelectric effect.[9]

Energy harvested from unintended mechanical vibration and abandoned heat can also be directed to low powered electronic devices. Conversion of energy from the ambient environment into electrical energy is known as vibration energy harvesting (VEH).[10]

[9] https://science.sciencemag.org/content/312/5771/242

[10] https://www.ingentaconnect.com/content/tandf/ginf/2019/00000201/00000001/art00010;jsessionid=35omsvrame402.x-ic-live-01

4.4. Thermal Energy Harvesters

Thermal energy harvester use "heat" and temperature variation as the main source to generate power. Two techniques are being used to harvest energy from a thermal source. The first is pyroelectric and the other is thermoelectric.

- The Seebeck effect[11] is used by the thermoelectric technique to convert the difference in temperature into usable energy forms directly. A 5–8% efficiency of harvesting is achievable by the thermoelectric technique of harvesting.
- In pyroelectric energy harvesting, the wasted heat energy is converted into electrical energy for battery-free IoT-based portable devices and wireless sensors. It can convert temperature fluctuations into electrical energy, and this makes it more attractive for harvesting energy.

5. Case Study

As discussed in the earlier section of this document, it is very important and critical that we innovate technologies that help us maintain the growing number of IoT devices and also provide a sustainable solution. Increasing battery usage creates challenges in recycling them. In the sections above we looked at how networking technologies, both PAN and WAN perform. As the number of endpoints and sensors grow, reliable communication at farther distances is a challenge. Increasing output power not only increases draw on the energy sources, but increases the noise in the system and makes it more susceptible to noise.

In this section, we look at technologies that can offer a low-power alternatives to some of the wireless PAN technologies. We will look at the network performance in terms of data throughput, range, efficiency, latency, and power consumption. As a case study, we look at RF backscatter technology to see if it has potential as a low-power, protocol-agnostic, secure network that could be used by IoT sensors to transmit data and stay connected.

5.1. RF Backscatter Network

In RF backscatter networks, multiple network topologies are possible. We will look at a few of those topologies, and one in detail, to evaluate its effectiveness in terms of data throughput, efficiency, and latency.

- Ambient - In the first network topology, the endpoints (backscatter devices) communicate with each other by reflecting either ambient signals or a dedicated carrier service, as shown in Figure 10. The challenge with this network topology is that the receivers exhibit poor sensitivity in the range of -40 to -60 dBm, which limits the operating distances between the endpoints.

[11] https://en.wikipedia.org/wiki/Thermoelectric_effect

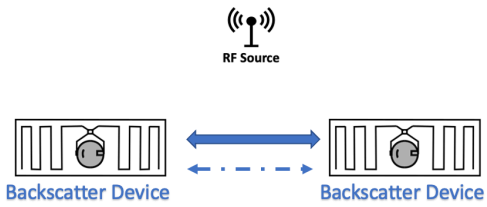


Figure 10 - Backscatter network using Ambient signals

- WiFi/LoRa: The second backscatter topology involves receiving backscattered packets on commodity radio receivers like WiFi or LoRa as shown in Figure 11. This topology is a significant improvement from the ambient backscatter, as the WiFi/Bluetooth/LoRa radios have better sensitivities. Using a LoRa receiver can further extend the range to 100s of meters.

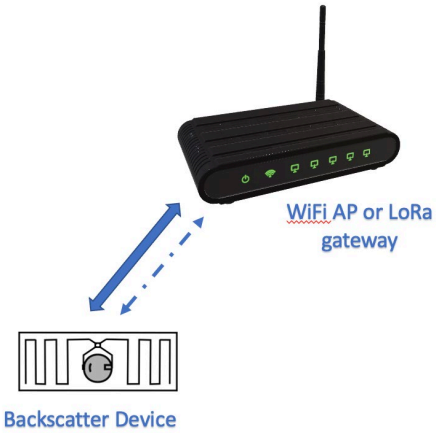


Figure 11 - Backscatter using Wi-Fi or LoRa signals

- Full Duplex: The third topology uses custom signal sources and custom receivers in a full-duplex configuration to operate the backscatter devices as shown in Figure 12. This configuration provides maximum flexibility.

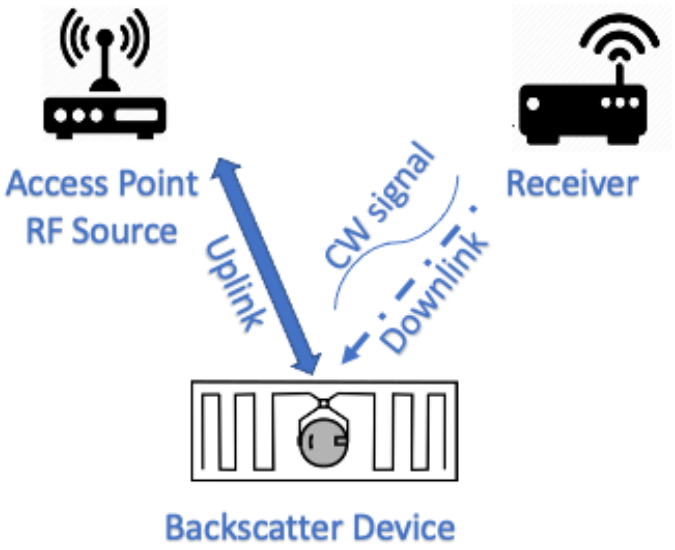


Figure 12 - Backscatter using Custom RF sources

While backscatter can be used with batteries, RF harvesting is still a popular choice for building battery-free systems since we can use the same antenna for both power and communication. Backscatter research is still in its infancy with primary focus on novel physical layer design. Recent research proposes solving some of the pressing networking and MAC-layer challenges of backscatter.[12]

Unlike a traditional wireless network that consists of solely a transmitter and a receiver, a RF Backscatter network consists of three distinct device roles: an Endpoint, a Gateway, and a Companion.

- ***Endpoint***
  The Endpoint a tiny, low cost, low power device that would be a sensor (such as a temperature sensor or smoke detector). The endpoint in a typical network (using WiFi/BLE) would have a wireless transmitter. However, in a backscatter network topology, the sensor achieves low power connectivity by using a backscatter uplink to transmit data packets to the gateway, as shown in Figure 12.

- ***Gateway/Access Point***
  The Gateway in this backscatter network takes the role of a traditional wireless receiver. It acts as a data sink for backscatter packets that are sent in from multiple endpoints. The gateway also serves as a bridge/hub between the backscatter network and another PAN or WAN or cloud server that can process the data generated by the endpoints. The gateway is shown as the RF source in Figure 12.

[12] M. Hessar, A. Najafi, and S. Gollakota. Netscatter: Enabling large-scale backscatter networks. In NSDI 19

- ***Companion/Receiver***
  The companion enables Backscatter communication by transmitting a brief CW (Continuous Wave) signal for each uplink transmission. The companion also coordinates Endpoint devices using a low power downlink signal.

In this network topology, the communication is initiated by the Companion. The companion first transmits a wake code to an Endpoint using On-Off Keying (OOK)[13] modulation. The Endpoint has an ultra-low power detector that helps it detect the wake code. Upon receiving the wake code, the Endpoint goes into a listening state/mode. An Endpoint in listening state would be able to receive data from the Companion via the downlink.

The Companion then transmits a downlink packet followed by a brief CW signal. The downlink packets instruct the Endpoint to transmit a backscatter packet during the subsequent CW signal. By selectively reflecting and absorbing the CW signal from the companion, the Endpoint is able to use backscatter to synthesize a wide variety of standard RF protocols (WiFi/LoRa) that can be received by the Gateway.

The Gateway can then bridge the data coming in from the backscatter network to a server that can process the data. Figure 13 shows the sequence followed by a backscatter network to transmit data from a sensor to an app on the customer's phone. Since the sensor/endpoint consumes low power, we could as well add any energy as discussed in the Energy Harvesting section of this paper.

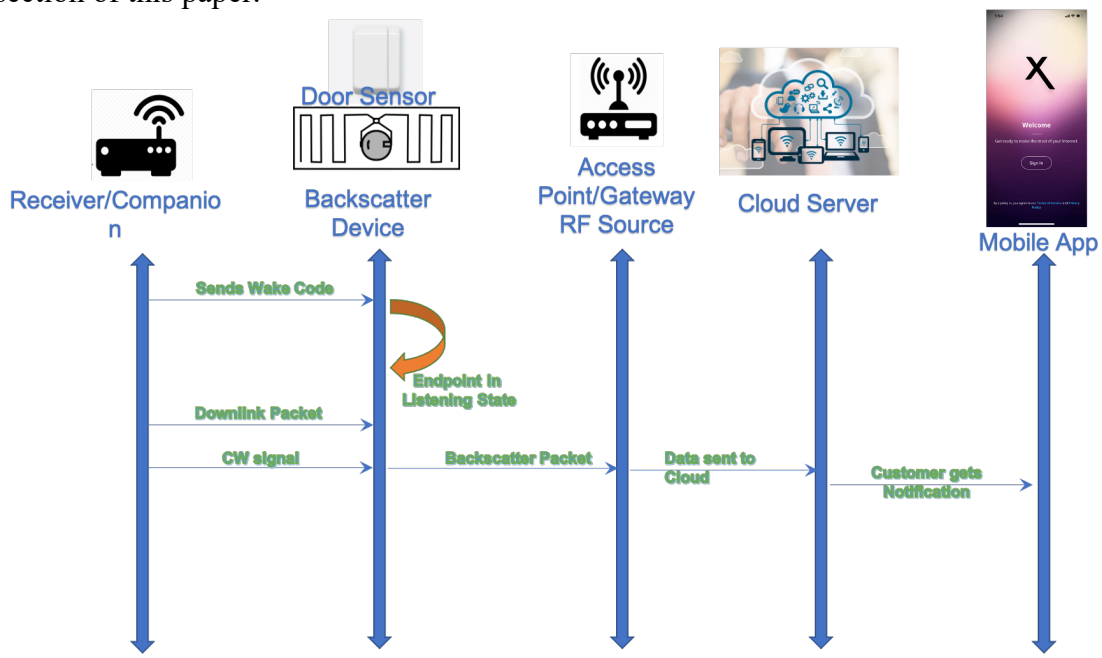


Figure 13 - Sequence diagram of a RF backscatter network packer transmission

[13] https://en.wikipedia.org/wiki/On%E2%80%93off_keying

RF Backscatter endpoints can achieve extremely low power to transmit the data generated by the sensor. A conventional radio actively generates and transmits RF signals for communication. Such system requires complex hardware and uses significant amounts of energy. See the comparison of power consumed, shown in Figure 7. Backscatter modulation utilizes reflected RF signals to communicate, which dramatically reduces all the systems needed to generate radio signals directly at an endpoint/sensor.

The reflections are produced by modulating the Endpoint's antenna impedance in the presence of a CW (Continuous Wave) signal, generated by the Companion. See Figure 14, when any type of wave encounters a boundary between two mediums that have different impedances, a portion of the wave is reflected. The reflection's magnitude is determined by the difference in impedance between those mediums. This is true for a mechanical wave traveling through a rope to a fixed point on a wall, or an electromagnetic wave encountering an antenna.

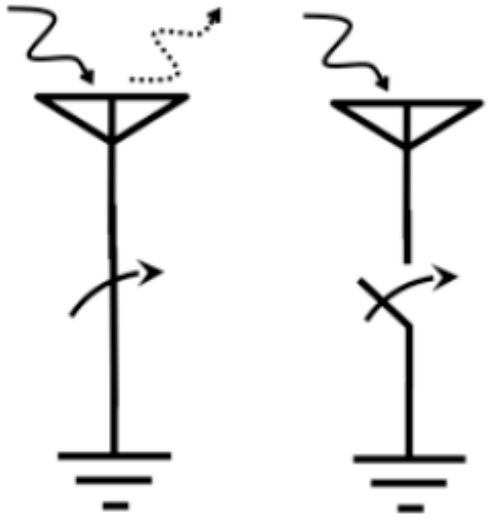


Figure 14: RF wave reflected/absorbed due to change in impedance

6. Conclusion

The world around us is becoming more connected on a daily basis; it is easier than ever today to grab information about the weather in your favorite destination, the air quality in your neighborhood, if you left a window open or a door unlocked and if you need to water your grass. In many of these cases, battery-powered IoT devices and their networks are supplying this information, but to keep the information flowing, these devices need to stay powered up and there's only so many options available. This paper has outlined some of the potential design considerations that could be implemented to extend the life of these devices either through improving the energy efficiency of the protocols used or additional technologies that could replace or extend the battery life of these IoT devices. In IoT applications and networks, where devices are deployed in a stand-alone fashion and expected to function for an extended period

normally measured in years as opposed to days. This is where changes in power consumption can have enormous impacts on the life of the device.

Innovation has been a constant in this industry, as we strive to bring in brand new technologies or improve existing ones to accommodate more use cases and help manage the ever growing number of battery-powered IoT devices.

The energy harvesting technologies mentioned in this paper have been deployed for many decades, and continue to evolve to address the pressing needs of a fast changing world.

The RF backscatter solution presented in this paper is promising and enables us to consider an ultra-low power, protocol-agnostic sensor portfolio.

Abbreviations

AP	access point
BLE	Bluetooth Low Energy
Bps	bits per second
CW	Continuous Wave
DVR	Digital Video Recorder
FEC	forward error correction
HD	high definition
Hz	Hertz
IoT	Internet of Things
ISBE	International Society of Broadband Experts
K	Kelvin
LAN	Local Area Network
LoRa	Long Range
LoRaWAN	Long Range Wide Area Network
LPWAN	Low Power WAN
nDVR	Network DVR
NFC	Near Field Communications
OOK	On Off Keying
OTT	Over The Top
PAN	Personal Area Network
PMU	Power Management Unit
RF	Radio Frequency
RF ID	Radio Frequency Identification
RX	Receive
SCTE	Society of Cable Telecommunications Engineers
SOC	System On Chip
TX	Transmit
VEH	Vibration Energy Harvesting
WAN	Wide Area Network
WLAN	Wireless LAN

Bibliography & References

ANSI C63.5-2006: *American National Standard Electromagnetic Compatibility–Radiated Emission Measurements in Electromagnetic Interference (EMI) Control–Calibration of Antennas (9 kHz to 40 GHz)*; Institute of Electrical and Electronics Engineers
The ARRL Antenna Book, 20th Ed.; American Radio Relay League
Code of Federal Regulations, Title 47, Part 76
Reflections: Transmission Lines and Antennas, M. Walter Maxwell; American Radio Relay League

A Latency Measurement System Using STAMP

with LMAP for large scale data collection

A Technical Paper prepared for SCTE by

Karthik Sundaresan
Distinguished Technologist
CableLabs
858 Coal Creek Circle, Louisville, CO
3036613895
k.sundaresan@cablelabs.com

1. Introduction

Low Latency is gaining importance amongst operators, and they are focused on reducing latency in each of part of the network including the Wi-Fi links in the home, DOCSIS links in the access network and core network segments. Providing lower latency and hence measuring the latency in each portion of network is a vital requirement for MSOs. Operators will need to troubleshoot latency issues and need the ability to identify latency within their networks vs. outside of their networks.

This paper aims to share the experience from developing a simple end-to-end latency measurement framework. A new measurement protocol defined in the IETF is STAMP (Simple Two-Way Active Measurement Protocol, RFC 8762). The paper will provide the lessons learnt from developing a proof of concept for latency measurement using STAMP. It will describe the high-level measurement architecture and locations for measurement agents and peers. A STAMP-reflector could be implemented in a gateway or a device behind it and a STAMP sender can be implemented somewhere in the network (e.g., in a hub, north of a CMTS). An operator can start with small number of measurement entities and scale up as needed. If a Session reflector can by dynamically instantiated in a gateway, then one can run measurements on-demand. The paper will also investigate methods to kick off different latency tests and have the measurement end points report latency data. It will also look into how latency data from various sources can be aggregated and reported. It will also discuss measurement control and reporting methods based on LMAP (Large-Scale Measurement of Broadband Performance). The paper will provide an understanding of MSO needs around latency measurement, an overview of the most appropriate metrics to report, and how to deploy measurement technologies to meet those needs. The paper also reports on a prototype STAMP measurement system which is deployed and collecting latency data.

1.1. Latency Metrics

One-way latency is the total time it takes for a packet of data to travel from the sender to the receiver, across one or multiple hops. Round trip time (RTT) or round-trip latency, is the total time it takes for a packet of data to travel from the sender to the receiver, across one or multiple hops, plus the total length of time it takes for receiver to send a packet back to the sender, through one or multiple hops.

Packet Delay Variation, PDV, is also derived from a sequence of latency measurements where a single reference latency is chosen from the stream based on specific criteria. The most common criterion for the reference is the packet with the minimum delay in the sample.

1.2. Latency Measurements

Active measurements are conducted by generating traffic between two end points for the sole purpose of measuring the latency. Passive measurements are done simply by observing normal host-host interactions. Instead of measuring the latency of specially created test packets like in active measurements, passive measurements are based on the normal user packets that traverse the network.

Active measurements are conducted by generating traffic between two end points for the sole purpose of measuring the latency. An Active measurement method depends on a dedicated measurement packet stream and observations of the packets in that stream. These packets are used to measure packet delay, and packet loss. One-way performance metrics need clock synchronization across the test points for measurements which are tough to implement. Active measurements can use protocols such as TWAMP/ TWAMP Light/STAMP.

1.3. Measurement Protocols

[IETF RFC 5357] TWAMP defines a standard for measuring round-trip network performance between any two devices that support the TWAMP protocols. TWAMP consists of two inter-related protocols: TWAMP-Control and TWAMP-Test. TWAMP Light is an alternative architecture which eliminates the need for the TWAMP-Control protocol and assumes that the Session-Reflector is configured and communicates its configuration with the Server through non-standard means.

Simple Two-way Active Measurement Protocol (STAMP) is a newer IETF standard [IETF RFC 8762] which provides a simpler mechanism for active performance monitoring. It separates the control functions (vendor-specific configuration or orchestration) and test functions. STAMP enables measurement of both one-way and round-trip metrics (delay, delay variation, and packet loss). It is intended to be used on production networks to enable the operator to assess service level agreements based on delay, delay variation, and loss.

2. STAMP

Simple Two-way Active Measurement Protocol (STAMP) is an IETF defined Standards Track RFC 8792[STAMP RFC]., which enables the measurement of both one-way and round-trip performance metrics, like delay, delay variation, and packet loss.

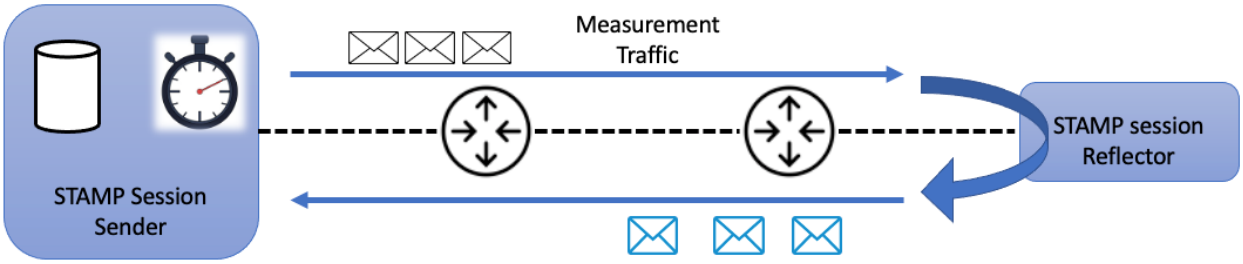


Figure 1 – Simple two way Active measurement protocol

A STAMP measurement session is a bidirectional packet flow between a Session-Sender and a particular Session-Reflector for a given time duration. A STAMP Session-Sender transmits test packets over UDP to the STAMP Session-Reflector. The STAMP Session-Reflector receives the Session-Sender's packet and sends a response per the configuration.

The configuration and management of the STAMP Session-Sender, Session-Reflector, and sessions are outside the scope of the STAMP RFC and can be achieved through various means. E.g., Operators could develop: Command Line Interface, Operational Support System (OSS) / Business Support System (BSS), SNMP, and NETCONF/YANG-based Software-Defined Networking (SDN) controllers.

2.1. Modes of operation

There are two modes of operation for the STAMP Session-Reflector, stateless and stateful per [IETF RFC 8762]

Stateless: The STAMP Session-Reflector does not maintain test state. It uses the value in the Sequence Number field of the received packet as the value for the Sequence Number field in the reflected packet. As a result, values in the Sequence Number and Session- Sender Sequence Number fields are the same, and only round-trip packet loss can be calculated while the reflector is operating in stateless mode.

Stateful: The STAMP Session-Reflector maintains the test state. This enables the Session-Sender to determine which direction the loss is happening by using the of gaps the Session Sender Sequence Number and Sequence Number fields. As a result, both forward path loss and return path packet loss can be computed.

STAMP supports two authentication modes: unauthenticated and authenticated. Unauthenticated STAMP-Test packets, ensure interworking between STAMP and TWAMP Light. As STAMP and TWAMP use different HMAC algorithms in authenticated mode interoperability is only in the unauthenticated mode.

2.2. Port number & Interop with TWAMP

A STAMP Session-Sender and STAMP Session-Reflector use UDP port 862 (same as TWAMP) as the default destination UDP port number. An implementation of the Session-sender and Session-Reflector can define the port number to receive STAMP-Test packets from User Ports and Dynamic Ports ranges.

One of the essential requirements to STAMP is the ability to interwork with a TWAMP Light device. For example, a TWAMP Light Session-Reflector may not support the use of UDP port 862, as specified in [RFC8545]. A STAMP Session-Sender is allowed to use alternative ports. If any of STAMP extensions are used, the TWAMP Light Session-Reflector will view them as the Packet Padding field.

2.3. Packet Format and Size

By default, STAMP uses symmetrical packets, i.e., the size of the packet transmitted by the Session-Reflector equals the size of the packet received by the Session-Reflector. The STAMP Session-Sender packet has a minimum size of 44 octets in unauthenticated mode (see Figure 2) and 112 octets in the authenticated mode (see [IETF RFC 8972]

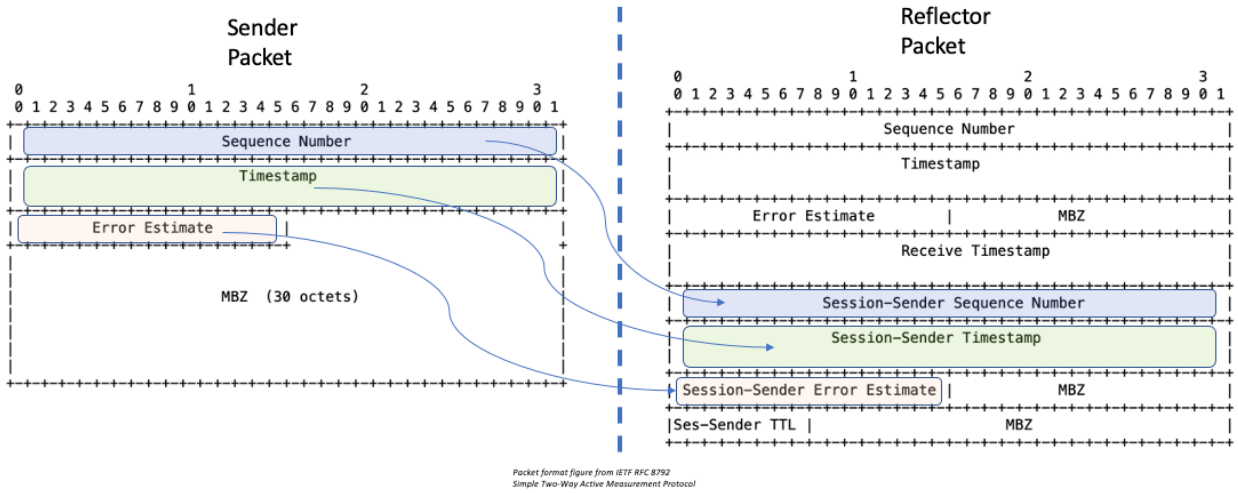


Figure 2 - STAMP Test Packet Format (Sender & Reflector)

STAMP supports symmetrical size of test packets, i.e., a reflected base test packet includes information from the Session-Reflector and, thus, is larger. To maintain the symmetry between base STAMP packets, the base STAMP Session-Sender packet includes the Must-Be-Zero (MBZ) field to match to the size of a base reflected STAMP test packet.

Generating variable length of a test packet in STAMP is defined in [IETF RFC 8972]

The field definitions for Authenticated mode are the same as the unauthenticated mode. The STAMP Session-Reflector test packet format in authenticated mode includes a HMAC hash at the end of the PDU. The detailed use of the HMAC field is in described in the [IETF RFC 8762].

2.4. STAMP Extensions

STAMP defines multiple extensions to give the operator additional functionality as related to latency measurement. Some of these extensions include functionality such as extra padding, location, timestamp information, class of service, direct measurement, access report, follow-up telemetry and HMAC. These are described in the [IETF RFC 8972].

Here is the format of the TLVs used within the existing stamp packet. [IETF RFC 8972]

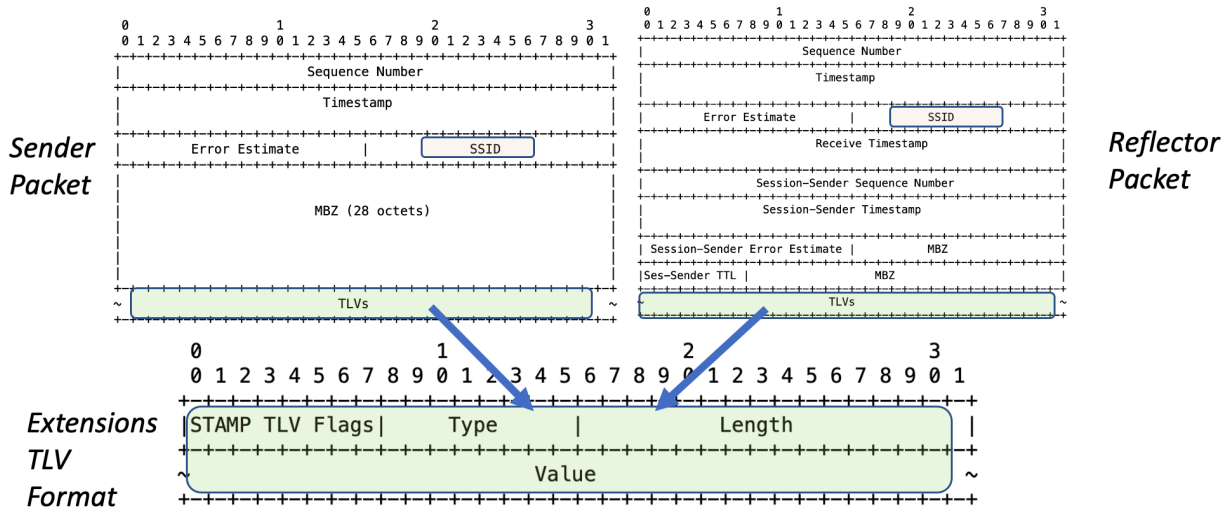


Figure 3 - STAMP Test Packet Extensions Format

Of the eight types of STAMP extensions, defined below are a couple of the extensions. The “Extra padding” TLV (Type 1) can extend the size of the STAMP packet, to allow an operator to test the network with different sized packets. The “Class of service” TLV (Type 4) can be used by an operator to check how the DSCP value of the IP test packet changes as it traverses different networks.

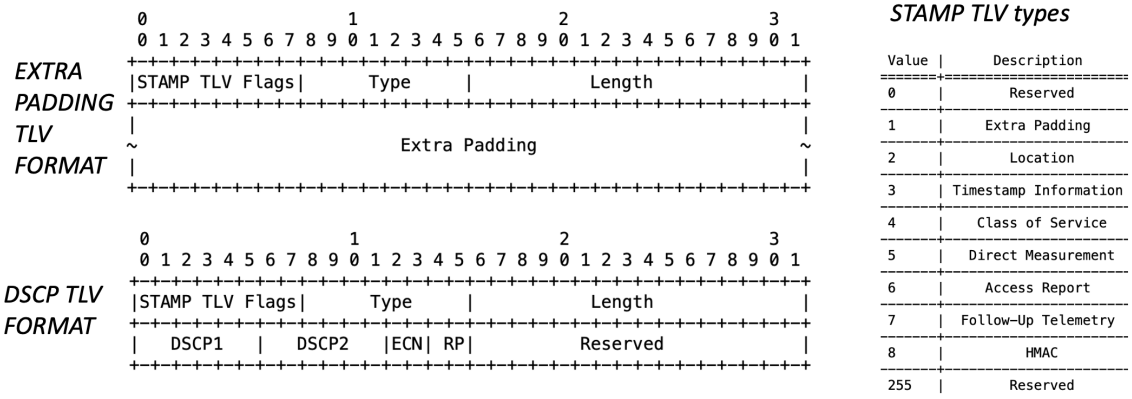


Figure 4 - STAMP Extra padding and DSCP Extensions

3. LMAP

The Large-Scale Measurement of Broadband Performance (LMAP) [IETF RFC 7594] developed by the IETF standardizes a measurement system for performance measurements of broadband access devices such as home and enterprise edge routers, personal computers, mobile devices, set top box, whether wired or wireless.

Measuring portions of the Internet on a large scale is essential for accurate characterizations of performance over time and geography, for network diagnostic investigations by providers and users. The goal is to have the measurements made using the same metrics and mechanisms for a large number of end points on the Internet, and to have the results collected and stored in the same form.

3.1. LMAP Architecture

A large-scale measurement platform involves basically three types of protocols, namely, a Control Protocol between a Controller and the Measurement Agent (MA), a Report Protocol between the MAs and the Collector(s), and several measurement protocols between the MAs and Measurement Peers (MPs), used to perform the measurements. In addition, some information is required to be configured on the MA prior to any communication with a Controller.

LMAP has defined the following

- A Control Protocol, from a Controller to instruct Measurement Agents what performance metrics to measure, when to measure them, how/when to report the measurement results to a Collector.
- A Report Protocol, for a Measurement Agent to report the results to the Collector.

The LMAP framework has three basic elements: Measurement Agents, Controllers, and Collectors.

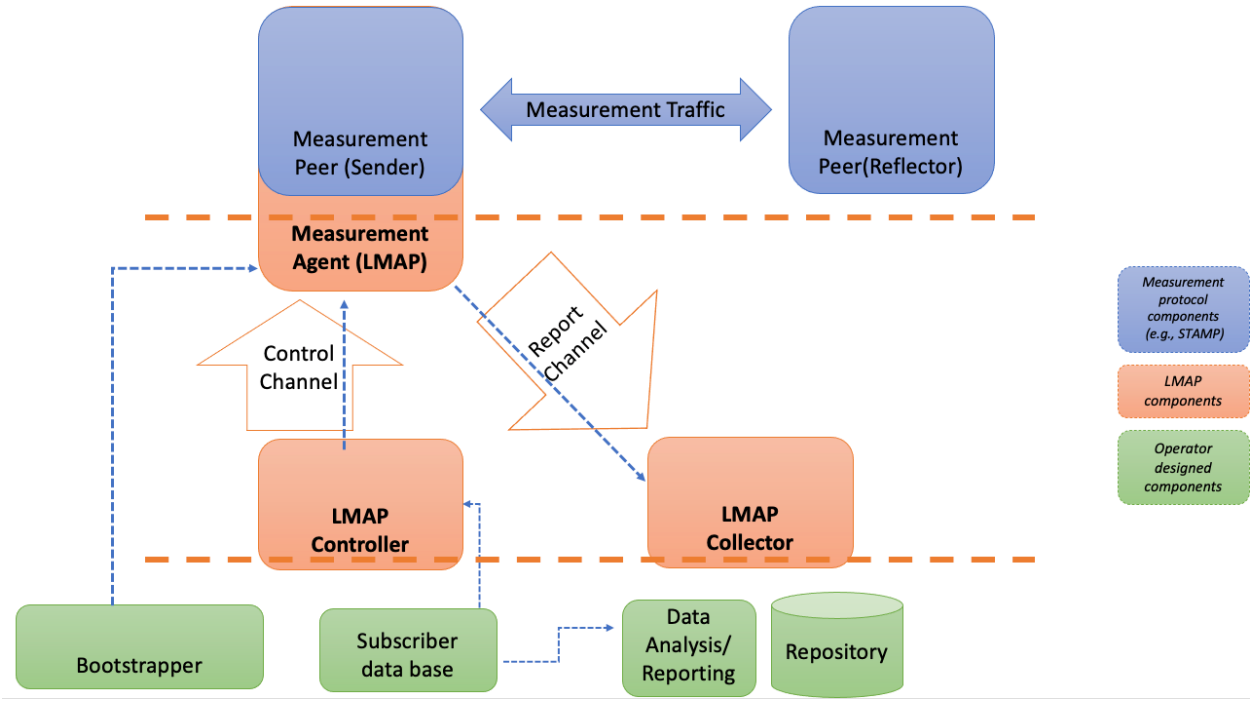


Figure 5 - Elements of an LMAP-based Measurement System

Measurement Agents (MAs) initiate the actual measurements, which are called Measurement Tasks in the LMAP terminology. In principle, there are no restrictions on the type of device in which the MA function resides.

The Controller instructs one or more MAs and communicates the set of Measurement Tasks an MA should perform and when. For example, it may instruct an MA at a home gateway: "Measure the 'UDP latency' with www.cableco.org; repeat every hour at xx.05". The Controller also manages an MA by instructing it on how to report the Measurement Results, for example: "Report results once a day in a batch at 4am" (a Report Schedule.)

The Collector accepts Reports from the MAs with the Results from their Measurement Tasks. Therefore, the MA is a device that gets Instructions from the Controller, initiates the Measurement Tasks, and reports to the Collector. The communications between these three LMAP functions are structured according to a Control Protocol and a Report Protocol.

The LMAP effort has specified an information model [IETF RFC 8193], the associated data models [IETF RFC 8194], and protocols for secure communication. Information Model applies to the Measurement Agent within an LMAP framework. It outlines the information that is configured on the Measurement Agent or exists in communications with a Controller or Collector within an LMAP framework. The purpose of such an Information Model is to provide a protocol- and device-independent view of the Measurement Agent that can be implemented via one or more Control and Report Protocols. The data models are extensible for new and additional measurements.

3.2. LMAP YANG model

The LMAP framework has three basic elements: Measurement Agents (MAs), Controllers, and Collectors. Measurement Agents initiate the actual measurements, called Measurement Tasks in the LMAP terminology. The Controller instructs one or more MAs and communicates the set of Measurement Tasks an MA should perform and when. The Collector accepts Reports from the MAs with the Results from their Measurement Tasks.

The YANG data model [IETF RFC 8194] for LMAP has been split into three modules: The common module (ietf-lmap-common.yang) provides common definitions such as LMAP-specific data types. The control module (ietf-lmap-control.yang) defines the data structures exchanged between a Controller and Measurement Agents. The report module (ietf-lmap-report.yang) defines the data structures exchanged between Measurement Agents and Collectors.

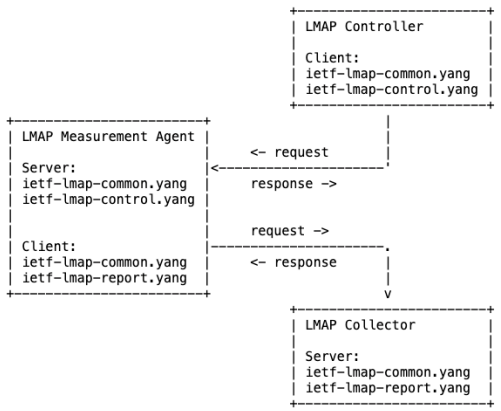


Figure 6 – High level view of the LMAP YANG model components

Below is a figure which describes the various components within the LMAP- Control and LMAP-Report YANG data modules.

Figure 7 – High level view of the LMAP-Control YANG model

(*) Figure 7 is built from views with the tools at https://yangcatalog.org/yang-search/yang_tree/ietf-lmap-control@2017-08-08

The LMAP Information Model [IETF RFC 8193] is divided into six functional parts mapped into the YANG data model as follows:

- Preconfiguration Information: bootstrapping information is outside the scope of the model
- Configuration Information: Modeled in the /lmap/agent subtree, the /lmap/schedules subtree, and the /lmap/tasks subtree

- Instruction Information: Modeled in the /lmap/suppressions subtree, the /lmap/schedules subtree, and the /lmap/tasks subtree.
- Logging Information: success/failure/warning messages in response to information updates from the Controller, will be handled by the protocol used to manipulate LMAP-specific configuration.
- Capability and Status Information: Capability is modeled in the /lmap/capability subtree. The list of supported Tasks is modeled in the /lmap/capabilities/task list. Status Information about Schedules and Actions is included in the /lmap/schedules subtree. Information about network interfaces can be obtained from the ietf-interfaces YANG data model [RFC 7223]. Information about the hardware and the firmware can be obtained from the ietf-system YANG data model [IETF RFC 7317]. A device identifier can be obtained from the ietf-hardware YANG data model [YANG-HARDWARE].
- Reporting Information: This is modeled by the report data model to be implemented by the Collector. Measurement Agents send results to the Collector by invoking an RPC on the Collector.

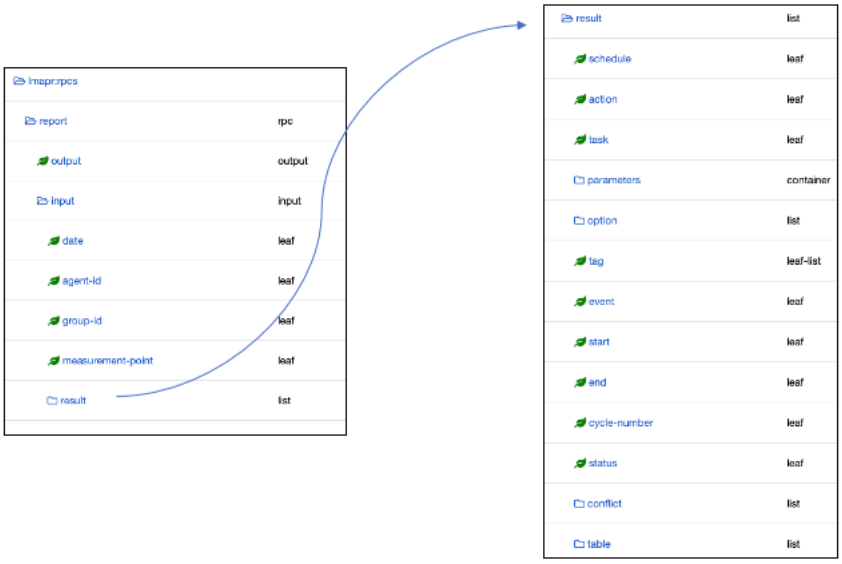


Figure 8 –Definition of the LMAP-REPORT YANG model

(*) Figure 8 is built from views with the tools at https://yangcatalog.org/yang-search/yang_tree/ietf-lmap-report@2017-08-08

4. Latency Measurement Architecture

The latency measurement architecture can be split into two parts as shown in the figure below:

Measurement Domain: This includes the measurement protocol itself and the measurement agent and the measurement peers. This domain performs the actual latency measurements/tests and calculate the latencies.

Large Scale control and data collection: This includes the large-scale latency measurement orchestration across the network by a controller/collector entity. The controller entity which coordinates the measurements across the various measurement agents in the network. The collector entity collects the data from the various measurement agents and then presents the aggregate data to the operator.

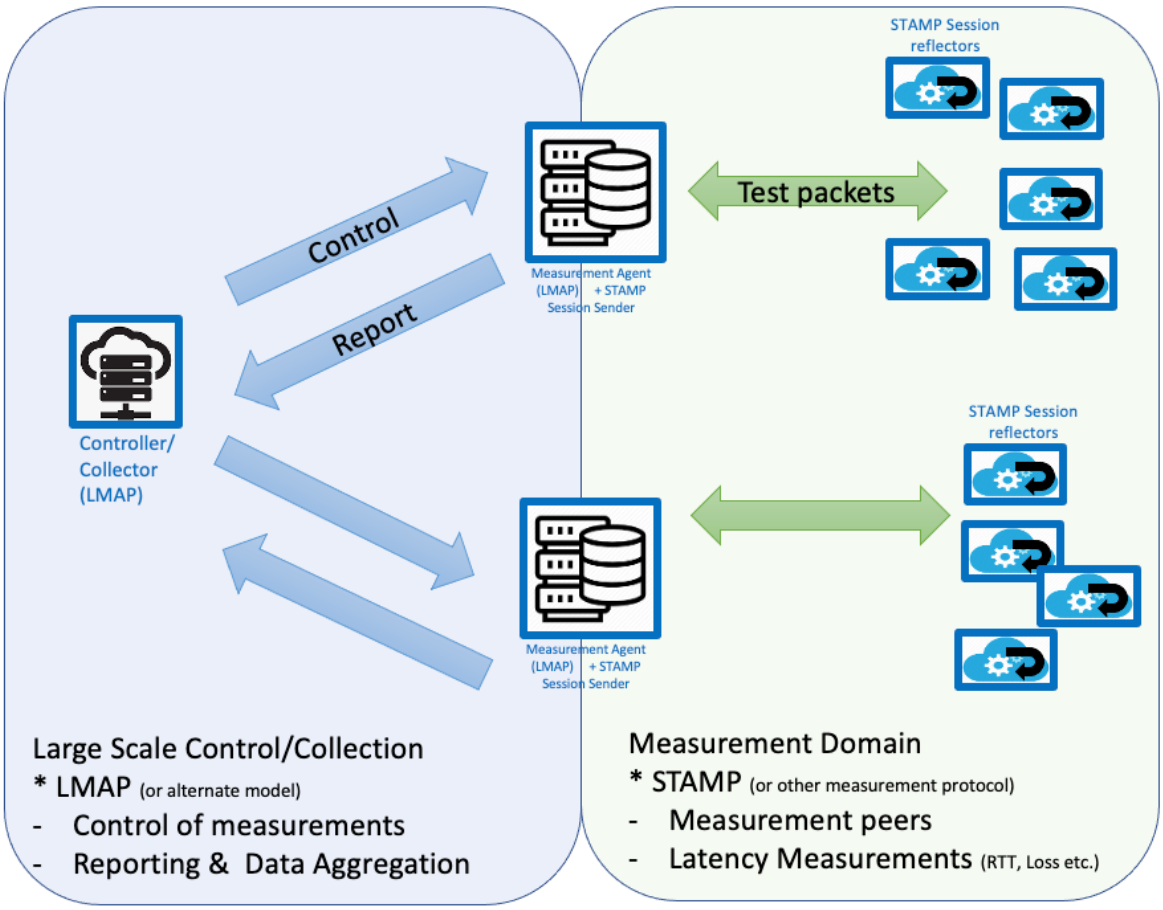


Figure 9 - Latency Measurement Architecture

Separating the data control and collection domain from the actual measurement domain allows the operator to make appropriate choices for both domains. The leading contender for large scale control and collection is the IETF LMAP model. For the measurement domain the IETF STAMP measurement protocol appears to be the best fit.

4.1. Measurement Architecture in a Cable Network

In a cable access network, the latency measurement architecture described above with and LMAP domain and the STAMP domain could be implemented as follows.

An operator could choose to deploy a measurement agent at each hub or headend location just north of the CMTS at that location. This way an operator can reliably measure the latency on the DOCSIS / access network portion of the network. STAMP session reflectors are lightweight entities and can be placed at the customer premise. This could be at the cable modem itself or on a gateway device which the operator installs at the customer premise.

To measure latencies in the core network and operator could also place the lightweight stamp session reflectors at locations close to the interconnection/ peering points to the Internet. Now an operator could deploy and LMAP collector and controller at a more centralized location for example the network operations center which oversees multiple hub/headend locations or perhaps even the whole network of measurement agents. See Figure below.

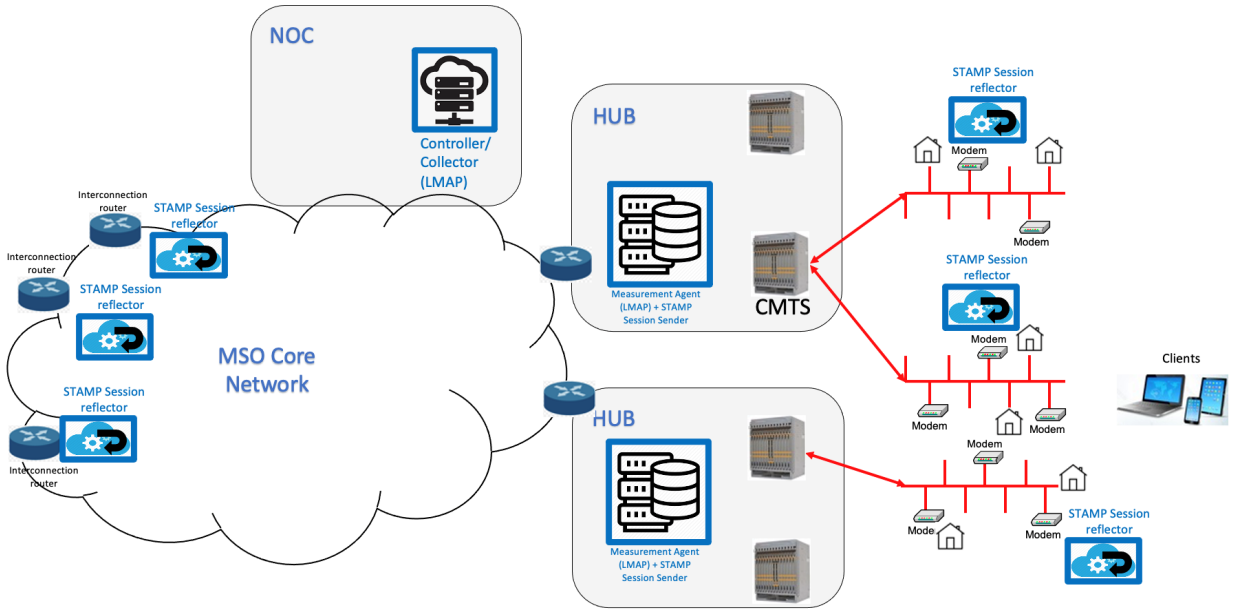


Figure 10 - Latency Measurement Architecture in a Cable Operator Network

4.2. Measurements in the Access vs Core vs Home Network

For the access network, each measurement agent at the hub locations would be responsible for running latency measurement tests to each of the session reflectors within its domain, i.e., within the part of the cable network to which it is connected.

For the core network, an operator could choose to run tests between every interconnection router and every measurement agent so that they can get a baseline understanding off the latencies across each of the potential paths across the core network.

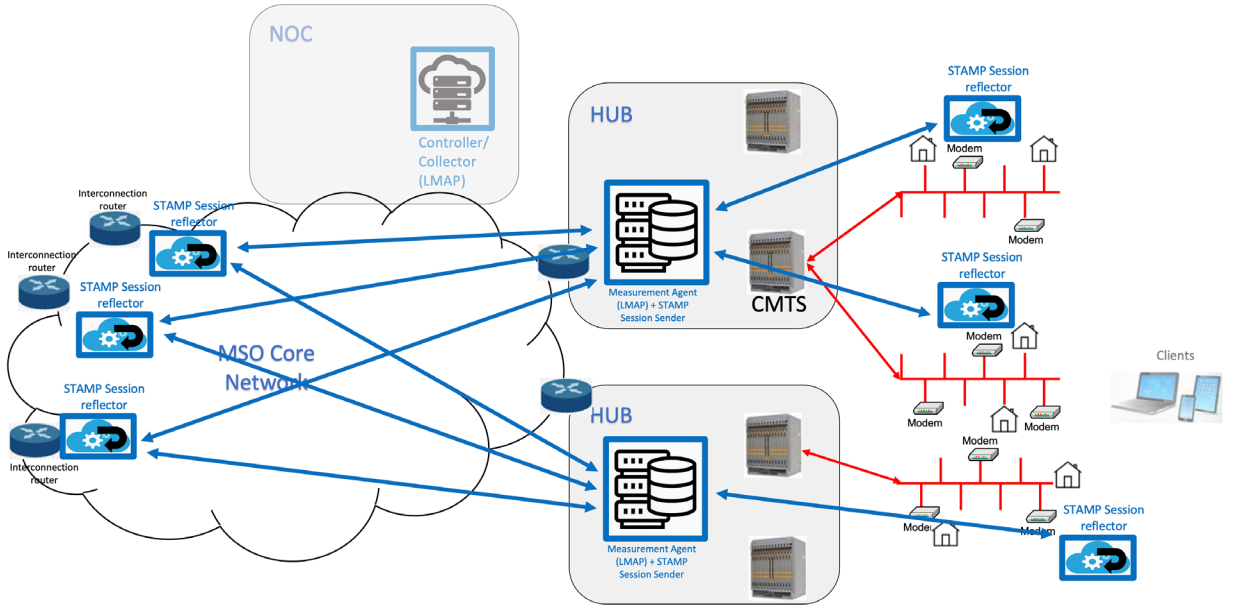


Figure 11 - STAMP Latency measurements (Access & core)

And operator can also instantiate a measurement agent within a client device for example a handheld device or a laptop which could be owned by a customer or by a technician. In this case this measurement agent within the customer's home can help measure latencies to each of the session reflectors within the network. In the case where this measurement agent talks to the session reflector within the gateway in their customer premise, this will this result in the operator understanding the latency in the home network (e.g., Wi-Fi). When the measurement agent in the handheld device runs latency tests with the session reflectors close to the interconnection points the operator can also get an understanding of the combined access and core network latencies from the customer location to the peering point.

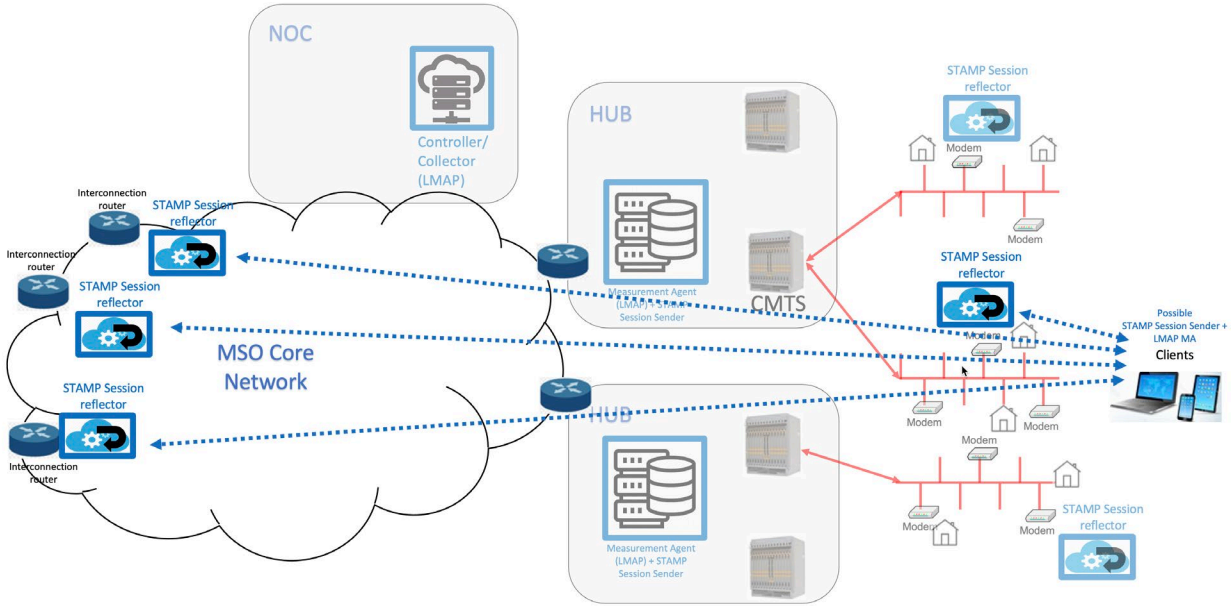


Figure 12 - STAMP Latency measurements from a Client-side MA

4.3. Measurement aggregation

Now for a large-scale measurement system the LMAP controller/collector will coordinate with each of the measurement agents in the network. The controller will command each of the measurement agents to run specific latency tests at a particular point in time or on a schedule. Each measurement agent will collect the set of latency measurements compute the requested statistics be it histogram counts or percentile data for that test. These results will be reported back to the collector. This way the LMAP controller/collector becomes the one central location where an operator can go to understand the latency performance across the whole network. All the data analytics and data aggregation off the latency measurements across the network both the access network and the core network and potentially the home network will be implemented at the collector.

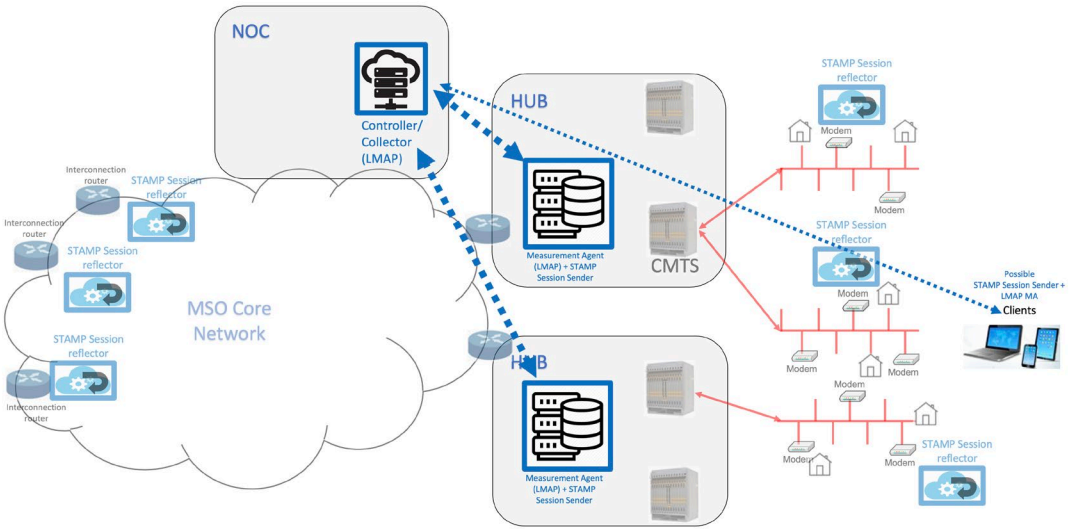


Figure 13 - LMAP Measurements control and reporting

4.4. Scaling Considerations

An operator would need to think through the scale of the measurement infrastructure that they deploy. The number of measurement agents and session reflectors needed will depend on the size of their network. An operator would need to account for the number of headend/hub locations they want to cover and the number of peering points they want to cover.

4.4.1. Core network Latency

As described above, for core network latency measurements, the operator would plan to have measurements from each hub or headend location to each of the Internet peering/transit points within MSO network. In an informal survey we found that each operator has on the order of 8 to 12 peering points and transit locations and while of the bigger operators having up to 30 interconnect locations (based on past mergers and acquisitions of cable properties). Typical operator networks vary from 100s to 1000s of CMTSs and this implies 10s to 100s CMTS hub/headend locations.

An operator would desire to get a full mesh of latency measurement of "CMTS" x "Interconnect" locations. For a small cable operator this would be in the order of 100 links and up to a 1000 links for a larger operator.

An operator will likely place one Session reflector at each interconnect location. An operator may choose to place a one Measurement Agent at each hub/headend location. Alternatively, they could place to near the Core/Aggregation router to reduce number of measurement agents, those this will lose the ability to truly isolate access network latency in the measurements.

4.4.2. Access network Latency

As per the previous section, the assumption here is that an operator places measurement agents at each hub/headend location. Now the question is how many section reflectors an operator wants in the access portion of the network. An operator needs to figure out if they want to perform latency measurements to every modem on the network or if they want to subsample the CM population.

If an operator decides to subsample the network, they need to decide what percentage of devices should be used in latency measurement, would it be 20%, 10%, or 1%? For e.g., for a mid-tier operator with 10 million broadband subscribers, a 10% choice implies: 1 million session reflectors and even a 1% sampling implies 100,000 session reflectors. If an operator decides to equally distribute these session reflectors across their CMTS footprint, let's assume a 1000 CMTSs and each CMTS supporting 50 nodes, then this is just about 2 session reflectors per node segment. These types of calculations give us an idea of the choices an operator will have to make in terms of the number of session reflectors the coverage of the modems that are needed and then start planning to scale the number of measurement peers accordingly. With Distributed CMTS architectures (Remote PHY or Flexible MAC architecture technology) with RPD and RMD devices in the network, an operator may choose to measure those links separately which means an additional layer of measurement agents at each of the nodes.

Once an operator gets comfortable collecting and understanding the latency measurements in a small part of the network then a phased approach to increasing measurement coverage across the network, will be the likely path than operators take.

5. Experimental results

To put all the latency measurement theory and protocols to test, we built an experimental latency measurement system prototype which was tested with measurement server and peer locations across the Internet.

5.1. Prototype Components

For the latency measurement system prototype, the following components were implemented: A Measurement agent (STAMP Session sender) and a STAMP session reflector. Additionally, an LMAP Controller + Collector along with adding LMAP functionality to the measurement agent are currently under development as well. These prototype components will be made available at [C3 CableLabs] after the development is complete.

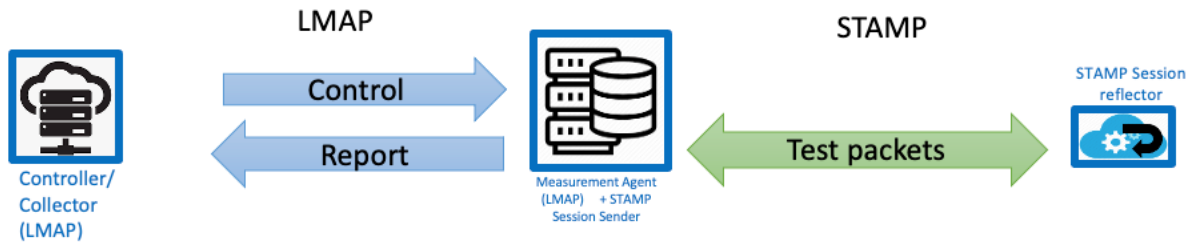


Figure 14 – Prototype components

5.1.1. Session-Reflector

The STAMP session-reflector software implementation was developed using C and built for a Unix environment. The Raspberry Pi platform was chosen for a session reflector. The idea is to create a lightweight stamp session reflector software which could potentially be embedded within a cable modem or a gateway device in the house or a Wi-Fi AP. Another option would be a stand-alone device which the operator installs as "probes" throughout their network.

5.1.2. Measurement Agent

The measurement Agent consists of two components, the STAMP Session sender and the LMAP Measurement Interface.

The STAMP session sender software implementation was developed using C and built for a Unix environment. The Measurement Agents were instantiated on AWS servers using Ubuntu Linux instances. The same version of the Measurement agent was also deployed onto the Raspberry Pi platform to run locally to test and debug. The idea here is to create a robust STAMP session sender software which could potentially be extended to support all the STAMP features and support a variety of latency tests for an operator. This Measurement Agent (STAMP session sender) will be controlled through the LMAP measurement Interface it implements.

The LMAP measurement interface was developed to support a NETCONF interface and support the LMAP YANG model. 'Netopeer2' is an opensource server for implementing network configuration management based on the NETCONF Protocol. The server uses 'sysrepo' (an opensource library for storing and managing YANG-based configurations for UNIX/Linux applications) as a NETCONF datastore implementation. The idea is that the measurement agent will interface with the LMAP controller using the NETCONF protocol.

5.1.3. LMAP Controller and Collector

The LMAP Controller and collector entity is being developed to support a NETCONF interface and support the LMAP YANG model. This communicates with the various Measurement Agents to configure tests and gather the results back. The additional data analytics and visualization happens hereat the controller/collector.

5.2. Test Metrics

There are various metrics which and operator could track when they are looking to understand the Latency performance of their networks, be it the access network, the home network, or the core network. each portion of the network needs to be measured to understand the current characteristics and then to improve on it.

The [LM SCTE 20] paper discusses the basics of latency and proposes some metrics to measure. The main measures which an operator should look at are

- Latency RTT (Round trip time)
- PDV (Packet Delay Variation)
- Packet Loss (and directionality of loss if available)

For the Latency and PDV measures, if an operator were to pick a number, the 99^{th} percentile value is a very good metric to track and is likely indicative of the customer experience especially in latency sensitive real time applications. To understand how the latency behavior changes over time it is also useful to track multiple percentile values, e.g. 0^{th} percentile (minimum), 25^{th} percentile, 50^{th} percentile (median), 75^{th} percentile, 95^{th} percentile, 99^{th} percentile, 99.9^{th} percentile, 100^{th} percentile (maximum).

To understand how the latency varies over time and visualizing it, a simple time series graph shows a lot of interesting patterns. Additionally, a histogram of the data set is a great place to start analyzing the latency performance. A cumulative distribution function (on a logarithmic scale) can show the operator the more interesting latency behavior regions.

5.3. Expermental Setup for Latency Measurement using STAMP

As a first step we decided to deploy the STAMP measurement agents and session reflectors in a network and work through the issues we would see with such a deployment on the Internet. Four different locations of the AWS servers were chosen, across the U.S, to give different sampling/variation of latency measurements. These server locations were in Oregon, N. California, Ohio, N. Virginia. There was a single session-reflector behind a DOCSIS 3.1 CM in a home in Denver. The session-reflector was connected via WiFi to the CM.

Figure 15 – Location of Measurement Agents & Session Reflector

5.4. Latency Measurement Test results

Each Measurement Agent ran latency tests to the Session reflector periodically. The interval chosen was every 5 mins, and tests were run over a period of a week. An individual 'test' consisted of 2000 to 5000 packets sent back-to-back.

Round trip latency and loss measurements were the focus of the experiments. The change in sequence numbers on the sender and received side also exposed packet loss in either direction.

As the clocks across the Session-sender and the Session-reflectors were not synchronized, one way latency measurements are not reported, though they were calculated.

5.4.1. Time Series view

The following figure shows a set of roundtrip latency measurements between a measurement agent in Oregon (STAMP Session sender) and a measurement peer (STAMP session reflector) in Denver. They're following figure shows about 75 hours of latency testing with 5000 test packets being sent every five minutes.

One can easily see that though the latency is around 35 ms for a lot of the time, there are some clear periodic larger spikes of latency up to 140 ms and then some even larger spikes of latency up to 2.5 seconds. The orange dots on the bottom of the graph indicate the number of packets lost during the test.

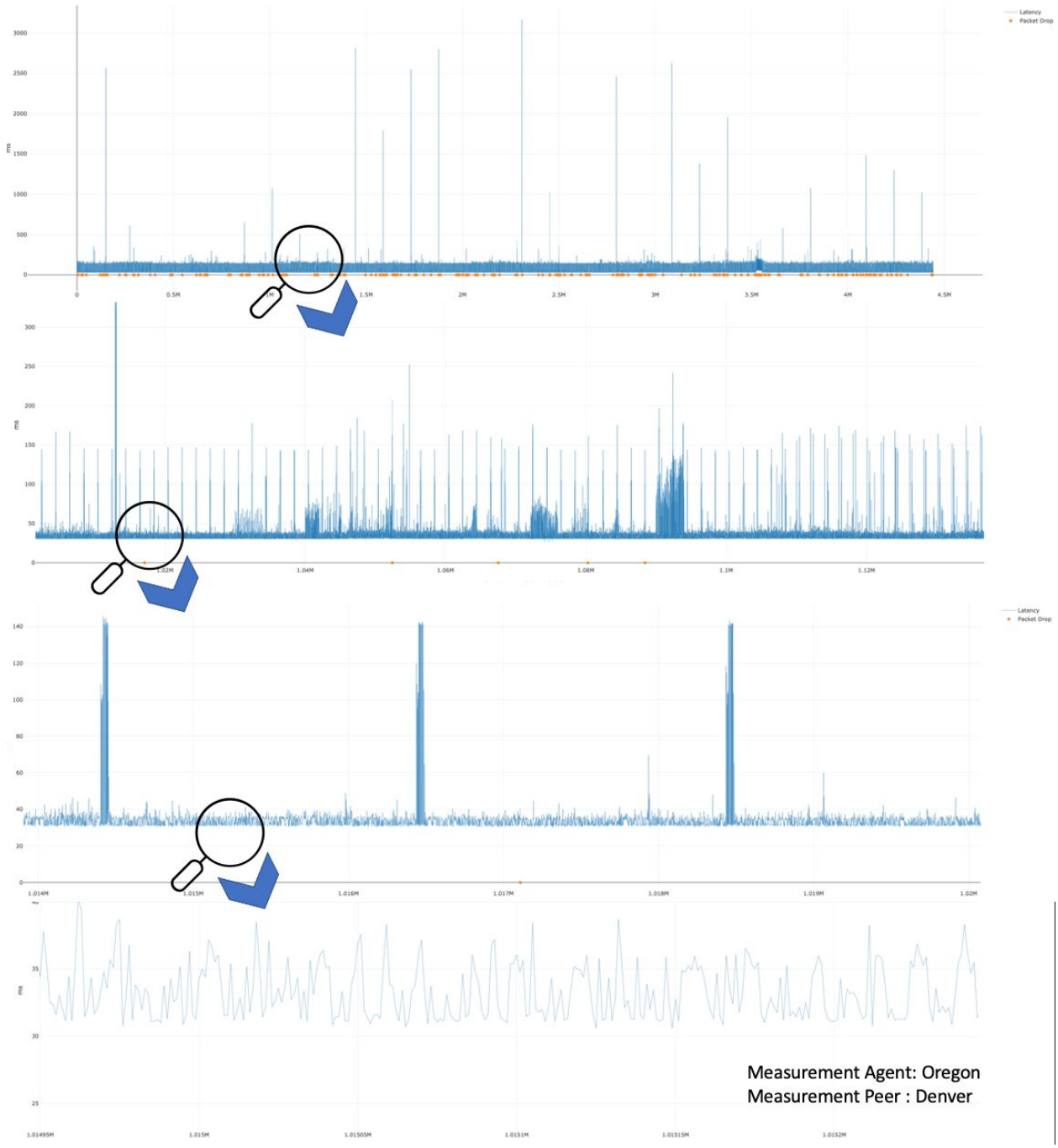


Figure 16 –Latency data - time series

5.4.2. Histogram view

The next figure below shows a histogram of latency, for each of the four links measured. The histogram bins are chosen to be one second each, do show with high granularity that different latency's that we observe to each of the measurement agents which are in different locations. As expected, the Ohio and N Virginia measurement agents have higher latencies (farther from Denver session reflector) compared to the measurement agents in California and Oregon (closer to Denver). As the histogram indicates, the

latency behavior is very different for different links and has multimodal distribution. Again, a reminder to network operators and users, to fight the urge to reduce latency to an “average” number or run a single test and take that to be latency number.

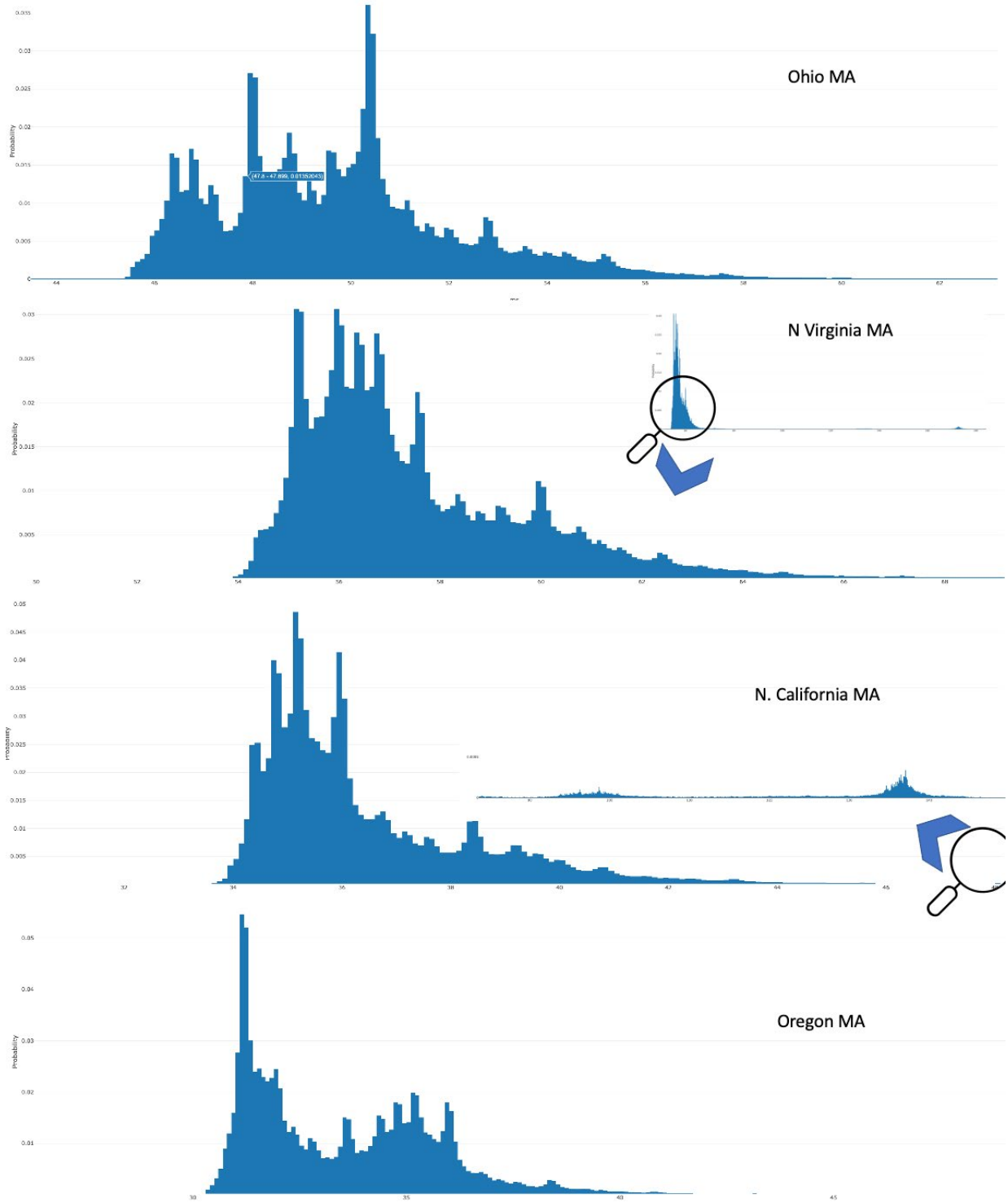


Figure 17 – Histogram of latencies from all 4 MAs

5.4.3. CDF view

The following figure shows a cumulative distribution function (CDF) graph of the latency data from each of the measurement agents. Here we look at the percentile values off Latency and use that as a metric to compare the different links. the scale on the Y-axis is logarithmic and the X-axis is zoomed in to the areas of interest. Operators are typically interested in the 99th percentile values of latency as that is shown to be indicative of the customer experience especially for real time applications such as gaming or real time communications. The table at the end of the section shows the 99th percentile, while the CDF graphs below show the 50th and the 95th percentile.

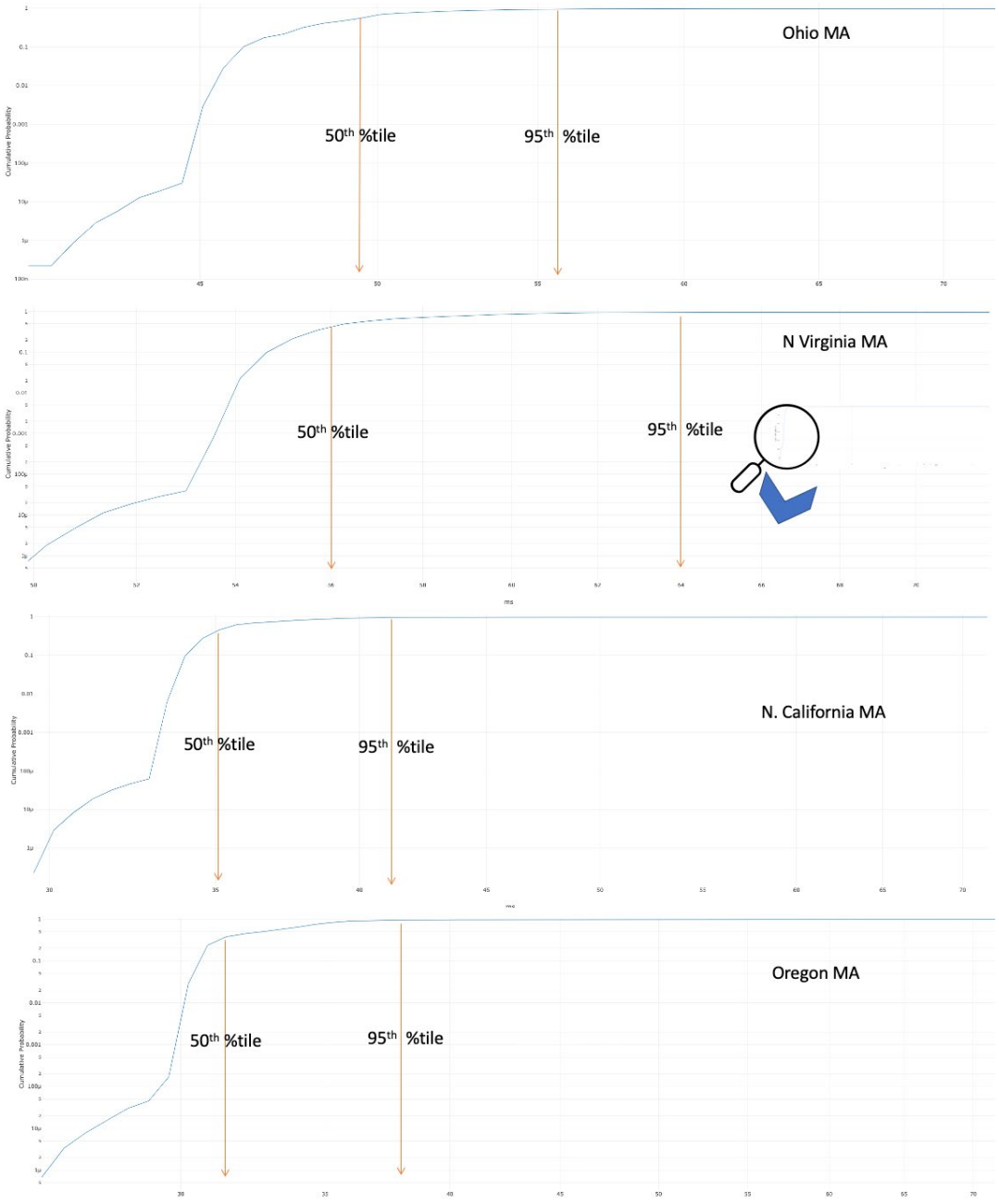


Figure 18 – CDF of latencies from all 4 MAs

5.4.4. Home Latency Testing

The next set of results below are for a measurement agent and a measurement peer located in the same home; this would be an example of testing the Wi-Fi latency within the home. the first graph below is a time series off a test run over 80 minutes, with 5000 packets being sent every five minutes. Again, while the nominal Latency hovers around the 5 ms mark, there are occasional spikes up to 14 to 16 ms and additional spikes of leading see up to 180 ms and some outliers of latency up to 1.2 seconds

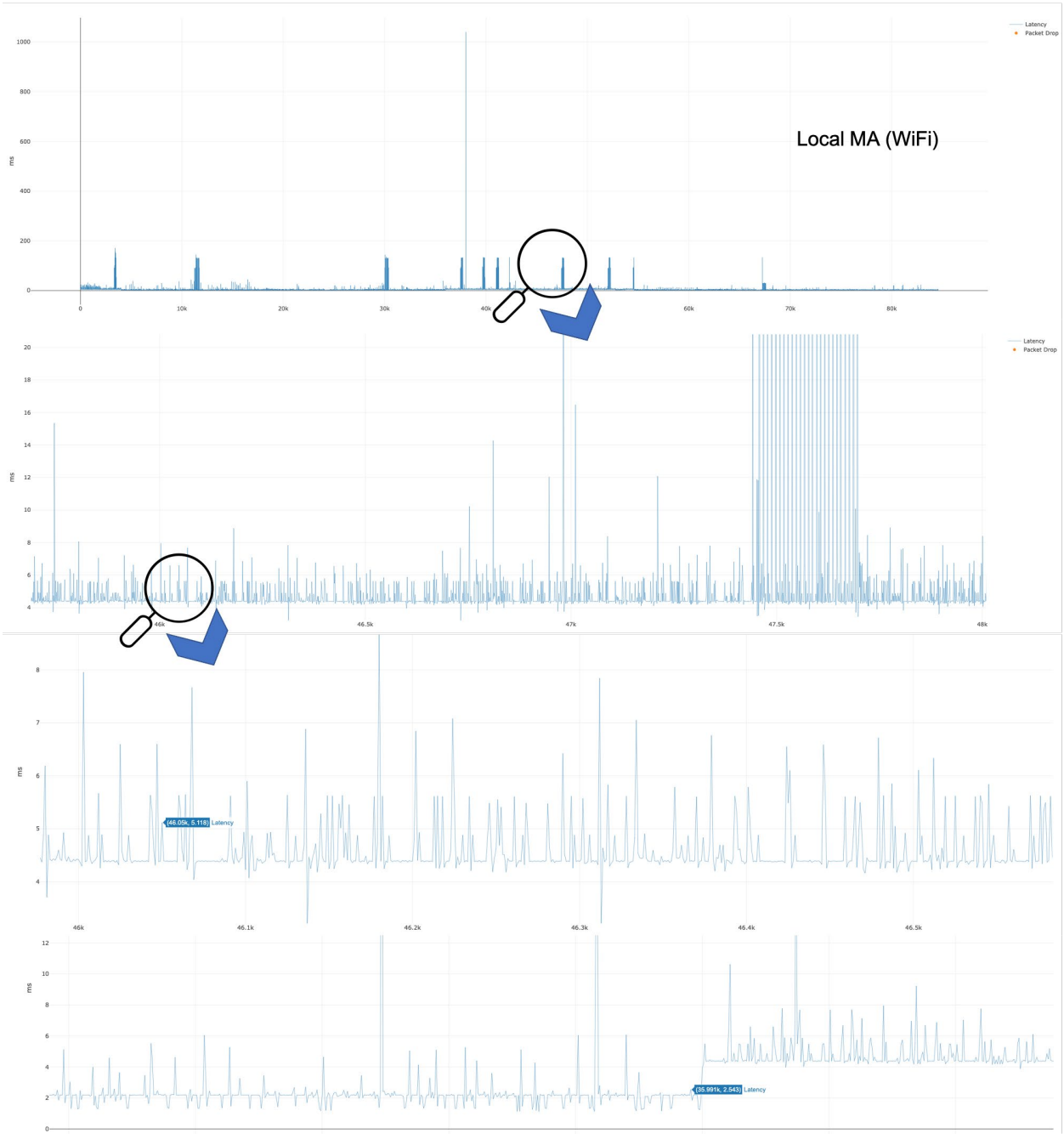


Figure 19 –LAN Wi-FI , Latency Time series

The LAN latency data, when distributed in two different bins and displayed as a histogram shows the 2.5 ms and 4.5 ms latency as the most common latencies observed in this test environment, and the CDF shows that the 50th percentile latency is around 2 milliseconds and the 95th percentile latency is around 5.3 milliseconds.

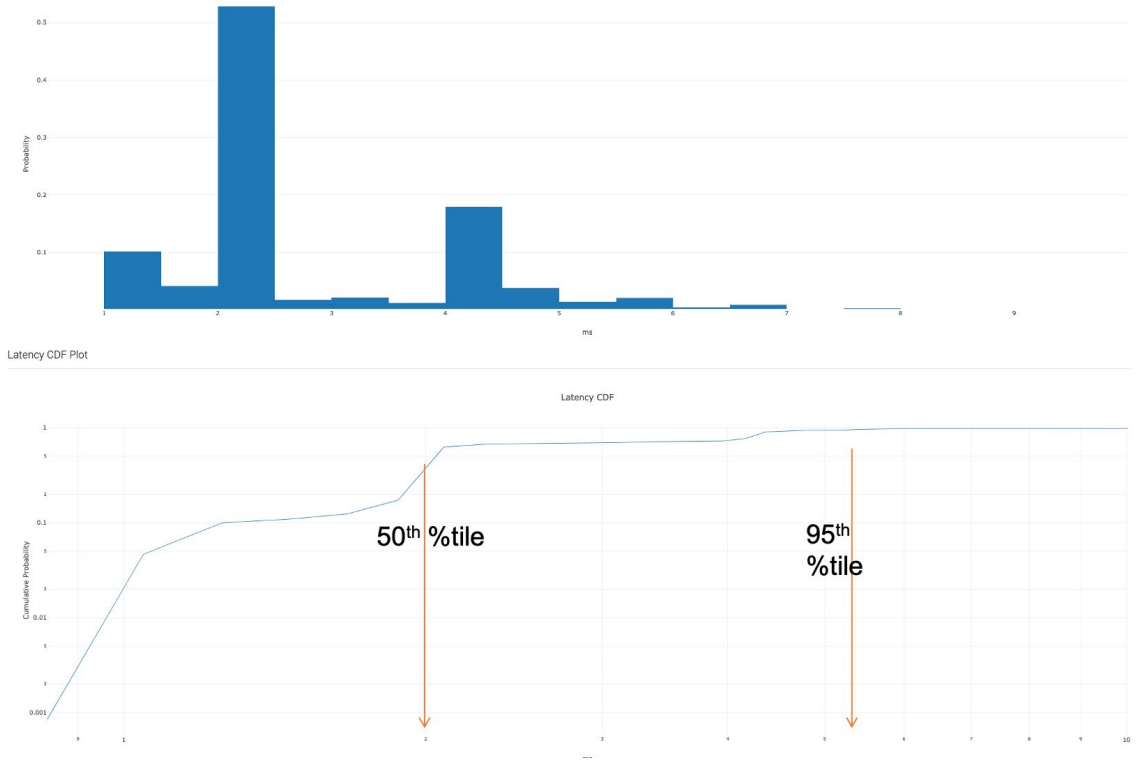


Figure 20 – LAN Histogram and CDF

5.4.5. Summary of results

The table below captures the Latency measurements percentiles of interest to the five different measurement agents four of which were spread across the country and the last was in the same LAN. Additionally, we also show the minimum and maximum latency observed across each of those five links.

Table 1 –Latency measurments to different locations

Server/Measurement Agent location (Measurement Peer in Denver)	**Min (0th percentile)**	**50th percentile**	**95th** percentile	**99th** percentile	**Max** (100th percentile)
N. Virginia	49.69 ms	56.3 ms	64.5 ms	172.2 ms	2.794 s
Ohio	40.63 ms	49.2 ms	56.6 ms	172.1 ms	2.812 s
California	29.56 ms	35.3 ms	42.5 ms	134.1 ms	2.808 s
Oregon	25.80 ms	32.7 ms	38.4 ms	116.8 ms	3.163 s
Local (LAN)	0.83 ms	2.02 ms	5.31 ms	13.41 ms	148.8 ms

6. Conclusion

Latency measurement is a vital requirement for operators deploying new low latency technologies going forward. Building a latency measurement system in hardware and software as a prototype can be relatively straightforward. Scaling it to production to measure the whole network needs planning and good engineering.

One main choice is that of a measurement protocol and here we successfully used STAMP across the Internet for latency measurement. STAMP is lightweight and easy to implement on existing hardware software platforms that it can be easily deployed by operators either in a standalone white box or as an add-on to existing devices (cable modems or gateways or APs). STAMP offers a variety of functionality (e.g., round trip and one-way measurements and loss, DSCP traversal, and different packet size testing) and can meet the needs of most Latency measurement requirements.

The second big choice is architecting the large-scale control and collection of data, and LMAP fits that bill quite well. The LMAP control and report architecture provide the operator with a well thought out set of information/data models to initiate latency measurement and collect data at scale. Understanding the CDF of the latency measurement and tracking a set of percentiles values should give an operator very good understanding of latency performance of their networks and how it changes as they deploy newer technologies.

Abbreviations

PDV	packet delay variation
RTT	round trip time
CDF	cumulative distribution function
DSCP	Diff Serv Code Point
ms	millisecond
LMAP	Large-Scale Measurement of Broadband Performance
TWAMP	Two Way Active Measurement Protocol
STAMP	Simple Two-Way Active Measurement Protocol

Bibliography & References

[IETF RFC 8762] Simple Two-Way Active Measurement Protocol https://www.rfc-editor.org/rfc/rfc8762.html IETF, RFC 8762, 2020

[IETF RFC 8972] STAMP Optional Extensions https://datatracker.ietf.org/doc/html/rfc8972 , IETF RFC 8972, 2021

[IETF RFC 7594] A Framework for Large-Scale Measurement of Broadband Performance (LMAP) https://datatracker.ietf.org/doc/html/rfc7594 , IETF RFC 7594, 2015

[IETF RFC 8193] Information Model for Large-Scale Measurement Platforms (LMAPs) https://datatracker.ietf.org/doc/html/rfc8193 IETF, RFC 8193, 2017

[IETF RFC 8194] YANG Data Model for LMAP Measurement Agents, https://datatracker.ietf.org/doc/html/rfc8194 IETF, RFC 8194, 2017

[IETF RFC 7223] A YANG Data Model for Interface Management https://datatracker.ietf.org/doc/html/rfc7223 , IETF RFC

[LM SCTE 20] Latency Measurement: What is latency and how do we measure it? Karthik Sundaresan, Greg White, Steve Glennon, SCTE 2020

[C3 CableLabs] CableLabs Common code community, https://community.cablelabs.com/wiki/display/C3

Augmented Reality and Artificial Intelligence Approaches for Inventory Synchronization

A Technical Paper prepared for SCTE by

Salvatore (Sam) Torrente
Director & Principal Solution Architect
Blue Planet, A Division of Ciena
5050 Innovation Drive, Ottawa, Ontario, K2K 0J2, Canada
+1 613 265 7614
storrent@blueplanet.com

Petar Djukic
Director, AI & Analytics
Office of the CTO, Ciena
385 Terry Fox Dr., Ottawa, Ontario, K2K 0J2, Canada
+1 613 670 3396
pdjukic@ciena.com

Dmitri Fedorov
Software Architect, Office of the CTO, Ciena
5050 Innovation Drive, Ottawa, Ontario, K2K 0J2, Canada
+1 613 670 2757
dfedorov@ciena.com

Mehran Bagheri,
AI Engineer, Office of the CTO, Ciena
5050 Innovation Drive, Ottawa, Ontario, K2K 0J2, Canada
+1 613 670 4931
mbagheri@ciena.com

Marco Naveda,
Senior Director, Network Architecture, Office of the CTO, Ciena
5050 Innovation Drive, Ottawa, Ontario, K2K 0J2, Canada
+1 613 670 2730
mnaveda@ciena.com

1. Introduction

Network operators understand that an increasing number of discrepancies between their network and inventory databases can lead to (1) greater service fallout rates, (2) impact to customer experience and (3) potential significant loss of revenue from dissatisfied customers. While networks are evolving toward more “addressable” (e.g., discoverable) devices, there is still a significant amount of resource elements in the network that are non-addressable (e.g., patch panels, fibers) thus requiring manual tracking in databases. Furthermore, most existing inventory systems either do not, or cannot, stay in lockstep with the “as-is” state of the network and continue to impose on network operators a significant amount of error-prone manual tasks. While network-to-inventory “manual audits” can provide a snap-shot update in time, each audit can cost network operators millions of dollars in consulting/staff costs and expenses they can no longer afford. And, with networks and services evolving quickly to keep up with customer demand this only increases the number of yearly audits needed to stay up to date with the network. As a result, network operators find it increasingly difficult to reconcile their physical network with their network inventory databases leading to (1) increasing number of discrepancies that is no longer manageable through traditional manual corrective means; and (2) generating an increasing amount of stranded assets that may no longer generate revenue to the network operator while still consuming space, power, and HVAC expenses.

In this paper, we will describe an architecture that uses augmented reality (AR) and artificial intelligence (AI) [1] as part of a planned network maintenance or upgrade to reconcile the physical network with planned inventory. We propose a new inventory architecture in which AR is used to capture images of network devices, seamlessly and unobtrusively, in the field during regular maintenance tasks, which are then analyzed using AI to update the inventory. It will also describe how technicians are assisted in verifying that a task was performed correctly. Visual examples will show, in tutorial style, how the technologies blend together.

2. Present Mode of Operations

The operations context in which we introduce the new inventory architecture begins with an operator planning changes to the physical network, as part of end-to-end service setup, and dispatches technicians to implement the change at a central office (CO). The technician will most likely have a workforce order (WFO) describing the changes; specifically, the order may require the technician to connect fiber “Z” to

Port “2” of a patch panel to complete connectivity, as illustrated in

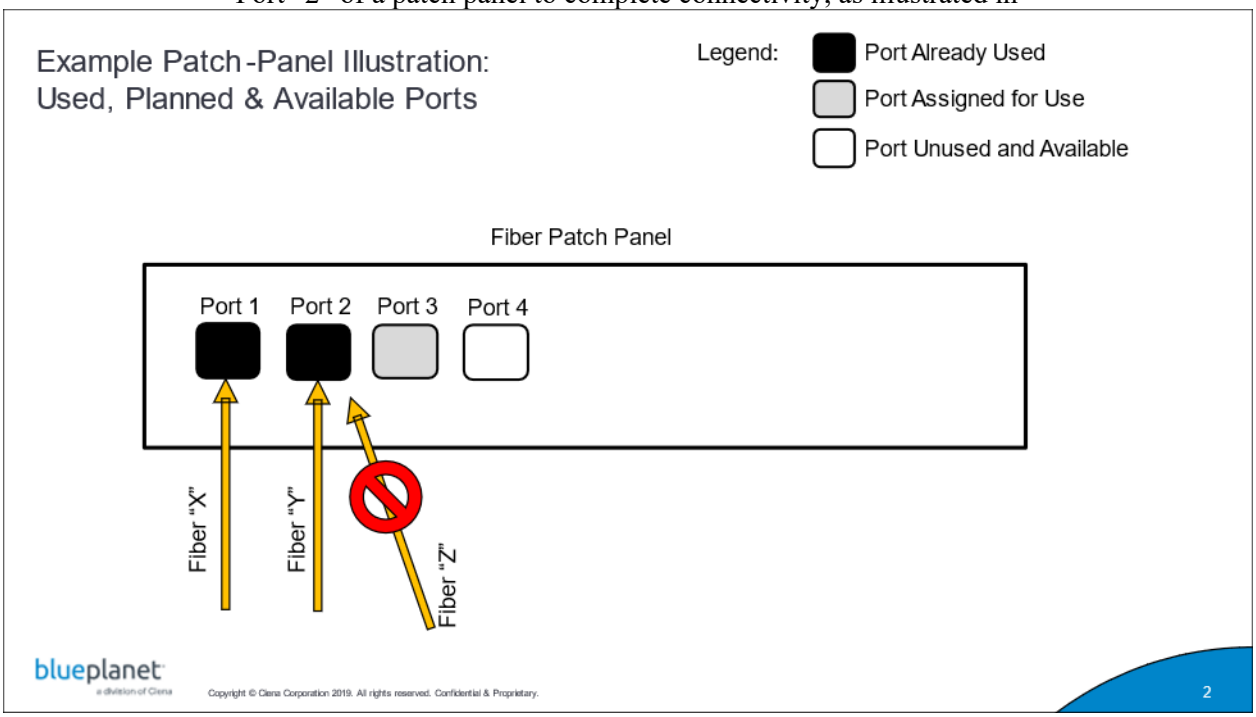


Figure 1. The problem begins that although the network operator planning team may have designed the end-to-end service based on inventory data showing port “2” available, the reality (“as-is” state) of the network shows that an existing fiber is already consuming this assigned port. Although this network discrepancy may have originated in different ways – e.g., terminated service but fiber not removed, previous workorder task assigned fiber “Y” to the wrong port, poor documented updates back into inventory… – the result is the dispatched technician will need to make a choice:

(1) Perform an ad-hoc change to the task, or
(2) Delay the task until the inventory and physical network match.

For most network operators, option (2) of first reconciling inventory with the network to then generate a new design and workorder for service introduces such delay and impact as to certainly jeopardize customer business, so option (1) becomes the de facto approach – Do what is necessary to deliver customer service now, update systems as best as possible later.

Present mode of operations (PMO) shows that network operators will define decision protocols for the technician to assist in option (1). For example, often the technician will not know if there is live traffic on the offending fiber (Fiber “Y”) consuming port “2” eliminating the possibility of disconnecting the fiber to release the port. As a result, a decision protocol may guide the technician that in case a port defined in

the workorder is already consumed, select the next available port on the panel, say port “3” in

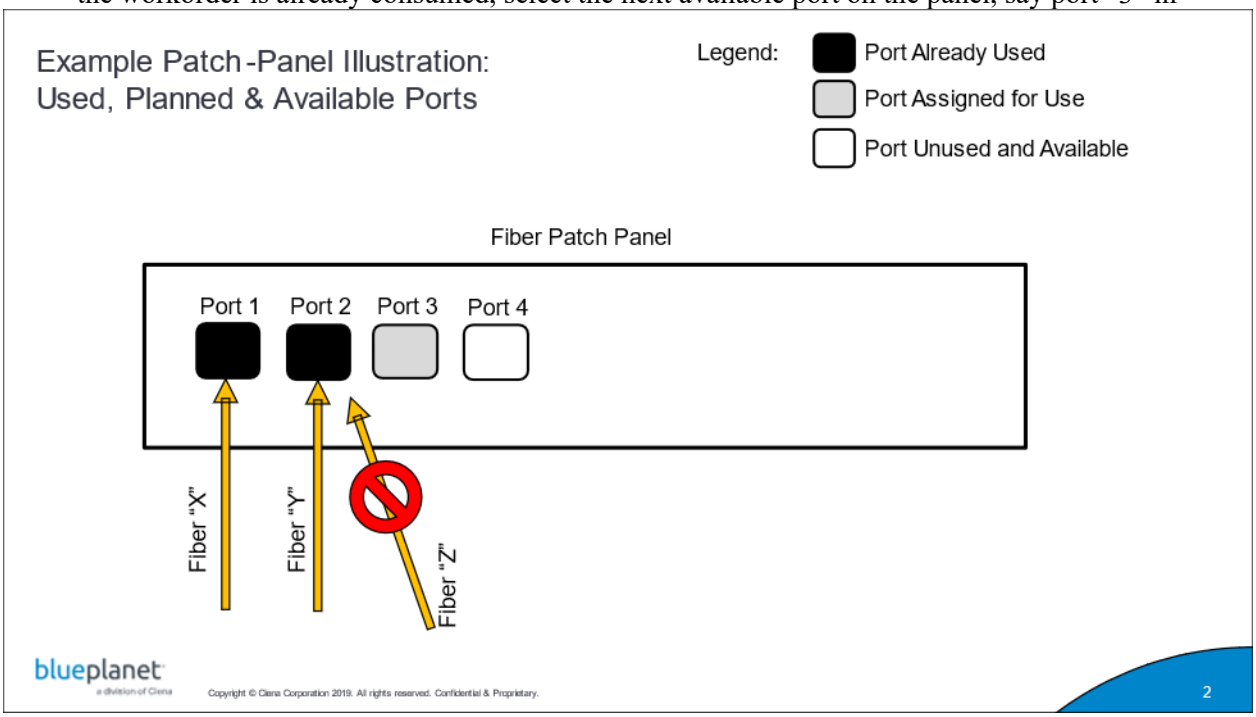


Figure 1, and document back to the planning team the newly selected port. What the technician will not know, and workforce order will not capture, is that port “3” may already be planned in inventory and assigned for another service to be connected in the field at a later time, making port “4” the next unused and viable port for service. This approach exacerbates the out-of-sync inventory problem and works against their efforts to plan network updates.

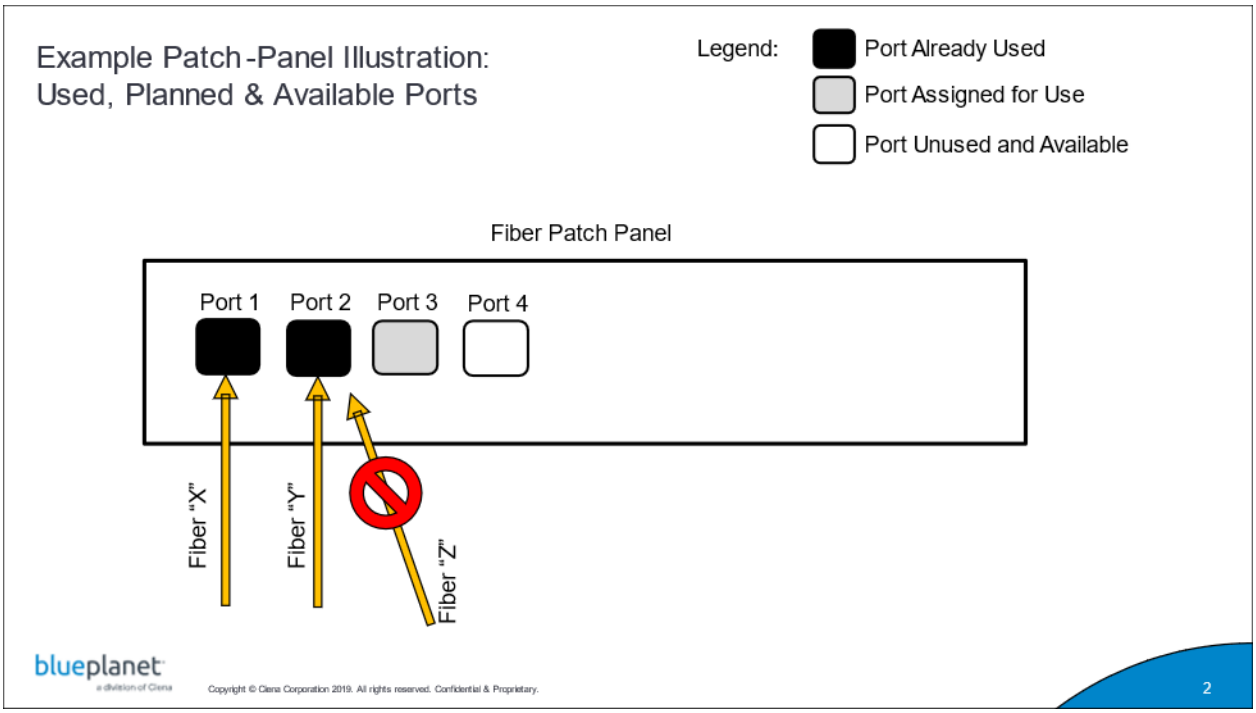


Figure 1: Patch Panel Illustration

Instead, what if we can use AR with physical visual devices such as phones, tablets, or "smart glasses" to superimpose real time digital records over the visuals of physical reality to guide operations and technicians to first time right service delivery [2]. And, if we can use AI with image processing and deep neural networks for object detection and image segmentation to convert physical reality into digital records ("as-is" network view) that can rapidly resolve discrepancies. Operators have a growing interest in using image capture, recognition, and AI [3] [4] to digitize inventory and track physical network assets while exposing any discrepancies in the network.

3. AR and AI for Inventory Synchronization

The first step in our proposed architecture approach is to enable the technician to quickly take stock of the "as-is" state of the resource elements and then compare with planned equipment for alignment before engaging any changes. Figure 2, provides an example network device as we would see in a CO, the device has several ports in use and cables dropping from the used connections. Here, some ports are used and cables that connect to the port overlap and sometimes cover available ports. The approach to capture and digitize the current port usage will need to determine which ports are used and which are available underneath the mesh of cables.

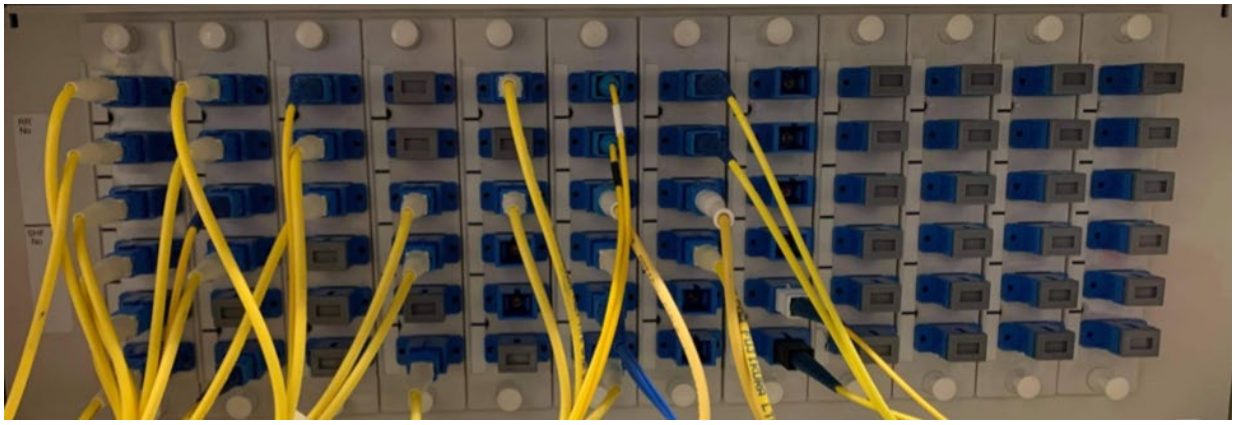


Figure 2: Example Network Device

For this we introduce the use of augmented reality (AR) combined with smart device enabling the technician to capture and digitize the target equipment as shown in Figure 3. The technician would aim the smart device camera toward the network device and ensure it fits within the boundary of digitization border of the application.

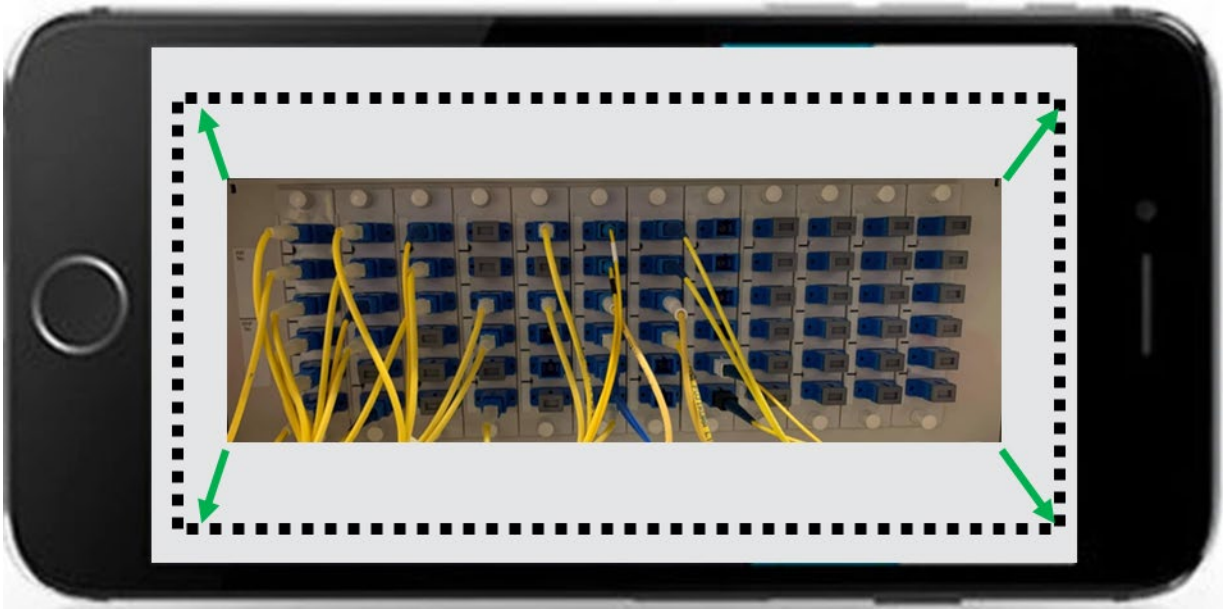


Figure 3: Device Capture with Augmented Reality

Because a single frontal picture would certainly not take a clear enough picture of ports through the cables, the augmented reality application would guide the technician to pivot the smart device in direction of the arrows shown on screen so the smart device can collect a series of picture or video from different angles creating a rich multi-perspective recording of the device before automatically closing the camera

indicating the capture process is complete. This image capture process would take less than a minute of the overall technician's time and have negligeable impact on his task schedule. The image capture has a dual purpose:

A) Digitize the "as-is" state of the device to be compared with the planned view of the device in inventory database and identify any potential discrepancies.
B) Become the reference image to guide the technician on the device tasks ahead.

Before we can identify any potential discrepancies, the digitized device must first be identified so the system knows what to compare with and what device to present back to the technician as operational guidance in next step tasks. For this, when the smart device captures the device image, it can also capture the device bar code, and/or the unique ID (combination of location, floor, shelf, etc.) many network operators stamp on devices today, and/or the physical characteristics of the device itself. The latter part means comparing the physical layout of the digitized device with digital reference pictures/models from device vendors themselves. When combining the above vectors of identifications with GIS (Geographic Information System) location the AI function can present back to the technician through their smart device an accurate determination of the network device discovered. In absence of enough identification vectors the smart device will ask the technician to enter a known ID code for the device to finalize selection.

The next step is for the image technology tied to digital image capture to determine what ports are used or available for use. This is done by extracting the features of the ports (used and available) and comparing them to models of available ports and computing their likely association to either available or used, as illustrated in Figure 4. This is where the importance of having created a stream of images or video by pivoting the smart device camera around the network device is key. As the image processing computes the likelihood across different image perspectives some ports will show consistently 100% available (green) or 100% used (red), while many of the ports that were "hidden" by cabling may be computed to higher accuracy as each perspective reinforces either "used" or "available". For ports that are not resolved as either 100% used or available, the smart device can allow the technician to "pinch-zoom" on the area of the device image they will work on and enable the technician to confirm by touch (yes/no) if the port is consumed or not. It is important to note that as technicians continue to take image captures over time as they work on the device the movement of cables, changes in connection and exposure to different angles will create an increasing accurate perspective of the as-is state of the device and phase out any potential differences across image captures.

Figure 4: Digital Image Map

At this point the resulting discovered digital image map is ready for comparison with inventory database to create a discrepancy map that will guide the network operator to correct and reconcile inventory to the network. The discrepancy map would list all the objects in database that do not match with the digital image map with associated probability based on the accuracy of the digital map. The network operator can take the discrepancy of highest to lowest probability and initiate reconciliation processes to correct discrepancies in future in-field work, or directly guide the technician on location to make the correction if in the

technician's remit. An example can be a cable is occupying a port that the inventory systems know the service was terminated and cable should have been removed but not completed by the last technician. By continually feeding image captures of network devices as changes are implemented the imaging technology and underlying AI engine creates a more accurate view of discrepancies and their associated probabilities. This enables the network operator to focus on correcting the discrepancies of greatest value to the business, all while the technicians continue their daily work orders with the addition of image capture.

While the AI system computes and updates the digital image map and probabilities of discrepancies continuously, the technician is guided in his workorder operations through the smart device without delay. Let's take the original example in

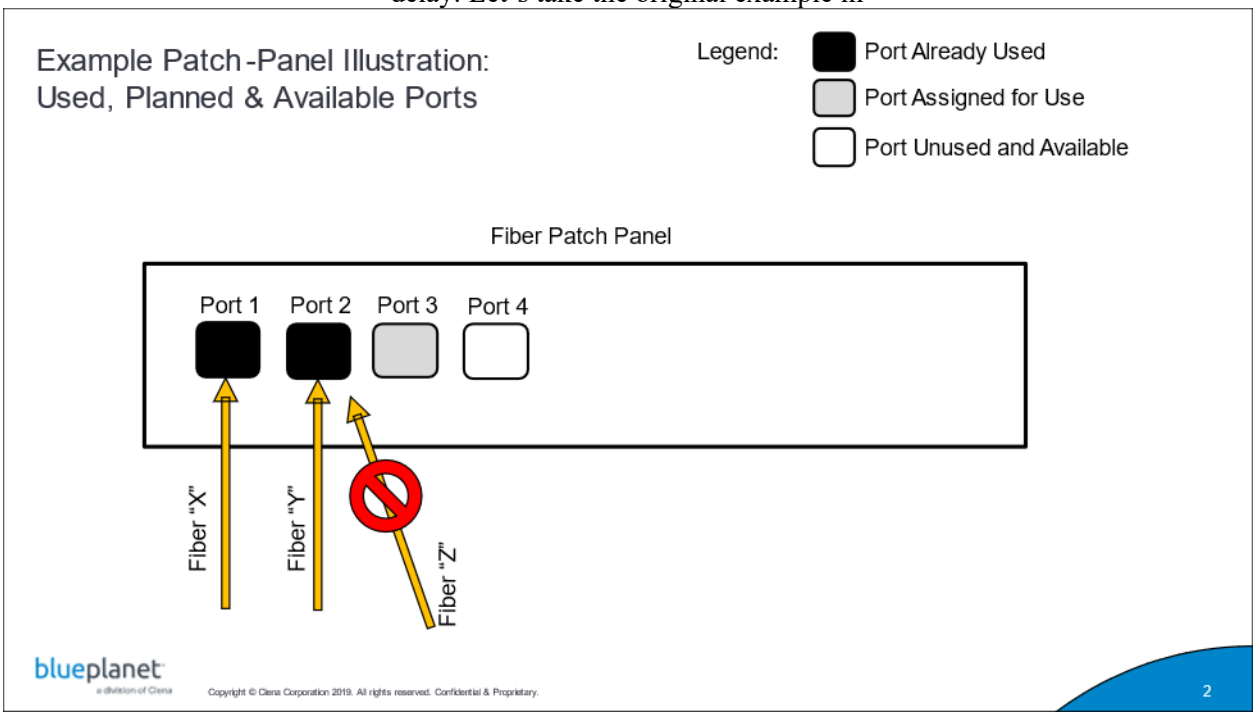


Figure 1: The technician has a workorder to connect a fiber cable to complete an end-to-end service. The workorder indicates that fiber "Z" should go to port "2" but the port is already used. Although the technician may have a protocol in place indicating he should use the next "available" port, the selection of the port will benefit from a guided operations approach from the combined AR/AI system. Because port "3" is already assigned for future service in the inventory database, the smart device will guide the technician to connect the fiber to port "4" as illustrated in Figure 5.

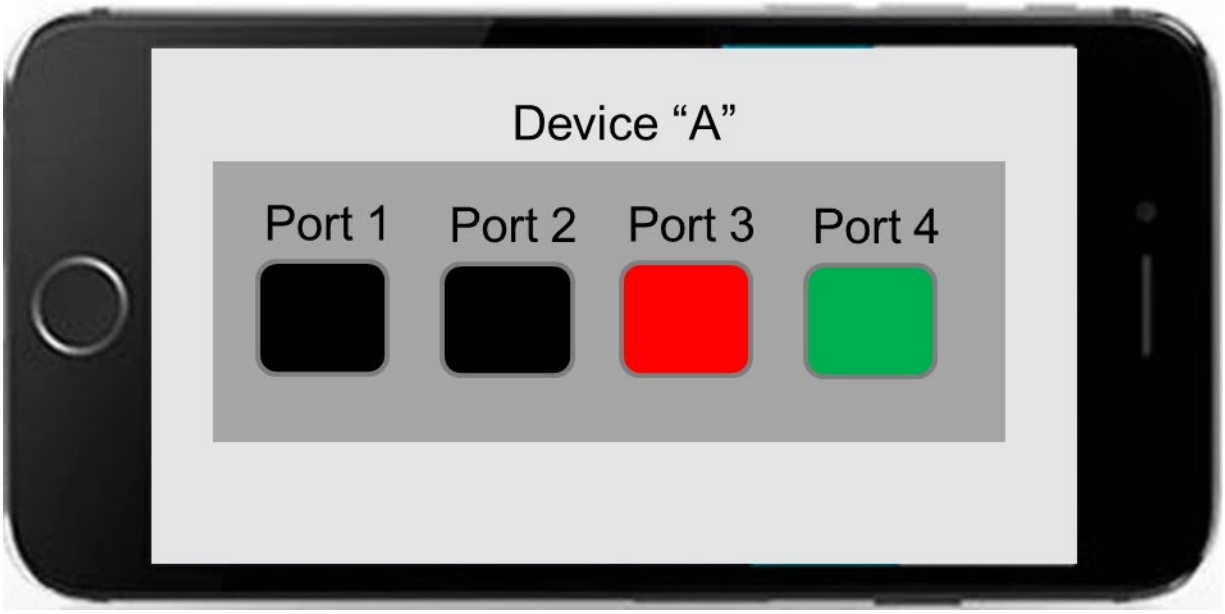


Figure 5: Guided Operations - Connections

Once the fiber connection is completed, the technician will capture another image of the device in its current “as-is” state which will update the inventory database with the new connection and can include communication to the planning team that an update to the original plan was required to complete the task.

The proposed architecture that captures the above AR and AI technology and guided operations can be illustrated as in Figure 6. At the heart of the architecture is the federation of data that that brings together under a common normalized information model, the following:

- Network operator defined device template used to model and identify the equipment to be analyzed for discrepancies.
- The Inventory “as-planned” view of the device to compare with the digital image map for discrepancies.
- The digital image map of the network device, generated by the imaging technology, from the captured smart device video/pictures.
- The federated data can also include vendor defined device templates that may be more accurate/complete than inventory models and when combined with GIS data enables very accurate determination of network devices across a wide range of resources in inventory.

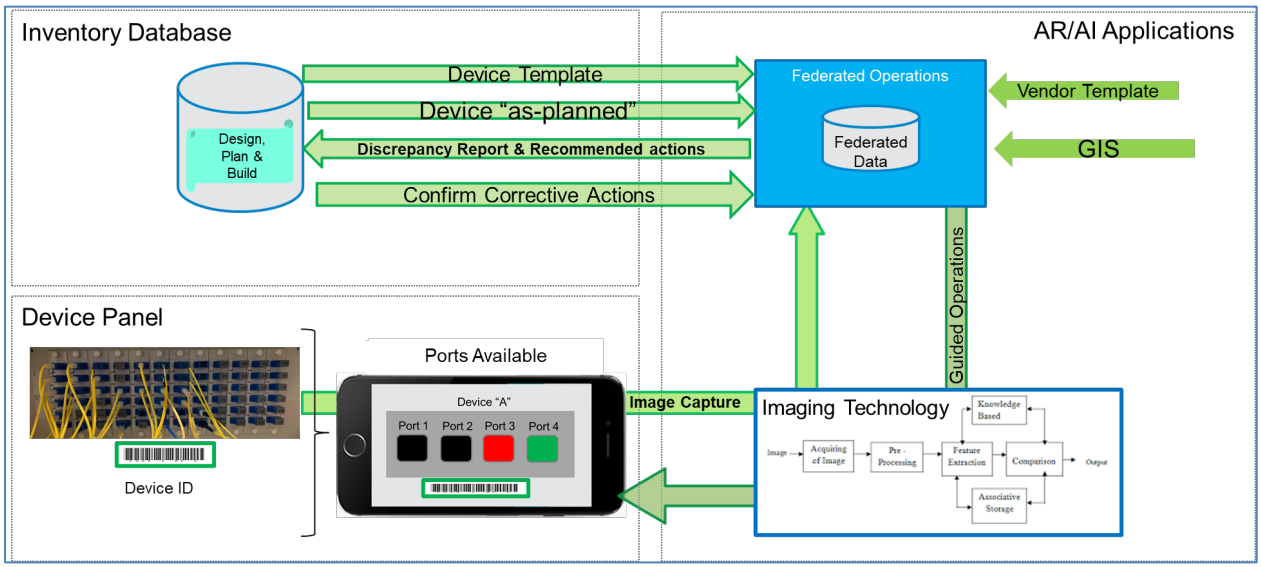


Figure 6: Proposed AR/AI based Architecture

Discrepancy detection rules are applied across the federated data to generate discrepancy reports where each discrepancy is mapped from highest to lowest level of accuracy and associated recommended action for remediation. Here the network operator can choose to either tabulate a series of reports into a dedicated workorder where a technician can be dispatched at a prescribed time and location to correct these discrepancies, or, the architecture allows the guided operations (as shown in Figure 6) where the corrective action is communicated back to the technician through the smart device and corrected in flight. Both approaches enable the network operator to achieve increased inventory and network synchronization (INS) toward eliminating device discrepancies that can impact their customer, their business, and their competitiveness in the market.

4. Future Mode of Operations

The AR/AI architecture presented in this paper is a proposed future mode of operations (FMO) to enable network operators in stopping and reversing the proliferation of network discrepancies that impact their business today. In this future mode of operations, the technicians in the field use readily available image capture tools like smart phones and/or tablets combined with existing AR software to provide a constant feed (crowd sourcing) of "as-is" state of devices as a constant stream of data complementing any traditional discovery process to keep the inventory database in lockstep with the network. The imaging technology and associated AI that creates the digital device map, and federated data infrastructure, can be cloud-based making the approach scalable to different business needs. Communications across image capture tools, imaging technology, federated data and Inventory databases can be achieved through open APIs for greater ease of integration. The result of this proposed architecture is to finally close, in as real-time as possible, the inventory and network synchronization loop that has eluded the industry for so long.

It is important to note that while the example used in this paper describes an approach to resolving discrepancies related to fiber connection, this same architecture can be used in wider context, for example:

- Equipment upgrade: Upgrading card, shelf, or device where the image technology is used to map the existing device to be upgraded against the new device characteristics to guide and confirm toward a successful upgrade.

- Initial equipment discovery: The digital image captured by the technician and resulting digital device map generated through AI can be quickly mapped to vendor templates that trigger the creation of accurate device object entries in inventory without lengthy or costly audit processes in the field.

Although the proposed architecture provides a promising approach to inventory and network synchronization, this future mode of operations would also need to fit within the following operational considerations:

- Recording devices: Not all network operators allow recording devices to capture network information. The off the shelf devices and software proposed with this architecture would need to be authorized on premises.
- Workforce: A technician's contracted responsibilities may not allow for the additional task of capturing images/video as part of workorder flow.
- Secure data/connection: The processing of images/video in the cloud would mean secure connection and possibly encryption between network operator and federated data.
- Inventory database access: Many existing and legacy inventory systems in network operators may not allow direct access to inventory data and could require multiple data exports that would be ingested into a federated environment.

5. Conclusion

This paper has proposed an AR and AI architecture approach to inventory synchronization. The approach shows that with off the shelf smart devices and the use of AR a technician can capture a rich set of multi-angle images/video data in very short time and effort that can be used by AI to quickly identify the device and generate an accurate digital image map for discrepancy detection between network and inventory. We also described how conversely AI and data federation can provide the technician, through AR, a guided operations approach to completing a workorder task without compounding network discrepancies and even correcting in real-time previously reported discrepancies. The proposed architecture can be either on-prem or cloud-based depending on the business needs of the network operator. The architecture approach applies to the connectivity example presented in this paper but also in wider context of device/equipment upgrade and maintenance. And the approach can be used as an initial network discovery and inventory setup method for network operators wanting to populate their new inventory system with a current view of network devices in the field.

Abbreviations

AR	augmented reality
AI	artificial intelligence
PMO	present mode of operations
CO	central office
WFO	workforce order
INS	inventory network synchronization
FMO	future mode of operations

Bibliography & References

[1] L. Blair, "Understanding the differences between virtual reality, augmented reality and mixed reality," 10 August 2016. [Online]. Available: https://www.networkworld.com/article/3106205/understanding-the-differences-between-virtual-reality-augmented-reality-and-mixed-reality.html.

[2] "A new, more efficient (augmented) reality for network operations," Ciena, 21 January 2020. [Online].

[3] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only Look Once: Unified Real-Time Object Detection," p. 10, 9 May 2016.

[4] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "SSD: Single Shot MultiBox Detector," p. 17, 29 December 2016.

Bandwidth Planning During the Age of CoVID

A Technical Paper prepared for SCTE by

Keith Alan Rothschild, Ph.D.
Senior Principal
Cox Communications
6305-B Peachtree Dunwoody Rd, Atlanta GA 30328
+1 404 269 8122
kar@cox.com

Andrew Robinson, Network Planning Engineer III, Cox Communications

Derek Bantug, Network Engineer II, Cox Communications

Kris McNally, Senior Technical Project Manager III, Cox Communications

1. Introduction

What initially started in early March 2020 as a multi-day trial to determine if the company could rapidly move to support a remote workforce, as part of emergency preparedness, quickly pivoted into a remote workforce for roughly 18-months in response to the pandemic. This was not only a radical change in the way we work, but also in the way our customers use our products. Disruptions in supply-chain, construction, and customer interactions added challenges to what would have otherwise required a tremendous-level of effort to ensure the customer experience remained as consistent and positive as possible.

Fortunately, we have a standing process for managing our bandwidth on an ongoing basis. This process has four components: (1) strategize, (2) model, (3) plan, and (4) deploy. This paper reviews each of these four components in general as well as how they enabled us to rapidly respond to the increased demand driven by the pandemic.

2. Strategize

The last decade has seen an evolution from planning for the co-existence of multiple products with discrete bandwidth requirements on the HFC network to all products utilizing IP transport on multiple access network technologies. Telephony evolved from circuit-switched telephony to IP/packet-switched. Video evolved first from analog to digital, then from MPEG-2 to MPEG-4, and dedicated QAM-based to shared IP-based delivery. This paper will focus on meeting the bandwidth needs resulting from the converged delivery of products over IP.

The general formula we use to determine the amount of bandwidth required (C, capacity) as being at least the sum of the Peak Traffic (we use P_{95}) and Max Tier (T_{Max}):

$$C \geq P_{95} + T_{Max}$$

The Peak Traffic is sometimes looked at as the product of the number of subscribers and the per-subscriber contribution to peak traffic, however, we use the measured value that the node is at or below for 95% of the time (P_{95}). When additional capacity is available, there are peaks in traffic that can be handled by the network without loss of data and which typically should not impact the customer experience, so we utilize this measure to prevent those peaks from driving unneeded capacity.

The maximum tier is typically the maximum advertised speed plus some overhead to ensure a positive quality of experience. Ulm & Cloonan (2017) list these as separate variables, and many argue that the overprovisioned we include in the tier is not the same as the quality of experience variable they intend. Cloonan (2014) refers to the QOE experience as K. Rather than treating these separately, we typically overprovision the customer by ~10% and use the overprovisioned value as T_{Max} in our calculations.

Capacity calculations for upstream and downstream are treated separately. For example, if the maximum tier on a node was 300/30 (300 Mbps downstream, 30 Mbps upstream), and these were overprovisioned by 10%, then we would use $T_{Max\text{-}D}$ of 330 and $T_{Max\text{-}U}$ of 33. In a node with 32 SC-QAM (37.5 Mbps) channels on the downstream and 4x6.4 MHz QAM (26.88 Mbps) channels on the upstream, then we can figure out the maximum amount of traffic that the nodes can bear before a node action is required as follows:

$$C \geq P_{95} + T_{Max} \xrightarrow{rearranged} P_{95} \leq C - T_{Max}$$

Downstream	Upstream
$P_{95} \leq C - T_{Max}$	$P_{95} \leq C - T_{Max}$
$P_{95} \leq (32 \times 37.5) - 330$	$P_{95} \leq (4 \times 26.88) - 33$
$P_{95} \leq 1200 - 330 \leq 870$	$P_{95} \leq 107.52 - 33 \leq 74.52$

For QAM, we use 7/8 of the raw throughput to estimate the available payload capacity. Although the traffic limits are based on these formulas, operations typically use a "70%" rule, which has worked as these numbers are close to 70% of the capacity, though in the future, this will need to be addressed as it is likely that this will not remain close to 70%. Today, nodes are considered targets for node-actions at 70%, congested at 80%, and heavily congested at 90%, with the latter two corresponding to increases in trouble-calls.

Approaching the target maximum P_{95} value in either the upstream or downstream will typically drive a node-action. In practice, we use a model to project when a node will approach the capacity limits to plan node actions much further in advance, and closely monitor actual data to make modifications to the plan based on changes that could not be predicted.

Downstream traffic is estimated to grow at a compound annual growth rate (CAGR) of 35% while upstream traffic grows at a 26% CAGR. It is not feasible for us to accommodate this type of growth in every part of the plant each year, instead we plan to address 20% of our footprint each year in anticipation of completing an upgrade cycle each 5-years. This means that our target design needs to result in an increased peak traffic capacity of about 4.5x on the downstream and 3.2x on the upstream. These targets do not include increases otherwise needed to support higher maximum advertised tiers.

Increases in product offerings don't follow the same growth patterns as usage. The maximum offered downstream speeds tend to increase at about 3.2x over a 5-year period. There appears to be pressure on the upstream to increase to symmetric speeds, but it is unclear if that will continue over the long term. It is probable that the primary driver is more likely to be around latency rather than sustained throughput, but marketing from competitors will likely keep the need for increased speed/bandwidth.

There are two types of bandwidth levers: optimize existing spectrum and increase available spectrum. Optimizing existing spectrum could be from optimizing the underlying technology, increasing the efficiency with regards to the payload, and/or by reallocating spectrum between services. Increasing available spectrum can come from repurposing spectrum, reducing the number of subscribers sharing the spectrum (segmentation/node-split), or increasing the frequency range (spectrum) that the plant is able to support.

If we assume that the downstream traffic was 870 Mbps and needs to accommodate growth of 4.5x over the next 5 years, and the maximum tier was 330 Mbps and needs to accommodate a growth of 3.2x over the next 5 years, we can apply our formulas as follows:

$$C \geq P_{95} + T_{Max}$$

$$C \geq (870 \times 4.5) + (330 \times 3.2) \geq 3915 + 1056 \geq 4971$$

This means that we would either need to increase the amount of downstream capacity to nearly 5 Gbps, segment the node, or do some combination of the two. Having the benefit of hindsight for this example, we knew that

over that timeframe we were looking to expand the downstream to 48-SC-QAMs and 192 MHz of OFDM, which would give us a downstream capacity of 3.6 Gbps, meaning that some degree of segmentation would be required. We also know that the T_{Max} was actually a 1 Gbps service with a 10% overhead, or T_{Max}=1100 Mbps. We assume that a simple node-split typically results in a 60/40 split in the traffic, a three-way split results in a 40/36/24 split, and a double split results in a 36/24/24/16 split in the traffic. T_{Max} remains unchanged.

$$P_{95} \leq C - T_{Max}$$

$$P_{95} \leq 3600 - 1100 \leq 2500$$

In order to achieve this the higher traffic leg of the segmentation would be targeted to grow to less than 2500 Mbps, which assuming a growth factor of 4.5x means it should be ≤ 555.5 Mbps, or 63.8% of current peak capacity, meaning that most simple node-splits should suffice.

2.1. CoVID Impacts

One of the major impacts of CoVID was to drive people to shelter at home, meaning all activities, including work and school, became dependent upon the residential network. Nearly overnight we noticed an increase in network utilization equivalent to what we would expect after a full year's growth, and at the 1-year mark we measured the expected 35% increase on the downstream but an unprecedented 55% increase on the upstream. The graph below shows a typical weekday before CoVID in orange and after CoVID in blue. It is also worth noting that while the downstream peak remains in the evening (and in-line with typical annual increases), the upstream peak moved to start in the late morning and usage remains significantly higher for roughly 10-hours.

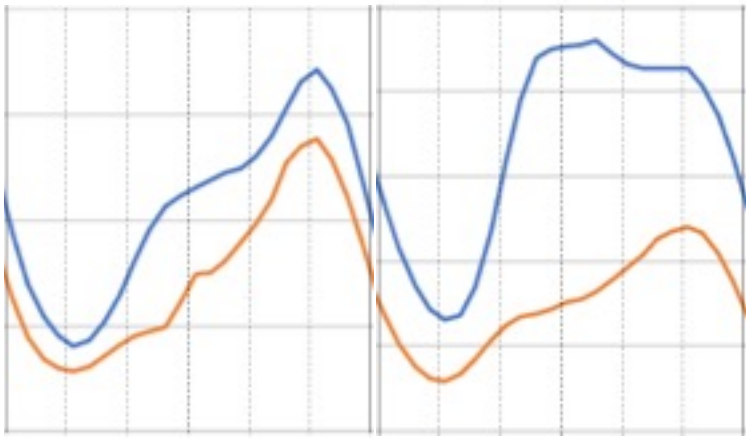


Figure 1 – Typical Weekday Bandwidth Impact, Downstream (left) and Upstream (right)

Our plans to accommodate downstream growth had already included reclaiming spectrum from other services and growing the amount of spectrum allocated to DOCSIS, especially focusing on DOCSIS 3.1 enabled devices that support OFDM capacity. Our plans to accommodate upstream growth were based on DOCSIS 3.1 and a migration to Mid-Split, including the use of OFDMA in the mid-split spectrum. We did not feel that accelerating the move to Mid-Split would enable us to respond as quickly as required to customer demand.

The first option that was proposed was to reallocate some of the ATDMA carriers to OFDMA to increase capacity. This is known internally as Sub-Split OFDMA (SOFA). The challenges with this approach were multi-fold, including requiring a higher penetration of DOCSIS 3.1 devices than was present at the time and a lack of maturity of OFDMA itself. While not implemented for this immediate circumstance, we kicked off a program that will allow us to implement SOFA should the upstream impact continue to last for many years.

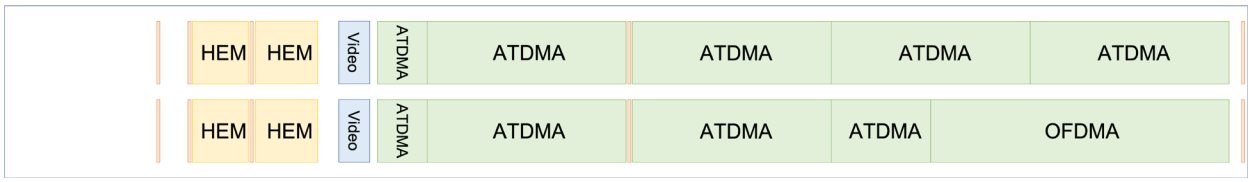


Figure 2 – Pre-CoVID Spectrum vs. Initially Proposed SOFA

We reached out to customers in heavily utilized nodes who were heavy users either intentionally (for example, to incentivize them to move to one of our Fiber products if that option was available) or unintentionally (for example, increasing outreach to customers whose devices may be impacted by malware or may be transmitting higher volumes of traffic due to configuration issues).

The option we ultimately implemented is known internally as fifth-carrier (5C) as it is the fifth carrier in the DOCSIS sub-split upstream that is intended to be used to deliver internet service (there is also a 1.6 MHz ATDMA carrier designated for DSG, which is why the initial state shows five items labeled ATDMA). This required us to work closely with all our markets to make alterations to how the upstream spectrum was utilized by non-Data products. Thanks to significant ongoing efforts resulting in moves from circuit-switched telephony to IP-Telephony, legacy video return to DOCISIS-based return, and dedicated telemetry return paths to using embedded DOCSIS modems in our plant gear, we had the opportunity to retire and/or relocate many legacy carriers, giving us enough spectrum to launch an additional 3.2 MHz upstream channel.

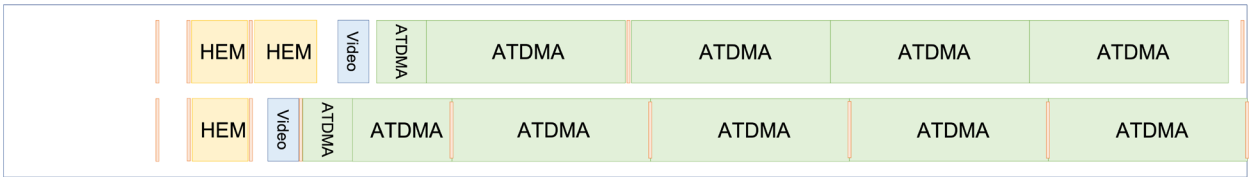


Figure 3 – Pre-CoVID Spectrum vs. Fifth Carrier

This brought along many challenges which will be discussed later in this paper, but this addition of 12.5% more raw spectrum gave us 18% additional capacity towards the P_{95} (as T_{Max} remained the same). This served to reduce the number of congested nodes back down to within 0.5% of the pre-CoVID levels, with an expectation that as this did not completely offset the 26% increase, the number of nodes approaching congestion could be expected to reach a level between 3% and 6% for the remainder of the 5-year upgrade cycle unless another activity brought additional relief. That other program is the previously mentioned SOFA program.

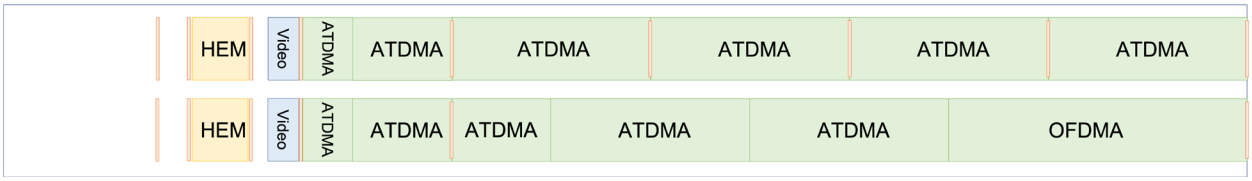


Figure 4 – Fifth Carrier vs. Fifth Carrier with SOFA

When coupled with the 5th Carrier, the newly proposed SOFA option has 2x3.2 MHz and 2x6.4 MHz ATDMA carriers in addition to the OFDMA in the subsplit spectrum region.

3. Model

Due to the sudden change in demand as customer behavior changed due to CoVID, it was crucial for our teams to quickly analyze the growing bandwidth concerns, predict the range of future growth patterns, and determine the impact of different strategies in handling each scenario. The automation of our planning tools allowed such a wide variance to be analyzed within days, and for solutions to be proposed and adjusted with similar speed.

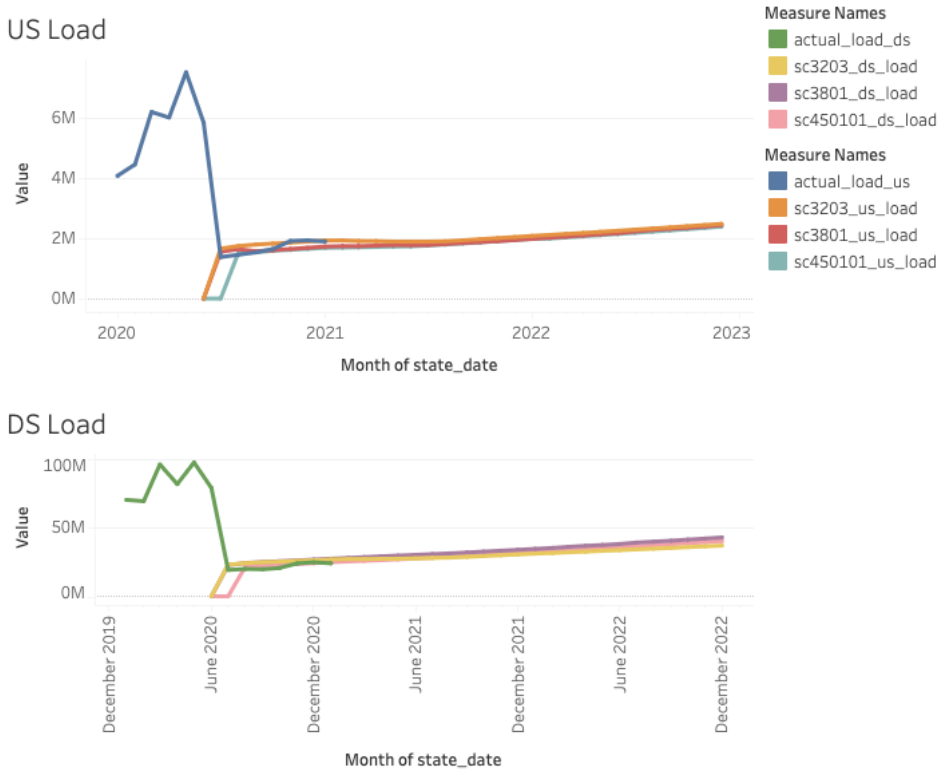


Figure 5 – Enterprise node load compared across forecast scenarios

Our forecasting model predicts network load based on historical node-level traffic data, weighted by seasonality and outlying factors, like the recent node action spike due to COVID. This predictive modeling is constantly updated to reflect new information, or to confirm existing patterns and narrow down the range of possible outcomes. Even with unexpected shifts to otherwise predictable data, the forecast gives a statistically sound expectation.

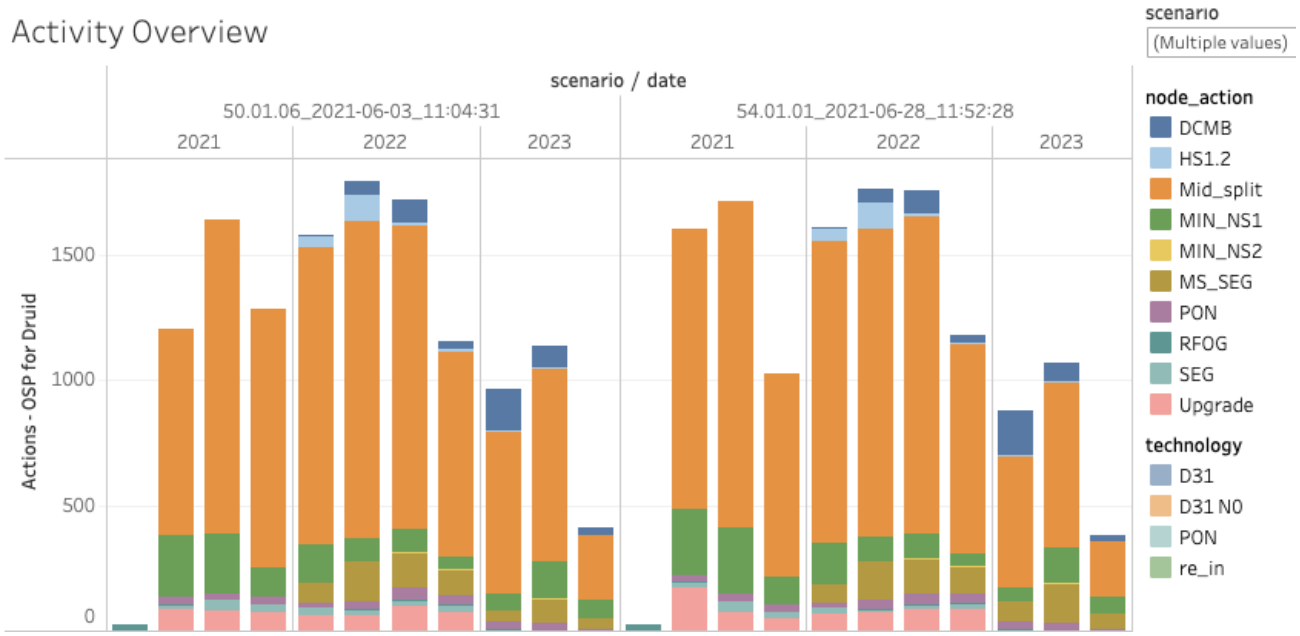


Figure 6 – Automated node action comparison example

Given any one set of expected network utilization, the Capacity Response module compiles a logical action priority based on a determined set of business requirements. As the strategy is adjusted by the planning team, the model us updated to reflect the appropriate action thresholds, technology definitions and restrictions, and deployment rates. The output of any given scenario provides an overview of the deployment schedule required at minimum to manage utilization, which is then processed through the Volume Model for the respectively required hardware and licensing per facility.

With these modeling capabilities, proposals for new technology considerations underwent rapid approval processes as we were immediately able to give relevant estimations of impact for each case. Ultimately, the abstraction of the underlying logistics to each solution allowed the planning teams to focus on solving these urgent problems at a much larger scale.

4. Plan

On Tuesday afternoon of March 10, 2020, we were told to pack up and head home early. The explanation was that the company wanted to conduct an emergency preparedness test with us working from home for a few days. By the end of the week, it was obvious that this was due to a virus with a peculiar name that was starting to get more headlines and an increasing sense of concern. By that weekend, it was clear that we would not be back in the office for at least a few weeks with few, if any of us, expecting that it would be nearly 18 months (or longer) before any of us were back at our desks.

By the following week, the node-action escalation team was seeing an uptick in node-action tickets as people started working from home and students began taking classes online and driving up upstream traffic utilization. Almost overnight, Zoom went from being a relatively obscure app to part of our lexicon. Daytime commercial usage shifted from high-capacity business infrastructure to residential nodes that until then were used primarily for daytime TV watching, personal email, and gaming. At nearly the same time, an internal node-action ticketing automation tool went live, and the node-action tickets went from a typical one or two a week to nine or ten per day, eventually averaging 24 per day in April. The small team responsible for processing the node-

action tickets was quickly overwhelmed, leading to recruiting other employees to sort and triage the incoming flood of tickets.

The months of March, April, and May of 2020 saw over 1,200 node action tickets processed in those three months alone, versus a total of 77 by the same team for all of 2019. The previously discussed traffic growth predictions had gone completely out the window as we experienced over a year's worth of demand increase in less than a month.

After it was decided to use a 5th upstream ATDMA carrier to provide more capacity, the initial thought was that we would add this to every node in all markets. This had the advantage of being a global response, however over two-thirds of our nodes were still below the node-action thresholds, and the 5th carrier would have simply been excess capacity. It also would have meant many nodes consuming an unneeded upstream license at a cost in millions for the unnecessary licenses. Thus, the decision was made to add the 5th carrier on a node-by-node basis.

Once that was decided, members of the bandwidth management team were tasked with reaching out to the markets, the video teams, and the regional DOCSIS teams to prepare them for the effort, while the EMO team began pulling the node utilization data from the ACOE group each week. As that was happening, trial nodes were being identified in multiple markets. Where needed, HEM channels to support the circuit-switched telephony were being moved or removed, and the non-DOCSIS based video-OOB signals were relocated (SCTE 55-1 and SCTE 55-2). This was across the enterprise and was a coordinated effort between multiple corporate and field teams and their boundary partners.

Figure 7 depicts a typical transition as it was implemented.

1. The top row represents a typical starting configuration prior to pandemic
2. Where applicable, plant sweep channels and telephony HEM channels were moved or removed
3. The legacy video upstream OOB was relocated
4. The DOCSIS carrier used for DSG was moved downward by the regional DOCSIS team to create a gap
5. New plant sweep channels were placed
6. Applicable nodes had the 5th DOCSIS ATDMA carrier added for use by cable modems and EMTAs

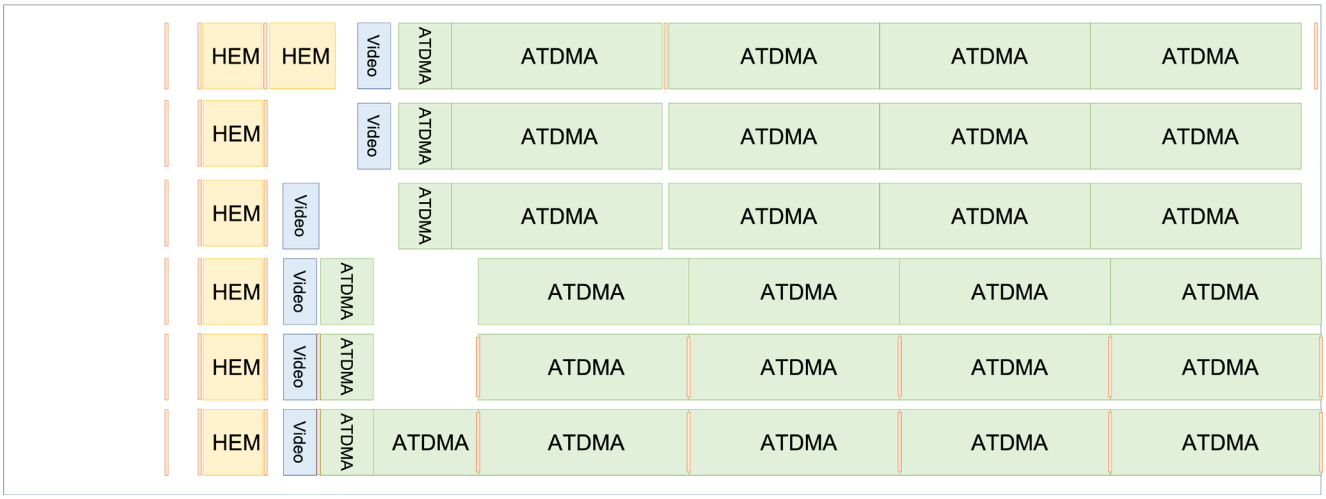


Figure 7 – Preparing to Implement 5th Carrier

The spectrum preparation (through Step 5) was implemented across all nodes in each market, leaving the space for the 5th Carrier to be inserted as required. As nodes hit the Upstream 70% utilization threshold, they were added on a weekly basis to a list to have the 5th carrier turned up. This response not only saved money but allowed for an agile response to prevent congestion where utilization approached node-action levels during the lockdowns. Such a measured response also kept the number of nodes per week manageable and prevented overwhelming the resources needed for deployment of the solution. For nodes not receiving a 5th Carrier, they will remain at the 5th row of configuration until such time as they are migrated to a new spectrum, such as mid-split or high-split. Current estimates show approximately half the enterprise nodes will eventually receive a 5th carrier.

Nodes continued to be added for the 5th Carrier throughout the year and the effort continues to this current day. At any point in time, the total number of 5th Carrier nodes is less than the total number of deployments as previously planned node actions have resulted in many 5th Carrier nodes being addressed with Mid-split.

The 5th Carrier solution is expected to run for the next several years but will gradually taper off as we execute on our plans to implement mid-split or higher capacity upstream solutions. Thus the 5th Carrier may have a fairly short duration within Cox network planning but is notable for the much-needed relief it provided at a critical time due to extraordinary and unforeseen circumstances.

5. Deploy

Typical pre-COVID deployments unsurprisingly involve high-level controls through program management and Engineering governance to standardize customer and employee impacting projects. The phase gates ensure cross-functional collaboration and awareness but also establishes metric baselines, success criteria, trial performance reviews, and early-life metrics to gauge deployment success. Our deployment process framework is internally referred to as Solution Certification. This framework integrates these cross-functional workstreams, fulfills governance requirements, and readies a service for production availability.

The Solution Certification process is used as an integration management tool to organize the Technology Readiness and Operational Readiness efforts through trials. The process' output is a scalable and sustainable deployment model to minimize operational exposures to critical activities. It combines the strategy, modeling, and spectrum planning outputs with the technology & operational readiness workstreams to culminate in production readiness of the deployment fulfillment and service guides. This process is used for large or small spectrum initiatives such as the plant and service enablement of Mid-Split with four ATDMA carriers or expanding the downstream OFDM from 96-192MHz. Overall, this process may take anywhere from 6 to 18 months depending on complexity and impact to field operations.

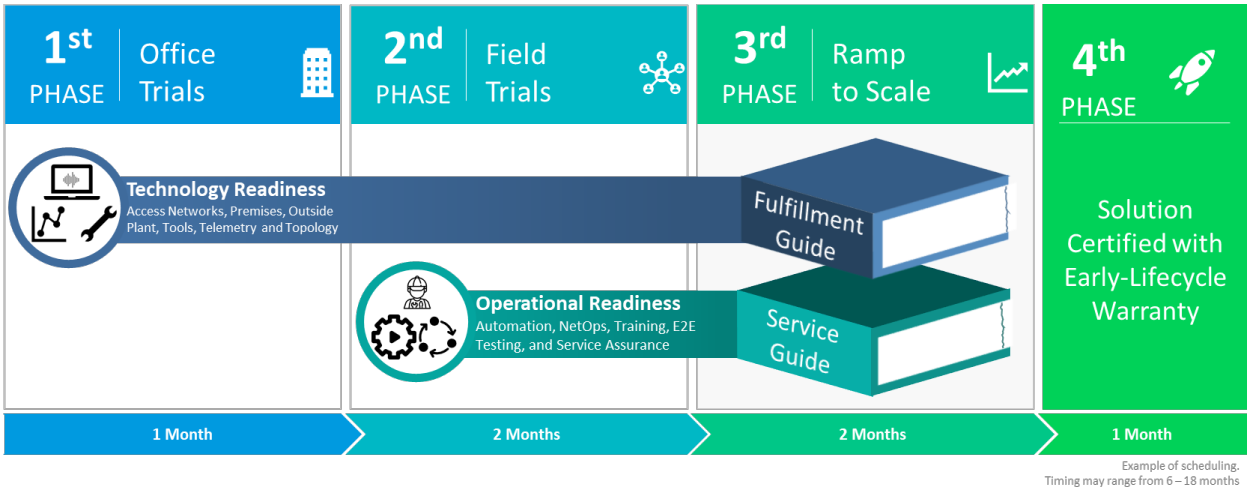


Figure 8 – Solution Certification Process

Office Trials: Technology Readiness' first major milestone that takes the lab-tested deliveries and converges multiple technical workstreams into an office test bed. Trial activities include end-to-end testing, speed tests, node validations, utilization rate checks, and validations for video, voice, and data devices. Pre-COVID, the office trial proved to be an ideal initial gate as employees quickly identify defects that could not be lab tested.

Field Trials: Ops Readiness converges on the certification path to include direct communication with local field teams to understand common issues, validate training, and identify defects that otherwise cannot be captured in analytics. Major activities include the second round of end-to-end testing, on-site node tests, residential & business test account validations, and automation initial testing iterations. Field trials may extend to soaks across multiple markets to widen the trial sampling as a final verification measure before scale.

Ramp to Scale: Handoff to Field Engineering & Operations begins when technology and ops readiness meet trial success criteria. This phase is focused to validate fulfillment and service guides operability in scaled deployment through iterative learning to refine operational processes and automation solutions. Additionally, it is used to test and improve reporting, management, and monitoring tool functionality before the solution is certified.

Solution Certified: Transition to the field is complete and the launch is functioning as designed while operating at full scale. The solution is warrantied for a minimum of 30 days by the program development team to monitor performance and address defects, however, a continuous feedback loop in place to ensure performance and enhance as the solution matures in lifecycle.

Dedicated analytics support is critical to the Solution Certification process for both technology and operational efforts. These reports are essential to trials in not only tracking utilization rates and customer transaction rates but also MER values, D3.1 penetration rate, ticketing, and even track devices in partial service. These trial-focused dashboards are continually modified to evolve it into an operational dashboard as the certification progresses in scalability. For demand-based programs, these dashboards are used as the input to identify node candidates, validate service changes, and also monitor performance to identify anomalies that require escalation to DOCSIS operations and Access engineers. This enables rapid response for our teams to review, troubleshoot, and action on nodes which may necessitate reverting nodes back to original configuration to ensure no harm is done. Ultimately, this dashboard is the key management and reporting tool through a project's lifecycle.

Automation is also a key operational readiness workstream in the effort to streamline changes and processes to meet the high operational tempo that are demanded by field engineering teams. Even small automation updates better enable enterprises to reach scale, improve efficiency, and stay ahead of the curve meet future bandwidth demand and ensure the quality of service expected by our customers. Early collaboration to standardize processes during Solution Certification with operations and engineering teams better enables automation effectiveness and increases agility of future deployments and automation updates.

5.1. Certifying 5th Upstream

COVID's demand for immediate bandwidth increase necessitated the prioritization of a technological solution. The executive leaders' *all hands on deck* call delivered the strategy, modeling, and spectrum plan but now required the development and deployment of more bandwidth. This top-down approach supported the teams by prioritizing this workstream over all others but also limited formalized project and program management as daily direct communication kept leadership abreast of 5th Upstream Carrier status.

Access engineers immediately configured 5th US Carrier in the lab, drafted a MOP in a half-day and worked alongside CPE engineers to form a test plan while outside-plant engineers tested thresholds to determine how to adjust for the low-end 3.2MHz ATDMA. Projects normally estimate 60 days for a single engineer dedicating 25% of their time for lab work, the COVID-team deployed 5th US at Cox Communications Atlanta office in three days. The greatest contributing factors for 5th US Carrier's development expediency were: 1) Leader-driven project crashing with multiple dedicated engineers, 2) standardized CMTS configurations enterprise-wide greatly limiting the volume of testing permutations, 3) and unoccupied offices permitting more aggressive entry to office trials. Labs and 5US carrier was ready for field trial in less than 10 days.

Field trial planning calls involved over a hundred engineers across Advanced Access, Outside Plant, and regional field DOCSIS teams. Each region volunteered to trial 5US, however, several markets required final spectrum alignment moves before it was 5US ready. Five markets across three regions were selected for field trials with Southern California heavily targeted due to exceptionally high utilization rates and heavy concentration of high-profile customers which would be a true test to 5US effectiveness. Overall, 5US was trialed on 15 nodes for a week before ramping to scale.

Concurrent with 5US Carrier's technology readiness workstream our analytics team, the ACoE, developed an MVP dashboard to assist teams to identify and prioritize 5US node candidates based off utilization rates and total customers in the early phases of COVID utilization. Progressive updates were added to develop a Node Prioritization dashboard to support the operational readiness to decisively target the most troublesome nodes with metrics that include high-speed subscription counts, average utilization rates ranging between 1-week to 3-months, and scheduled node actions.

By April 6th, 5US carrier graduated to the Ramp-phase and was soaking on 59 nodes. Engineering teams continued to improve 5US performance by adjusting thresholds but also addressed other issues like CSR VOIP performance, however, 5US full deployment management transitioned to Sustaining Engineering and the field's Bandwidth Management team.

By mid-April, 5US was at full-launch in spectrum-ready markets while the Bandwidth Management team coordinated with local teams to move HEMs, consolidate SCTE 55-1, and move the OOB/Upstream. They established a weekly cadence to review the Node Prioritization reports and queue nodes for the regional DOCSIS managers to review, schedule, and service enable with 5US. On April 20th, 5US had reached scale enterprise-wide with 2,291 nodes enabled by the end of the month.

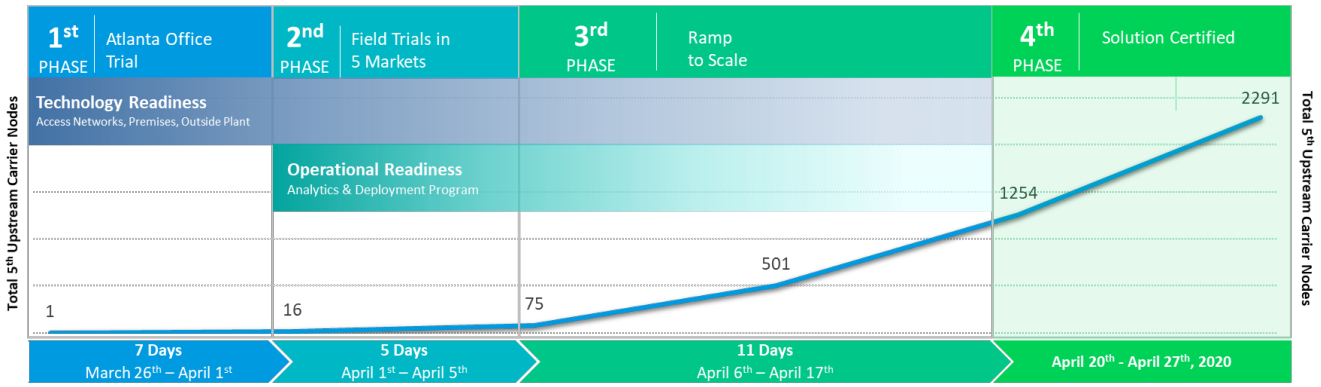


Figure 9 - 5th Upstream Carrier Solution Certification

5.2. 5th Upstream Carrier Deployed

By mid-May, the node volume utilization rates decreased to more manageable levels with teams gradually shifting focus back to the pre-COVID strategic roadmap with small contingents still focused on pandemic related demand. A total of 4,539 nodes were 5US enabled by end of May reducing the total volume of congested nodes (P_{95} ≥80%) from the 7.1% peak down to 2.6%.

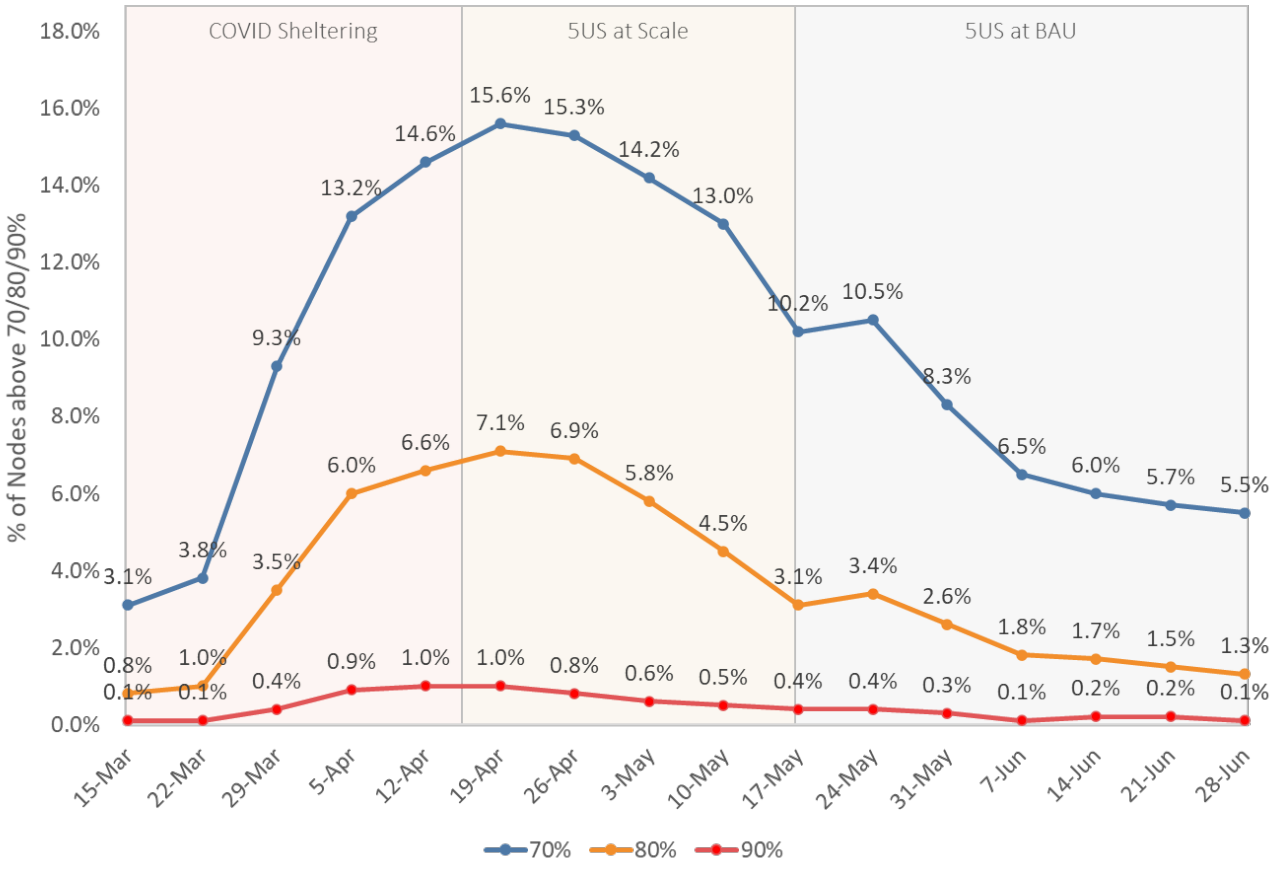


Figure 10- COVID-Driven Utilization and 5US Carrier

By Summer 2020, most nodes that could be addressed by 5US were already enabled and the team standardized the 5US enablement threshold to target nodes with a 4-week ≥70% P_{95}. Deployment rates sharply decreased to approximately 300 enablements per month but spiked to 645 once the remote-learning school year commenced. At the one-year Work from Home anniversary, 5th Upstream Carrier had been deployed to 9,958 nodes with over 42% of deployments occurring within the first 60 days of COVID sheltering.

As of Summer of 2021, the rate of 5th Upstream actions tapered off with some regions going weeks without enabling a fifth ATDMA channel. The total number of active 5US carrier nodes have progressively decreased over recent months as they continue to be consumed by mid-split migrations.

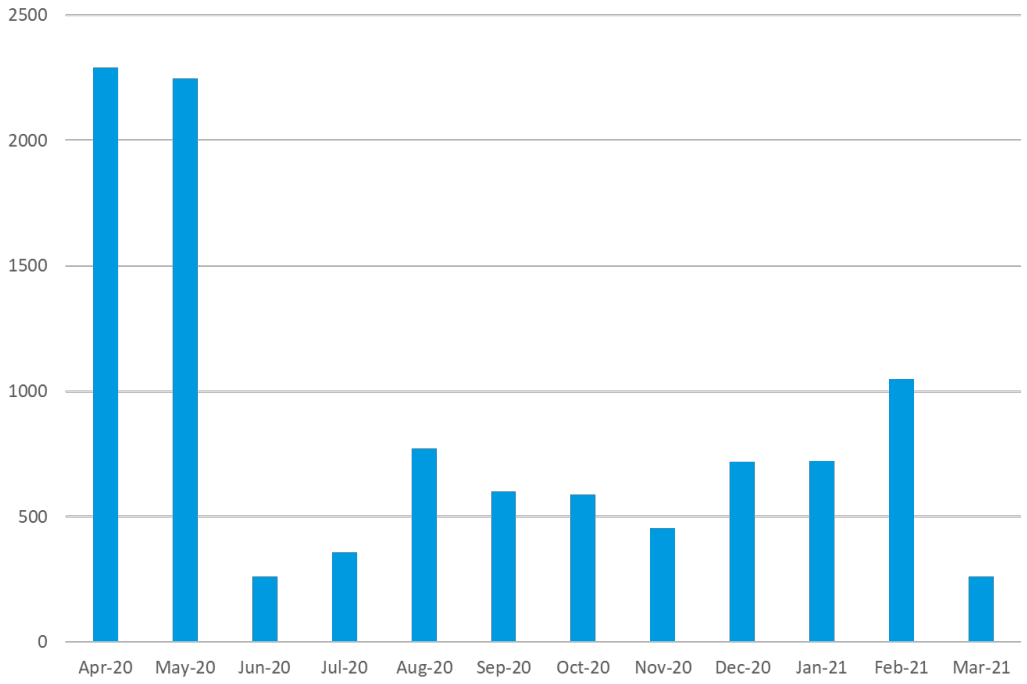
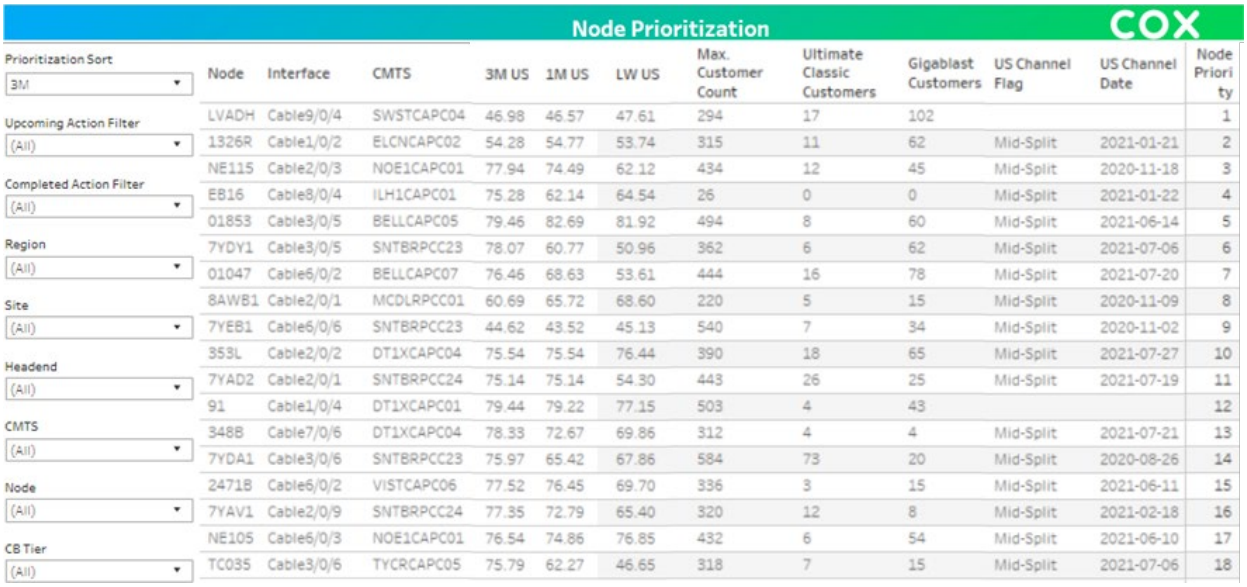


Figure 11- 5th Upstream Carrier Enablements

Operational readiness updates progressively rolled out to support the enablement process like Network Automation's service change update for 5US on RPD to enable scheduling and automated MOP execution. Additionally, the ACoE updated the Node Prioritization report to validate 5US enablements via utilization monitoring of the extra ATDMA channel. This removed the manual spreadsheet tracking of 5US nodes by both the deployment PM and the DOCSIS Operations managers. The Node Prioritization dashboard has since been expanded to also monitor downstream utilization rates and identify OFDM carrier widths to prioritize nodes that require expanded D3.1 capacity.



Node	Interface	CMTS	3M US	1M US	LW US	Max. Customer Count	Ultimate Classic Customers	Gigablast Customers	US Channel Flag	US Channel Date	Node Priority
LVADH	Cable9/0/4	SWSTCAPC04	46.98	46.57	47.61	294	17	102			1
1326R	Cable1/0/2	ELCNCAPC02	54.28	54.77	53.74	315	11	62	Mid-Split	2021-01-21	2
NE115	Cable2/0/3	NOE1CAPC01	77.94	74.49	62.12	434	12	45	Mid-Split	2020-11-18	3
EB16	Cable8/0/4	ILH1CAPC01	75.28	62.14	64.54	26	0	0	Mid-Split	2021-01-22	4
01853	Cable3/0/5	BELLCAPC05	79.46	82.69	81.92	494	8	60	Mid-Split	2021-06-14	5
7YDY1	Cable3/0/5	SNTBRPCC23	78.07	60.77	50.96	362	6	62	Mid-Split	2021-07-06	6
01047	Cable6/0/2	BELLCAPC07	76.46	68.63	53.61	444	16	78	Mid-Split	2021-07-20	7
8AWB1	Cable2/0/1	MCDLRPCC01	60.69	65.72	68.60	220	5	15	Mid-Split	2020-11-09	8
7YEB1	Cable6/0/6	SNTBRPCC23	44.62	43.52	45.13	540	7	34	Mid-Split	2020-11-02	9
353L	Cable2/0/2	DT1XCAPC04	75.54	75.54	76.44	390	18	65	Mid-Split	2021-07-27	10
7YAD2	Cable2/0/1	SNTBRPCC24	75.14	75.14	54.30	443	26	25	Mid-Split	2021-07-19	11
91	Cable1/0/4	DT1XCAPC01	79.44	79.22	77.15	503	4	43			12
348B	Cable7/0/6	DT1XCAPC04	78.33	72.67	69.86	312	4	4	Mid-Split	2021-07-21	13
7YDA1	Cable3/0/6	SNTBRPCC23	75.97	65.42	67.86	584	73	20	Mid-Split	2020-08-26	14
2471B	Cable6/0/2	VISTCAPC06	77.52	76.45	69.70	336	3	15	Mid-Split	2021-06-11	15
7YAV1	Cable2/0/9	SNTBRPCC24	77.35	72.79	65.40	320	12	8	Mid-Split	2021-02-18	16
NE105	Cable6/0/3	NOE1CAPC01	76.54	74.86	76.85	432	6	54	Mid-Split	2021-06-10	17
TC035	Cable3/0/6	TYCRCAPC05	75.79	62.27	46.65	318	7	15	Mid-Split	2021-07-06	18

Figure 12– ACoE Node Prioritization Dashboard Example

In contrast to our typical Solution Certification which normally takes 6-18 months, 5US for the integrated chassis fast tracked the process in 23 days after entering the lab with 5US for RPD following three weeks later. This acceleration was due to a multitude of factors highlighted above, however, the greatest factor to the lightning-fast deployment was due to the strategic decision to use ATDMA. This proven and mature technology was developmentally and operationally less burdensome than OFDMA and could be deployed with minimal risk. Of particular note is that the 5th Carrier was implemented without any measurable impact to call volumes or truck rolls. To implement such a project without the customers noticing was an achievement in itself, and a testament to how the compressed schedule did not sacrifice attention to customer experience.

6. Conclusion

For the near future, Cox network planning will continue per capacity calculations as described in the early sections of this paper. As part of our efforts to accommodate constantly rising demands on utilization, we will be applying capacity levers such as 5th Upstream, SOFA, and Mid-split. In parallel will be the continued march toward the phase-out of legacy video to enable downstream OFDM expansions and transition to all-IP. The pandemic created an unforeseen spike in demand that likely will never be repeated, if only because post-pandemic trends indicate that many people will continue working from home, which will be factored into future growth models. However, for that once in a lifetime event, Cox had the resources, leadership, and corporate agility to able to react to the unprecedented growth and increase our network's resiliency to take care of our customers' needs.

Abbreviations

5US	Fifth Upstream ATDMA Carrier
ACoE	Analytics Center of Excellence (Cox Specific)
ATDMA	Asynchronous Time-Division with Multiple-Access
BAU	Business as Usual
DOCSIS	Data Over Cable Service Interface Specification
HEM	Head-End Modem (Circuit-Switched Telephony)
MER	Modulation Error Ratio
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division with Multiple-Access
OOB	Out of Band
SCTE	Society of Cable Telecommunications Engineers
SOFA	Sub-Split OFDMA
TNPM	(IBM) Tivoli Netcool Performance Manager

Bibliography & References

Ulm, John, and Cloonan, Tom. (2017) *Traffic Engineering in a Fiber Deep Gigabit World.* A Technical Paper prepared for the 2017 Fall Technical Forum of SCTE-ISBE/NCTA/CableLABS.

Cloonan, T., Emmendorfer, M., Ulm, J., Al-Banna, A., & Chari, S. (2014). Predictions on the evolution of access networks to the year 2030 and beyond. *The Cable Show NCTA/SCTE Technical Sessions, Spring*, *38*.

Cable And Rural Broadband

How Cable Plays a Critical Role in Closing the Digital Divide

A Technical Paper prepared for SCTE by

Kevin A. Noll
Principal Access Architect
Vecima

Jay Rolls, Pacband

Patrick Ryan, Esri

1. Abstract

Cable was birthed in rural America. Cable grew up in rural America. Cable is still rooted in rural America with many rural American households serviced by a cable operator. The US is facing a broadband equity crisis in which nearly 20% of Americans, primarily in rural communities, do not have access to reliable broadband Internet. Cable operators are uniquely positioned to solve this problem once and for all.

This report will explain why rural broadband, like rural electrification in the 1930's and universal telephone service in the 1950's, is the most important societal technology issue of our time.

The report will further explore why delivering broadband to rural Americans is perceived as a challenging engineering and fiscal problem and will explain how innovative approaches to engineering, materials, construction practices, and business modeling can overcome those challenges. The advantages and disadvantages of these innovative approaches will be discussed.

Today, funding sources have expanded to include federal and state grants, broadband funds, and county and municipality initiatives which, when coupled with public / private partnerships, can easily tip the balance to a faster return on investment. Finally, the paper will discuss the financial opportunities and incentives that enable universally accessible broadband Internet to become a reality.

2. Introduction

What began as a pragmatic tool to enable researchers to communicate and collaborate more effectively has, today, become an essential tool for everyone's everyday life. Access to the Internet is as essential as electricity or the telephone and is more widely used for news and information than television and radio. With 90% of employment applications being submitted online, the Internet is an essential tool for seeking employment. For those currently employed, the number of people using the Internet to work from home has increased dramatically during the pandemic. While many private education programs were already exclusively using the Internet to deliver their curriculum, during the COVID-19 pandemic, the Internet became the essential tool for public education institutions. COVID-19 also brought about large increases in the use of telemedicine

However, an estimated 14 million to 160 million homes and businesses in the US do not have adequate access to the Internet. With an echo from the 1940's and rural access to electricity, 90% of those without access to the Internet live in rural areas.

In the 1930s electricity providers refused to deliver electricity to rural Americans. In the 1940s, telephone companies, and television broadcasters also refused. The cost was perceived to be too high to build each of these essential utilities and communication tools to rural communities. Cable, though, is rooted in bringing service to those that would not otherwise have it. Beginning in 1949, the pioneers of cable created cable to deliver television programming to rural America when no one else would.

Today, Cable faces another opportunity to serve its communities and to expand its networks. Many cable operators have been reluctant to build broadband services into rural communities due to poor internal rates of return (IRR). Conditions have changed, though, and Cable is well positioned to help eliminate this gap.

3. The Broadband Gap

The literature and media make frequent references to the “Digital Divide”, but what is this Digital Divide?

Generally, the Digital Divide refers to the gap between those that have access to information and communications technologies (ICT) and those that do not. In the past, the discussion about the Digital Divide was focused on access to computers and technology education. Recently, this is most often applied to describe lack of access to the Internet at broadband speeds, which is also referred to as the Broadband Gap.

The Broadband Gap, though, is not simply a technical or technological one. Those that lack access to broadband Internet are directly and negatively impacted because they are unable to fully benefit from educational, economic, and healthcare resources available on the Internet and they are unable to fully participate in the political and social aspects of their community, nation, and the world.

These are compounding disadvantages. Those that do not have access to the Internet are most likely to already be economically and educationally disadvantaged and their inability to reap the benefits of unencumbered access to the Internet creates a situation in which they are unable to pull themselves up and, instead, are simply drawn deeper into the chasm.

Some might be skeptical about the impacts of this Digital Divide. Consider that in 2015, 80% of Americans used the Internet to search for and apply for employment [1], but in 2020 reliable broadband Internet access was not available to tens of millions of Americans. Further, rural populations, less educated populations, low-income populations, and minority populations are less likely to have broadband access [2] and, therefore, do not have equal access to employment resources.

Education was in the spotlight during the COVID-19 pandemic. Prior to the COVID-19 pandemic, it was already well known that academic performance was better among those that have ready and reliable access to the Internet. According to a study by the Quello Center [3], students with no access to the Internet at home or who rely on a mobile phone for home Internet access will typically be ½-letter grade behind those students that have reliable Internet access at home. According to the US Census Bureau’s Household Pulse Survey [4], an estimated 2.5 million households with school-age children reported their Internet access was not reliably available for education purposes. With a virtually nationwide switch to online education in the United States, lack of reliable broadband Internet access caused significant impacts on academic achievement and disproportionately impacted rural populations and low-income and minority families.

Estimates range from 14 million to 160 million [5] Americans that lack Internet access at broadband speeds. This is a wide range that reflects the lack of accurate and standardized methods to identify the unserved.

Download and upload speeds are the most used standards for fixed broadband service. In the United States, the Federal Communications Commission (FCC) definition is the most referenced standard for broadband. The FCC’s definition has evolved over time, as demonstrated in Table 1, and as of July 2021 the FCC defined broadband Internet access as a service that delivers 25Mbps downloads and 3Mbps uploads.

Table 1 - History of Broadband Definitions in the US

Year Published	Source	Download Speed	Upload Speed
2021	US Treasury Department (minimum build-to, proposed)	100 Mbps	20 Mbps
2021	US Treasury Department (Eligibility)	Less than 25 Mbps[1]	Less than 3 Mbps
2018	USDA ReConnect (Build-To)	25 Mbps	3 Mbps
2018	USDA ReConnect (Eligibility)	Less than 10 Mbps	Less than 1 Mbps
2015	FCC	25 Mbps	3 Mbps
2010	FCC	4 Mbps	1 Mbps
1996	US Telecommunications Act	200 Kbps	200 Kbps

The FCC launched the Rural Digital Opportunity Fund (RDOF) program in 2019 and used Form 477 data, self-reported by Internet providers, to identify census blocks eligible for funding. This analysis resulted in the maps shown in Figure 1, and an estimate of 2,552,251 households without 25/3Mbps Internet access using a terrestrial network.

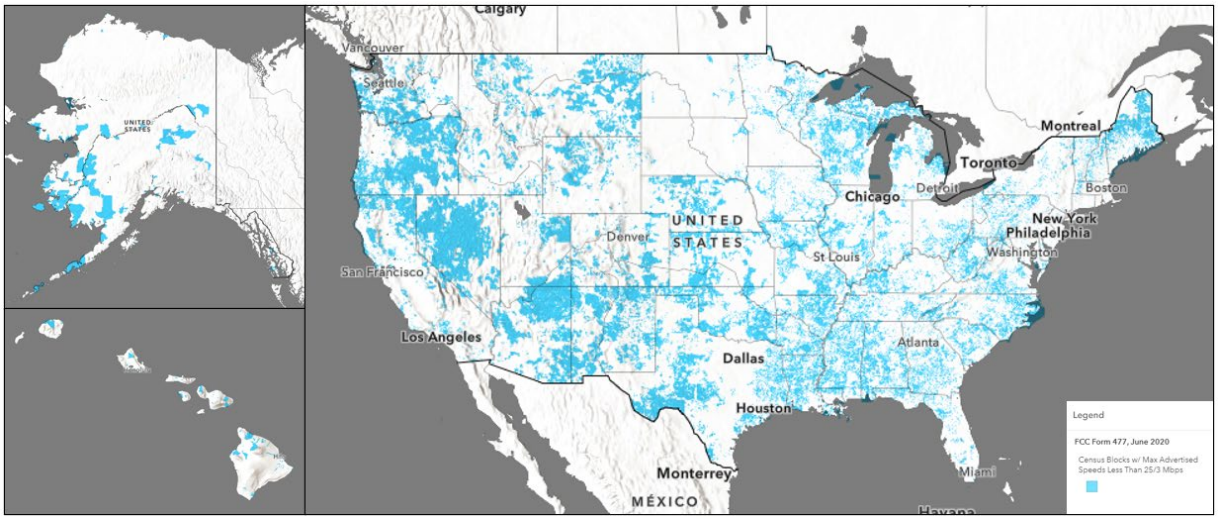


Figure 1- Census Blocks with maximum advertised terrestrial speeds less than 25/3 Mbps(FCC Form 477, June 2020)

With the introduction of the broadband requirements, including allowed overbuild of DSL networks[1], from the US Treasury and proposed legislation, the US has implicitly set the standard of broadband to be fiber-based or cable-based. This significantly changes the picture to show an estimated 10.6 million households without access to broadband.

[1] US Treasury has authorized the use of ARPA funds to overbuild DSL and older DOCSIS networks even though they might deliver 25Mbps/3Mbps.

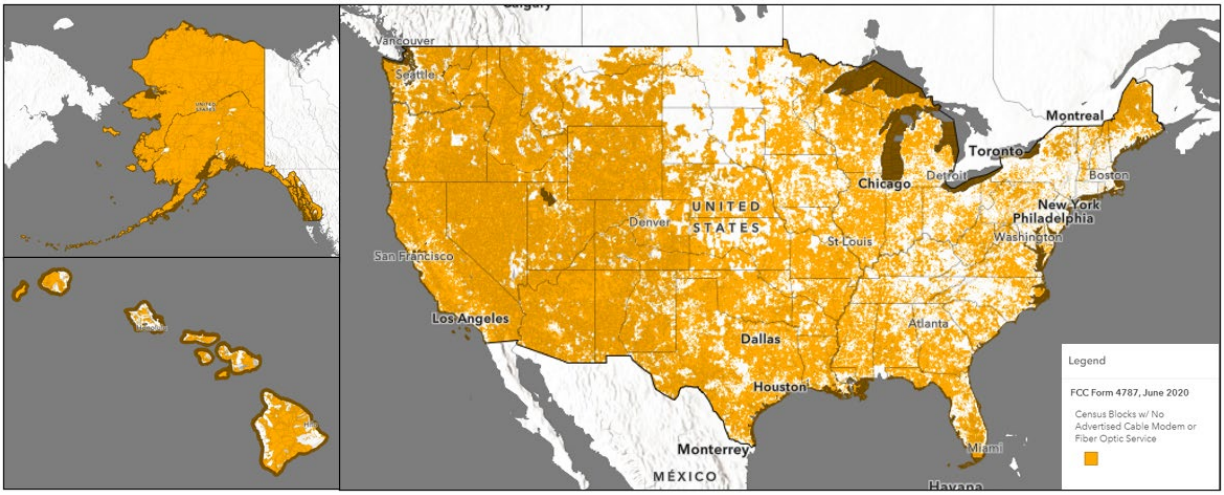


Figure 2 – Census Blocks with no advertised Cable Modem or Fiber Optic Internet Service (FCC Form 477, June 2020)

The US Treasury, in its guidance on use of American Rescue Plan Act (ARPA) funds, also empowered communities and governments to consider a wide range of information other than the FCC Form 477 as evidence that broadband networks were inadequate. These include data from other federal agencies like the National Telecommunications and Information Administration (NTIA), which recently released the Indicators of Broadband Need mapping application [6]. The NTIA map brings datasets together from the US Census (the American Community Survey), Department of Education, NTIA, Ookla Speed test results, Measurement Lab (M-Lab) speed test results, and Microsoft Broadband Usage Statistics.

A recent non-profit study [7] of Loudoun County, Virginia is an example of the significant difference these new data sources can make. Using Form 477 data, the FCC disqualified the entire county from the Auction 904: Rural Digital Opportunity Fund, presuming that all locations in the county have access to the Internet at 25/3Mbps (see Figure 3).

However, the local study shows that nearly 9000 households and businesses do not have access to the Internet at 25Mbps or higher (see Figure 5). Their study further shows that, on average, those households receive download speeds at 15Mbps or lower while paying almost 3 times more per month for Internet service than those that have access to cable-modem or fiber-based Internet access. This disparity between published coverage data and actual broadband accessibility is just one example of many that indicates that broadband coverage is significantly overestimated in the US.

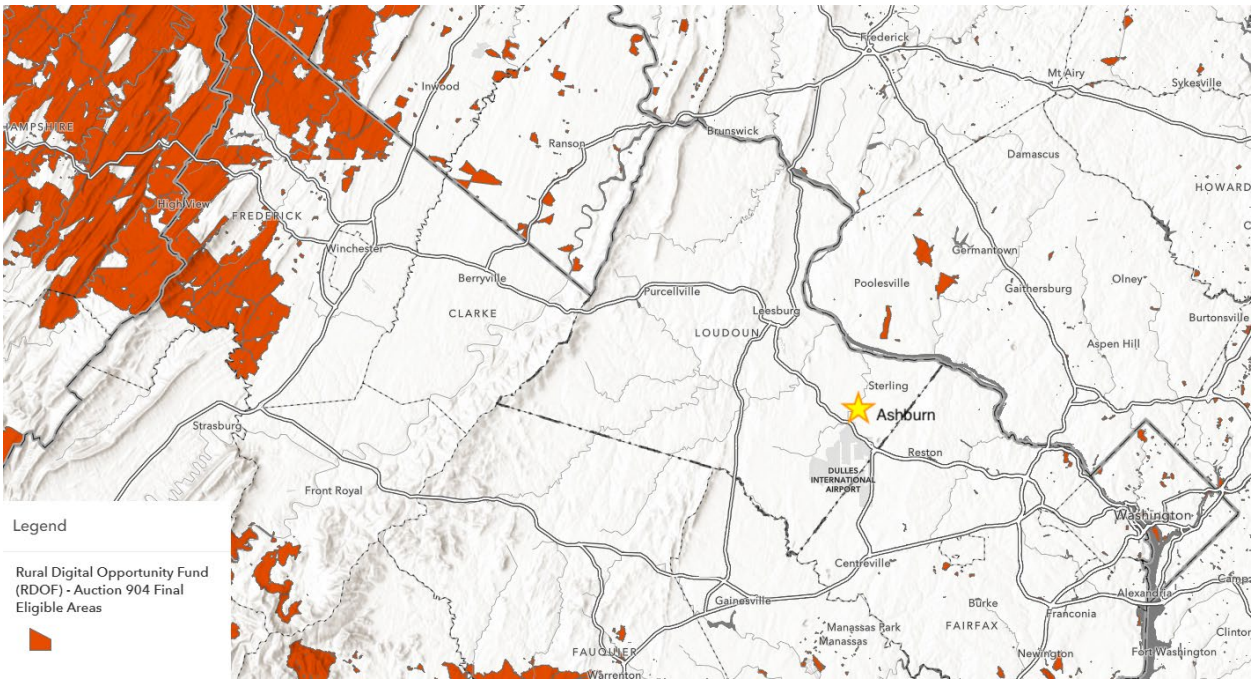


Figure 3 - RDOF Eligibility for Loudoun County, VA and Surrounding Areas

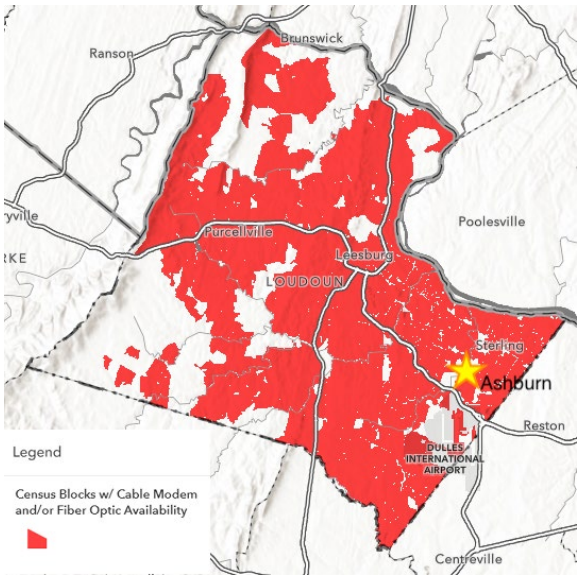


Figure 4 - Form 477 Data indicating accessibility to fiber or cable-based broadband in Loudoun County, Va.

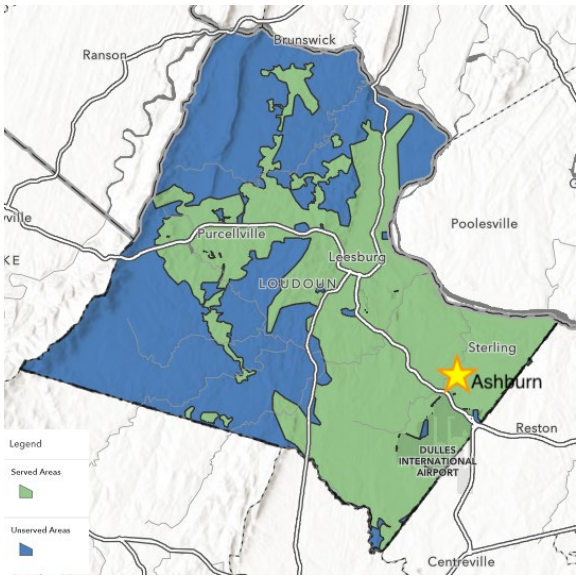


Figure 5 - Locally curated research indicating no broadband coverage (blue shaded areas) in Loudoun County, Va.

With estimates ranging from 14 million to 160 million homes and businesses that lack adequate access to broadband Internet, the available data clearly supports the claim that broadband access is severely lacking in rural areas.

4. Enabling Cable to Expand its Rural Footprint

Cable and other Internet providers have traditionally avoided serving rural areas due to technical or financial criteria. Based on business requirements and goals at the time, those criteria might have been perfectly valid. Conditions have changed and qualification criteria need to be re-evaluated.

Existing US cable franchise areas have very high penetration rates. Subscriber growth is slowing and is often driven only by churn between providers.

For US operators, there are four primary areas of company growth:

1. Growing penetration/stealing share for existing products
2. Introducing new products (e.g., mobile)
3. New home construction in existing footprint
4. Expanding footprint (mostly into underserved markets)

The most often cited reason for not building rural Internet access is that it is too expensive, but what are the factors that drive cost, and what are the characteristics of "rural" networks?

4.1. Construction dominates cost

It's important to understand when constructing rural broadband networks that build costs are dominated by construction labor. Materials and electronics, even sophisticated networking equipment, make up a very small portion of the network deployment costs. Construction is commonly outsourced to third-party contractors. Entering into these third-party relationships, it's important to obtain competitive pricing while at the same time not compromising on build quality. Often cable operators will be able to leverage numerous such relationships that they already have in place as part of their ongoing needs.

4.1.1. Logistical Factors

Logistics related to construction contribute significant overhead to the construction process. Before beginning construction, the operator must acquire permits from right-of-way owners such as local and state departments of transportation, easements where no existing access rights exist, and utility permits to connect to utility poles. Construction and safety permits must also be acquired. The application and permitting processes and requirements can be unique for each jurisdiction and can be tedious and require expertise in the specific jurisdiction.

These factors can be mitigated by considering cost sharing opportunities, reducing the number of times to "go to the well", building relationships with the permitting agents and agencies, and by coordinating construction with other utilities.

Previously, construction of new network infrastructure was planned on a 1-year cycle. This is no longer an effective strategy.

A one-year cycle was predicated on a "do-it-yourself" approach in which the operator designs, plans, and constructs the network as an independent and autonomous entity. This do-it-yourself approach is not cost effective for rural networks. Cost effective construction will require the operator to share the cost of trenching and boring. Municipalities are loath to have their roads disrupted and having each utility or communications provider running independent digs means multiple disruptions and more risk to the municipal infrastructures (like water and sewer).

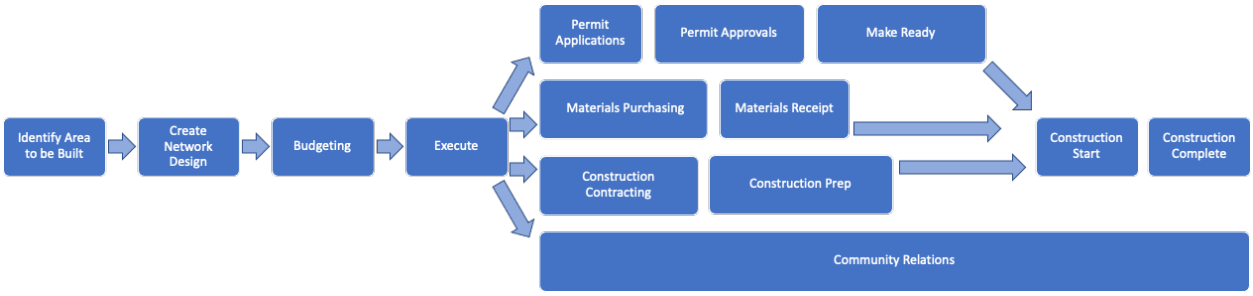


Figure 6 - Typical 1-year planning and construction cycle

Another factor to consider is availability of labor and materials. The industry is already experiencing shortages in skilled labor and materials, especially for fiber builds. As federal and state funding for broadband expansion ramps up in 2021 through 2024, these shortages will only get worse. The cable operator will need to order materials at least one year out, perhaps entering risk-buy situations. Recent experience has shown that even when pre-ordered, materials might not be delivered on time. Most publicly supported broadband expansion programs impose 2-year deadlines, so without flexibility in the cable operator's processes, network designs, materials choices, and construction practices, operators could find themselves facing penalties due to supply-chain issues.

All of this means that a 5-year planning cycle will be a necessary cost reduction strategy. This allows the operator to be opportunistic in construction – keeping an eye open for infrastructure projects where constructions costs can be shared (e.g. laying conduit alongside new water and sewer infrastructure, or coordinating water and road crossings with bridge and overpass replacements). With the federal funding for infrastructure being distributed in 2021 through 2024, these opportunities will be at a peak.

Becoming aware of potential cost-sharing opportunities will require planners to build relationships. Relationships with jurisdictional personnel are already a necessity to streamline permitting processes. Cable operators will need to foster new relationships in local and state planning offices and with planning teams in the various utility companies and possibly competitive broadband providers. These relationships will be critical to discovering future infrastructure actions that can be leveraged for the benefit of the cable operator through cost reduction and building goodwill among the participants. In fact, such relationships and coordination might become the new regulatory norm due to legislation introduced in 2021 [8, 9, 10].

Operators will often have choices to make in selecting which markets they chose to build. When a strong partnership is formed with the localities (both local and state), it can help foster an environment that facilitates and streamlines network construction. In the same manner, uncooperative localities can be a major obstacle to a successful build – and operators use this local cooperation criteria when selecting the areas that they chose to invest in.

Another factor will be the availability and willingness to take advantage of symbiotic relationships. As an example, Central Virginia Electric Coop (CVEC), responding to requests from its members to deliver broadband Internet access, started an initiative to offer zero-fee pole attachments to Internet providers. Through this program, CVEC formed a strong relationship that brought fiber-based broadband to 37,000 co-op members in rural Virginia [8] and also accelerated its smart-grid deployment. While the partner was not a cable operator, opportunities like this will be available to those cable operators that are alert and willing.

4.1.2. Improved Construction Techniques and Tools

Once the planning is complete construction can start, but it's not so simple as hanging or burying some cable. The art of network construction includes a variety of skills and requires an equally varying set of tools.

Advances in fiber cable have made it possible to fit more fiber strands into smaller and smaller cables. Ribbon cable makes it possible to fit, for example, 144 fibers into the same size cable as 72 fibers of loose-tube fiber. Unarmored cables are also available and can be a measured-risk for cable operators. These options can save space and reduce weight and wind loads on poles. Ribbon cable, though, has its own challenges since it typically requires special handling and splicing tools. In the past, splitting a ribbon cable in mid-span has not been practical. New tools, though, change this dynamic. Ribbon separation tools are available from multiple sources and enable crews to easily separate fibers from the ribbon. Ribbonizing tools and adhesives are also available to dress and complete ribbon splices. With new splicing kits, it is also now possible to splice whole ribbons or individual fibers.

Tools to bury cable have also advanced. Availability of tools like the vacuum excavator means that manually digging to avoid existing utility lines is no longer necessary. This is a significant time and cost saving tool.

Figure 7 - Examples of Vacuum Excavators[2]

Directional boring has been in use for many years. In the past, though, safely completing a directional bore was dependent on good documentation of existing utility lines (which is often incomplete or inaccurate) and could be quickly thwarted by unexpected geological features. They were also highly dependent on the skill of the operator to know the location and direction of the bit. Today's boring efforts, though, benefit from advanced tools like electric strike indication systems and proximity detection systems that significantly improve accuracy of the drill, but also reduce risks of injury and unexpected utility damage.

These are just a couple of examples of advancements in tools and techniques in construction.

[2] Photos courtesy of Vermeerand Vac-Con

4.2. Technology Considerations

Unless an area is directly adjacent to current hybrid fiber-coax (HFC) infrastructure, a service area expansion is typically going to best be supported by the deployment of an all-fiber network. Build costs are equivalent to HFC, and the result is a much more extensible, more fault tolerant, and less expensive to operate network.

For cable operators, while making a change to an all-fiber delivery, it will be important to continue to offer video services and leverage the substantial investments that have already been made in the video space. IP Video technology has matured, enabling cable operators to deliver nearly equivalent video services over fiber, albeit with some new challenges in service delivery. Video CPE and service activation, to name two, can look quite different from current practices.

Just as cable operators have started to embrace the push of some electronics deeper in the network (i.e. Distributed Access Architecture), the same movement is occurring within all fiber network architectures.

From a timing perspective (circa 2021), 10Gbps symmetrical PON (XGS-PON, or 10G-EPON) has become the technology of choice. Previously, the Optical Line Terminal (OLT) device was a monolithic modular chassis designed to be deployed in an environmentally stable environment like a data center. Recently, though, the remote OLT has become an enabling technology for lower-cost network builds. The OLT is now available and commonly used in several remote configurations, including node OLTs, outdoor cabinet OLTs, as well as compact OLTs designed for multiple dwelling unit (MDU) and high-rise buildings. Such configurations allow for the elimination of high fiber count trunks that must span the long distances to a hub or headend. These well-placed remote devices can also eliminate reach restrictions, since PON networks have distance limitations, often as low as 20km. Lastly, remote OLTs can eliminate or limit the needs for headend space or avoid the need for small electronics huts. All of these serve to save on construction cost.

Figure 8 - Traditional OLT Deployed Remotely

Figure 9 - Modern Remote OLT

Another enabler for lower-cost network builds is availability of many types of data, electronically, from so many sources. These include geocoding resources, address verification databases, geocoded census and consumer demographics, electronic network inventories (locating cable routes), satellite imagery and detailed topographical and soil-type maps.

In the past, with so much data available, subject matter experts would need to be employed to analyze the data and planners would need to interpret summaries provided from several sources to make build-out decisions. Often in-the-field surveys would need to be performed to verify the findings or to collect data that was otherwise unavailable. These processes often took months to complete. Today's modern planning and design tools, though, allow designers and planners to aggregate massive amounts of data, perform measurements and calculations, perform address density calculations, measure distances, and quickly run cost estimates simply by snapping out polygons on their computer desktop.

4.3. Workforce Challenges

Maybe one of the most significant concerns for expanding broadband service into rural areas is finding skilled labor. Especially with increased funding and mandates from the federal and state governments, demand for a workforce capable of building outside plant will be increasing and the existing workforce is already stretched.

Most operators no longer maintain dedicated construction crews. Instead, they might retain crews responsible for plant maintenance, and often will retain a contractor for repairs and hire other contractors to perform initial builds.

It will not be enough to have access to workers skilled simply in the art. Those workers will need to be trained in sound safety practices. Without this, the operator will be at risk for code violations, on-the-job injuries, or liability risks post-project completion. Skilled labor will be easiest found with a contractor.

With the growing emphasis on building fiber networks to previously unbuilt areas, a dearth of pop-up fiber construction outfits has surfaced. Often these crews are learning on-the-fly and have no first-hand experience nor quality training from the industry or from the manufacturer of the equipment they have chosen to use.

A reputable and quality network construction contractor will maintain continuous training in the art and safety for its employees. The contractor will have experience constructing infrastructure in the regions of interest. They will also be familiar with the safety protocols and permitting protocols of those regions and have the financial backing and stability to absorb changes in project schedules and unforeseen circumstances "on-the-ground".

For these reasons, choosing the least expensive contracting option on a per-job basis is not necessarily the best decision. Since longer-term planning is of growing importance, part of that planning will be to build relationships and retainers that ensure quality contractors are available and that they are able to plan their workforce development to match the operator's projected demand. These relationships are secured by committing to 5-year build plans and by forming long-term contractual agreements with contractors and suppliers.

Construction is not the only area of concern for the workforce, though. There are many connected processes that support construction. For example, prior to beginning construction, locators must be dispatched to locate and mark existing infrastructure (electrical, water, sewer, etc.). City, county, and state regulating offices must keep up with permitting and inspection volumes. These agencies might need to increase staff and provide education related to broadband to support increased volume of builds.

Another issue that operators must be prepared to manage is variations in regulations among jurisdictions and variations in enforcement within the same jurisdiction. The former is much more manageable and can be mitigated during the planning and permitting phases of a project. Variations in enforcement are caused by varying interpretation of regulations among inspectors or region-specific leadership. Such variations, though, often do not appear until after the project has begun and can result in construction activity being halted (while workers continue to be paid) and the project being delayed.

Anecdotes abound about such situations. One story related to the authors described a fiber project that required construction along a state right-of-way. The state regulating authority's (department of transportation, or DOT) safety rules made the use of an attenuator truck (crash truck) optional. During construction in one region, the local DOT personnel allowed the project to proceed without an attenuator truck being present. As the project moved into a neighboring region, the DOT personnel in that region would not allow construction to continue until an attenuator truck was present. The impact was a delayed project and additional cost for the truck and personnel to operate it.

A consistent and accurate inventory of broadband infrastructure is as important to bridging the digital divide as a complete record of serviceable locations (e.g., the locations of households and businesses). Historically, telecommunications companies have used an assortment of information systems – some developed in-house and some purchased – that were never designed to work together. When these systems were implemented, there was no perceived requirement for information sharing. Today telecommunications companies operate networks that have equipment from multiple vendors, lease bandwidth and antenna sites from other companies, and manage federal and state funding requirements. Mergers with, or acquisition of, other companies require the incorporation of different systems.

The provision of broadband connectivity is closely tied to geography. Location intelligence is fundamental to all communication services. A geographic information system (GIS) is a system that creates, manages, analyzes, and maps all types of data. It provides for data interoperability. GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there). This foundation provides for mapping and analysis that helps users understand patterns, relationships, and geographic context. The benefits include improved communication and efficiency as well as better management and decision making. Using the open data standards of a modern GIS promotes transparency, data sharing, and collaboration.

Dirty data and non-existent data that should have been collected (but was not) adds substantial cost to telecommunication projects. Mobile GIS workflows location-enable field activities, modernize data collection, and facilitate near real-time updates between the field and office. Authoritative geospatial data is key to support the work of broadband development and programs. Telecommunication providers and governmental organizations have a bottom-line interest in ensuring it's collection.

4.4. Financing Challenges

For many years, research has shown those that lack access to broadband Internet are directly and negatively impacted because they are unable to fully benefit from educational, economic, and healthcare resources available on the Internet and they are unable to fully participate in the political and social aspects of their community, nation, and the world. During the COVID-19 pandemic the world came to realize that access to the Internet can no longer be a benefit enjoyed only by urban and higher income populations, but it is an essential service that should be available to everyone.

Unfortunately, by their very nature, unserved and under-served areas are often rural and are expensive areas to build a network. A greenfield fiber build will often cost between $900 and $1200 per home passed, and even costlier for low density areas with costs up to $8000 per home passed. These costs have traditionally not passed the internal rate of return (IRR) tests of an operator. When building a greenfield network, fiber will likely be the best choice, particularly if in a low-density area. If cable operators build in an area adjacent to their current HFC network (aka "edge-outs"), it might behoove an operator to extend their existing HFC architecture.

As cable operators look for ways to grow subscribers, investing in under-served areas begins to be more attractive as a way to capture the pent-up demand for quality broadband. An operator will often enjoy significant service uptake in newly built areas because they have been inadequately served for so long. This helps support the business model around investing in the network build. Additionally, an all-fiber network will benefit from better economics from a total cost of ownership (TCO) point of view.

A major change in the environment for building broadband to rural areas is the availability of funding. Private investors have long been interested in building broadband to support their overall mission, and the changing environment over the past couple of years has made investment in rural broadband much more likely to pay off. Firms like Searchlight Capital [9] [10] [11] and GTCR [12] have been prominently investing in rural broadband because of this potential.

Recently available public funding has played a significant role in the decisions from private investors. It has also played a big role in decisions by some cable operators. For example, Charter Communications won $1.22B from the FCC's RDOF Auction 904 and plans to build broadband to up to one million currently unserved homes and businesses[3].

At the federal level, in the year 2020, the Coronavirus Aid, Relief, and Economic Security Act (CARES) act set aside $100M to be spent on broadband expansion and RDOF allocated up to $20.4B (to be distributed over 10 years) toward broadband expansion. In 2021, the US Congress passed the American Rescue Plan Act (ARPA), a $1.9 trillion economic stimulus program which allocates, through seven different programs, at least $20.3 billion and up to $265 billion that can be used for digital equity programs including broadband physical network buildouts. In addition to ARPA funding, it appears that another $65 billion will be made available in the infrastructure bill that is making its way through the US Congress in August of 2021 [13].

At the state level, 38 out of 50 states in the US have created funding programs specifically to support expansion of broadband service.

[3] Charter, and other Auction 904 winners, are revamping these estimates because the RDOF eligibility maps were based on incomplete data.

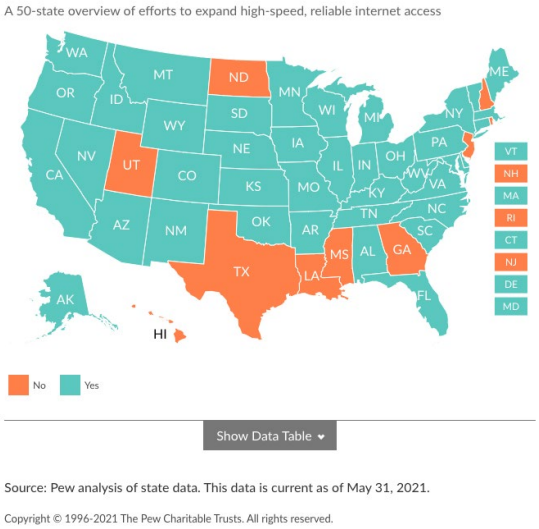


Figure 10 - US States with Broadband Funding Programs [14]

The structure and scope of these funding programs varies from state to state. The ARPA distributes much of its broadband funding directly to states and localities, so states' programs are currently undergoing changes to adapt to this influx of money. For example, the Commonwealth of Virginia's FY2022 budget initially allocated almost $60 million, but with the authorization of broadband funding from ARPA, Virginia will expand this to allocate $700 million over the next 3 years [15]. Further, many counties, towns and cities will receive ARPA funding and have the option to use that money to expand broadband infrastructure.

This influx of funding will change the way cable operators calculate the financial models for building to unserved and rural areas. In fact, these new funding options will directly impact the construction planning cycle because these funding sources impose delivery deadlines with penalties. Cable operators that develop, hire or contract expertise in broadband grant writing will have a distinct advantage over competitive providers that do not.

5. Conclusion

The utility of broadband Internet access has been studied at length and, by all measures, those that have access to broadband Internet are better off economically, educationally, socially, and health-wise. Those that do not have access to broadband Internet are disproportionately affected because they are already more likely to have lower incomes, more likely to live in rural areas that have fewer employment options, less access to basic healthcare, and access to fewer educational support resources.

Lack of broadband Internet access affects tens of millions of people in the United States, and approximately 90% of those affected live in rural areas. This is not a unique story. Rural broadband access has followed a similar storyline as rural electrification, rural telephone, and rural television. Internet access has simply failed to keep up with urban areas because Internet providers have failed to deploy newer technologies into rural areas.

Birthed in the mountains of Pennsylvania, the river towns of Oregon, and the plains of Wyoming, cable's history is rooted in rural communities. Over the last 30 years, cable has stepped up to the plate and delivered advanced services to the communities they serve. Modern-day economics has held back many cable operators from deploying their networks into rural communities.

Any incumbent network operator will be able to leverage advantages in building out rural broadband networks. They bring resources and knowledge to the challenge. Cable operators can bring this and more to the table. This is most profound when a new build area is adjacent to their current networks/footprint. Access networks (or "backhaul") can be a significant portion of both a network build and operation (if using leased fiber connections). As well, simply using existing people, processes, and resources will assist in building and running these networks.

Using a variety of new and even unconventional financial vehicles, these networks can easily be built and operated profitably. As discussed, there are a large variety of options available to enter into public-private partnerships, that not only make for good business for a cable operator, but also help to enable an essential service for these under-served communities.

Bibliography & References

[1] Pew Research Center, "Searching for Work in the Digital Era," 19 November 2015. [Online]. Available: https://www.pewresearch.org/internet/2015/11/19/searching-for-work-in-the-digital-era/. [Accessed 12 July 2021].

[2] Pew Research Center, "7% of Americans don't use the internet. Who are they?," 2 April 2021. [Online]. Available: https://www.pewresearch.org/fact-tank/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/. [Accessed 12 July 2021].

[3] Michigan State University Quello Center, "BROADBAND AND STUDENT PERFORMANCE GAPS," March 2020. [Online]. Available: https://quello.msu.edu/wp-content/uploads/2020/03/Broadband_Gap_Quello_Report_MSU.pdf. [Accessed 12 July 2021].

[4] US Census Bureau, "Week 32 Household Pulse Survey: June 9 – June 21," June 2021. [Online]. Available: https://www.census.gov/data/tables/2021/demo/hhp/hhp32.html. [Accessed 12 July 2021].

[5] Microsoft, "United States Broadband Usage Percentages Dataset," October 2020. [Online]. Available: https://github.com/microsoft/USBroadbandUsagePercentages. [Accessed 15 July 2021].

[6] NTIA, Office of Public Affairs, "NTIA Creates First Interactive Map to Help Public See the Digital Divide Across the Country," 17 Jun 2021. [Online]. Available: https://www.ntia.doc.gov/press-release/2021/ntia-creates-first-interactive-map-help-public-see-digital-divide-across-country. [Accessed 12 Aug 2021].

[7] Loudoun Broadband Alliance, "LBA Maps the Broadband Unserved In Loudoun County," 2 May 2021. [Online]. Available: https://loudounbroadbandalliance.org/education/lba-maps-the-broadband-unserved-in-loudoun-county/. [Accessed 12 Aug 2021].

[8] Conexon, "CVEC Listens to the Echoing Sentiment From Its Members and Moves Forward to Build a FTTH Network," [Online]. Available: https://conexon.us/case-studies/central-virginia-electric-cooperative/. [Accessed 12 Aug 2021].

[9] Searchlight Capital, "SEARCHLIGHT CAPITAL PARTNERS ANNOUNCES APPOINTMENT OF AJIT PAI AS PARTNER," 26 April 2021. [Online]. Available: https://www.searchlightcap.com/news/searchlight-capital-partners-announces-appointment-of-ajit-pai-as-partner/. [Accessed 10 Aug 2021].

[10] Searchlight Capital, "SEARCHLIGHT CAPITAL PARTNERS MAKES STRATEGIC INVESTMENT IN ALL POINTS BROADBAND," 6 July 2021. [Online]. Available: https://www.searchlightcap.com/news/searchlight-capital-partners-makes-strategic-investment-in-all-points-broadband/. [Accessed 10 Aug 2021].

[11] Searchlight Capital, "CONSOLIDATED COMMUNICATIONS ANNOUNCES STRATEGIC INVESTMENT FROM SEARCHLIGHT CAPITAL PARTNERS; INITIATES REFINANCING," 14 Sep 2020. [Online]. Available: https://www.searchlightcap.com/news/consolidated-communications-announces-strategic-investment-from-searchlight-cap/. [Accessed 10 Aug 2021].

[12] GTCR, "Point Broadband Announces Strategic Investment from GTCR," 16 Apr 2021. [Online]. Available: https://www.gtcr.com/point-broadband-announces-strategic-investment-from-gtcr/. [Accessed 11 Aug 2021].

[13] K. Snell, "The Senate Approves The $1 Trillion Bipartisan Infrastructure Bill In A Historic Vote," 10 Aug 2021. [Online]. Available: https://www.npr.org/2021/08/10/1026081880/senate-passes-bipartisan-infrastructure-bill. [Accessed 12 Aug 2021].

[14] Pew Research, "How Has Your State Designed Its Broadband Program?," 28 Jun 2021. [Online]. Available: https://www.pewtrusts.org/en/research-and-analysis/articles/2021/06/28/which-states-have-dedicated-broadband-offices-task-forces-agencies-or-funds. [Accessed 11 Aug 2021].

[15] Office of the Governor of the Commonwealth of Virginia, "Governor Northam Announces Virginia to Invest $700 Million in American Rescue Plan Funding to Achieve Universal Broadband by 2024," 16 Jul 2021. [Online]. Available: https://www.governor.virginia.gov/newsroom/all-releases/2021/july/headline-898837-en.html. [Accessed 11 Aug 2021].

[16] Obama Whitehouse Council of Economic Advisors, "THE DIGITAL DIVIDE AND ECONOMIC BENEFITS OF BROADBAND ACCESS," March 2016. [Online]. Available: https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160308_broadband_cea_issue_brief.pdf. [Accessed 14 July 2021].

[17] Pew Research Center, "Mobile Technology and Home Broadband 2019," 13 June 2019. [Online]. Available: https://www.pewresearch.org/internet/2019/06/13/mobile-technology-and-home-broadband-2019/. [Accessed 15 July 2021].

[18] Time Warner Cable Inc., Making Connections: Time Warner Cable and the Broadband Revolution, New York, NY: Time Warner Cable Inc., 2011.

[19] US Federal Communications Commission, "FOURTEENTH BROADBAND DEPLOYMENT REPORT," US Federal Communications Commission, Washington, D.C., 2021.

[20] US Department of Education, Office for Civil Rights, "Education in a Pandemic: The Disparate Impacts of COVID-19 on America's Students," 9 June 2021. [Online]. Available: https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf. [Accessed 12 July 2021].

[21] Pew Research Center, "As schools close due to the coronavirus, some US students face a digital 'homework gap'," 16 March 2020. [Online]. Available: https://www.pewresearch.org/fact-tank/2020/03/16/as-schools-close-due-to-the-coronavirus-some-u-s-students-face-a-digital-homework-gap/. [Accessed 12 July 2021].

[22] Pew Research Center, "34% of lower-income home broadband users have had trouble paying for their service amid COVID-19," 3 June 2021. [Online]. Available: https://www.pewresearch.org/fact-tank/2021/06/03/34-of-lower-income-home-broadband-users-have-had-trouble-paying-for-their-service-amid-covid-19/. [Accessed 13 July 2021].

[23] Loudoun Broadband Alliance, "LBA Maps the Broadband Unserved In Loudoun County," 2 May 2021. [Online]. Available: https://loudounbroadbandalliance.org/education/lba-maps-the-broadband-unserved-in-loudoun-county/. [Accessed 26 July 2021].

Cable and Wireless Subscriber Management Convergence

A Common Approach to Identity Management

A Technical Paper prepared for SCTE by

Pablo Stalteri
Master Solution Architect
Hewlett Packard Enterprise
Mississauga ON, Canada, L4W 5G2
pablo.stalteri@hpe.com

1. Introduction

Most operators offering Cable and Wireless services have been implementing their Operations Support System (OSS) platforms using a silo approach, where different access technologies and business units resulted in the deployment of dedicated systems. Thus, to support wireline, broadband, cable TV, IPTV or wireless services, for residential or enterprise customers, multiple Billing, CRM, subscriber and identity management systems are being deployed. Furthermore, technology evolution and tactical execution methodologies when launching new services have contributed to produce different subscriber identification flows for the authentication and authorization required to access each service, as well as a plethora of subscriber data bases and data sources focused on solving the associated entitlement problem. Then the Cable and Wireless operator end up managing and maintaining multiple systems performing similar functions, with increased OPEX, offering different experience to the end customer depending on the service to be used.

In the next paragraphs we will describe the Subscriber Management Convergence solution that has been implemented in North American Triple and Quad Play operators, offering a virtualized consolidated 360-degree view of the subscriber profile for any type of access technology, type of device or type of service offered, encompassing both enterprise and residential customers, making use of legacy data sources to avoid the need of costly migrations.

2. Subscriber Management Convergence Solution

In the telecommunications market every operator tries to differentiate from competition by offering new services to their customers in the hope that they will adopt them if there is a real benefit for the subscriber and user experience is consistent with services already purchased. Most of the times, the real benefit is measured in terms of cost, and for that purpose operators try to bundle the new service with others the subscriber already has, to offer a cross-product discount, but this bundling is sometimes difficult to do from the OSS systems perspective, as there may be a subscriber database for Wireless services, another for Cable and another for Wireline services, making impossible for the operator to link all services of the subscriber during purchasing time. We have seen time and again cases where the CSP has defined an Account for Wireless services, another for Cable and another for Wireless, joining all this information together only during invoicing time, so the end customer receives only one bill. Therefor cross-product discounts are generally limited to services within the same technology: Cable services only (if you have HBO and Netflix service, then bundle package includes x% discount), or Wireless services only, etc. This silo approach ends up penalizing the customer that uses different access technologies, and so the real benefit is not materialized.

With regards to the end user experience, this one is also difficult to implement when the operator has different subscriber databases for each technology and business unit (residential, enterprise), as query entitlements to indicate if a particular service can be used by the subscriber generally

provides specific information associated to the silo that has been reached, without a full view of all services the subscriber has, and the application receiving this data differs from one another depending on the type of device the subscriber is holding. Some of these device applications require user identification and passwords to grant access, which are difficult to remember and maintain by the end user.

Some operators have tried to overcome these issues by consolidating all subscriber profile data into one big database, taking months and years to migrate data from legacy systems to the new one, which added to their OPEX and slow down the speed to which new services can be launched to the market. In these cases, changes done on legacy systems, such as Product catalogue QoS upload/download bytes, service monthly usage limit, throttling, etc, must be replicated in the centralized DB in a timely fashion to ensure consistency, adding complexity to the solution.

These challenges can be summarized as follows:

1. Integrate Subscriber's Data and Entitlements of all users, from cable, wireline and wireless into a single platform. Store identifiers available in user devices
2. Changes done on legacy systems (such as Product catalogue QoS upload/download bytes, etc) must be immediately available to external applications

Subscriber Management Convergence solution provides an answer to each of these requirements, by increasing flexibility in the creation of new services and fast integration with legacy data sources, without the need for long migrations or data consolidation implementation. The concept is based on "use the data that you already have" from legacy applications, accessing subscriber information via user equipment identifications readily available, and presenting each external application the data they need so that end subscriber always has the same user experience. The main advantages of this approach are:

- Abstracts data models from external applications, offering dedicated views to each of them
- Supports entitlement queries for all type of subscribers: cable, wireline and wireless
- Flexible business logic enabling sequential and/or parallel requests to data sources
- Easy integration with legacy data bases and platforms, with more than 100 protocols available
- No need of data consolidation or data migration
- Adaptable to Operators ecosystem: no impact on existing applications or legacy systems

The solution is based on two products that are integrated using 3GPP protocols, deployed in a virtualized environment and benefits from a micro service architecture, to pave the way to 5G implementation.

- **Data Federation** allows applications to query account, subscriber, service and device information from a set of downstream data sources, being the main one Unified Identity Repository (UIR). It federates external data sources via different protocols, to build a **consolidated XML response** and supports AuthN and AuthZ to allows the subscriber to access premium content (Entitlements)
- **UIR** provides an LDAP database which stores identification values (keys), that are used to access legacy data sources to obtain detail information

The outcome can be summarized as follows:

- Rapid adaptation to new business cases
- The data federation **abstracts query requests** so that applications do not need to care about data models and/or where data is retrieved
- Reduce implementation and maintenance costs: single platform across all access types, business units and services
- When legacy IDs are maintained, service catalogue **changes** are **reflected immediately**

3. Solution Architecture

Figure 1 shows the generic LDAP subscriber profile data model that has been implemented in several Quad Play operators using UIR.

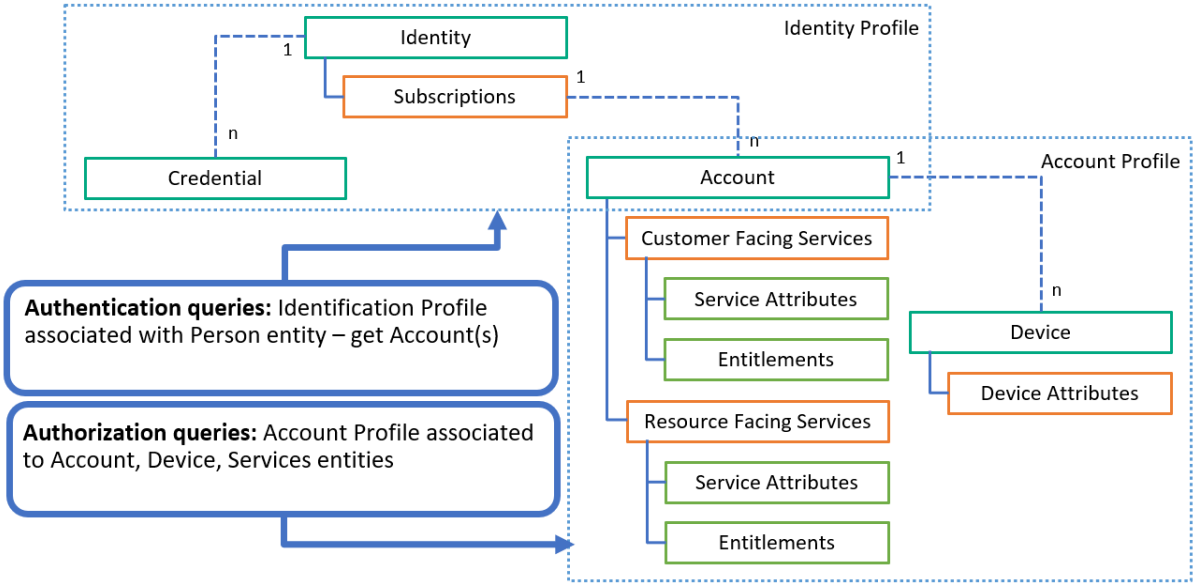


Figure 1 - UIR – Subscriber Profile Data Model

For every subscription, there is an Identity profile, storing different credentials, and an Account profile, which can be made up of multiples accounts (one for Cable, another for Wireless, another for Wireline, etc), which are keys to the corresponding legacy OSS systems the operator has. LDAP ensures the extensibility of the data model and we have found that it can adapt to the realities of any type of CSP.

The Identity profile is used in entitlement queries submitted by external applications, such as Video On Demand, wireless application or WiFi, to authenticate the physical subscriber (person) using a key that is provided by the device used. Thus, in the case of cable this can correspond to the MAC Address of the Set top box, IMSI for wireless, eSIM for Tablet, etc. In this way the subscriber does not need to memorize any userid / password to access the service: the device itself provides this key automatically when the service is selected. The result of the authentication query is the Account number.

The Account profile is accessed via the Account number previously retrieved for the authorization query, which with the input of type of service accessed and device used, allows to validate if the subscriber is entitled to use the service/device. Please note that in one account you may have multiple services (each one with several instances) and allow several types of devices/instances. Examples:

- A wireline Account may have Internet service, IPTV service, Digital TV service, Home Phone service, Home Alarm service, etc.
- A wireless Account may have Share data plan service made up of 2 smartphones (father-mother), each one with each own MSISDN, one Individual data plan service (son) with another MSISDN, etc

- Devices can be classified into access device (i.e GPON, FTTH, etc), consumer devices (set up box, smartphones, smartwatch, tablet, etc), each one with its own instance ID

The second component of the solution, Data Federation, is implemented in a layered architecture to provide flexibility in the northbound and southbound interfaces, as displayed in Figure 2. It consists of:

API Broker Layer

Logical component that hosts one or more API broker components. It provides HTTPS/REST based interfaces to which North Bound applications can authenticate, connect and query the system

Business Logic Layer

Based on API request type it triggers the flow selecting the Data Sources to consolidate the dedicated response:

1. Flow execution order in which each corresponding query needs to be executed according to business logic
2. Query in parallel and / or sequential order the different Data Sources using the appropriate protocol
3. Receive response from all Data Sources and keep them in memory
4. Compose all answers into a virtual data model to create the XML response (Payload)

Data Source Layer

Set of out of the box connectors to interface with each Southbound Data Source (ex. LDAP, JDBC, WebServices, HTTP/REST, SOAP/XML, etc) to acquire relevant information from keys retrieved from UIR and / or external data sources.

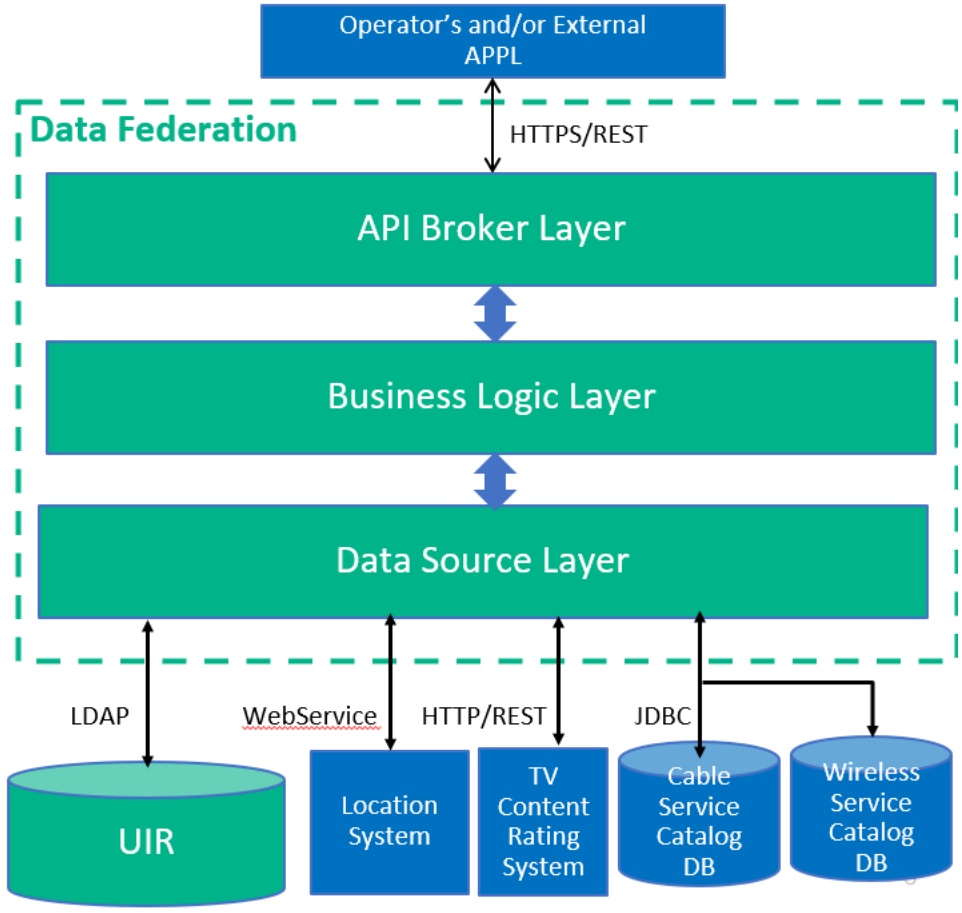


Figure 2 - Data Federation Architecture

It is important to mention that Data Federation is built using micro services technology provided by the enhanced Interactive Usage Manager (eIUM) platform. The rich set of out of the box connectors give the desired flexibility in both northbound and southbound interfaces, while its library of java classes provides the foundation for the business logic layer. Each of the layers are easily configurable, adapting to the realities of each operator in a matter of weeks. The Figure 3 presents an example of API implementation for Cable Enterprise application. Please note the sample list of external applications, some of them from the operator, some from partners (such as content partners), which make use of the authentication and authorization flows supported by the combination of Data Federation and UIR.

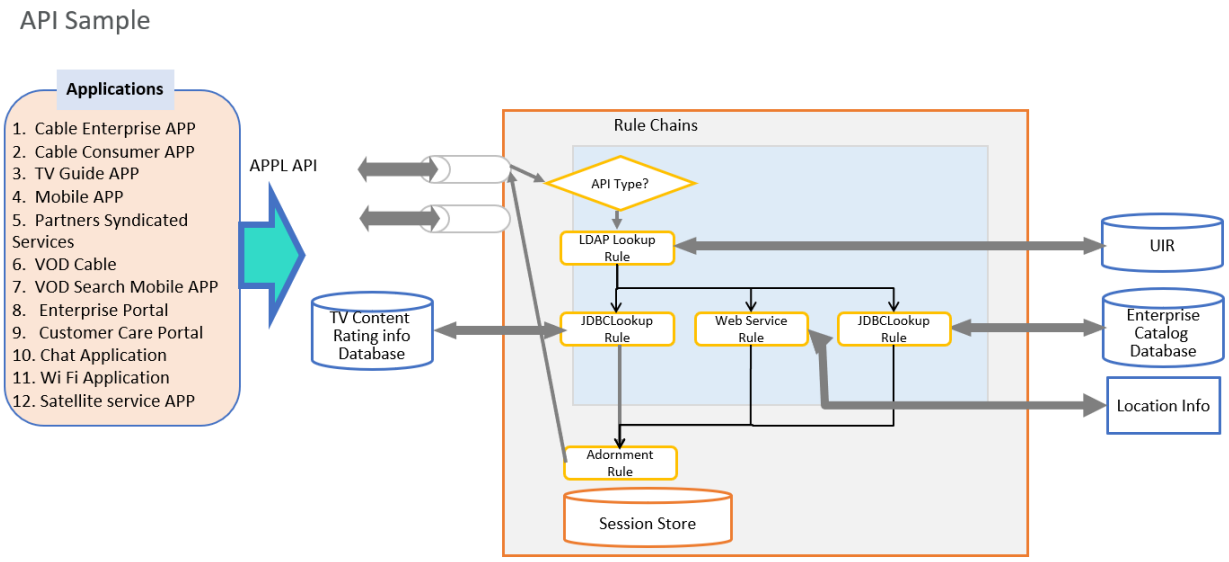


Figure 3 - Data Federation – API Sample

UIR data is populated automatically by the different legacy provisioning systems of the operator, using LDAP protocol. However, we have implemented in the Data Federation business logic layer of one CSP the ability to auto-provision new services in UIR if certain eligibility criteria are fulfilled for the same Person (Identity Profile), allowing cross-product discounts and loyalty programs. The case in question required to provision the entitlement to access for free a specific Sport channel from any wireless device if the subscriber has Service 1 (Wireless Data Plan of 10GB), Service 2 (Internet with no data limit), Account is active and residential and lives in states X,Y,Z. Please note that none of the legacy systems in the CSP were able to register this condition, as they were deployed with the typical silo approach, and only in UIR the operator was able to see for the first time all services the subscriber has purchased, and so offer a new service for free to create stickiness. This auto-provisioning is triggered by the provisioning of any Service 1, Service 2, Account update and/or location update.

4. Features

Data Federation key features are:

- Real Time Protocols for both Northbound and Southbound interface: HTTPS, REST, SOAP, JDBC, LDAP, Diameter, RADIUS client/server/proxy with advanced access control, TLS support
- Messaging format supported: XML, JSON, HTML
- Supports synchronous and asynchronous requests for both northbound and southbound interfaces
- Onboarding of Consumer Applications: Who access What attribute with What API

- OOB GUIs: Data Modelling Studio, Configuration Editor, File Service Workflow, Rule Chain Visualization, App Deployment Visualization, Operations Console, Deployment Manager, O&AM Workflow, Common Codec Framework
- NFV ready and cloud scalability: Level 3 Scale in/out
    - Dataless Micro service architecture, paving the way to 5G and Wi-Fi 6
    - VNF Manager (new component for NFV/MANO integration) and O&AM Workflow
    - EMS Management enhancement: NFV Templates and Operations Console Extensibility
    - Load Balancer for Real Time
- Management enhancement: statistics threshold with alarm, resource utilization chart, SNMP v2c and v3 support, Management HA support
- Reference Data Manager and Services enhancement: extend control for admin to manage reference data that drives business logic, advanced querying and caching

In addition, Data Federation provides optionally an Access Feature including the following two functions:

- **Access Control:**
    - Allow only **trusted clients/applications** to access UIR DF by providing mutual authentication using HTTPS Two-Way SSL
    - Authenticate and Authorize the users accessing UIR DF by checking against its LDAP database where user information, credentials and user's **access control list** (user, roles and privileges objects) data is stored
    - Provide **secure communication channel** between the clients and UIR Data Federation platform: TLS 1.2
- **Access Filtering:**
    - Ability to include service types / privileges / instructions as URL parameters, which enable a client to ask for a subset of the data that it is entitled to (e.g. {BASE_URL}/{RELATIVE_PATH_TO_ACCOUNT_VIEW}?pname=qamTV&pname=xyz)
    - Ability to include instructions, as URL parameters, which request UIR Federation to ignore one or more external data sources (e.g. {RELATIVE_PATH_TO_ACCOUNT_VIEW}?pname=qamTV&pname=xyz&exclude=LOCATION_SYSTEM&exclude=TV_RATINGS)

Access Control function:

- Allows only **trusted clients/applications** to access UIR DF by providing mutual authentication using HTTPS Two-Way SSL
- Authenticates and Authorizes the users accessing UIR DF by checking against its LDAP database where user information, credentials and user's **access control list** (user, roles and privileges objects) data is stored

- Provides **secure communication channel** between the clients and UIR Data Federation platform: TLS 1.2

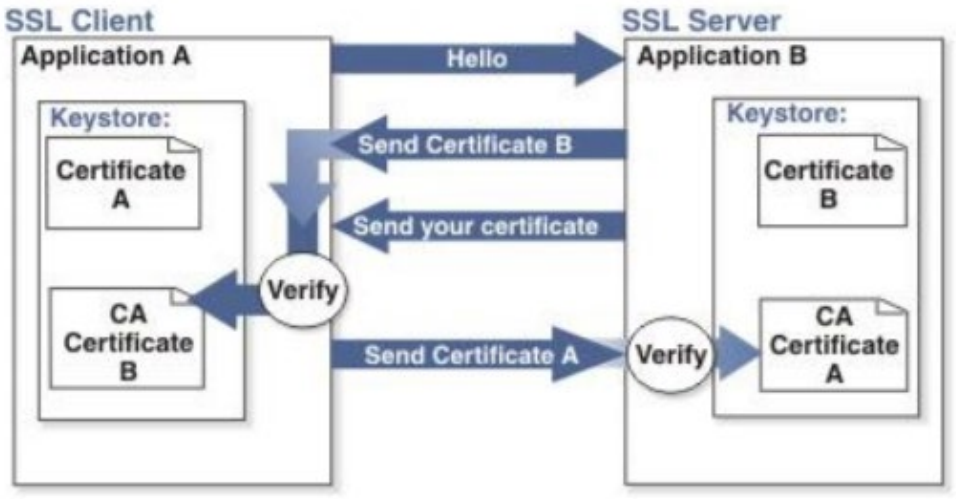


Figure 4 - Data Federation ACL Feature

Data Federation internally will leverage UIR LDAP Database where the Access Control List for a User is stored. The Access Control list would consist of User Role and Its Privileges and other artifacts. Data Federation will provide only those information elements to a user allowed by Access Control List, using the data model shown in Figure 5.
Data Access is enabled for the Account View to:

- Access to services, by type.
- Access to Service attributes, by name
- Access to Devices, by type.
- Access to Device attributes, by name
- Access to Person Role Map. By Id

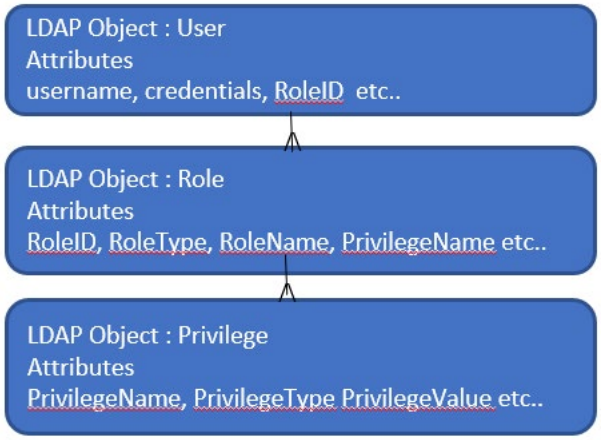


Figure 5 - Data Federation Access Control Data Model

5. Sample Deployment

This section describes a production implementation of the solution in one Quad Play operator in North America with 3 Million Subscribers. Here different provisioning, CRM and billing systems are implemented, for Cable, Wireless and Wireline access technologies, with dedicated flows for residential and enterprise customers. The main goal was to provide to any external application a consolidated view of the subscriber profile, accessed using different keys. Depending on the entitlement query type, defined by the external APP, a different view is presented.

Table 1 - Performance Metrics: Operator with 3M Subscribers - Payload

View	Average TPS	Max TPS	Average Response Time (ms)	Payload size
Account Lite View	5	80	75	Average – 405KB
Account View	19	300	200	Max- 3.3MB
Person View	7	20	10	Average – 1KB Max- 30KB
Mobile App View	4	73	188	Average – 2.5KB Max- 8KB
TV Guide View	9	290	200	Average – 18KB Max- 224KB
VOD Auth View	24	47	140	Average – 3.5KB Max- 17KB
Wifi View	2	26	2	Average – 1.3KB Max- 5KB
TOTAL	**70**	**500**		

As indicated in Table 1, two Account views were implemented, as there are APP that only require a “Lite” version of the account, indicating list of accounts associated to a subscriber ID, while the other version (called “Account View” in the table) lists all accounts, for each account all services and devices, for each service all instances, for each device all instances, etc. Therefor you can see that the payload for this full view is quite high, up to 3.3MB, and its data is retrieved in an average of 200 milliseconds. On the other side of the spectrum, you have the WiFi Application, that is just concerned about QoS associated to the subscriber, and so payload is just 5KB maximum. The platform was sized for 500 TPS queries, measured from the northbound interface, and each query can trigger multiple queries to the UIR, which in turn, according to the business logic of the entitlement query selected by the APP, can query one or multiple external data sources, some of them using JDBC protocol (i.e Oracle DBs), or webservices, for location information.

Figure 6 displays UIR LDAP reads and writes over time, where the number of TPS for reads is greater than the northbound entitlement requests, as the first query to LDAP is associated to Authentication, and the second and onwards with Authorization to retrieve details of the account, service and device.
Provisioning requests, coming from legacy systems, are classified into Inserts (ADD), Updates (MODIFY) and Delete (DEL).

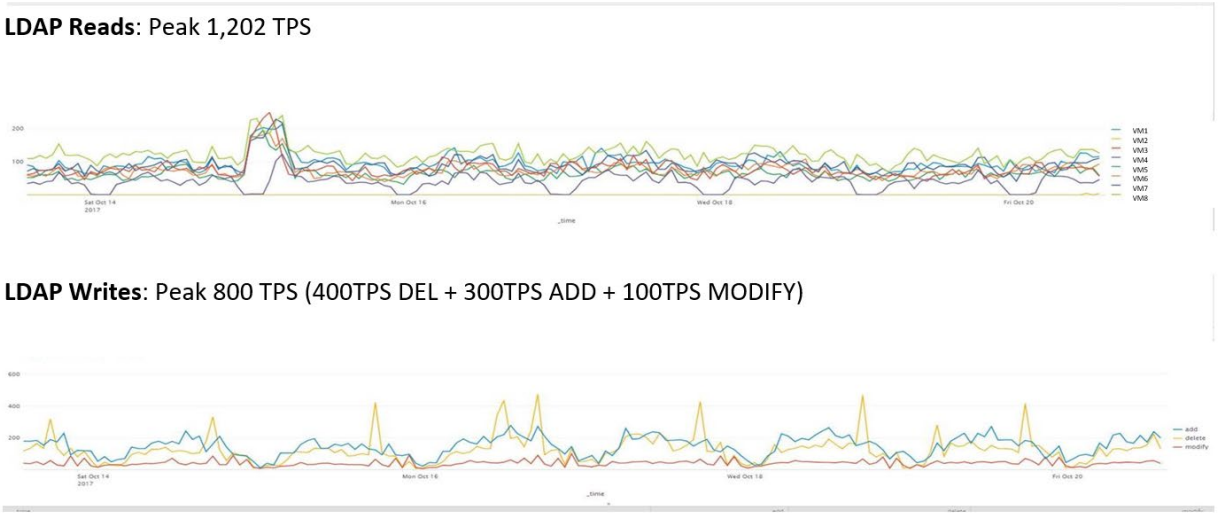


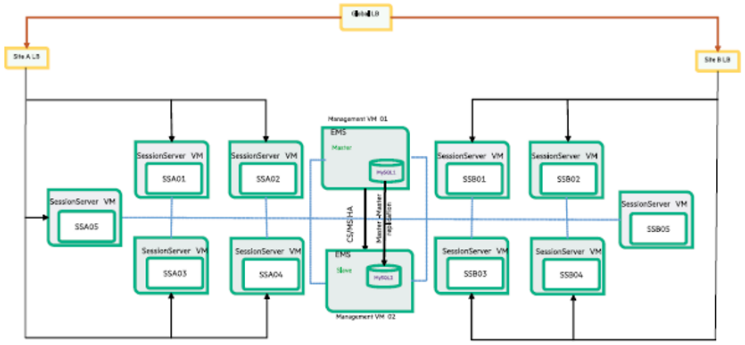
Figure 6. Performance Metrics: Operator with 3M Subscribers – LDAP

Finally, Figure 7 details the production deployment in two identical sites, where each site has five Virtual Machines with Data Federation, each one supporting up to 100 TPS query. The EMS component in each site is made up of an Operations Console, a web-based application designed for operators who need to monitor and manage micro services deployment on a daily basis. Among its capabilities, it can give you an at-a-glance, global view of the health of your deployment, view and monitor all the processes, as well as alarms and problems with your processes. You can also use it to create process groups for easier monitoring and management, and view history charts to show process activity over time. Operators can also perform routine

management operations, for example, starting and stopping processes and groups. The Operations Console also provides role-based security to limit operations capabilities for different users.

- Two identical sites
- Each site with 5 Session Servers VMs and 1 EMS VM, supporting up to 500TPS query
- Each VM with one Session Server containing the 3 layers: API Broker, Business Logic and Data Source Layer

	vCPU	vRAM (GB)	Disk (GB)
VM supporting 100 TPS	8	16	200
OS	RHEL 7.5 (x86-64)		
Hypervisor	VMWare		



• Multi site deployment is achieved with built in HA, that is, there is no need for third party clustering software (i.e. Red Hat Clustering)

• EMS: Single Config Server (CS) and Management Server (MS) run on HA mode with a replicated master-master MySQL database

Figure 7. Sample Production Deployment: Operator with 3M Subscribers

High availability is achieved by having each real time component (called Session Server, in eIUM) in active-active mode in both sites.

6. Conclusion

This paper described the HPE Subscriber Management Convergence solution, which leverages existing legacy data sources and streamlines the identity management process to produce a consistent customer experience across different access technologies and services involved. The solution introduces a centralized Subscriber Profile and Identity Repository along with Data Federation capabilities, so that when there is a need to check the authorization of a customer accessing a specific service from a particular device and location, this entitlement information can be retrieved from a single point. In addition, the authentication flow required to access the service will use the data already provided by the device making the request (i.e. IMSI in smartphone, eSIM in tablet, etc), without the need for the end customer to introduce user id and password whenever possible. These two improvements result in a better customer satisfaction and fast time to market to launch new services.

Abbreviations

HPE	Hewlett Packard Enterprise
CTG	Communication Technology Group
eSIM	Evolved Subscriber Identity Manager
eIUM	Enhanced Interactive Usage Manager
DB	Database
HTTP	Hypertext Transfer Protocol Secure is an extension of the Hypertext Transfer Protocol. It is used for secure communication over a computer network, and is widely used on the Internet
IPTV	Internet Protocol Television
JSON	Java Script Object Notation
REST	Representational State Transfer. Simple HTTP-based protocol that enables to contact the message broker through a Web browser
LDAP	Lightweight Directory Access
OPEX	Operating Expenses
OSS	Operation Support System
SOAP	Simple Object Access Protocol specification for exchanging structured information in the implementation of Web Services
TPS	Transactions Per Secong
UIR	Universal Identity Repository
XML	eXtended Markup Language

Bibliography & References

(1) HPE eIUM UIR Data Federation Delivery Guide

(2) HPE Universal Identity Repository Overview Manual

Cluster-Based Network Traffic Prediction Pipeline For Big Data Time Series

A Technical Paper prepared for SCTE by

Wei Cai
Network Planning Engineer IV
Cox Communication
6205-B Peachtree Dunwoody Rd, Atlanta, GA 30328
678-645-0000
wei.cai@cox.com

1. Introduction

Upstream broadband usage and network capacity have been increasing sharply in recent years, particularly under global pandemic lockdowns since March 2020 [NCTA/covid-19-overview]. The spread of COVID-19 around the entire world has placed upstream traffic growth on an extremely irregular pattern with fluctuations, posing great challenges to use conventional methods such as Auto-Regressive Integrated Moving Average (ARIMA) and other classical statical models to ensure accurate forecasts because those classical methods have been proven to be weak and inadequate for modeling non-stationary traffic flows.

In practical forecasting applications, the use of different modeling methods has become a popular research by many scholars [Chen, 2011; He and Zeng, 2021]. However, with the rapid development of cable companies' network, network traffic data scale is increasingly important in modern network traffic world. Just combining forecasts from different models with weighs allocated does not satisfy the need to model big traffic flow data more effectively.

In this study, we propose a clustering-based two-stage approach to build network traffic forecasting models using ARIMA, eXtreme Gradient Boosting (XGBoost) and Long Short-term Memory (LSTM) that includes data preprocessing and model building, as shown in Fig. 1.

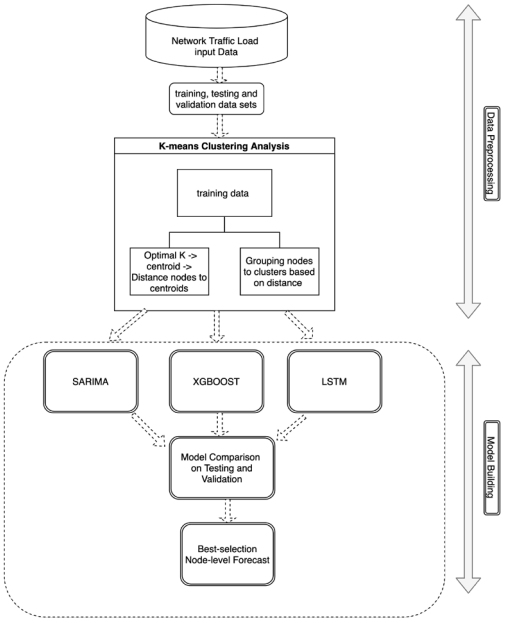


Figure 1 - The flowchart of the proposed best-selection forecasting

In the data preprocessing stage, using the training subset, we apply the K-means clustering method to find the K cluster centers to group a collection of network nodes into homogenous subsets based on intra-cluster similarity in traffic data [Vujicic et al., 2006]. K-means clustering is one of the most-commonly used data clustering algorithms, originally from signal processing. It aims to partition n observations into k clusters in which each observation belongs to the cluster with nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster [Anderberg, 1973, Kaufman and Rousseeuw, 1990]. A key assumption of applying forecasting models on clusters is the similarity of behavior patterns within a cluster. This algorithm is helpful to obtain computational efficiency when it comes to big data time series forecasting.

The paper is structured as follows: Section 2 explores the related studies and research gaps in existing literatures. Section 3 describes the methodology. Section 4 discusses the experimental results and Section 5 concludes the paper.

2. Background

Network traffic-related data are collected as time-series, hence time series analysis and forecasting techniques such as ARIMA modelling, XGBoost and LSTM can be employed for the traffic forecasting. In this section, we first review three modeling techniques applied: ARIMA, XGBoost and the deep learning-based Long Short-Term Memory (LSTM) model, then followed by review of related work.

2.1. Learning methods

2.1.1. ARIMA

The ARIMA model was pioneered by Box and Jenkins [Pankratz, 2008]. It is one of the most classical statistical models in time series prediction. This model can be presented as a linear regression function, in which lags and the lagged forecast errors are used for prediction. A standard ARIMA model can be expressed by three terms: autoregressive order (AR, p), moving average part (MA, q), and difference order (differencing, d). A seasonal ARIMA is an expanded version of a standard ARIMA model with information extracted from the seasonal parts that cannot be processed by the standard ARIMA model. To build a seasonal ARIMA model, finding the non-seasonal part (p, d, q) of ARIMA is the first step, and tuning the seasonality components (P, D, Q) is the second. Diagnostic checking of stationary and residual uncorrelation is vital for a good fit model. A simple ARIMA structure is illustrated in Figure 2 as follows:

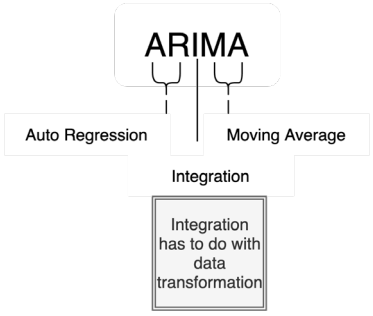


Figure 2 – ARIMA Structure

2.1.2. XGBoost

Extreme Gradient Boosting as one of the boosting algorithms in ensemble learning was first proposed by Tianqi Chen in 2015 and Carlos Guestrin in 2011. It is proved in the literature [Chen and Guestrin, 2016] that the XBGoost model has the characteristics of being efficient and scalable [Saeed, 2016]. The basic idea of the Boosting algorithm is to combine many weak learners to form a strong model that can predict accurately. Extreme gradient lifting tree [Wang et al., 2019] essentially is an integrated learning algorithm in which the number of decision tree keeps being added with each iteration till a strong classifier is found [Fabricius, 2000]. The XGBoost model has proved to have many advantages in model prediction, such as the lack of a need to preprocess the data, a fast operation speed, complete feature extraction, a good fitting effect and high prediction accuracy [Alim et al., 2020]. A simple XGBoost structure is illustrated in Figure 3 as follows:

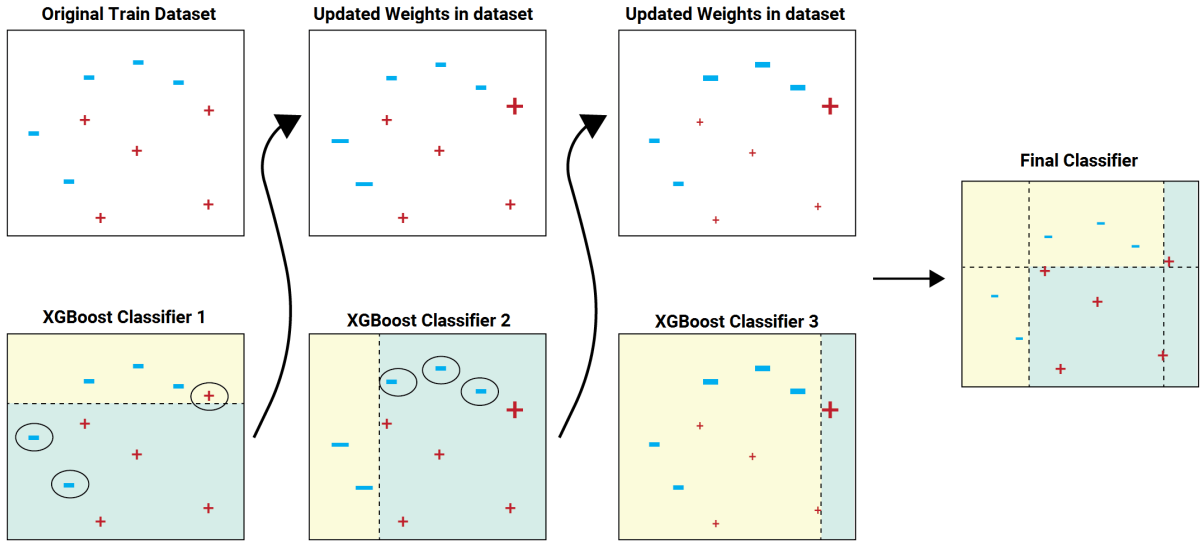


Figure 3 - XGBoost structure (https://blog.quantinsti.com/xgboost-python/)

2.1.3. LSTM

Long short-term memory networks (LSTM), proposed by (Hochreiter & Schmidhuber, 1997) is a special kind of recurrent neural network (RNN). Different from other RNN models that could easily suffer from the problems of getting small gradient loss (i.e., less than one), and multiplication of those gradient losses would vanish during training process, LSTM overcomes these problems by memorizing the prior stages' internal trend through a few different gates (i.e., input gate, forget gate, control gate and output gate) and optionally let data pass through or dispose of based on a sigmoidal neural network layer to predict future patterns. A simple LSTM network is illustrated in Figure 4 as follows:

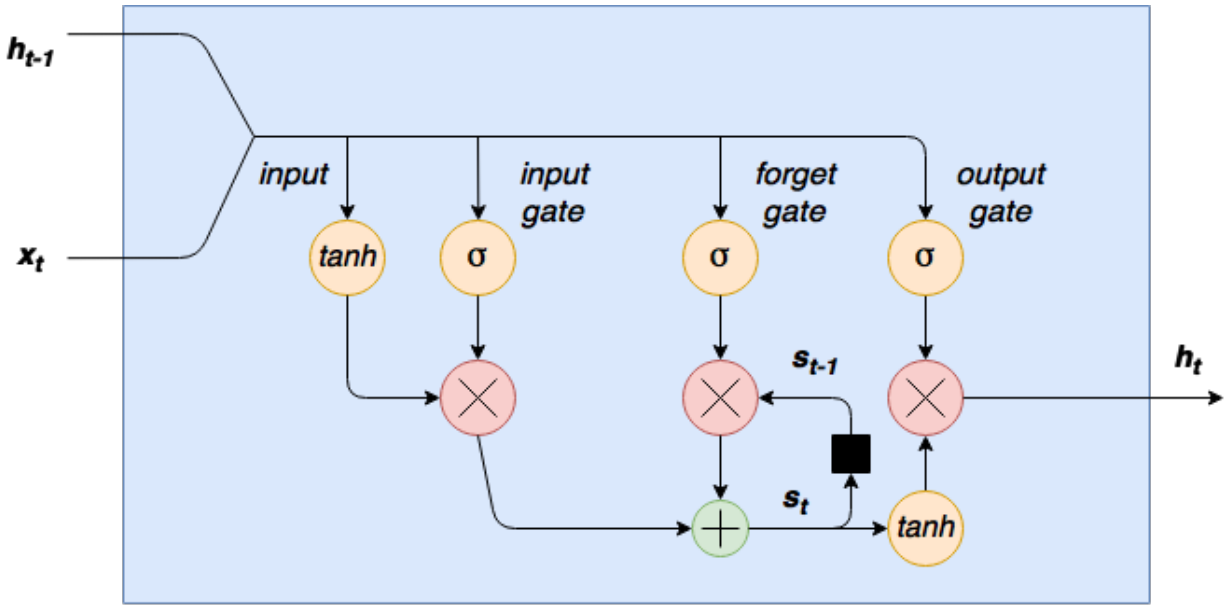


Figure 4 - LSTM structure (https://adventuresinmachinelearning.com/keras-lstm-tutorial/)

2.2. Related work

There are several works that have focused on comparing the performance of different forecasting techniques. He and Zeng (2021) have put forward a forecasting method of XGBoost-LSTM combination model based on a weighting method in forecasting product sales. The two years' sales time series data are modeled by using XGBoost and LSTM neural network models, respectively. They prove that the performance of using XGBoost-LSTM model to deal with time series is much higher than that of the original XGBoost single model, which maximizes the advantages of the two prediction models. Chen (2011) proposed to use the combination models that combine the linear model and the nonlinear model to predict tourism demand time series data. The results showed that the combination is superior to the individual models for the test cases of tourism demand time series. Alim et al. (2020) compared the performance of the XGBoost model a seasonal ARIMA model for human brucellosis in mainland China and used them to make short-term predictions. The results showed that the prediction accuracy of the XGBoost model was much better than that of the ARIMA model.

2.3. Research gap

From the above-mentioned research works, we have identified a couple of major research gaps. To the best of our knowledge, first major research gap is that there are very few research papers that have employed a clustering approach to perform classification of large quantities of network traffic data, grid search parameters or train model within a cluster. Clustering-based modeling approach can be very effective in terms of reducing computational cost for big data but still obtain decent prediction accuracy. Researchers in one study prove training the classical SARIMA models on clusters of public safety network users identified by the K-means algorithms performs well compared to the prediction based on the overall aggregate traffic [Vujicic et al., 2006]. With the data we have, at an individual node level, time series are very volatile, include a high amount of noise, and are unevenly and nonlinearly affected by different effects. However, most nodes in this study follow similar pattern and many series move in tandem, suggesting possibility to group those nodes into homogeneous clusters for hyperparameter search and model training.

Secondly, different from many relevant research studies recommending using combination of forecasting model: selecting appropriate weights to weight and average the results obtained from several different model, we propose to build different models separately and let three models compete to always select the best individual-level forecast in practical forecasting application. This will provide more reliable traffic prediction to the network management and the network capacity planning.

3. Methodology

11,738 network nodes' weekly upstream traffic load time series were selected into input for this study to test the performance of three models: ARIMA, XGBoost and LSTM. After replacing missing values with linear interpolation for each time series, 174 weekly data polls between 2018 and May 2021 were collected as the input data. During model development, 70% of the input data were used to train models, 20% is used as the testing data set, and the remaining 10% is used as the verification data set.

3.1. Data description and Pre-processing

The upstream traffic dataset used in the study was obtained from a cable company's network. It contains information regarding *site id*, *node description*, and *traffic loads* in weekly intervals, gathered from 11, 738 nodes. *Site id* and *node description* were concatenated to a string to label each network node in this study.

The significant problem forecasting is facing is the presence of missing values and outlier values in the observed historical traffic load data. In the dataset we collected, there exist missing values in weekly timestamp and traffic loads due to occurrences of data collection failures. Therefore, those missing values were linear interpolated based on historical trend. Linear interpolation is illustrated in Figure 5. In Figure 5, the left plot shows raw data weekly trend for one sample node but with missing polls highlighted with yellow arrows. After interpolation, those missing polls were filled with imputed values which can be seen on the right plot.

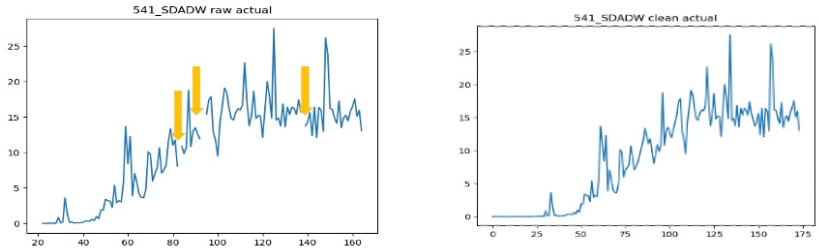


Figure 5 - Missing value linear interpolation

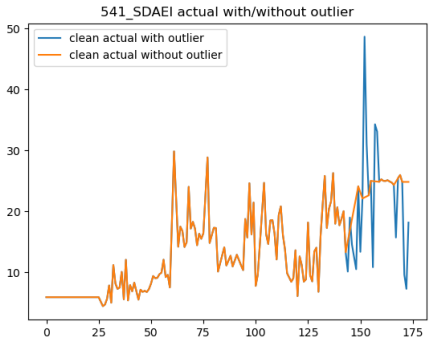


Figure 6 - Data with/without outliers

Outliers in time series can be grouped into two types: outliers affecting a single observation and outliers affecting a single observation and subsequent observations, which is true in our input data. In this study, we use a combination of statistical method to detect those anomalies and then apply linear interpolation based on time to replace them with imputed values. The effect is depicted in Figure 6.

3.2. Forecasting model set-up

Three different models based on the same input are proposed and fitted to predict future traffic loads. 70%-20%-10% of data partition is used for training, testing and validation purposes. The performance of all three models is presented based on the criteria of Mean Absolute Percentage Error (MAPE) as shown below:

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \quad (1)$$

where A_t is the actual value, F_t is the forecast value mean, n is the number of times the summation iteration happens. M stands for mean absolute percentage error used to evaluate prediction accuracy and is expressed in percentage. In this study, a value lower than 10 is said to be a fit model. The evaluation of all three models is done by preparing a model on a training dataset and by making predictions on a test dataset and a validation dataset. Training, test and validation MAPEs are calculated to measure model performance.

Before training each model, K-means clustering method was used to partition 11,738 nodes into 4 optimal clusters. In this study, we use Euclidean as a distance metric because we have normalized each time series to have the same length for all nodes selected after preprocessing. the node count distribution by cluster is summarized as follows: Cluster 0: 10,082 nodes; Cluster 1: 392 nodes; Cluster 2: 71 nodes; Cluster 3: 1,193 nodes. To intuitively observe the data characteristics of each cluster's traffic load, the sum of traffic loads by each cluster is plotted in the following figure:

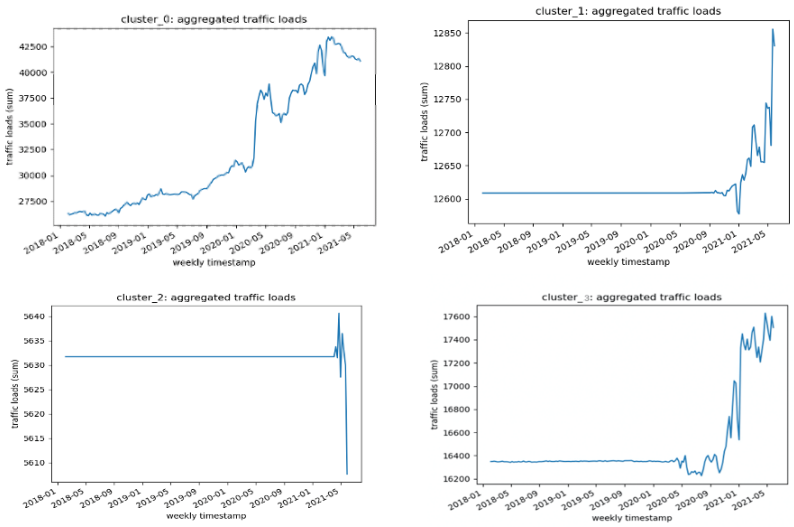


Figure 7 - Upstream Traffic Loads Trend by Cluster

As we can see form Figure 7, Cluster 0 shows much more volatilities and fluctuations since the start of COVID-19 compared to three other clusters. Clusters 1 and 3 experience more seasonal effect compounded with COVID-19 since late 2020. Cluster 2 has a relatively more peaceful trend compared to three others.

After K-means clustering, the main steps followed for the ARIMA model fitting as follows:

- Grid search was used to automatically discover the optimal order of non-seasonal and seasonable parameters at a cluster level. The combination of (p, d, q) that returns the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values are selected as the most optimal parameters for each cluster.
- Dataset is split into 70%, 20%, 10% training, test and validation sets, respectively.
- Stationarity of the series are checked using the Adfuller function with the P-value along with the autocorrelation function (ACF) and the partial autocorrelation (PACF) plots. Training data are transformed based on the stationarity check.
- SARIMAX models are trained and obtained to make estimation on fresh test data.
- To check the goodness of fit of the ARIMA model, MAPE values for both training dataset and testing data set are calculated.

After K-means clustering, the main steps followed for the XGBoost model fitting as follows:

- To obtain the best performance, the grid search algorithm is used to optimize the parameters: min_child_weight, gamma, subsample, colsample_bytree, max_depth and learning rate at a cluster level.
- Dataset is split into 70%, 20%, 10% training, test and validation sets, respectively.
- XBGoost models are trained and obtained to make estimation on fresh test data.
- To check the goodness of fit of the XGBoost model, MAPE values for both training dataset and testing data set are calculated.

After K-means clustering, the main steps followed for the LSTM model fitting as follows:

- Scale data using MinMaxScaler to speed up the learning process and help model.

- Hyperparameter such as number of layers, layer depths, activation functions, dropout coefficients are repeatedly tuned at a cluster. Manual tuning of LSTM hyperparameter is used instead of using automated tuning package such as SMAC3 to avoid common failure to find well-defined global minima with automated packages.
- Dataset is split into 70%, 20%, 10% training, test and validation sets, respectively.
- LSTM models are trained and obtained to make estimation on fresh test data. During the model training, the validation loss was monitored by early stopping call back function to halt the training if there is an increment observed in loss values. The number of epochs for the training was tested with different values.
- To check the goodness of fit of the LSTM model, MAPE values for both training dataset and testing data set are calculated.

4. Results and Discussion

In order to see the prediction effect of all three selected models, training, test and validation MAPEs are matched on 11,542 nodes for comparison. All three groups of MAPEs were classified into four groups: <= 5%, between 6 and 10%, between 11% and 15% and greater than 16%. Number of nodes were summarized for each MAPE group.

Tables 1 and 2 respectively present the comparison of the training and test prediction accuracy for three selected models.

Table 1 - Training MAPE by Models

Model	Training MAPE				Total
MAPE Range	<= 5%	>=6% & <= 10%	>=11% & <= 15%	>=16%	
SARIMA	1,360	2,828	2,059	5,295	11,542 nodes
XGBoost	11,333	55	20	134	11,542 nodes
LSTM	3,573	2,979	2,909	2,081	11,542 nodes

From Table 1, for the training set, the XGBoost model has the highest number of nodes with the MAPE value less than or equal to 10% than two other models.

Table 2 - Test MAPE by Models

Model	Test MAPE				Total
MAPE Range	<= 5%	>=6% & <= 10%	>=11% & <= 15%	>=16%	
SARIMA	982	746	1,175	8,639	11,542 nodes
XGBoost	2,244	2,228	2,870	4,200	11,542 nodes
LSTM	2,299	3,996	2,753	2,494	11,542 nodes

In Table 2, for the test set, the LSTM model has a higher number of nodes with the MAPE value less than or equal to 10% than two other models. Majority of the nodes with the LSTM showed the MAPE value less

than 15%, which shows the potential benefit of using the LSTM to predict network traffic. The ARIMA has the lowest number of nodes with a good fit, which might indicate the weakness of using conventional statistical models to predict for non-linear data.

To further evaluate model performance of all three selected models, we chose nodes with both training MAPE and test MAPE less than 10% to compare training MAPEs with test MAPEs. If the test MAPE is lower than training MAPE, we consider that node's forecast is a good fit. Vice versa, if the test MAPE is higher than training MAPE, we label that mode as overfitting training data but does not perform well on the evaluation data. Node counts for good fit and overfitting by different models are summarized in Table 3:

Table 3 - Node Counts with <= 10% MAPE by Models

Model	Node counts with good fit (Mape <= 10%)	Node counts with overfit (Mape <= 10%)
SARIMA	956	586
XGBoost	1,195	3,248
LSTM	2,129	2,278

As shown in Table 3, it is clearer that LSTM overall has less overfitting issues in network traffic forecasting than two other models. It further indicates that LSTM in this paper has higher fitness and predictive performance in network traffic prediction compared to ARIMA and XGBoost. XGBoost model has the most overfitting issues in this study.

To map model performance for the validation dataset, a comparison of 596 nodes validation MAPEs is shown in Table 4.

Table 4 - Node Counts MAPE by Models

Model	Node counts with Mape <= 10%)	Node counts with Mape >10%)
SARIMA	101	495
XGBoost	139	457
LSTM	346	250

Table 4 shows that LSTM has the most of number of nodes with validation MAPE less than 10%, followed by XGBoost and SARIMA.

According to the overall results, the performances of the ARIMA models were the lowest and the LSTM models performed the best. LSTM models produced the best results.

5. Conclusion

The author proposed to use K-means analysis to partition network traffic nodes to different cluster and apply seasonal ARIMA model (SARIMA), XGBoost model, and the Long Short-Term Memory (LSTM) model on clusters to reduce computational cost and predict effectively.

Three selected models are compared with respect to performance. According to the comparison, the performance of the LSTM network is better than ARIMA and XGB model. XBGoost model shows a reasonable performance but showed serious overfitting issues. Moreover, the classical data analysis model ARIMA has more obvious forecast error, which shows the disadvantage of classical parameterized approach faced with tremendous traffic data.

As for future work, the SARIMA model could be improved using one-step ahead forecast method. At the same time, LSTM models can also be improved by hyperparameter tuning such as the number of layers, learning rate, optimizers, etc. Overfitting with the XGBoost model could be avoided using the delta difference between time interval (i.e., lag term of the time series) as the input and let the input lag term predict the univariable time series. L1 and L2 regularisation can be introduced to the LSTM and XGBoost to address overfitting. Cross-validation can be tried with all three selected models to obtain more reliable forecasting models.

Abbreviations

ARIMA	Auto-Regressive Integrated Moving
SARIMA	Seasonal Auto-Regressive Integrated Moving
XGBoost	Extreme Gradient Boosting
LSTM	Long Short Term Memory
AR	Autoregressive Order
MV	Moving Average
RNN	Recurrent Neural Network
MAPE	Mean Absolute Percentage Error
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
ACF	autocorrelation function
PACF	partial autocorrelation

Bibliography & References

Alim, Mirxat, Ye, Guo-Hua, Guan, Peng, et al. *Comparison of ARIMA Model and XGBoost Model for Prediction of Human Brucellosis in Mainland China: a Time-series Study*, England: BMJ Publishing Group LTD BMJ open, 2020-12-07, Vol.10 (12), p.e039676-e039676

Anderberg, M. R., 1973. Cluster Analysis for Application. Acdemic, New York

Babajide Mustapha I, Saeed F. *Bioactive Molecule Prediction Using Extreme Gradient Boosting Molecules* 2016;21. doi:10.3

Chen T, Guestrin C. *XGBoost: A Scalable Tree Boosting System* CoRR abs, 2016, pp. 1603-02754.

De'Ath G, Fabricius K E. *Classification and Regression Trees: a Powerful Yet Simple Technique for Ecological Data Analysis Ecology*, vol. 81, no. 11, 2000, pp. 3178-3192.

He, Wei; Zeng, QingTao. *Research on sales Forecast based on XGBoost-LSTM algorithm Model* Journal of Physics: Conference Series; Bristol Vol. 1754, Iss. 1, (Feb 2021). DOI:10.1088/1742-6596/1754/1/012191

Hochreiter, S., Schmidhuber, Jürgen (1997). *Long short-term memory* Neural Computation, 9(8), 1735–1780.

Kaufman, L., Rousseeuw, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis* New York, NY: Wiley-Interscience, 1990

NCTA/covid-19-overview: https://www.ncta.com/whats-new/ncta-launches-covid-19-internet-dashboard

Pang L,Wang J, Zhao L, et al. *A novel protein subcellular localization method with CNN-XGBoost model for Alzheimer's disease*. Frontiers in Genetics (S1664-8021), 2019, 9:751.

Vujicic, B., Chen, H., Trajkovic, L. *Prediction of traffic in a public safety network* June 2006, SourceIEEE Xplore Conference: Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on Project: Collection, Characterization, and Modeling of Network Traffic

Configuring and Deploying Low Latency DOCSIS Networks

A Technical Paper prepared for SCTE by

Greg White
Distinguished Technologist
CableLabs
g.white@cablelabs.com

Karthik Sundaresan
Distinguished Technologist
CableLabs
k.sundaresan@cablelabs.com

CableLabs
858 Coal Creek Circle, Louisville, CO,80027
303-661-9100

1. Introduction

Now that Low Latency DOCSIS (LLD) gear is certified and available, operators are beginning lab testing and field trials, and are working out how to deploy and configure the equipment to provide the best performance for their users. This paper provides guidance and best practices for network operators in the configuration and management of Low Latency DOCSIS services. It is essentially a cheat-sheet on the different provisioning modes available in LLD as well as a run-down of the various control knobs available to the operator - what they do, how they interact with one another, and what effect they have on the service characteristics. It introduces each of the features that make up LLD and lays out a template of recommended configuration settings, the reasons behind those choices, and how those choices might evolve over time. It provides MSOs with a step-by-step approach to turning on each of the LLD features to ultimately ensure that customers can enjoy the benefits of the technology. In addition, this paper discusses some of the end-to-end aspects of delivering on the low latency promise, such as DSCP and ECN traffic marking, and performance monitoring.

2. Brief overview of Pre-LLD latency management features

Applications today that use a transport protocol like TCP seek out as much bandwidth as possible and use a "congestion control" algorithm (most commonly "Cubic") to adjust to the available bandwidth at the bottleneck link through the network. Typically, this will be the last mile link—the DOCSIS link for cable customers—where the bandwidth available for each application can vary rapidly as the activity of all the devices in the household varies. The behavior of the Cubic congestion control algorithm is that it increases its sending rate (or, more precisely, the number of bytes in flight) until it experiences packet loss (generally due to the bottleneck link buffer overrunning), then it pauses sending for a short time (allowing the bottleneck link buffer to drain a bit) before ramping up its sending rate again. The result typically is that the bottleneck link buffer is kept nearly full whenever a TCP file transfer is occurring. This in turn results in latency (and latency variation) for all of the applications that are sharing that bottleneck link.

There are two features available in DOCSIS equipment that pre-date LLD, and that can help manage the latency caused by congestion controlled traffic. For reference, these two features are briefly described here.

2.1. Buffer Control

Past implementations of cable modems had upstream buffer sizes that were found to be quite large, much larger than they needed to be to ensure good TCP throughput. CMTSs similarly were built with over-sized downstream buffers. In many cases, these buffers could hold multiple seconds worth of packets. This phenomenon was referred to as "bufferbloat", and was not unique to DOCSIS equipment. The result was that applications would see highly variable latency, with huge latency spikes occurring whenever a sufficiently large TCP file transfer occurred.

DOCSIS 3.0/3.1/4.0 specifications allow a cable operator to tune the transmit buffer size for cable modems and CMTSs in their networks. The feature controls upstream buffering in the cable modem, and downstream buffering in the CMTS for each Service Flow.

The Buffer Control parameters limit the maximum queue depth of a Service Flow. The Service Flow buffer holds the packets that are enqueued for transmission and this parameter sets an upper limit on the amount of data that can be enqueued for transmission. By providing the ability to control per-Service Flow buffers, the Buffer Control parameters provide a means of balancing throughput and latency in a standardized and configurable manner.

In order to accommodate implementation differences (e.g., varying amounts of memory available for buffering) and to allow an optimized partitioning of buffering memory based on the number of active Service Flows, the main Buffer Control parameter for a Service Flow is referred to as Target Buffer (TLV [24/25].35.2), which the operator can use to set the desired buffer size in bytes. The implementation (CM or CMTS) is required to set the actual buffer as close to the target value as is reasonably possible.

In cases where more precise control of the buffer size is desired, there are two additional parameters: Minimum Buffer (TLV [24/25].35.1) and Maximum Buffer (TLV [24/25].35.3), both also specified in bytes. The implementation is required to reject the configuration if it cannot comply with the buffer size limits configured by these two parameters.

Alternatively, an operator can configure a Default Upstream Target Buffer Configuration (TLV 68) parameter that applies to all upstream Service Flows in the absence of the Buffer Control setting for the Service Flow. This parameter is only applicable to the upstream buffers in the CM. It is specified in milliseconds, and the CM is required to size each relevant Service Flow buffer by calculating the appropriate buffer size based on the Service Flow's Maximum Sustained Traffic Rate value.

CableLabs previously published guidelines on configuration of the Buffer Control feature [Buffer Control]. Here we provide some updated recommendations on setting buffer sizes.

In order to ensure that a single TCP connection can fully utilize the SF rate, the SF buffer size needs to be greater than or equal to the base RTT (i.e. the round-trip-time in absence of queuing delay) of the TCP connection, yet to minimize the latency and latency variation caused by TCP traffic, the buffer should be kept as small as possible. Thus setting the buffer size involves a tradeoff between ensuring good TCP throughput for long distance file transfers, and good quality of experience for latency sensitive applications.

Since many TCP connections are between a client and a local CDN node (perhaps 10-20 ms base RTT), and in many areas RTTs greater than 50 ms are fairly rare, it is recommended that operators set the buffer size to 50 ms, as a balance between good TCP performance for most file transfers and good QoE for latency sensitive applications. But, this value is region dependent. In regions where typical RTTs are significantly shorter (or longer) than 50 ms, a smaller (or larger) value can be used.

Table 1 – Recommended Buffer Control Parameters

Option 1: Per SF configuration		
Parameter	**Upstream Settings**	**Downstream Settings**
Buffer Control: Target Buffer	TLV 24.35.2 = 50 ms * MSR / 8 *MSR = Maximum Sustained Rate for the Service Flow*	TLV 25.35.2 = 50 ms * MSR / 8
Option 2: Default per CM or CMTS		
Parameter	**Upstream Settings**	**Downstream Settings**
Default Target Buffer Configuration	TLV 68 = 50 ms	CMTS vendor proprietary = 50 ms

2.1. AQM

While the Buffer Control feature allows the operator to set the buffer size to a more appropriate level, Active Queue Management (AQM) can work to reduce the buffering latency even further without sacrificing TCP throughput. AQM algorithms monitor the queue depth (or queue delay) in the buffer, and automatically make intelligent decisions to drop packets at appropriate intervals in order to send a congestion signal to the traffic senders (e.g. TCP senders) that results in the queuing delay being kept relatively low, while allowing the link to be fully utilized. Thus, a network device can support enough buffering so that it can absorb bursts of packets on the ingress but doesn't let a standing queue build up.

The DOCSIS 3.1 specification requires cable modems to implement a particular AQM algorithm called DOCSIS-PIE that is described in [RFC8034]. The DOCSIS 3.1 specification also requires CMTS equipment to implement an AQM algorithm, although it leaves the choice of that algorithm up to the implementer. Active Queue Management in DOCSIS is further discussed in [DOCSIS AQM].

By default, AQM is turned on for all Service Flows. It is a best practice to ensure that the CM Upstream AQM disable (TLV 76) in the configuration file is set to Enable, or absent (the default is enable). The Service Flows also have AQM encodings within the Service Flow set of parameters. It is also recommended to ensure that the SF AQM Disable (TLV [24/25].40.1) in the configuration file is set to Enable, or absent (the default is enable) for both upstream and downstream Service Flows. Also it is recommended to leave the Classic AQM Latency Target (TLV [24/25].40.2) at its default value of 10ms, unless experimentation is done to validate another setting.

3. Brief Overview of LLD Features

The LLD functionalities that were recently added to the DOCSIS 3.1 specifications recognize that even AQM isn’t enough to provide a consistent low latency experience for latency sensitive applications like multiplayer online gaming and cloud gaming. For these applications, the bursts of packets that AQM occasionally allows to queue up in the bottleneck link buffer, are disruptive to the Quality of Experience for the user.

The key insights that led to the development of LLD were: a) while the majority of application traffic uses traditional TCP and thus causes latency variation and packet loss, other applications (in particular many latency and loss sensitive ones) aren’t causing these degradations, but nonetheless suffer as a result of them; and b) there are TCP congestion control algorithms that don’t cause these degradations, but they are nearly impossible to deploy in the internet because they don’t work well with existing bottleneck link technology.

LLD seeks to address these two insights by allowing applications that don’t cause latency variation & loss to avoid being subjected to those degradations, and by introducing support for a new class of congestion control algorithm that can adjust to the available capacity of the bottleneck link without causing those degradations in the first place.

Further details can be found in [SCTE LLD], but the core of LLD consists of support for two queues in each direction, a “Classic” queue for traditional congestion-controlled traffic, and a “Low Latency” queue for traffic that doesn’t cause latency, latency variation and loss. The Classic queue has a comparatively deep buffer (along with AQM) that allows traditional congestion controllers to achieve high throughput while keeping latency reasonably under control. The Low Latency queue has a very shallow buffer along with some other features that enable “well behaved” applications to achieve ultra-low delay.

The key to LLD is that it doesn't differentiate between these two categories of traffic based on any kind of subjective judgement as to the importance of one application vs. another, or even the relative latency/loss sensitivity of one vs. another. Instead, the distinction is made based on the application sender's behavior: does it send data in a smooth and/or sparse manner that doesn't materially contribute to queuing delay and packet loss, or does it send data in such a way that it over-drives the link, causing a queue to build up, and backs off only when it senses severe congestion. A related aspect of LLD is that it doesn't prefer one type of traffic over the other. Instead the goal is to allow both types to coexist and share the link capacity in a fair manner, and to allow the application to make the decision as to how it wishes to behave, and thus which queue it should utilize.

DOCSIS equipment, going all the way back to DOCSIS 1.1 gear, supports the concept of Service Flows, which are uni-directional logical pipes set up on the DOCSIS link between the CM and CMTS. Each Service Flow is described by Quality of Service parameters that govern aspects like data rate, and packet classifiers can be set up to direct packets into the appropriate Service Flows. Low Latency DOCSIS uses this existing Service Flow (SF) functionality to create a logical pipe (SF) for the Classic queue and another for the Low Latency queue, and uses the existing classifier mechanism to direct packets into the appropriate SF. LLD then adds some new functionality that ties this pair of Service Flows together, referred to as a Low Latency Aggregate Service Flow (LL ASF), as well as some additional features that enable the LL ASF and its constituent LL SF and Classic SF to provide great performance for all of the different applications that make use of it.

One of those additional features is support for a new congestion control mechanism that allows applications to dynamically adjust their sending rate in order to fully utilize the available capacity in a fair manner, but without causing queuing delay or loss. This mechanism is referred to as the "Low Latency, Low Loss, Scalable Throughput" (L4S) architecture, and it involves the bottleneck link providing real-time congestion signals to applications via a single bit in the header of each IP packet. This "Explicit Congestion Notification" mechanism allows the sender to react to the initial onset of congestion, without triggering packet drops, and thus send at a data rate that keeps the congestion level (i.e. queuing delay) to an absolute minimum.

4. Low Latency Service Configuration

Low Latency DOCSIS service requires configuration by the operator in order for the feature to be enabled. Some of the service parameters require direct configuration by the operator, and others are specified to have useful default values, and may only require modification in certain exceptional conditions. This section describes the service parameters (configuration file TLVs) and the effect that those parameters have on the low latency service offering.

4.1. Service Rate Configuration

In traditional (single upstream/downstream SF) DOCSIS configurations, the tier of service is defined via three Service Flow rate shaping parameters (TLVs) in each direction: Maximum Sustained Traffic Rate (TLV [24/25].8), Maximum Traffic Burst (TLV [24/25].9), and Peak Traffic Rate (TLV [24/25].27). With an LLD configuration, the same three rate shaping parameters are used, however they are configured on the ASF (TLV [70/71]) rather than the Service Flows underneath the ASF. The CMTS scheduler will rate shape the aggregate of traffic in both the LL SF and the Classic SF to meet the limits set by those TLVs.

The individual Service Flow parameters for rate shaping are not expected to be configured in an LLD configuration, and equipment may or may not support configurations in which those TLVs are set.

4.2. Low Latency Classifiers

4.2.1. Packet Classifiers in LLD ASFs

Classification of traffic into the two SFs under the LL ASF is configurable by the operator using traditional DOCSIS classifiers. The DOCSIS spec does not allow definition of classifiers that direct traffic to an ASF, but rather only to a SF.

In cases where an operator would have traditionally defined a single SF in each direction, classifiers were not needed since all traffic would by default use the single (primary) SF. When that configuration is updated to LLD, Classifiers will be needed to direct a subset of traffic to the LL SF, with the Classic SF serving as the primary SF.

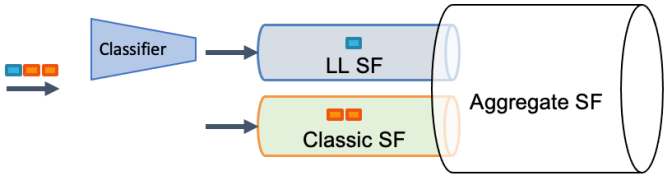


Figure 1 – Classifier Setup for US and DS Service Flows

In cases where an operator would have traditionally defined multiple SFs in each direction to carry traffic for different services (e.g. Community WiFi or voice signaling traffic), when that configuration is updated to support LLD (by replacing one or more of those SFs with ASFs), classifiers need to be defined - with appropriate Classifier Rule Priorities - such that the correct subset of traffic will be directed to the appropriate LL and Classic SFs.

The LLD-capable CMTS provides a feature referred to as "Classifier Merge" that can assist in the creation of these classifiers in certain cases. This feature is described in Section 5.5 of this paper.

4.2.2. Marking of Non-Queue-Building and L4S Traffic

There are two categories of traffic that are compatible with the LL SF: sparse "Non-Queue-Building" traffic (such as VoIP and traditional online games), and capacity-seeking (i.e. high data rate) L4S traffic (such as cloud gaming and video conferencing). These two categories use two different fields in the IP header to identify their traffic. Non-Queue-Building (NQB) traffic uses the 6-bit Diffserv field, and L4S traffic uses the 2-bit Explicit Congestion Notification (ECN) field. These two fields together form an octet that was formerly referred to as the Type of Service (ToS) octet, with the DS field being the upper 6 bits, and the ECN field being the lower 2 bits.

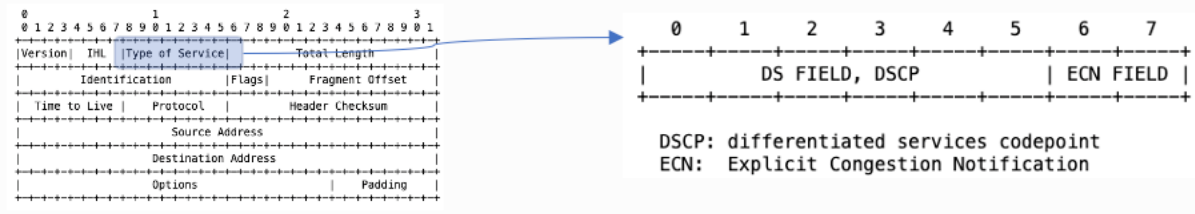


Figure 2 – The differentiated Services and ECN fields in the IP Header

4.2.2.1. NQB and DSCP

The DS field encodes a value between 0 and 63, where each specific value is known as a Diffserv Code Point (DSCP). There are several DSCPs that are used by applications today to mark sparse traffic that is generally compatible with the LL SF. It is expected that operators will wish to classify all such traffic into the upstream LL SF.

While multiple DSCPs are in use by applications in the home network, there is no requirement or expectation that an operator will carry these DSCPs into their core network. Rather it is expected that operators will select a single DSCP to be used as the NQB codepoint within their core network. Thus the configuration in the upstream is likely to include classification on multiple DSCPs and use of the DOCSIS ToS Overwrite feature to ensure that all LL traffic gets a consistent DSCP in the core network, just as the operator likely will want to configure DOCSIS ToS Overwrite on the Classic SF to ensure that all classic traffic gets a consistent default DSCP (e.g. 0) in the core network.

In terms of DSCPs that are compatible with LL, the Expedited Forwarding (EF) DSCP (46) is defined by the IETF to denote voice traffic, and is used by default by some VoIP applications, as well as for the audio stream of some video conferencing applications. Further, Windows 10 allows applications to mark "AudioVideo" or "ExcellentEffort" traffic as CS5 (40). Several multiplayer online games utilize this value for their game state update packets. Windows 10 also allows "Voice" or "Control" traffic to be marked as CS7 (56). (Note: other operating systems do not limit the choices of DSCP that applications can choose). The IETF is in the process of finalizing the assignment of an "NQB" DSCP (45) specifically for applications to denote Non-Queue-Building behavior (see [IETF NQB]). Additionally, the IETF recommends that network operators re-mark upstream NQB packets to the DSCP value 5 prior to an interconnection with another network, and that they re-mark downstream NQB packets to DSCP 45 prior to a user's home network. The rationale for this recommendation can be found in [IETF NQB].

As a result, it is recommended that upstream traffic marked by the source as CS7(56), EF(46), NQB(45), or CS5(40) be classified into the LL SF. It is also recommended that the LL SF be configured (via DOCSIS ToS Overwrite) to change the LL DSCP to the value 5, and the Classic SF be configured to change the DSCP on classic traffic to the value 0. The four recommended DSCPs for upstream classification (56,46,45,40) can be selected using two DOCSIS ToS Range and Mask classifiers, one that matches values 45 & 46 (0xB4 B8 FC), and one that matches 40 & 56 (0x28 28 2F). These classifier encodings are shown in the example in Section 4.2.3.

For downstream traffic, it is recommended that operators classify the NQB (5) DSCP into the LL SF, and use the DOCSIS ToS Overwrite feature to change it to 45. Note that the value 5 above is only a recommendation, an operator could choose a different value for use in their core network (even the value 45 if that works well). If a different value is chosen, it would be necessary to either re-mark NQB to 5 at interconnection, or negotiate with the interconnection partner on the use of the different value for NQB traffic.

4.2.2.2. L4S and ECN

L4S compliant senders mark their packets with the value ECT1 (0b01) in the ECN field. Additionally, it is possible that ECN traffic gets re-marked to CE (0b11) by a network element. As a result, it is recommended that operators classify both of these ECN values to the LL SF. These can be selected via a single DOCSIS ToS Range and Mask Classifier (0x01 01 01).

4.2.3. Recommendations for LL classifiers

Table 2 – Low Latency Classifiers

Classifier	Upstream	Downstream
IPv4 Classifier for DSCP 45 & 46	22.9.1 = 0xB4 B8 FC	23.9.1 = 0xB4 B8 FC
IPv4 Classifier for DSCP 40 & 56	22.9.1 = 0x28 28 2F	23.9.1 = 0x28 28 2F
IPv4 Classifier for ECN (ECT1 & CE)	22.9.1 = 0x01 01 01	23.9.1 = 0x01 01 01
IPv6 Classifier for DSCP 45 & 46	22.12.1 = 0xB4 B8 FC	23.12.1 = 0xB4 B8 FC
IPv6 Classifier for DSCP 40 & 56	22.12.1 = 0x28 28 2F	23.12.1 = 0x28 28 2F
IPv6 Classifier for ECN (ECT1 & CE)	22.12.1 = 0x01 01 01	23.12.1 = 0x01 01 01

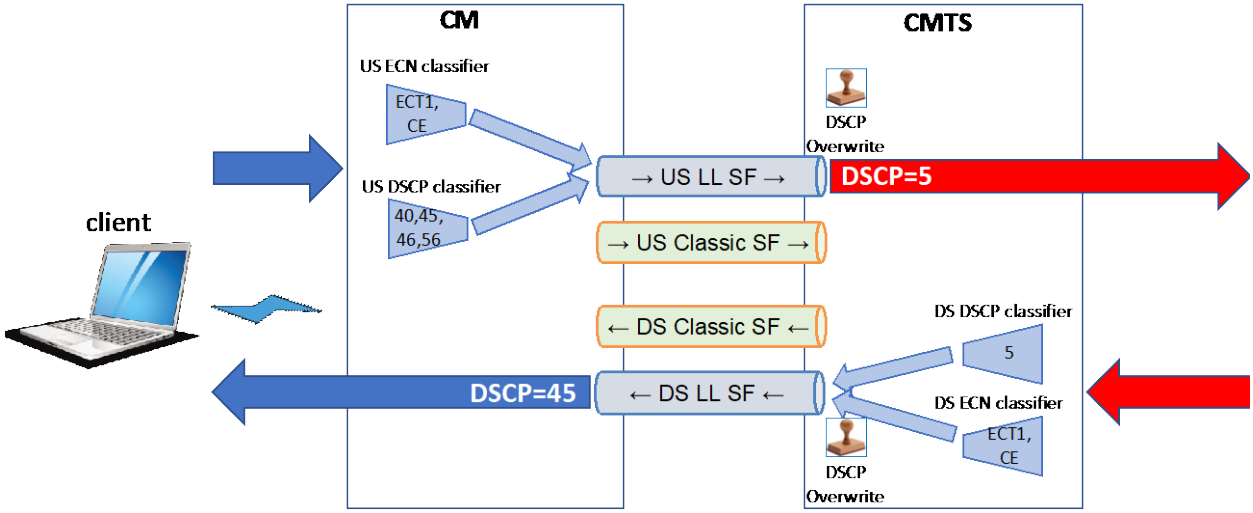


Figure 3 – Classifier Usage on CM & CMTS

4.2.4. DSCP Overwrite

Operators typically wish to manage the DSCP values that are utilized within the core network, and commonly mark all customer Internet traffic with a single DSCP (usually either CS0(0) or CS1(8)) so that it can be identified and handled appropriately in core network routers. This DSCP marking is usually set as packets ingress into the core network, both at interconnections (via router configuration), and at the access network (via the DOCSIS ToS Overwrite feature).

To support low latency end-to-end (including when one or both endpoints are outside the MSO's network) it is recommended that the operator use two DSCP values in their core network, one for default Internet traffic and the other for NQB Internet traffic. An operator can choose any pair of DSCPs that they wish, but as discussed in Section 4.2.2.1, the value 5 is recommended by IETF for use across interconnections. Here we will assume that the operator uses 0 for default Internet traffic and 5 for NQB Internet traffic in their core network.

For upstream SF configuration, it is recommended to use the ToS Overwrite feature to mark all traffic on the LL SF as NQB(5) and to mark all traffic on the Classic SF as CS0(0). Note that this will mark L4S traffic as NQB(5) as well, which is unintended. An alternative, in the case that the operator controls the home gateway software, is to selectively re-mark the four DSCPs 40,45,46,56 to 5 in the home gateway, and then just use DSCP 5 and ECN classifiers for the upstream LL SF in DOCSIS.

Current Wi-Fi gear does not yet support the same low latency features as DOCSIS (i.e. L4S, NQB isolation, Queue Protection). However, low latency traffic can be given a separate queue from classic traffic via the use of the WMM QoS Video Access Category. By default, most consumer Wi-Fi gear will map DSCPs in the range 32-47 to the Video Access Category. For upstream traffic, three of the four DSCP values that are recommended for low latency treatment (40,45,46) already get mapped into the Video Access Category by default, and the fourth DSCP (56) gets mapped into the Voice Access Category. It is recommended that operators mark downstream NQB traffic with the NQB(45) DSCP so that it is mapped to the Video Access Category by default in current Wi-Fi gear, and so that it can get full NQB treatment in future Wi-Fi gear that supports all of the features defined for the NQB PHB. Note that this will mark L4S traffic as NQB(45) as well, which is unintended. An alternative is to use selective re-marking in a router within the MSO core network (possibly even the CMTS router functionality) to change DSCP 5 to 45.

To summarize, the following DOCSIS ToS Overwrite settings are recommended.

Table 3 – DSCP TOS Overwrite Settings

Service Flow	Overwrite Setting	TLV
Upstream LL SF	set DSCP= 5 & don't modify ECN	24.23 = 0x03 14
Upstream Classic SF	set DSCP= 0 & don't modify ECN	24.23 = 0x03 00
Downstream LL SF	set DSCP= 45 & don't modify ECN	25.23 = 0x03 B4
Downstream Classic SF	set DSCP= 0 & don't modify ECN	25.23 = 0x03 00

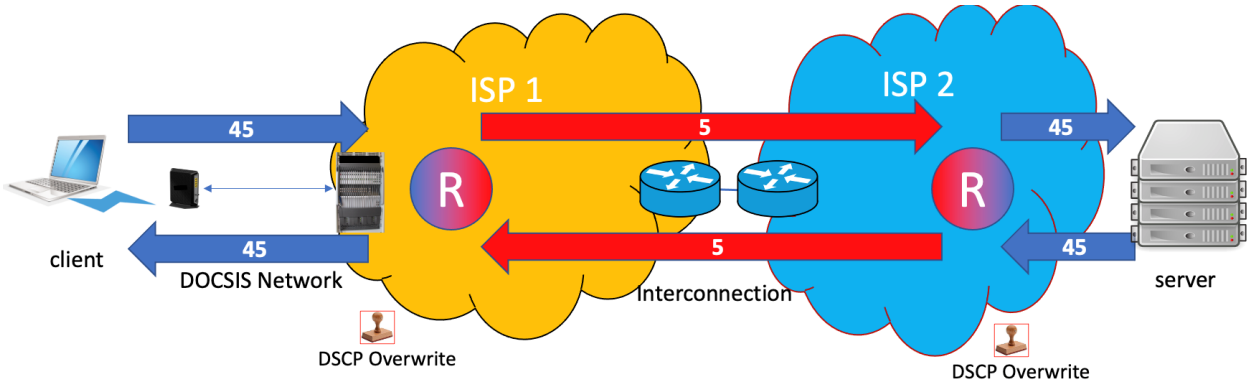


Figure 4 – NQB DSCP changes through the network

In the cases where the DOCSIS ToS Overwrite feature is used, it is noted above that this feature will overwrite the DSCP on L4S traffic as well. The implication of this is that in current Wi-Fi gear, the L4S traffic will be sent in the Video Access Category, which will give it higher priority across the Wi-Fi link than classic traffic by default. This may or may not be desirable. The IETF recommends that WiFi gear

be configured such that the Video Access Category is given equal priority to the Best Effort access category, by adjusting the "EDCA" parameters for AC_VI to match those for AC_BE.

4.3. Queue Protection

LLD supports a Queue Protection (QP) function that monitors traffic classified into the LL SF in order to ensure that those traffic flows are compatible with the low latency queue. The QP function identifies flows that are misbehaving (i.e. causing queue build-up) and redirects the packets of those flows into the Classic SF.

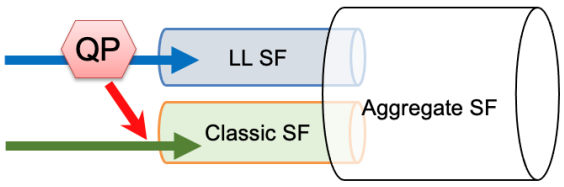


Figure 5 – Queue Protection

4.3.1. Algorithm Details

The three core concepts to understand about the queue protection function are:

1. What is a "microflow" in the eyes of the QP function?
2. What is the "congestion rate" of a microflow?
3. How does the QP decide whether to re-direct (sanction) packets?

The QP function defines a "microflow" as a stream of packets that share a common "flow ID" composed of elements of the packet header. The details of this functionality are described in Annex P.3 of [DOCSISv3.1 MULPI], but it can be summarized by the following. Implementations utilize the "5-tuple" of source IP address, destination IP address, IP Protocol (e.g. udp or tcp), and then the source port and destination port (in the case of TCP/UDP/UDP-Lite/SCTP/DCCP) or the Security Parameters Index (in the case of IPsec). In cases of un-encrypted tunnels (e.g. v4-in-v4, v4-in-v6, v6-in-v4, v6-in-v6, and GRE) implementations will parse into the inner header to find the appropriate fields. For protocols that lack a layer-4 header with meaningful flow identifiers (e.g. ICMP or IGMP), implementations will just use the 3-tuple of source-IP, dest-ip & protocol. Note that an **encrypted tunnel** (e.g. a VPN), obscures the microflows inside the tunnel, and so appears to the QP function as a single microflow.

As each packet enters the QP function, the algorithm calculates the packet's flow ID, updates the "queuing score" for that microflow, and then decides whether to put the packet in the low latency queue or the classic queue.

The queuing score for a microflow originates from a concept referred to as the "congestion rate" of the microflow. As a packet enters the QP function, the algorithm estimates how much queuing delay this packet would experience if it were to be enqueued in the low latency queue. That queue delay estimate is then used to select a "weight" for that packet between 0 and 1, using the ramp function shown in Figure 6 below.

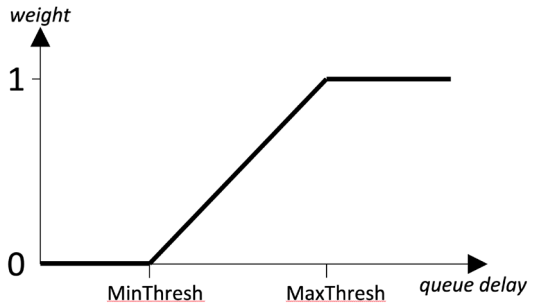


Figure 6 – Queue Protection Congestion Weight Function

This packet's size in bytes, times this weight value, is referred to as the "congestion bytes" for this packet. For example, if the packet's queue delay estimate is greater than MaxThresh, then all of the bytes of this packet are considered congestion bytes. Or, if the queue delay estimate is less than MinThresh, then this packet has zero congestion bytes. Finally, if the queue delay estimate is somewhere in between MinThresh and MaxThresh, then some fraction of the bytes of the packet are considered congestion bytes. The "congestion rate" for a microflow is then a measure of how quickly its congestion bytes are arriving. This is clearly a function of how bursty the microflow is (i.e. how much queue delay is it causing on its own) and any queue delay caused by other microflows.

The QP function mathematically 'filters' the arriving congestion bytes for a microflow through something that resembles a token bucket filter in order to generate the queuing score for that specific packet. In other words, each microflow is allowed to send at a certain congestion rate (default 4 Mbps), and any congestion bytes that arrive in excess of the allowed congestion rate 'queue up' virtually and form the basis for the queuing score. In fact the queuing score is simply the calculated time it will take for any queued-up congestion bytes (including the newly arrived congestion bytes for this packet) to drain out of this virtual queue.

The decision as to whether to re-direct the packet to the classic SF is made based on the current queuing score and the current queue delay estimate taken together. If the product of these two values (both are in units of time) exceeds the product of the two QP parameters: QPLatencyThreshold (default is equal to MaxThresh) and QPQueuingScoreThreshold (default 4 ms), and the current queue delay is greater than the QPLatencyThreshold, the packet will be re-directed. The significance of these two criteria are as follows. The 'product' criteria allows each flow to "burst" above the allowed congestion rate a little bit, but, if the queue delay grows, the amount that a flow is allowed to burst diminishes, thus causing the algorithm to more strictly enforce that allowed congestion rate. The queue delay criteria ensures that a packet arriving to a nearly empty queue won't be re-directed, even if it has a high queuing score.

This all may seem like a lot of state for the algorithm to track for the dozens or even hundreds of microflows that could exist at any point in time. But, the algorithm actually only needs to hold on to state for the microflows that currently have a queuing score. As soon as that queuing score expires (drains out), the algorithm forgets about the microflow.

Of note is the fact that the weight value used to calculate the congestion bytes for a packet is identical to the CE-marking probability used in the Immediate AQM algorithm described in the next section. This was not done simply for convenience or to reduce processing load. The result of using the same ramp function for both purposes is that the rate of CE-marked data arriving at the microflow receiver is a close approximation to the congestion rate calculated internally by the Queue Protection algorithm. As the next section describes, a responsive microflow will monitor the rate of CE-marked data arriving, and

automatically adjust its sending rate in order to (in effect) keep the congestion rate low, thus ensuring that its packets remain in the low latency SF and are not re-directed to the classic SF.

4.3.2. Configuration Parameters

The behavior of the QP function is affected by six configurable parameters:

- MinThresh – the queue delay threshold below which a packet's bytes are not considered to be congestion bytes. This threshold is identical to the MinThresh used for IAQM explicit congestion notification marking. See the IAQM section of this paper for a discussion of how this value is configured and its default value.
- MaxThresh (TLV [24/25].40.3) – the queue delay threshold above which all of a packet's bytes are considered to be congestion bytes. This threshold is identical to the MaxThresh used for IAQM explicit congestion notification marking. See the IAQM section of this paper for a discussion of the default value.
- Queue Protection Enable (TLV [70/71].42.7) – a Boolean value that can be used to disable the Queue Protection function for an ASF. The default is true (enabled).
- Drain Rate Exponent (TLV [70/71].42.10) – the "drain rate" (aka "allowed congestion rate") sets the congestion rate allowance for microflows. Microflows that maintain a congestion rate that is less than this value will never have packets sanctioned. Microflows that exceed this congestion rate can experience sanctioning. This parameter is expressed as a power-of-two exponent, which allows the divide operation to be implemented as a simple bit-shift. The default value is currently 19, which corresponds to 2^{19} Bytes per second, or 4.2 Mbps. The reason for this choice of default value is that, in steady-state operation, L4S flows are expected to aim for a maximum of 2 CE marks every 15ms. Assuming 1500 byte MTU-sized packets, this equates to a maximum congestion rate of 1.6 Mbps. The exponent value 18 (2.1 Mbps) may be sufficient to allow well behaved L4S flows to avoid sanctioning, but the value 19 gives a bit more cushion. As L4S congestion control designs evolve, it may be worthwhile to experiment with the value of this parameter. Additionally, for ASF configurations where the Aggregate Maximum Sustained Traffic Rate (AMSR) value is less than 4.2 Mbps / scheduling_weight (4.66 Mbps using the default scheduling weight – scheduling weight is discussed in Section 4.6), this default value will be too high to trigger any packet sanctioning. In these scenarios, it will be necessary to configure the Drain Rate Exponent to a lower value if Queue Protection functionality is desired.
- QPLatencyThreshold (TLV [70/71].42.8) – the queue delay threshold below which QP sanctioning is suppressed, regardless of the queue score for the microflow. This value is also multiplied with the QPQueueingScoreThreshold to form the threshold for packet sanctioning described above. The default value of QPLatencyThreshold is equal to MaxThresh. Setting this to a lower value may be beneficial to better protect low latency traffic from an unresponsive microflow flooding the LL SF.
- QPQueuingScoreThreshold (TLV [70/71].42.9) – the nominal queuing score threshold for packet sanctioning. This parameter provides a congestion burst allowance for each microflow. Setting a larger value will allow microflows to cause larger bursts of queuing delay without being sanctioned, and setting a lower value will enforce the allowed congestion rate more strictly. As described above, the queue delay of a packet is multiplied by the microflow's queuing score, and this product is compared against the product of QPLatencyThreshold and QPQueuingScoreThreshold. The default value is 4ms.

4.3.3. Implications for applications

So, what does this all mean to an application or microflow?

Since the congestion rate of a microflow is always less than or equal to its total data rate, any microflow that maintains an **instantaneous** data rate (i.e. packet size divided by packet interarrival time) that is less than the Drain Rate will **never** have its packets re-directed to the classic queue. Conversely, any microflow that bursts (or continuously sends) at a rate greater than the Drain Rate, could be subject to having its packets re-directed, depending on how much queue delay it and the other flows sharing the low latency queue are creating.

Any microflow that supports L4S ECN congestion signaling can (and should be), in effect, monitoring its congestion rate and using closed-loop feedback to adjust its sending rate to avoid QP re-direction.

4.4. IAQM & Coupled AQM

The Low Latency SF supports an Active Queue Management function which performs Explicit Congestion Notification marking of packets in accordance with the L4S Architecture. Packets that pass the QP function (and thus are enqueued into the LL SF queue) and are marked by the sender as L4S ECN Capable Transport (ECT1) can potentially be marked with a Congestion Experienced (CE) mark by the LL AQM function. The LL AQM function takes no action on packets that are not marked by the sender as ECT1.

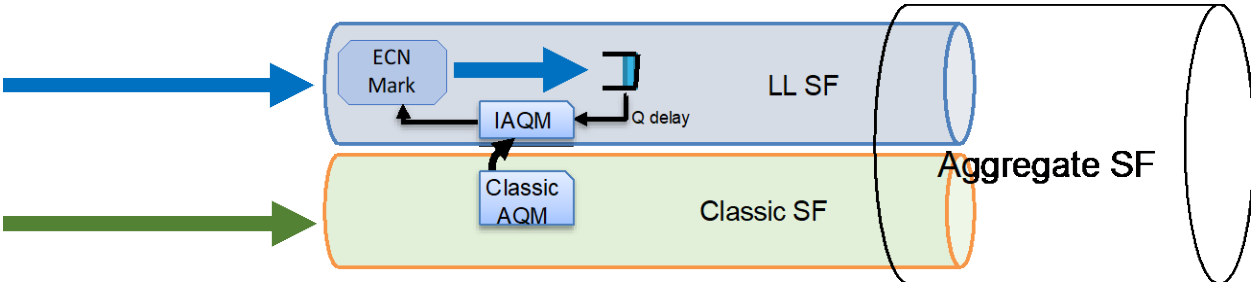


Figure 7 – IAQM in Low Latency Service Flow

4.4.1. Algorithm Details

The decision to CE mark a packet is driven by a coupling between two independent AQM functions, the "Immediate AQM" (IAQM) function and the Classic queue AQM function. The IAQM function uses the estimated queue delay of the packet to calculate a marking probability (probNative), the value of which is identical to the weight value used in QP.

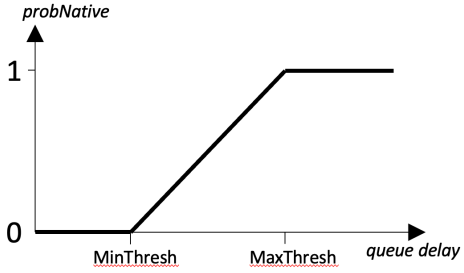


Figure 8 – IAQM Calculation of probNative

The two thresholds, MinThresh and MaxThresh, are actually configured by setting MaxThresh and the "range" of the ramp function (i.e. the difference between MaxThresh and MinThresh), where the range of the ramp function is represented by the base-2 log of the range in nanoseconds. This allows implementations to calculate probNative (and QP weight) using a simple bit-shift instead of a divide.

The Classic AQM function generates a separate probability value (drop_prob) as described in the next section. The LL AQM function calculates a CE-marking probability (probL) for the arrived packet using these two probabilities and a configurable coupling factor.

$$probL = \max\left(\text{probNative}, \min\left(1, \text{coupling\_factor} * \text{sqrt}\left(\text{drop\_prob}\right)\right)\right)$$

The purpose of this calculation, and of the LL AQM algorithm in general, is to send congestion signals to L4S capable transport protocols so that they can modulate their sending rates and bytes-in-flight to maintain both low queuing delay and a fair allocation of the link bandwidth.

When the drop_prob value is zero (i.e. there is very little traffic or queue build-up in the classic queue), probL is simply equal to probNative, and thus the LL AQM function sends congestion signals based on the instantaneous queue delay in the LL queue. By using low thresholds for the probNative ramp function, the AQM enables L4S senders to maintain low queuing delay.

When a queue forms in the Classic SF (e.g. due to classic congestion controlled traffic such as TCP cubic) the Classic AQM function will calculate a drop_prob that is appropriate for classic flows, and this value then dominates the probL equation and drives the CE-marking decisions. The result is that congestion in the Classic SF queue induces higher CE-marking in the LL SF queue, causing the L4S flows to slow down and yield capacity. The goal of this function is to balance the congestion signals for Classic flows and L4S flows such that all flows achieve approximately equal throughput.

4.4.2. Configuration Parameters

The LL AQM function is controlled by four configurable parameters:

- MaxThresh (TLV [24/25].40.4) – the queue delay above which the LL AQM would always trigger a CE-mark. The default is 1 ms.
- Range Exponent of Ramp Function (TLV [24/25].40.5) – the queue delay difference between the MinThresh value (below which the IAQM component of the LL AQM would not trigger a CE-mark) and the MaxThresh value, expressed as the base-2 log of the range in nanoseconds. The default value is 19, which equates to 524 μs.
- Coupling_factor (TLV [70/71].42.5) – the weight used to couple congestion signals from the Classic SF to the LL SF, expressed in units of tenths. Setting this to a higher value will result in classic congestion controlled flows getting greater throughput than L4S flows. Setting this to a lower value will result in L4S flows getting greater throughput than classic flows. Setting the value to 0 disables coupling. The default value is 2.0, which aims to achieve approximate fairness between classic and L4S flows in a range of conditions (see the Scheduling Weight section for further details).
- SF AQM Disable (TLV [24/25].40.1) – when set to True, disables CE-marking of packets by the LL AQM function. The default is false.

When the IAQM MinThresh value derived from the configurable parameters above equates to a value that is less than 4000 bytes divided by the AMSR, implementations will adjust the MinThresh and MaxThresh (leaving the range as configured) to ensure that MinThresh is equal to 4000 bytes divided by AMSR.

When using the default values of MaxThresh and Range Exponent, this occurs whenever the AMSR is less than ~67 Mbps.

4.5. Classic AQM

The Classic SF supports an Active Queue Management (AQM) algorithm that sends congestion signals to flows that utilize the classic SF. For cable modems, the required algorithm is DOCSIS-PIE [RFC8034], for CMTS/CCAP equipment, the choice of Classic AQM algorithm is left to the vendor. In DOCSIS-PIE, packet drops are used as congestion signals, regardless of whether the sender marks its packets as ECN capable.

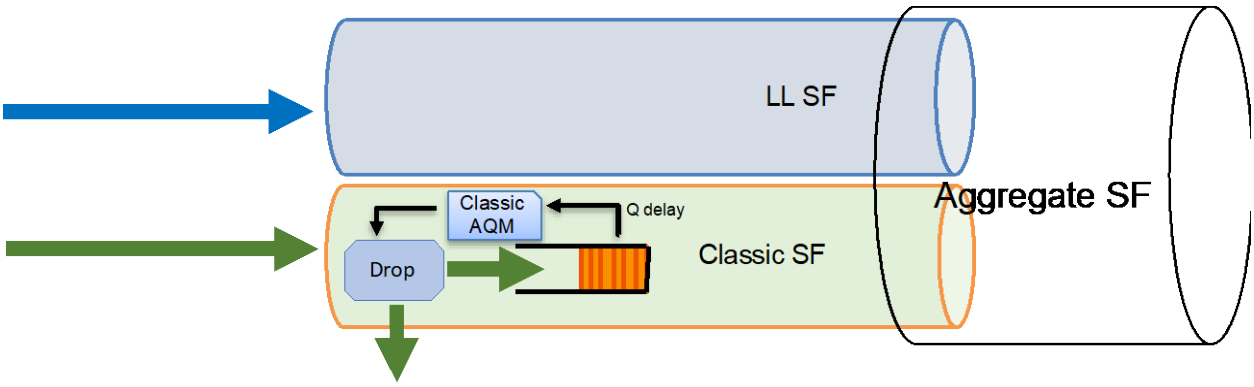


Figure 9 – DOCSIS PIE AQM in Classic Service Flow

The Classic AQM algorithm is controlled by two configurable parameters:

- Classic AQM Latency Target (TLV [24/25].40.2) – the average queue delay that the AQM algorithm is targeting. When the actual queue delay is less than the target, the AQM algorithm will reduce its drop probability, thus allowing classic congestion-controlled senders to increase their sending rate or bytes-in-flight. When the actual queue delay is greater than the target, the AQM algorithm will increase its drop probability, triggering classic senders to reduce their sending rate or bytes-in-flight, thus reducing queue delay. This value should be set to approximately one half of the expected base RTT (i.e. RTT in the absence of queue delay) that the majority of classic flows experience. The default for cable modems is 10ms. The default for CMTS/CCAP equipment is vendor-defined.
- SF AQM Disable (TLV [24/25].40.1) – when set to True, disables the Classic AQM function. The default is false.

4.6. Scheduling Weight

The two constituent SFs in the LL ASF compete for access to the bandwidth provided by the ASF rate shaper implemented in the CMTS. The CMTS scheduler is responsible for enforcing a configurable scheduling weight between the two SFs, both for downstream transmissions and for grants provided in the upstream. The scheduling weight is configured (TLV [70/71].42.6) as a value W between 1 and 255, where the LL SF is given W/256 share, and the Classic SF is given (1-W)/256 share of the bandwidth. The default value for W is 230, which equates to a scheduling weight of ~90% for the LL SF and ~10% for the Classic SF.

At first, it may seem that this parameter sets a hard partition in the capacity available to the two constituent SFs, and that the default value of (90/10) for LL vs Classic would result in LL traffic getting significantly greater throughput than Classic traffic (nine times as much!). Also, it may seem like a 50/50 ratio would be a more fair allocation. But this is not the case. The scheduling weight favors the LL SF, but is counter-balanced by the coupling mechanism described in Section 4.4 to enable per-flow fairness for all of the flows, both Classic and L4S. More detail on the coupled AQM mechanism and weighted scheduling can be found in [IETF dual-queue].

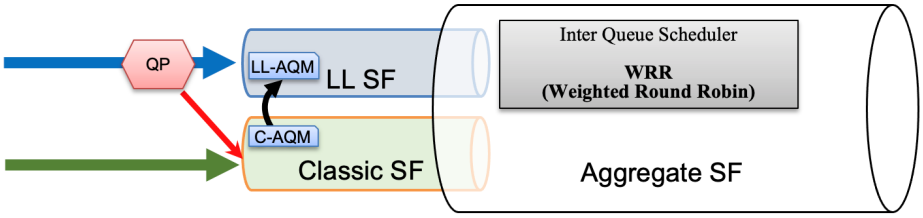


Figure 10 – WRR Scheduler

Without the coupling mechanism (e.g. if the operator disables coupling by setting Coupling Factor to 0), an operator wishing to provide "fair" allocation of bandwidth to flows utilizing the two SFs would need to predict the number of flows of each type that would likely be present, and set the weight accordingly. In essence, they would be forced to decide what fraction of the AMSR should be reserved for the LL SF vs the Classic SF.

So, when there is a mix of traffic utilizing both SFs, the default scheduling weight allows LL SF flows to consume up to 90% of the channel capacity, only as long as the Classic SF flows aren't pushing back via the coupled AQM. Thus per-flow fairness is provided as long as the ratio of L4S flows to Classic flows remains below 90/10. If the ratio of flows exceeds this value, then the L4S flows will be forced to share 90% of the capacity, and the much smaller number of Classic flows will share the remaining 10%, thus giving a throughput advantage to the Classic flows. It is expected that the ratio of L4S to Classic flows will grow over time as more operating systems and applications adopt an L4S congestion control algorithm, and thus there may be a point in time when a value for W greater than 230 would become more appropriate.

The above discussion assumes that the traffic sharing the ASF is responsive (i.e. it responds to congestion signals). However, one might wonder what would happen if unresponsive traffic were sent into one or the other of the SFs. If the unresponsive traffic is sent only to the Classic SF, it would consume whatever bandwidth that it consumes, and the remaining responsive flows (both the Classic and LL) would share the remaining bandwidth approximately equally. If the total Classic SF traffic (unresponsive and responsive) exceeds 10% of the AMSR (assuming the default scheduling weight), this would result in the Classic AQM generating a non-zero drop probability, which means that the unresponsive flow would experience a small amount of packet loss. On the other hand, if the unresponsive traffic is sent only to the LL SF, it would similarly consume whatever bandwidth that it consumes, and the remaining responsive flows (both Classic and LL) would share the remaining bandwidth approximately equally to a point. Due to the weighted scheduling, the total LL SF traffic cannot consume more than 90% of the AMSR (assuming the default scheduling weight) whenever there is capacity-seeking traffic in the Classic SF. So, if the unresponsive flow's traffic rate increases, it will eventually first squeeze the responsive LL SF traffic to zero throughput. If the unresponsive flow continues to increase its rate beyond 90% of AMSR, it will begin to have its packets sanctioned to the Classic SF, and only then begin to squeeze the responsive Classic SF traffic to zero throughput. So, in comparison, responsive traffic in the Classic SF

has slightly more protection than LL SF traffic does from an unresponsive flow that occupies the LL SF. If two unresponsive flows are sent, one to the Classic SF and one to the LL SF, and together they are sending at a rate greater than AMSR, the LL SF flow would be able to consume up to 90% of the AMSR, and the Classic SF would be constrained to 10%. If the LL SF flow's rate exceeds 90% of AMSR, it would have its packets sanctioned to the Classic SF, where it would compete with the other unresponsive flow for access to the 10% of AMSR allocated to the Classic SF. Overall, this situation isn't tremendously different from the traditional single-queue situation, where an unresponsive traffic flow can squeeze the responsive traffic down to zero throughput if it wishes.

4.7. Buffer Sizing

The configuration of buffer sizes for both the LL SF and the Classic SF is done via the Buffer Control TLVs. By default, LL SFs have a buffer size that equates to 10 ms at the configured AMSR (i.e. the buffer size in bytes is equal to AMSR in bps, times 10 ms divided by 8 bits/byte), whereas Classic SFs have a default buffer size equal to at least 50ms at the configured AMSR. Note that when both queues are fully utilized, the effective tail drop limit for the LL SF would be 10ms / scheduling_weight (i.e. 11ms), and for the classic SF would be 50ms / (1 – scheduling_weight) (i.e. 500 ms).

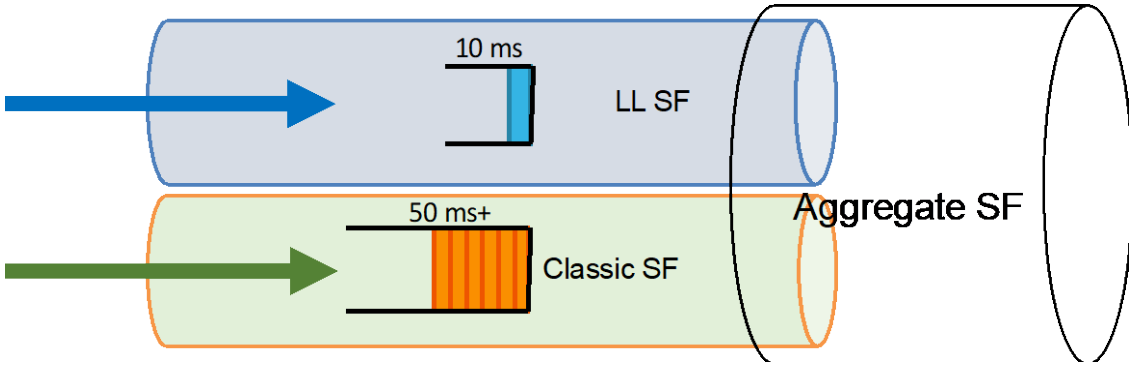


Figure 11 – Buffer sizing

For most deployment conditions, these default buffer sizes are expected to be sufficient. In many cases, the AQM and QP functions will prevent the buffers from reaching this tail drop limit. In particular, the QP function will generally keep the LL SF queue from significantly exceeding the MaxThresh value (i.e. 1 ms in the default configuration). For the Classic SF, the AQM algorithm will adjust drop probability in order to keep the steady-state queue delay at the target value (10ms).

4.8. Proactive Grant Service

The Proactive Grant Service (PGS) is designed to provide low latency for US Service Flows that may carry variable size data packets with random packet arrivals. With PGS a CMTS proactively schedules a stream of grants to the Service Flow. The intention is that the stream of grants is scheduled at a constantly adjusted rate that attempts to match or exceed the instantaneous demand. In doing so, the system ensures that the vast majority of packets carried by the Service Flow can be transmitted without being delayed by the Request-Grant process. If the traffic arrival rate exceeds the rate at which the CMTS is proactively providing grants, the CM will automatically piggyback requests for additional grants.

The bare minimum implementation of a PGS scheduler as described in the specification is the following.

There are three parameters for an upstream PGS Service Flow:

- Guaranteed Grant Interval (GGI) specifies the maximum interval (microseconds) between successive data transmission opportunities.
- Guaranteed Grant Rate (GGR) specifies the minimum granting rate, in bits/sec.
- Guaranteed Request Interval (GRI) specifies the maximum interval (microseconds) between successive request opportunities (unicast and piggyback).

If traffic activity is detected on a PGS Service Flow, the CMTS provides unsolicited data transmission opportunities at a minimum rate of GGR with an interval less than or equal to the GGI. The algorithm for detecting traffic activity on a PGS Service Flow is CMTS vendor specific.

The efficiency of a PGS service is highly dependent on the CMTS with its ability to detect inactivity and reduce the bandwidth usage to only the requests and shutting down any grants. There will always be a bit overhead in bandwidth lost, as an CMTS implementation may not be able to perfectly size the PGS grants to the varied type of traffic seen on the upstream.

Beyond this basic implementation, it is expected that CMTS schedulers will implement more sophisticated algorithms that can adjust the PGS grant rate automatically, in real-time, as demand for capacity fluctuates. Thus, the GGR value becomes a lower limit for the PGS granting, and when activity is present the CMTS grants at or above this level.

It is recommended that all Low Latency SFs be configured with PGS scheduling. For systems where the upstream capacity is constrained, it might be appropriate to set GGR to 0 (or a very low value), and allow the CMTS to dynamically adjust the actual grant rate without restriction. For systems which have sufficient upstream capacity, (and especially if the CMTS scheduler implementation is a basic one) an operator may choose to allocate some amount of bandwidth to the PGS Service Flows. GGR settings of 1-2 Mbps shows a lot of effectiveness in reducing they request-grant delay seen by the small packets that are typical of upstream voice services and online games.

Keeping in mind that the request-grant delay for an LLD-capable system might be in the range of 2-6 ms, setting the GGI value to something toward the upper end of that range may not provide much benefit. As a result, it is recommended that the GGI be set to a value less than 2 ms.

A PGS configuration of GGR = 2 Mbps and GGI = 1 ms will provide a 250 byte grant every millisecond, and thus can forward small packets (less than 250 bytes) with at most 1 ms of media access delay, and slightly larger packets (250 – 500 bytes) with at most 2 ms of media access delay. If a large packet (say 1500 bytes) arrives, the first 250 bytes will be sent in the first grant, along with a piggybacked request for 1250 bytes. It will take a full request-grant delay for the additional data grant to arrive, and in the meantime, multiple proactive grants may have been provided, which the CM uses to send further fragments of the original packet. As a result, the total media access delay for isolated packet arrivals is never greater than the request-grant delay.

The GRI value can be set by the operator to enable unicast polling (i.e. request opportunities that are dedicated to this Service Flow) that continue even during inactivity. This can be used to enable a SF to quickly restart PGS grants after a period of inactivity, even in the presence of heavy service group congestion. A value of ‘0’ for GRI disables this polling feature.

4.9. CMTS MAP interval

LLD lowers the request-grant delay by requiring support for a shorter MAP Interval and a shorter MAP processing time. The MAP interval is the amount of upstream time that each MAP message contains grants for and is also the time interval between consecutive MAP messages. Reducing the MAP interval means that the CMTS processes incoming requests more frequently, thus shortening the amount of time that a request might wait at the CMTS before being processed. A shorter MAP interval also means that grants are not scheduled as far into the future within each MAP message. Thus, for every millisecond that the MAP interval is reduced, the request-grant delay improves by 2 milliseconds.

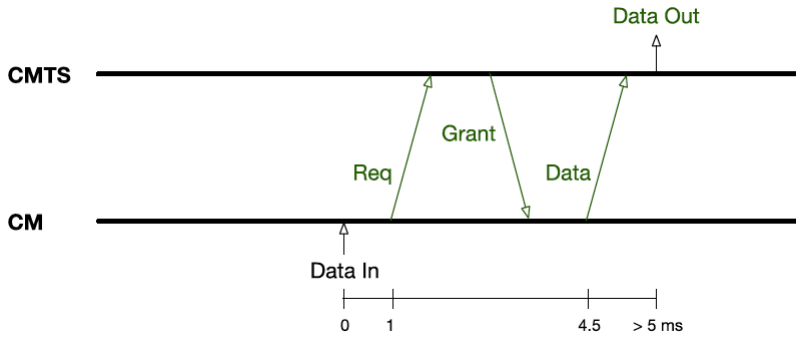


Figure 12 – CMTS Req-Grant Delay

DOCSIS MAP intervals have been in the 2 ms (typically) to 4 ms range for a long time. With LLD, CMTSs support a MAP Interval of 1 millisecond or less. Choosing a lower MAP interval for each upstream channel will reduce the latency experienced by all traffic, and is an important consideration for an operator.

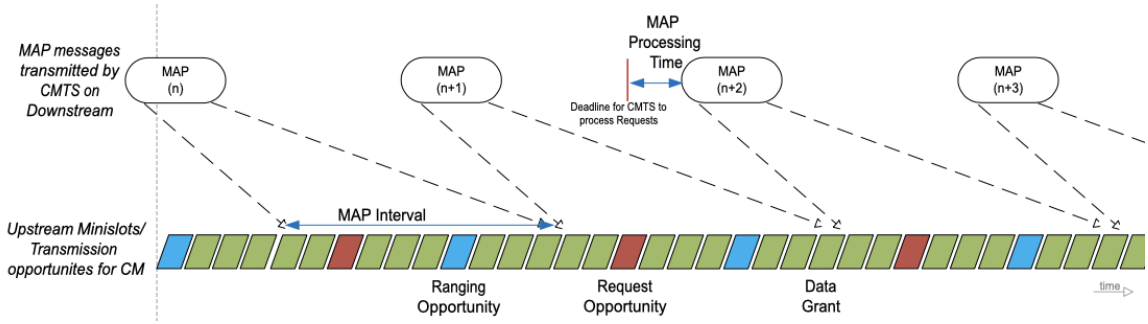


Figure 13 – CMTS MAP Interval

Decreasing the MAP Interval does increase the amount of downstream MAP message traffic, but in many systems the latency benefit outweighs the small impact on available downstream capacity.

5. LLD Configuration Mechansims

Low latency service is configured in the upstream and/or downstream direction by enabling an Aggregate Service Flow with two underlying individual Service Flows, a Low Latency Service Flow (LL SF) and a Classic Service Flow. There are a few different ways an operator can configure Low Latency DOCSIS services on a DOCSIS 3.1 or 4.0 CM.

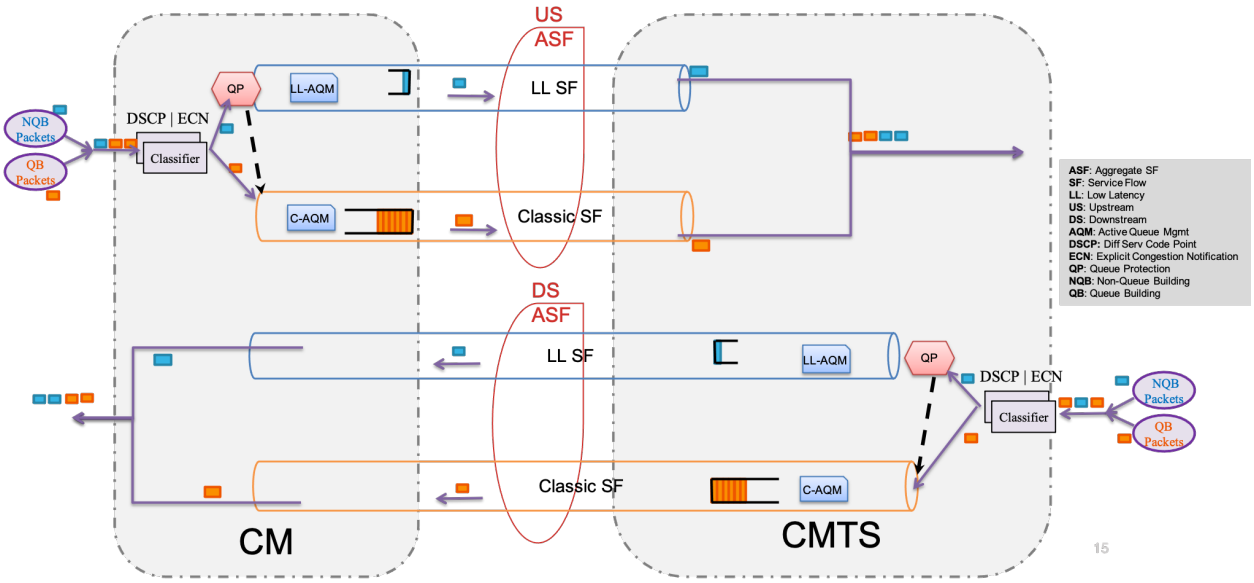


Figure 14 – Low Latency Provisioining : CM Service Flows

Enabling Low Latency services on a DOCSIS CM and CMTS involves the following components:

- Creating an ASF: An Aggregate Service Flow can be created by using explicit configuration file TLVs or by defining a template entry in the Aggregate QoS Profile (AQP) table that is then referenced in the configuration file. A separate ASF is needed for both the upstream and the downstream.
- Creating the individual SFs: Low Latency Service Flows and Classic Service Flows can be configured via explicit config file TLVs or by defining template entries in the Service Class table and then referencing those entries either in the configuration file or in an Aggregate QoS Profile (AQP) table entry. The Low Latency Service Flows and Classic Service Flows can be configured with different Quality of Service parameter sets. For example, in the upstream, the LL SF may use a scheduling type of PGS while the classic may use a scheduling type of Best Effort.

The following section describes the approaches an operator could take to configure these components.

5.1. AQP table with SF/SCN in configuration file

The primary goal with approach is to make use of existing config files and try to minimize changes to those provisioning process. This is possible in cases where the configuration file specifies a Service Class Name (SCN) for each Service Flow that the operator wishes to enable low latency services for. In this case, the current CM config file can be used as-is and the needed Low Latency DOCSIS feature configurations are made on the CMTS via the AQP table, using the Service Class Name from the config file as the AQP Name.

The provisioning system would need no changes and all CMs will boot up with the same config file as they did before the Low Latency DOCSIS features were turned on. For cable modems that support Low Latency DOCSIS, the CMTS will look up the Service Class Name in the AQP table first and expand the single Service Flow from the config file into an Aggregate Service Flow and the component Low Latency Service Flow and Classic Service Flow. For CMs that do not support Low Latency DOCSIS the CMTS will look up the Service Class Name in the Service Class Table and provision the appropriate single Service Flow.

Here is an example of the Service Flow parameters in a config file for the modem. In an actual configuration file, many other parameters will be present, this paper only shows the TLVs relevant to the LLD configuration.

Table 4 – CM Configuration File with TLV24/25 SCN

Config file parameter	TLV/Sub-TLV	Name /Value
Upstream Service Flow Encoding (TLV 24)	(Type 24.1) (Type 24.6) (Type 24.4)	Service Flow Reference = 1 QoS Parameter Set. = 07 Service Class Name = USBronze
Downstream Service Flow Encoding (TLV 25)	(Type 25.1) (Type 25.6) (Type 25.4)	Service Flow Reference= 2 QoS Parameter Set= 07 Service Class Name= DSBronze

Below are examples of an AQP & SCN definition on the CMTS, again we are focusing on the relevant LLD parameters.

Table 5 – AQP Definition on the CMTS

MIB Attributes docsQosAqpTable	US Profile	DS Profile
AQPName	USBronze	DSBronze
Direction	Upstream	Downstream
MaxAggregateTrafficRate	50 (Mbps)	100 (Mbps)
PeakTrafficRate	55 (Mbps)	110 (Mbps)
MaxTrafficBurst	50,000 (Bytes)	100,000 (Bytes)
DataRateUnitSetting	mbps	mbps
LowLatencyAsf	true	true
ClassicSfScn	USClassicSF	DSClassicSF
LatencySfScn	USLLSF	DSLLSF
AqmCouplingFactor	20 (value of 2)	20 (value of 2)
SchedulingWeight	230	230
QpEnable	0x01	0x01
QpLatencyThreshold	1000	1000
QpQueuingScoreThreshold	2000	2000
QpDrainRateExponent	19	19
LowLatencyClassifierList	0x16 0x07 0x09 0x05 0x01 0x03 0xB4 0xB8 0xFC 0x16 0x07 0x0C 0x05 0x01 0x03 0xB4 0xB8 0xFC 0x16 0x07 0x09 0x05 0x01 0x03 0x28 0x28 0x2F 0x16 0x07 0x0C 0x05 0x01 0x03 0x28 0x28 0x2F 0x16 0x07 0x09 0x05 0x01 0x03 0x01 0x01 0x01 0x16 0x07 0x0C 0x05 0x01 0x03 0x01 0x01 0x01	0x17 0x07 0x09 0x05 0x01 0x03 0xB4 0xB8 0xFC 0x17 0x07 0x0C 0x05 0x01 0x03 0xB4 0xB8 0xFC 0x17 0x07 0x09 0x05 0x01 0x03 0x28 0x28 0x2F 0x17 0x07 0x0C 0x05 0x01 0x03 0x28 0x28 0x2F 0x17 0x07 0x09 0x05 0x01 0x03 0x01 0x01 0x01 0x17 0x07 0x0C 0x05 0x01 0x03 0x01 0x01 0x01

Table 6 – SCN Definition on the CMTS

MIB Attributes docsQosScnTable	US SCN1	US SCN1	DS SCN 2	DS SCN 2
Service Class Name	USClassicSF	USLLSF	DSClassicSF	DSLLSF
TrafficPriority	0	0	0	0
MaxTrafficRate	0	0	0	0
MaxTrafficBurst	50,000	50,000	100,000	100,000
MinReservedRate	0	0	0	0
MinReservedPktSize	500	500	500	500
GuaranteedGrantRate	0	0	0	0
GuaranteedGrantInterval		1000		
GuaranteedRequestInterval		1000		
SchedulingType	bestEffort (2)	proactive GrantService (7)	bestEffort (2)	bestEffort (2)
Direction	upstream	upstream	downstream	downstream
AqmDisabled	False	False	False	False
ClassicAqmLatencyTarget	10		10	
AqmAlgorithm	docsisPIE (1)	Immediate Aqm (2)	docsisPIE (1)	Immediate Aqm (2)
ImmedAqmMaxThreshold		1000		1000
ImmedAqmRangeExponent RampFunc		19		19
DataRateUnitSetting	bps (0)	bps (0)	bps (0)	bps (0)

5.2. AQP table with ASF/AQP Name in configuration file

The goal with approach is to make use of the new AQP definitions as defined in the Low Latency DOCSIS technology. The idea here is to align the changes with the new provisioning TLVs. The idea is to use the current CM config files with the new AQP definitions which now will be used only by the CMs which support Low Latency DOCSIS. The operator will also make the needed the Low Latency DOCSIS feature configurations on the CMTS side. An operator will use the new as the profile names for the AQP definition on the CMTS. The config file will point to the same profile name as defined on the CMTS.

The provisioning system would need to now accommodate these new configuration files and CMs with LLD support will boot up with the new config file. This makes things explicit for the operator and in the network. For cable modems that support Low Latency DOCSIS the CMTS will look up the AQP table first and provision the Aggregate Service Flow and the component Low Latency Service Flow and Classic Service Flow. For CMs that do not support Low Latency DOCSIS, would receive a different configuration file and the CMTS will look up the service class name table and provision the appropriate Service Flow.

Here is an example of the ASF parameters in a new config file for the modem that reference an AQP definition on the CMTS.

Table 7 – CM Configuration File with TLV70/71 AQPName

Config file parameter	TLV/Sub-TLV	Name /Value
Upstream Aggregate Service Flow Encoding (TLV 70)	(Type 70.1) (Type 70.4)	Aggregate Service Flow Reference= 1 ASF QoS Profile (AQP) Name= USBronze
Downstream Aggregate Service Flow Encoding (TLV 71)	(Type 71.1) (Type 71.4)	Service Flow Reference= 2 ASF QoS Profile (AQP) Name = DSBronze

The same parameters in the QP and SCN table can be configured as per previous section.

5.3. Explicit TLVs in CM Configuration file

This approach requires no configuration on the CMTS, rather it includes all low latency configuration including the ASF and the constituent SFs directly in the CM config file. These config files will only be usable by CMs that support Low Latency DOCSIS, so the provisioning system will need to ensure that such a config file is only given to CMs that report support for LLD in their CM Capabilities encoding in the DHCP Discover.

Here is an example of a new config file for the LLD CM.

Table 8 – CM Configuration File with TLV24/25 SCN

Config file parameter	TLV/Sub-TLV	Name / Value
Upstream Parameters		
Upstream Aggregate Service Flow Encoding (TLV 70)	(Type 70.1) (Type 70.8) (Type 70.9) (Type 70.42.1)	Aggregate Service Flow Reference= 1 Maximum Sustained Rate= 50,000,000 Maximum Traffic Burst= 50000 Low Latency SF Reference= 3
Upstream Service Flow Encoding (LL SF) (TLV 24)	(Type 24.1) (Type 24.36) (Type 24.6) (Type 24.15) (Type 24.40.1) (Type 24.45) (Type 24.44) (Type 24.46)	Service Flow Reference= 2 Aggregate Service Flow Reference= 1 QoS Parameter Set= 07 Service Flow Scheduling Type= 07 (PGS) SF AQM Disable = 0 (enabled) Guaranteed Grant Rate = 0 Guaranteed Grant Interval = 1000 Guaranteed Request Interval = 1000
Upstream Service Flow Encoding (Classic SF) (TLV 24)	(Type 24.1) (Type 24.36) (Type 24.6) (Type 24.15) (Type 24.40.1)	Service Flow Reference= 3 Aggregate Service Flow Reference= 1 QoS Parameter Set= 07 Service Flow Scheduling Type= 02 SF AQM Disable = 0 (enabled)
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.9.1)	Classifier Reference = 1 Service Flow Reference = 2 IPv4 ToS Range and Mask= tos-low=0xb4, tos-high=0xb8, tos-mask=0xfc

Config file parameter	TLV/Sub-TLV	Name / Value
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.9.1)	Classifier Reference = 2 Service Flow Reference = 2 IPv4 ToS Range and Mask= tos-low=0x28, tos-high=0x28, tos-mask=0x2f
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.9.1)	Classifier Reference = 3 Service Flow Reference = 2 IPv4 ToS Range and Mask= tos-low=0x01, tos-high=0x01, tos-mask=0x01
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.12.1)	Classifier Reference = 4 Service Flow Reference = 2 IPv6 TC Range and Mask= tos-low=0xb4, tos-high=0xb8, tos-mask=0xfc
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.12.1)	Classifier Reference = 5 Service Flow Reference = 2 IPv6 TC Range and Mask= tos-low=0x28, tos-high=0x28, tos-mask=0x2f
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.12.1)	Classifier Reference = 6 Service Flow Reference = 2 IPv6 TC Range and Mask= tos-low=0x01, tos-high=0x01, tos-mask=0x01
Downstream Parameters		
Downstream Aggregate Service Flow Encoding (TLV 71)	(Type 71.1) (Type 71.8) (Type 71.9) (Type 71.42.1)	Aggregate Service Flow Reference= 11 Maximum Sustained Rate= 200000000 Maximum Traffic Burst= 200000 Low Latency SF Reference= 12
Downstream Service Flow Encoding (LL SF) (TLV 25)	(Type 25.1) (Type 25.36) (Type 25.6) (Type 25.40.1)	Service Flow Reference= 12 Aggregate Service Flow Reference= 11 QoS Parameter Set= 07 SF AQM Disable = 0 (enabled)
Downstream Service Flow Encoding (Classic SF) (TLV 25)	(Type 25.1) (Type 25.36) (Type 25.6) (Type 25.40.1)	Service Flow Reference= 13 Aggregate Service Flow Reference= 11 (Type 25.36) QoS Parameter Set= 07 SF AQM Disable = 0 (enabled)
Downstream Pkt Classifier Encoding (TLV 23)	(TLV 23.1) (TLV 23.3) (TLV 23.9.1)	Classifier Reference = 11 Service Flow Reference = 12 IPv4 ToS Range and Mask= tos-low=0xb4, tos-high=0xb8, tos-mask=0xfc

Config file parameter	TLV/Sub-TLV	Name / Value
Downstream Pkt Classifier Encoding (TLV 23)	(TLV 23.1) (TLV 23.3) (TLV 23.9.1)	Classifier Reference = 12 Service Flow Reference = 12 IPv4 ToS Range and Mask= tos-low=0x28, tos-high=0x28, tos-mask=0x2f
Downstream Pkt Classifier Encoding (TLV 23)	(TLV 23.1) (TLV 23.3) (TLV 23.9.1)	Classifier Reference = 13 Service Flow Reference = 12 IPv4 ToS Range and Mask= tos-low=0x01, tos-high=0x01, tos-mask=0x01
Downstream Pkt Classifier Encoding (TLV 23)	(TLV 23.1) (TLV 23.3) (TLV 23.12.1)	Classifier Reference = 14 Service Flow Reference = 12 IPv6 TC Range and Mask= tos-low=0xb4, tos-high=0xb8, tos-mask=0xfc
Downstream Pkt Classifier Encoding (TLV 23)	(TLV 23.1) (TLV 23.3) (TLV 23.12.1)	Classifier Reference = 15 Service Flow Reference = 12 IPv6 TC Range and Mask= tos-low=0x28, tos-high=0x28, tos-mask=0x2f
Downstream Pkt Classifier Encoding (TLV 23)	(TLV 23.1) (TLV 23.3) (TLV 23.12.1)	Classifier Reference = 16 Service Flow Reference = 12 IPv6 TC Range and Mask= tos-low=0x01, tos-high=0x01, tos-mask=0x01

5.4. Parameter Overrides

When using an Aggregate QoS Profile encoding or a Service Class Encoding, it is possible to override the values of individual parameters configured on the CMTS by specifying the override values in the configuration file. When the CMTS finds a match for the SCN or AQP Name in the AQP Table or Service Class Table and there are overriding parameters in the CM Configuration file, the CMTS utilizes the parameter values from the config file in place of the values that were specified in the AQP table or Service Class Table.

In the case that the CMTS is expanding a Service Flow encoding into an ASF and two SFs as described in Section 5.1, the individual Service Flow parameters from the config file are distributed to the ASF and SFs as discussed in Section 7.7.4.2 of [DOCSISv3.1 MULPI].

5.5. Classifier Merge Operation

When using AQP expansion, as described in Section 5.1 and 5.2 above, the AQP Table includes an attribute that allows the operator to configure packet classifiers to select traffic that is directed to the low latency SF. When the associated Classic SF is the Primary SF, these classifiers may be sufficient for directing traffic into the LL SF, and the Classic SF itself would have no classifiers. In other cases, for example when the operator wishes to create two upstream Low Latency ASFs, it is necessary to craft classifiers that can appropriately select traffic for each of the three non-Primary SFs.

The CMTS supports a functionality called Classifier Merge that makes this process automatic and straightforward.

For each configuration file classifier that points to an ASF or points to a Service Flow that is expanded into an ASF via an AQP Name match (like discussed in Section 5.1), the CMTS will apply that classifier to the Classic SF under the ASF, and will merge in the classifier fields from the AQP table to build appropriate classifiers for the Low Latency SF.

As an example, consider the following upstream portion of a configuration file definition, which includes two upstream SFs and a classifier that directs traffic for the WAN subnet 11.12.13.0/24 into the secondary SF.

Table 9 – CM Configuration File with TLV24/25 SCN

Config file parameter	TLV/Sub-TLV	Name /Value
Upstream Service Flow Encoding 1 (TLV 24)	(Type 24.1) (Type 24.6) (Type 24.4)	Service Flow Reference = 1 QoS Parameter Set. = 07 Service Class Name = USBronze
Upstream Service Flow Encoding 2 (TLV 24)	(Type 24.1) (Type 24.6) (Type 24.4)	Service Flow Reference = 21 QoS Parameter Set. = 07 Service Class Name = USBronze
Upstream Pkt Classifier Encoding (TLV 22)	(TLV 22.1) (TLV 22.3) (TLV 22.9.5) (TLV 22.9.6)	Classifier Reference = 1 Service Flow Reference = 21 IPv4 Destination Address = 11.12.13.0 IPv4 Destination Mask = 255.255.255.0

In this configuration file, both SFs are configured with a Service Class Name that matches an AQP entry on the CMTS, so each of the two SFs will be expanded into an ASF and two SFs. Let's refer to these using reference numbers 1, 2, 3 for the ASF, Classic SF & Low Latency SF expanded from Service Flow Reference 1 (the primary SF), and reference numbers 21, 22, 23 for the ASF, Classic SF & Low Latency SF expanded from Service Flow Reference 21 (the secondary SF).

Referring to Table 5 in Section 5.1, the AQP entry for USBronze has six low latency classifiers defined (three IPv4 and three IPv6). The CMTS Classifier merge would result in the following set of ten classifiers:

Table 10 – Classifier Merge Example

Classifier for	Classifier reference	Service Flow Reference	Classifier TLVs
IPv4 DSCP 45 & 46	1	3 "Primary" Low Latency SF	IPv4 ToS Range and Mask= 0xb4, 0xb8, 0xfc
IPv6 DSCP 45 & 46	2	3 "Primary" Low Latency SF	IPv6 ToS Range and Mask= 0xb4, 0xb8, 0xfc
IPv4 DSCP 40 & 56	3	3 "Primary" Low Latency SF	IPv4 ToS Range and Mask= 0x28, 0x28, 0x2f
IPv6 DSCP 40 & 56	4	3	IPv6 ToS Range and Mask= 0x28, 0x28, 0x2f

		"Primary" Low Latency SF	
IPv4 ECN	5	3 "Primary" Low Latency SF	IPv4 ToS Range and Mask= 0x01,0x01, 0x01
IPv6 ECN	6	3 "Primary" Low Latency SF	IPv6 ToS Range and Mask= 0x01,0x01, 0x01
IPv4 DSCP 45 & 46 & WAN subnet 11.12.13.0/24	7	23 "Secondary" Low Latency SF	IPv4 ToS Range and Mask= 0xb4, 0xb8, 0xfc IPv4 Destination Address = 11.12.13.0 IPv4 Destination Mask = 255.255.255.0
IPv4 DSCP 40 & 56 & WAN subnet 11.12.13.0/24	8	23 "Secondary" Low Latency SF	IPv4 ToS Range and Mask= 0x28, 0x28, 0x2f IPv4 Destination Address = 11.12.13.0 IPv4 Destination Mask = 255.255.255.0
IPv4 ECN & WAN subnet 11.12.13.0/24	9	23 "Secondary" Low Latency SF	IPv4 ToS Range and Mask= 0x01, 0x01,0x01 IPv4 Destination Address = 11.12.13.0 IPv4 Destination Mask = 255.255.255.0
IPv4 WAN subnet 11.12.13.0/24	10	22 "Secondary" Classic SF	IPv4 Destination Address = 11.12.13.0 IPv4 Destination Mask = 255.255.255.0

Note that the classifier provided in the config file was an IPv4 classifier, so the CMTS does not merge it with the three IPv6 Low Latency classifiers from the AQP table, since this "merge conflict" would create invalid classifiers.

The classifier merge process is described in Section 7.7.4.3 and Annex Q of [DOCSISv3.1 MULPI].

5.6. Primary Service Flow

Independent of the method of Low Latency DOCSIS provisioning, the CM and CMTS continue to activate the Primary Service Flows at registration time. Low Latency DOCSIS introduced new rules that the CMTS uses to determine which upstream and downstream Service Flows are the primaries based on the contents of the CM configuration file, and it introduced a new Registration Response TLV that the CMTS sends to inform the CM of this selection. The rules that the CM & CMTS use to select the primary upstream and downstream SF are as follows.

When the CM sends its registration request, it is required to send the configuration file TLVs in the order that they appear in the config file. The CMTS then selects the first SF or ASF encoding (either TLV 24/25 or 70/71) for upstream and for downstream, from the registration request. If the first selected Service Flow is an individual Service Flow (TLV 24/25), then this becomes the Primary Service Flow. If the first selected Service Flow is an ASF (TLV 70/71), then the associated Classic SF is chosen to be the primary Service Flow (even if that Classic SF TLV encoding comes much later in the registration request). If the first selected Service Flow is a TLV 24/25 that gets expanded via an AQP expansion, then the associated Classic SF after AQP expansion is chosen to be the primary Service Flow. For a CM indicating LLD support, the CMTS sends the Primary Service Flow Indicator TLV to the CM in the registration response, to identify the primary upstream and downstream Service Flow.

5.7. Device Capabilities

Configuration of Aggregate Service Flows and individual Service Flows for low latency services happens during the Registration process or can be dynamically initiated by the CMTS post-registration.

The CMTS will support at least one upstream and one downstream Low Latency ASF instance per CM. The CMTS optionally can support more than one Low Latency ASF instance in each direction per CM. A CM will support at least two Low Latency ASFs in the upstream direction. This is conveyed by the CM during the registration process, via the CM Capabilities encoding (TLV 5), in the Low Latency Support (sub-TLV 5.76). A value of 0 indicates Low Latency features are not supported and a value of 1 or more indicates the number of ASFs supported by the CM.

5.8. Compatibility Features with CMs Lacking Low Latency Support

A subset of the Low Latency features can be utilized in cases where the CM does not indicate support for Low Latency in its Modem Capabilities encoding. For example, Downstream Low Latency ASFs (and their constituent Service Flows) can be instantiated in order to provide isolation between queue-building and non-queue-building traffic in the downstream direction.

In the upstream direction, it is not possible to configure a Low Latency ASF for a CM that lacks support for Low Latency. However, it is possible to configure separate upstream Service Flows along with classifiers to direct NQB traffic to one of the two Service Flows. Keep in mind, this configuration lacks all of the functionality associated with a Low Latency ASF, i.e. aggregate rate shaping, IAQM with ECN marking, Coupled AQM, default buffer sizing, and Queue Protection, and so should be used with caution.

It is possible to configure PGS scheduling for CMs that don't support LLD. PGS scheduling is only enforced by the CMTS. For CMs that do not indicate Low Latency Support in CM capabilities, or for which Low Latency is disabled (TLV 91), the CMTS replaces the PGS scheduling type with the BE scheduling type and removes the GGI, GGR and GRI parameters in Service Flow definitions in the Registration Response that it sends to the CM. Regardless of whether the CMTS communicates the PGS configuration to the CM, the CMTS is expected to enforce the GGI, GGR and GRI parameters as configured for the Service Flow. This enables configuration of proactive scheduling on CMs that do not indicate Low Latency Support in CM capabilities. Note that the CM will report the SF as having BE scheduling type.

5.9. IPDR Ramifications

Some operators utilize Internet Protocol Detail Records (IPDR) to track per-customer data utilization stats (byte counts) for usage based billing or monthly byte cap accounting. With a Low Latency configuration, each customer will have two SFs in each direction instead of one. As a result, operators will need to consider what updates are needed in their IPDR collection and data processing implementations in order to identify both SFs and sum their byte counts.

In some cases, IPDR post processing functions key off of the Service Class Name to identify the broadband service for a customer. If the operator is using the explicit configuration file approach described in Section 5.3 above, they can include the same Service Class Name for both the LL and Classic SF (so, effectively, all of the explicit config file parameters are overrides to a basic Service Class configuration), and it may then be that no changes are needed in the IPDR post processing.

Alternatively, if the operator is using one of the AQP expansion options described in Sections 5.1 & 5.2, they could still use the same SCN for both LL & Classic, as long as they are ok with BE scheduling on the LL SF (as opposed to PGS), and default values for all of the buffering and AQM parameters.

Another approach could be to use a special character in the SCN to delimit between the “service” name and the Service Flow role, e.g. a colon as in “MyServiceClassUpstream:L” and “MyServiceClassUpstream:C”, and update the IPDR post processing function to look for a match up to the first instance of the delimiter.

6. Performance Monitoring

6.1. Service Flow Statistics

Low Latency DOCSIS technology also has added support for reporting statistics on Low Latency and Classic Service Flows. An operator may want to keep an eye on the IAQM Marking Rate, the number of packets undergoing sanctioning due to queue protection, AQM drop and tail drop statistics for the Service Flows when an operator has deployed LLD technology. These statistics are summarized in the table below and it would be good operational practice for an operator to track these statistics.

Table 11 – Service Flow statistics

Device	MIB Name	MIB Entries
CM & CMTS	DocsQosSfCongestionStatsEntry	docsQosSfCongestionSanctionedPkts docsQosSfCongestionTotalEct0Pkts docsQosSfCongestionTotalEct1Pkts docsQosSfCongestionCeMarkedEct1Pkts
CM & CMTS	DocsQosServiceFlowStatsEntry	docsQosServiceFlowPkts docsQosServiceFlowOctets docsQosServiceFlowTimeCreated docsQosServiceFlowTimeActive docsQosServiceFlowPHSUnknowns docsQosServiceFlowPolicedDropPkts docsQosServiceFlowPolicedDelayPkts docsQosServiceFlowAqmDroppedPkts

An operator using these statistics could be able to understand how Queue Protection is effectively working on the modems (upstream) or CMTS (downstream), and to help debug any issues that come up. Also once L4S traffic becomes more prevalent on the internet, the TotalEct1Pkts counter on the upstream/downstream will give the operator views into how much of the traffic is L4S compliant and the CeMarkedEct1Pkts counter will give the operators confidence that packets are being marked in the DOCSIS network and the network is supporting the L4S functionality appropriately.

6.2. Latency Histograms

Downstream Service Flows and Upstream Service Flows configured for BE or PGS scheduling support Active Queue Management (AQM) algorithms. As part of their operation, these AQMs generate estimates of the queuing latency for the Service Flow. The Latency Histogram Calculation function exposes these estimates to the operator in order to provide information that can be utilized to characterize network performance, optimize configurations, or troubleshoot problems in the field.

The ‘Latency Histogram Encodings’ parameter, when present, enables latency histogram calculation for the given Service Flow. The latency estimates from the AQM are represented in the form of a histogram as well as a maximum latency value. The operator configures the bins of the histogram, and the CM or the CMTS logs the number of packets with recorded latencies into each of the bins. The CM implements

histograms for upstream Service Flows, and the CMTS implements histograms for downstream Service Flows. While the latency histogram calculation function utilizes the latency estimation algorithm from AQM, the latency histogram calculation function can be enabled even for Service Flows for which the AQM algorithm is disabled.

6.2.1. Enabling Latency Histogram via CM Config file

The histograms can be enabled in the CM config file. The table below shows an example of a portion of a configuration file with the explicit histogram TLVs included within the SF definitions.

Table 12 – CM Configuration File Portion with Histograms Enabled

Config file parameter	TLV/Sub-TLV	Name / Value
Upstream Parameters		
Upstream Aggregate Service Flow Encoding (TLV 70)	(Type 70.1) (Type 70.8) (Type 70.9) (Type 70.42.1)	Aggregate Service Flow Reference= 1 Maximum Sustained Rate= 50000000 Maximum Traffic Burst= 50000 Low Latency SF Reference= 3
Upstream Service Flow Encoding (LL SF) (TLV 24)	(Type 24.1) (Type 24.36) (Type 24.6) (Type 24.15) (Type 24.45) (Type 24.44) (Type 24.46) (Type 24.40.1) **(Type 24.40.6)**	Service Flow Reference= 2 Aggregate Service Flow Reference= 1 QoS Parameter Set= 07 Service Flow Scheduling Type= 07 Guaranteed Grant Rate = 0 Guaranteed Grant Interval = 1000 Guaranteed Request Interval = 1000 SF AQM Disable = 0 (enabled) **Latency Histogram Encoding** (in unit of 0.01ms) = 12,20,29,37,45,53,61,69, 100,200,400,800, 1000, 1200, 1500 (Enables 16 histogram bins from 0.12ms – 15ms)
Upstream Service Flow Encoding (Classic SF) (TLV 24)	(Type 24.1) (Type 24.36) (Type 24.6) (Type 24.15) (Type 24.40.1) **(Type 24.40.6)**	Service Flow Reference= 3 Aggregate Service Flow Reference= 1 QoS Parameter Set= 07 Service Flow Scheduling Type= 02 SF AQM Disable = 0 (enabled) **Latency Histogram Encoding** (in unit of 0.01ms) = 100, 300, 500, 600, 700, 800, 900, 1200, 1500, 1800, 2500, 5000, 7500, 10000, 12500 (Enables 16 histogram bins from 1 – 125ms,)
Downstream Parameters		
Downstream Aggregate Service Flow Encoding (TLV 71)	(Type 71.1) (Type 71.8) (Type 71.9) (Type 71.42.1)	Aggregate Service Flow Reference= 11 Maximum Sustained Rate= 200000000 Maximum Traffic Burst= 200000 Low Latency SF Reference= 12

Config file parameter	TLV/Sub-TLV	Name / Value
Downstream Service Flow Encoding (LL SF) (TLV 25)	(Type 25.1) (Type 25.36) (Type 25.6) (Type 25.40.1) **(Type 25.40.6)**	Service Flow Reference= 12 Aggregate Service Flow Reference= 11 QoS Parameter Set= 07 SF AQM Disable = 0 (enabled) **Latency Histogram Encoding** (in unit of 0.01ms) = 50, 75, 100, 125, 150, 175, 200, 300, 400, 500, 600, 800, 900,1200, 1500 (Enables 16 histogram bins from 0.5ms – 15ms)
Downstream Service Flow Encoding (Classic SF) (TLV 25)	(Type 25.1) (Type 25.36) (Type 25.6) (Type 25.40.1) **(Type 25.40.6)**	Service Flow Reference= 13 Aggregate Service Flow Reference= 11 (Type 25.36) QoS Parameter Set= 07 SF AQM Disable = 0 (enabled) **Latency Histogram Encoding** (in unit of 0.01ms) = 100, 300, 500, 600, 700, 800, 900, 1200, 1500, 1800, 2500, 5000, 7500, 10000, 12500 (Enables 16 histogram bins from 1 – 125ms)

6.2.2. Enabling Latency Histogram via SCN

The histograms calculation can also be enabled using the service class name definition on the CMTS. the service class name MIB object (see [DOCS-QOS3-MIB]), can be configured with the appropriate settings. This service class name is then referenced from the config file either within the AQP or the SCN definitions.

Table 13 – SCN Latency Histogram definition on CMTS

CMTS MIB Name/entry	Value
docsQosServiceClassTable -- docsQosServiceClassLatencyHistBinEdges	The attribute is formatted as a string of unsigned 16-bit integers, each representing a histogram upper bin edge in units of 10 microseconds. (This matches the Config file encoding for the same)

6.2.3. Enabling Latency Histogram via SNMP

Once a cable modem is initialized the latency histogram calculations can also be enabled after registration. This is done by setting the histogram bin edges in the docsQosSfLatencyHistCfgTable, for a particular Service Flow on a particular modem. This is useful for testing on the fly where an operator would like to understand the latencies on a particular CM.

Table 14 – Enable Latency Histogram Calculation - Configuration

Device	MIB Name	MIB Entries	Example 1	Example 2
CM	docsQos SfLatency HistCfgTable	DocsQosSfLatencyHistCfgEntry ::= docsQosSfLatencySfLabel	 myUSLLSN	 US2LLSN

Device	MIB Name	MIB Entries	Example 1	Example 2
	(for Upstream Service Flows)	docsQosSfLatencyBin1UpperEdge	12,	10,
		docsQosSfLatencyBin2UpperEdge	20,	50,
		docsQosSfLatencyBin3UpperEdge	29,	100,
		docsQosSfLatencyBin4UpperEdge	37,	300,
		docsQosSfLatencyBin5UpperEdge	45,	600,
		docsQosSfLatencyBin6UpperEdge	53,	800,
		docsQosSfLatencyBin7UpperEdge	61,	1000,
		docsQosSfLatencyBin8UpperEdge	69,	1500,
		docsQosSfLatencyBin9UpperEdge	100,	--
		docsQosSfLatencyBin10UpperEdge	200,	--
		docsQosSfLatencyBin11UpperEdge	400,	--
		docsQosSfLatencyBin12UpperEdge	800,	--
		docsQosSfLatencyBin13UpperEdge	1000,	--
		docsQosSfLatencyBin14UpperEdge	1200,	--
		docsQosSfLatencyBin15UpperEdge	1500,	--
		docsQosSfLatencyBinEdgeNum	15	8

Note: This MIB table includes an SfLabel entry that the operator can set to a text string of their choosing. This string is used to identify the histogram data for a particular Service Flow in the histogram data files that are uploaded by the TFTP method described in Section 6.2.5. In the case that histogram calculation is turned on via the config file, the CM/CMTS will automatically populate the SfLabel field with the Service Class Name for that Service Flow (if one exists).

6.2.4. Querying Histogram Data via SNMP

Once the histogram calculations have been enabled either via the config file or via SNMP the operator can view the histogram of latencies via SNMP in the docsQosSfLatencyStatsTable. See [DOCS-QOS3-MIB] for the detailed definitions.

Table 15 – Latency Histogram Statistics

Device	MIB Name	MIB Entries
CM & CMTS	docsQosSfLatencyStatsTable for US & DS Service Flows	DocsQosSfLatencyStatsEntry ::=
		docsQosSfLatencyMaxLatency
		docsQosSfLatencyNumHistUpdates
		docsQosSfLatencyBin1Pkts
		docsQosSfLatencyBin2Pkts
		docsQosSfLatencyBin3Pkts
		docsQosSfLatencyBin4Pkts
		docsQosSfLatencyBin5Pkts
		docsQosSfLatencyBin6Pkts
		docsQosSfLatencyBin7Pkts
		docsQosSfLatencyBin8Pkts
		docsQosSfLatencyBin9Pkts
		docsQosSfLatencyBin10Pkts
		docsQosSfLatencyBin11Pkts
		docsQosSfLatencyBin12Pkts
		docsQosSfLatencyBin13Pkts
		docsQosSfLatencyBin14Pkts
		docsQosSfLatencyBin15Pkts

Device	MIB Name	MIB Entries
		docsQosSfLatencyBin16Pkts

6.2.5. TFTP Reporting of Histogram Data

The CMTS and CM features and capabilities can be leveraged to enable measurement and reporting of latency estimates through each of the Service Flows. With this information, operations personnel can monitor latency trends and adjust network configurations as appropriate. Latency statistics include histogram counts, maximum latencies, etc., for each enabled Service Flow.

An operator can view the instantaneous statistics via SNMP as described in the previous section or they can have the CM & CMTS store the historical latencies observed and upload that data as a file to an external TFTP server. The following objects show how to enable this style of bulk data uploads on both the CM and the CMTS.

See [CCAP OSSI] and [CM OSSI] for the detailed definitions.

Table 16 – TFTP report of Latency Histogram - Configuration

Device	MIB Name	MIB Entries	Example Values
CM	docsCmLatency RptCfgTable for Upstream Service Flows	docsCmLatencyRptCfgSnapshotDuration docsCmLatencyRptCfgNumSnapshots docsCmLatencyRptCfgNumFiles docsCmLatencyRptCfgMeasStatus docsCmLatencyRptCfgFileName	300 288 7 ready CMUpstreamLLHist
CMTS	docsCmtsLaten cyRptCfgTable for Downstream Service Flows	docsCmtsLatencyRptCfgCmMac docsCmtsLatencyRptCfgSnapshotDuration docsCmtsLatencyRptCfgNumSnapshots docsCmtsLatencyRptCfgNumFiles docsCmtsLatencyRptCfgMeasStatus docsCmtsLatencyRptCfgFileName	CM-MAC-Addr-xxx 600 144 7 ready CMDownstreamLLHist

The operator can configure the duration of a snapshot, the measurement duration of Service Flow latency estimates. One row of statistics per Service Flow is captured during this snapshot interval and stored in the file. The number of Snapshots thus batched into a file can also be configured. E.g. If the SnapshotDuration is set to 300 seconds (5 mins), and NumSnapshots is set to 288, when enabled, the CM returns the latency data samples recorded once per 5-minute interval over the next 24-hour period. The NumFiles attribute controls the number of such files the CM uploads (0 disables, 255 enables unlimited number of files and 1-254 are other valid values). A FileName prefix can call be configured to customize the name of the files.

Statistics files can be enabled for TFTP /bulk file upload via the bulk data transfer mechanisms defined in [CCAP OSSI] and [CM OSSI].

Table 17 – TFTP /bulk file upload Configuration

Device	MIB Name	MIB entry
CM	docsPnmBulkData docsPnmBulkDestIpAddrType docsPnmBulkDestIpAddr docsPnmBulkDestPath	

Device	MIB Name	MIB entry
	docsPnmBulkUploadControl	

CMTS	docsPnmCcapBulkDataControlTable
	DocsPnmCcapBulkDataControlEntry ::=
	docsPnmCcapBulkDataControlServerIndex
	docsPnmCcapBulkDataControlDestIpAddrType
	docsPnmCcapBulkDataControlDestIpAddr
	docsPnmCcapBulkDataControlDestPath
	docsPnmCcapBulkDataControlUploadControl
	docsPnmCcapBulkDataControlPnmTestSelector
		docsPnmCcapBulkDataControlPnmTestSelector {
		other(0),
		dsOfdmSymbolCapture (1),
		dsOfdmNoisePowerRatio(2),
		cmtsUsOfdmaActiveAndQuietProbe(3),
		usImpulseNoise(4),
		usOfdmaRxMerPerSubcarrier(5),
		upstreamHistogram(6),
		usOfdmaRxPower(7),
		usTriggeredSpectrumCapture(8),
		latencyRpt(9)
		}

6.2.6. Using the Reported Histogram Data

When commanded by the operator to upload histogram data the CM or the CMTS will upload a file to a TFTP server with the latency histogram metrics captured as per the snapshot configuration in the previous section.

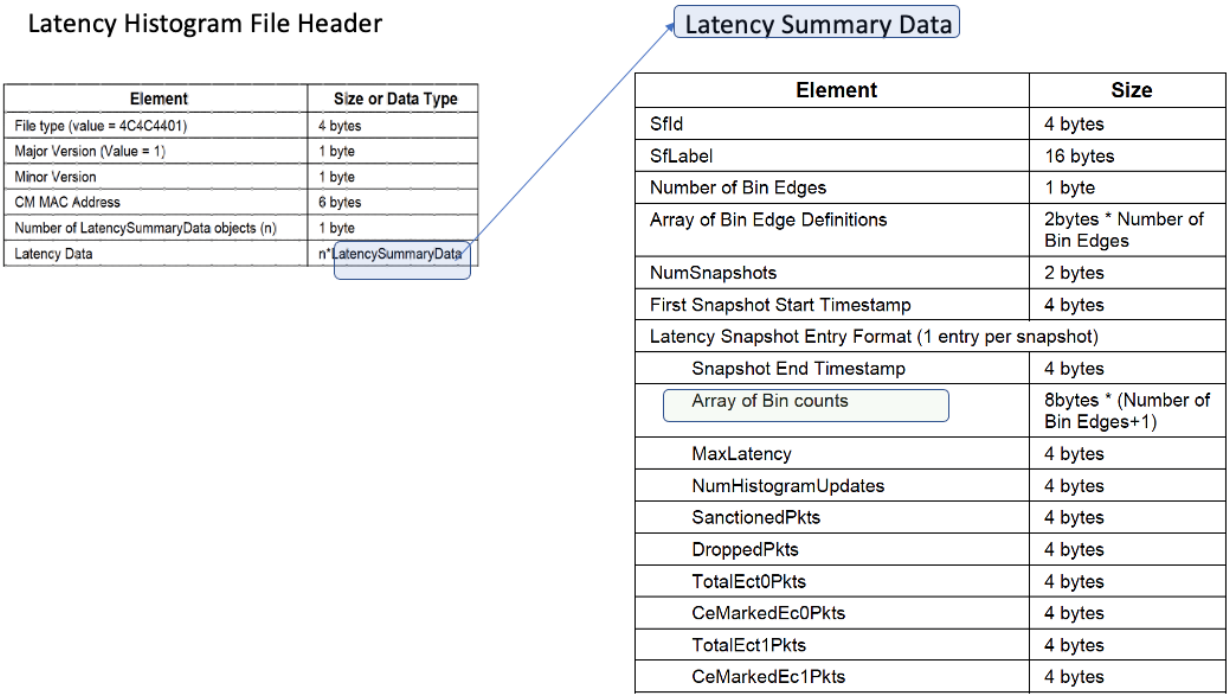


Figure 15 – Latency Histogram File format

Each latency histogram file from a CM or CMTS will have a standard file header as shown in the figure and will have a number of latency summary data entries. Each latency summary data entry will identify the Service Flow for which the latency histogram is being reported, the bin edge definitions and the number of snapshots with their timestamps, and for each snapshot will provide the count of packets in each of the bins and also some additional statistics such as the maximum latency, the number of

sanctioned packets within that snapshot, etc. CableLabs has developed a tool to decode these LLD Histogram files and this is available at the CableLabs [C3 Repository].

The histogram bin counts and bin edge definitions can be used to visualize the latencies seen by the Service Flow as shown below. The two graphs below show the latency data for an upstream Aggregate Service Flow which includes a Low Latency Service Flow and a Classic Service Flow. We can see that the majority of the packets in the Low Latency Service Flow have a latency between 1~2 milliseconds. In the Classic Service Flow we can see that the majority of the packets have a latency between 9 ~ 12 milliseconds. Building histograms like this with well-crafted bin edges will help an operator understand the latency performance of each of the CMs/CMTSs. this data can then be aggregated across the network to develop a baseline of latency performance

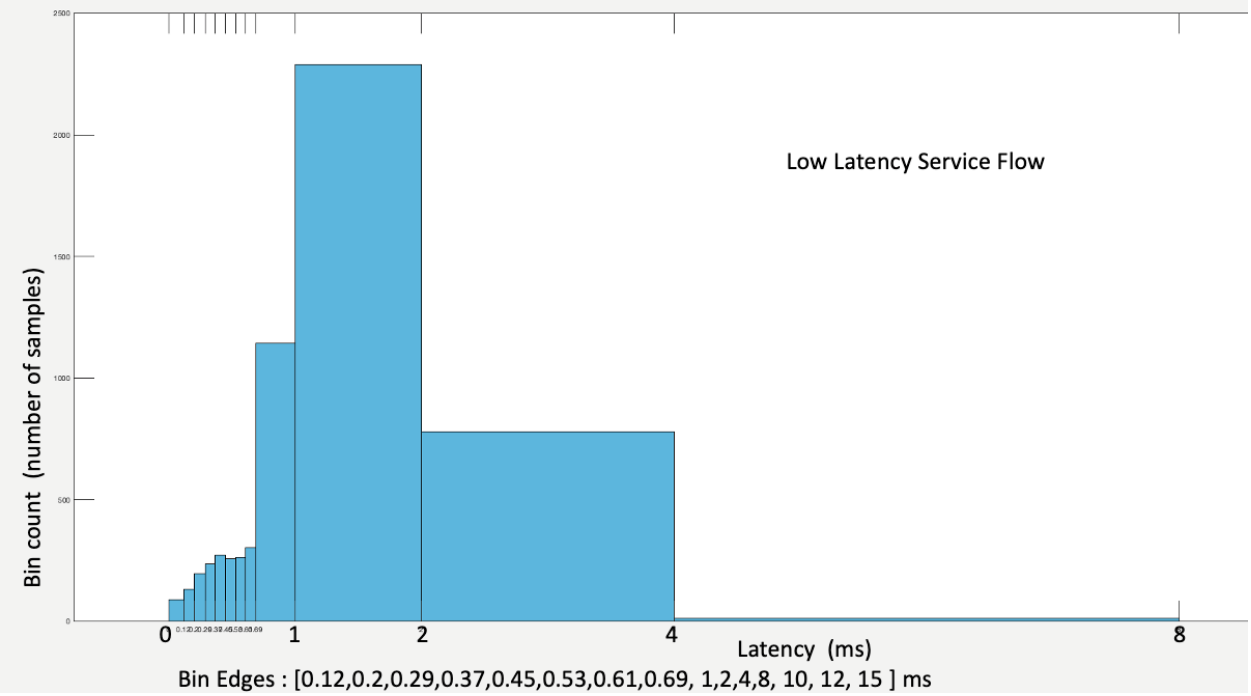


Figure 16 – Histogram plot Upstream Low Latency Service Flow

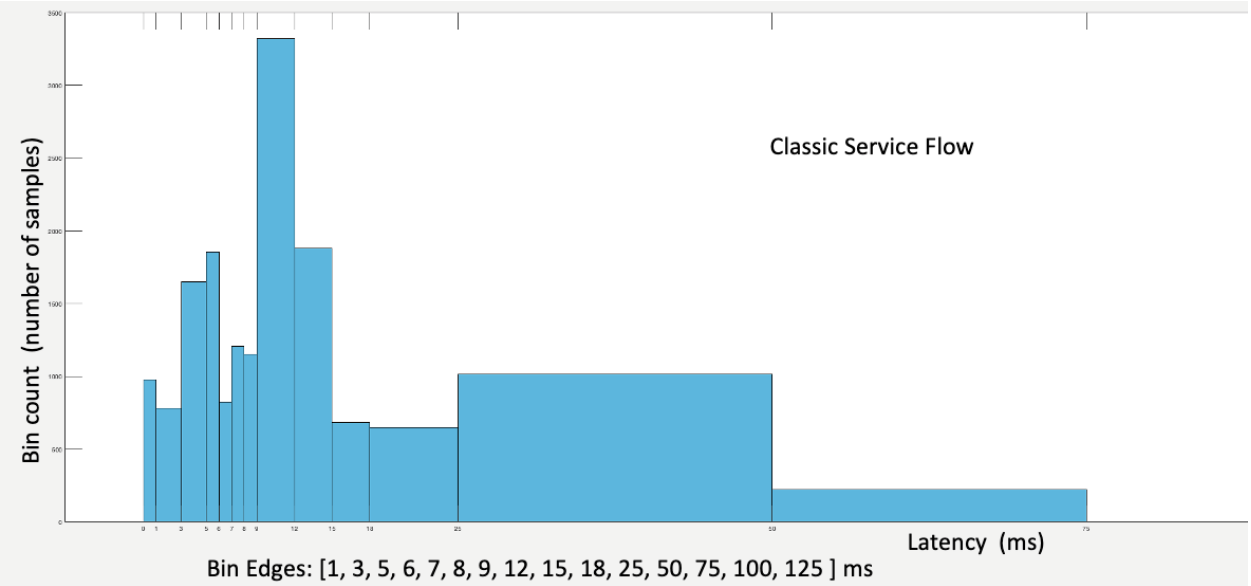


Figure 17 – Histogram plot Upstream Classic Service Flow

7. Conclusion

Low Latency DOCSIS technology (LLD) tackles the two main causes of latency in the network: queuing delay and media acquisition delay. In LLD, data traffic from applications that aren't causing latency can take a different logical path through the DOCSIS network without being stuck behind data from applications that are causing latency, as is the case in today's Internet architectures. In addition, LLD improves the DOCSIS upstream media acquisition delay with a faster request-grant loop and a new proactive scheduling mechanism.

While the LLD parameters can be tweaked to achieve the behavior that an operator wants, the specification already chooses default values for each the parameters based on the combined judgement of the LLD working group.

LLD can be deployed by field-upgrading DOCSIS 3.1 cable modem and cable modem termination system devices with new software. The technology includes tools that enable automatic provisioning of these new services. It allows for multiple ways to provision and enable the low latency services onto the CM/ CMTS. This ranges from methods which minimize the config file changes to methods which minimize the CMTS side configuration, and an operator can choose one of the methods to initiate field trials and ultimately finalize the configuration when deploying LLD across the footprint.

In addition, it also introduces new tools to report statistics of latency performance to the operator, which can be useful to validate configuration and functionality of implementations, as well as to monitor performance over time.

Abbreviations

AC_BE	Access Category - Best Effort
AC_VI	Access Category - Video
aka	also known as
AQM	Active Queue Management
AQP	Aggregate QoS Profile
ASF	Aggregate Service Flow
B	Byte
bps	bits per second
CCAP	Converged Cable Access Platform
CDN	Content Distribution Network
CE	Congestion Experienced
CM	Cable Modem
CMTS	Cable Modem Termination System
CS	Class Selector
DCCP	Datagram Congestion Control Protocol
DHCP	Dynamic Host Configuration Protocol
DOCSIS	Data-Over-Cable Service Interface Specification
DS	Diffserv
DS	Downstream
DSCP	Diffserv Code Point
ECN	Explicit Congestion Notification
ECT	ECN Capable Transport
EDCA	Enhanced Distributed Channel Access
EF	Expedited Forwarding
GGI	Guaranteed Grant Interval
GGR	Guaranteed Grant Rate
GRE	Generic Routing Encapsulation
GRI	Guaranteed Request Interval
IAQM	Immediate Active Queue Management
ICMP	Internet Control Message Protocol
ID	Identifier
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IP	Internet Protocol
IPDR	Internet Protocol Detail Record
L4S	Low-Latency Low-Loss Scalable Throughput
LL	Low Latency
LLD	Low Latency DOCSIS
MAC	Medium Access Control
MAP	Map
max	Maximum
Mbps	Megabits per second
MIB	Management Information Base
min	Minimum
ms	millisecond
MSO	Multiple-System Operator

MSR	Maximum Sustained Traffic Rate
MTU	Maximum Transmission Unit
MULPI	MAC and Upper Layer Protocols Interface
NQB	Non-Queue-Building
OSSI	Operations Support System Interface
PGS	Proactive Grant Service
PHB	Per-Hop Behavior
PIE	Proportional Integral Enhanced
QB	Queue-Building
QoE	Quality of Experience
QoS	Quality of Service
QP	Queue Protection
RFC	Request For Comments
RTT	Round-Trip Time
SCN	Service Class Name
SCTE	Society of Cable Telecommunications Engineers
SCTP	Stream Control Transmission Protocol
SF	Service Flow
SNMP	Simple Network Management Protocol
sqrt	Square root
TCP	Transport Control Protocol
TFTP	Trivial File Transfer Protocol
TLV	type-length-value encoding
ToS	Type of Service
UDP	User Datagram Protocol
US	Upstream
VPN	Virtual Private Network
WRR	Scheduling weight
WAN	Wide Area Network

Bibliography & References

[IETF NQB] *A Non-Queue-Building Per-Hop Behavior (NQB PHB) for Differentiated Services*, Internet Engineering Task Force draft-ietf-tsvwg-nqb-07, Work In Progress, July 2021.

[DOCS-QOS3-MIB] DOCSIS Quality of Service MIB Module, http://mibs.cablelabs.com/MIBs/DOCSIS/DOCS-QOS3-MIB-2021-06-24.txt, CableLabs

[DOCS-PNM-MIB] DOCSIS PNM MIB Module, http://mibs.cablelabs.com/MIBs/DOCSIS/DOCS-PNM-MIB-2021-06-17.txt, CableLabs

[C3 Repository]. CableLabs Common Code Community, C3, https://code.cablelabs.com

[IETF L4S] *Low Latency, Low Loss, Scalable Throughput (L4S) Internet Service: Architecture*, Internet Engineering Task Force, https://datatracker.ietf.org/doc/draft-ietf-tsvwg-l4s-arch/, Work In Progress, July 2021.

[Buffer Control] DOCSIS® Best Practices and Guidelines: Cable Modem Buffer Control CM-GL-Buffer-V01-110915 https://www.cablelabs.com/specifications/cable-modem-buffer-control

[IETF dual-queue] *DualQ Coupled AQMs for Low Latency, Low Loss and Scalable Throughput*, Internet Engineering Task Force, https://datatracker.ietf.org/doc/draft-ietf-tsvwg-aqm-dualq-coupled/, Work In Progress, July 2021.

[RFC8034] Active Queue Management (AQM) Based on Proportional Integral Controller Enhanced (PIE) for Data-Over-Cable Service Interface Specifications (DOCSIS) Cable Modems https://datatracker.ietf.org/doc/html/rfc8034

[DOCSISv3.1 MULPI] DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification https://www.cablelabs.com/specifications/CM-SP-MULPIv3.1

[DOCSIS AQM] Active Queue Management in DOCSIS 3.X Cable Modems, Greg White, Dan Rice, CableLabs, May 2014, https://www-res.cablelabs.com/wp-content/uploads/2019/02/28094021/DOCSIS-AQM_May2014.pdf

[SCTE LLD] Low Latency DOCSIS: Overview And Performance Characteristics, SCTE 2019 https://www.nctatechnicalpapers.com/Paper/2019/2019-low-latency-docsis

[CCAP OSSI] DOCSIS 3.1 CCAP Operations Support System Interface Specification, https://www.cablelabs.com/specifications/CM-SP-CCAP-OSSIv3.1

[CM OSSI] DOCSIS 3.1 CM Operations Support System Interface Specification, http://www.cablelabs.com/specifications/CM-SP-CM-OSSIv3.1

Considerations for Moving Your Access Network to the Cloud (and Back)

A Technical Paper prepared for SCTE by

Michael O'Hanlon
Senior Principal Engineer
Intel Corporation – Network Platforms Group
Dromore House, Shannon, Co. Clare, Ireland
+353 (86) 354 3536
michael.a.ohanlon@intel.com

Eric Heaton
Platform Solutions Architect
Intel Corporation – Network Platforms Group
2200 Mission College Blvd., Santa Clara, CA
1-408-765-3447
eric.d.heaton@intel.com

Randy Levensalor, CableLabs

Brendan Ryan, Intel Corporation – Network Platforms Group

Richard Walsh, Intel Corporation – Network Platforms Group

David Coyle, Intel Corporation – Network Platforms Group

Pavel Belitskiy, Intel Corporation – Network Platforms Group

Subhiksha Ravisundar, Intel Corporation – Network Platforms Group

1. Introduction

The Cable industry continues to innovate on both technical and business fronts in their Access and Edge networks to meet and exceed ever-ascendant customer demands. Currently, Distributed Access Architecture, DOCSIS 4.0, Generic Access Platform, fiber deep, Wireless Access, and virtualization are all being deployed in various permutations in order to deliver on the promises of the 10G era – high bandwidth, low latency, seamless fail-over, single user profiles, and so forth. Looking at the evolution of the Cable Modem Termination System (CMTS) specifically, what was once a fixed, vertically integrated, HFC CCAP appliance from a single vendor has become a virtualized CMTS (vCMTS) appliance running on commercial off the shelf (COTS) servers. This transformation into the software domain has brought the agility needed to keep up with other upgrades to the network.

Along these lines, the vCMTS is now being broken into smaller component functions packaged in software containers that now can be deployed, managed, and scaled individually. In fact, some of those functions may be shared across other network functions or Access technologies to achieve some aspects of convergence and cost savings – for example, a DHCP server or a network telemetry database. The flexibility of such a decomposed, cloud-ready, vCMTS solution brings a lot of opportunities for the MSO in terms of hardware or software consolidation, placement of a workload to the "best" place in the network, individual scaling, and general positive reactivity to the real time needs of users or the operator.

These characteristics are also the promise from Cloud Service Providers (CSPs) when deploying a given application on their infrastructure. That is, they will manage the hardware, furnish a consistent pane of glass for management, and transparently provide redundancy and scale. This certainly sounds appealing! In fact, the Communications Service Provider (CoSP) industry is already familiar with moving enterprise workloads to the Cloud; but will this approach work for vCMTS – and the DOCSIS MAC dataplane in particular?

Figure 1 shows how a vCMTS solution could be deployed to Cloud infrastructure[1] and tie into the general MSO network. This approach moves the vCMTS functionality from equipment in the Headend/Hub in the RPD case or the Node in the RMD case owned by the network operator to servers and switches deeper in the network owned by a CSP. The upside is that the CSP handles the deployment and management of this hardware and provides a common software infrastructure, distributed backup capability, and other operational benefits to the network operator.

[1] Cloud can refer to Cloud Service Provider hardware located in the cloud providers datacenter or within a Telco or MSO providers premises. Section 1.1**Error! Reference source not found.** outlines various possible locations for cloud hardware.

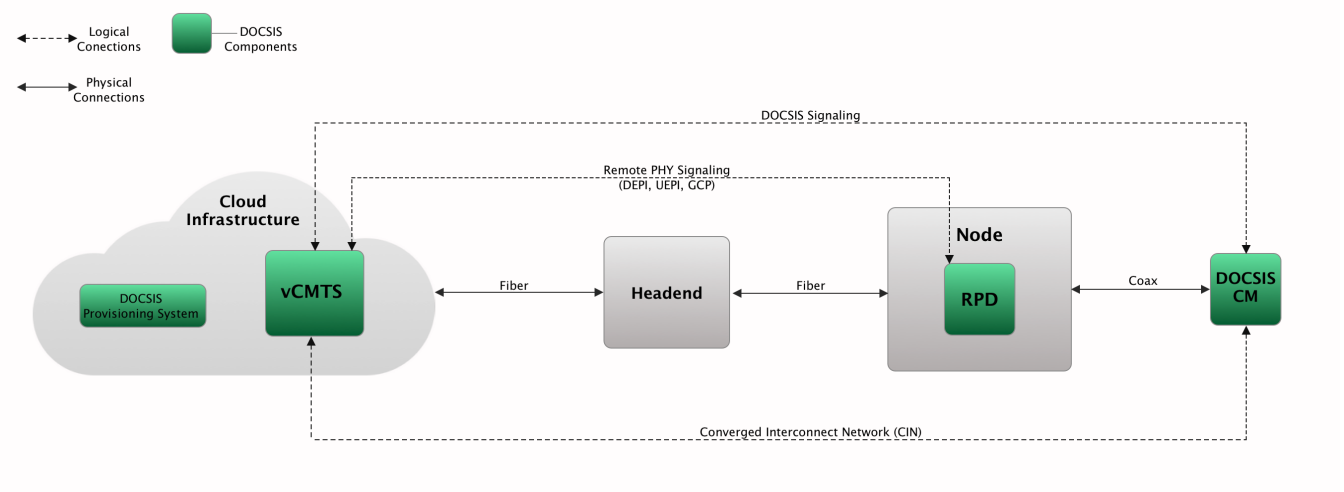


Figure 1 – vCMTS in the Cloud

For all the benefits, there are some trade-offs versus having the vCMTS running on local servers and switches in the Headend/Hub. While there are business and strategic considerations for moving a workload to the cloud – for example, costs, core competencies, and competitive circumstances – this paper focuses on technical questions of expected vCMTS dataplane performance and comparison to the same applications running on on-premise equipment. Just to be clear, the benchmarking and subsequent analysis described in this paper only looks at considerations for the DOCSIS dataplane as opposed to control plane elements of a full vCMTS solution (ex. availability of IEEE 1588), which is, itself, a topic ripe for further discussion.

- Section 1 lays out the test infrastructure and process used to evaluate the current Cloud offerings specifically for networking workloads like a vCMTS dataplane
- Section 2 presents the initial single Service Group benchmarking results across several cloud instance types and makes comparisons to on-premise solutions
- Section 3 describes considerations for effective and performant scaling of the vCMTS dataplane within the Cloud to work-around any per-instance performance limits
- Section 4 summarizes this testing and infrastructure analysis and recommends an “ideal cloud instance type” for the vCMTS dataplane or other similar Access workloads

In short, this paper will describe how a vCMTS dataplane workload can be run optimally in the Cloud, lay out benchmarks comparing various cloud products to on-premise solutions, provide a recommendation for an ideal instance type, and detail the full process from start to finish to allow this work to be generalized across other CSP offerings and workloads in the future.

1.1. Definition of the Cloud

A cloud instance is essentially a virtual machine that is hosted and managed by a CSP and can be rented to users. Each instance is defined by a set of properties, such as the number of virtual CPUs (including those with specific capabilities or instruction sets), the amount of memory or storage, accelerator

options (ex. GPUs, custom ASICs, etc.), and network connectivity. This analysis focuses solely on the offerings from Amazon Web Services (AWS) as being representative of the offerings from other CSPs.

The cloud products from AWS only continue to grow in number and to help customers determine the right match for their needs, AWS groups them into the families shown in Figure 2. Each family, and in many cases, sub-family (not shown), has a general set of workload requirements in mind, such as being optimized for machine learning algorithms or having large storage but low compute requirements. These attributes then translate into the actual instance configuration for compute, hardware accelerators, storage capability, number and size of network interfaces, and so forth.

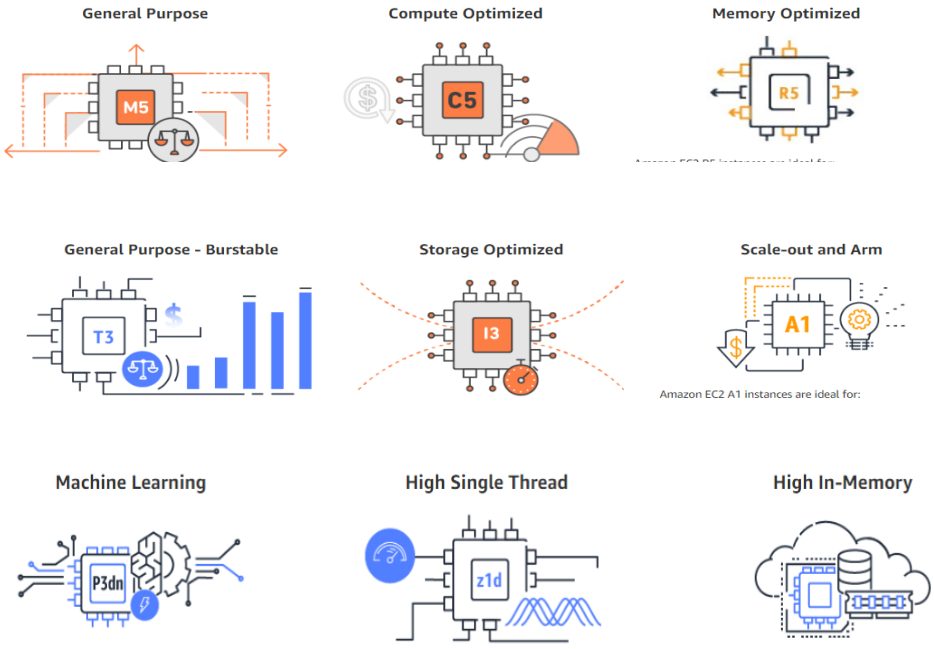


Figure 2 – AWS Instance Families[2]

The first criteria to filter the selection is knowing that vCMTS dataplane reference software used for this analysis was written and optimized for an x86 architecture and thus it will run best on instances based on the newest possible x86 processors with AVX-512 and AES-NI instructions. The second criteria is knowing that an on-premise server running this application can saturate 100Gbps per socket of DOCSIS traffic and thus would like to see instances with very high network connectivity.

Reviewing the options – and performing some baseline testing – the Compute Optimized family is most suitable for NFV generally and the vCMTS dataplane specifically because of its powerful compute capabilities. Within the Compute Optimized family is a sub-family known as "c5n" (the "n" is for networking), which pairs x86 compute from 1st Gen Intel® Xeon® Scalable Processors with up to 100Gbps of network bandwidth per instance. The instance configurations within this sub-family are shown in Figure 3.

[2] https://aws.amazon.com/ec2/instance-types/

Model	vCPU	Memory (GiB)	Instance Storage (GiB)	Network Bandwidth (Gbps)	EBS Bandwidth (Mbps)
c5n.large	2	5.25	EBS-Only	Up to 25	Up to 4,750
c5n.xlarge	4	10.5	EBS-Only	Up to 25	Up to 4,750
c5n.2xlarge	8	21	EBS-Only	Up to 25	Up to 4,750
c5n.4xlarge	16	42	EBS-Only	Up to 25	4,750
c5n.9xlarge	36	96	EBS-Only	50	9,500
c5n.18xlarge	72	192	EBS-Only	100	19,000
c5n.metal	72	192	EBS-Only	100	19,000

Figure 3 – Instance options within the c5n family[3]

All c5n instance types run on a single c5n server type that includes dual 18 core 1st Gen Intel® Xeon® 8124M Scalable Processors running at 3.0GHz (with up to 3.5GHz Turbo operation; Turbo cannot be disabled), 192GB of DDR4 memory, and a 100GBE network interface, as shown in Figure 4. The largest instances, c5n.18xlarge and c5n.metal, consume a complete server and hence are not co-resident with other c5n instances. Smaller instances consume sub-portions of the server resources and may be co-resident with other instances. Figure 5 shows the inverse relationship to the instance vCPU count to the number of such instances that will fit on a given c5n physical server. In other words, divide the total vCPU count of the hardware (72) by the number of vCPUs in an instance to calculate how many of the same instance can theoretically be deployed there.

[3] https://aws.amazon.com/ec2/instance-types/

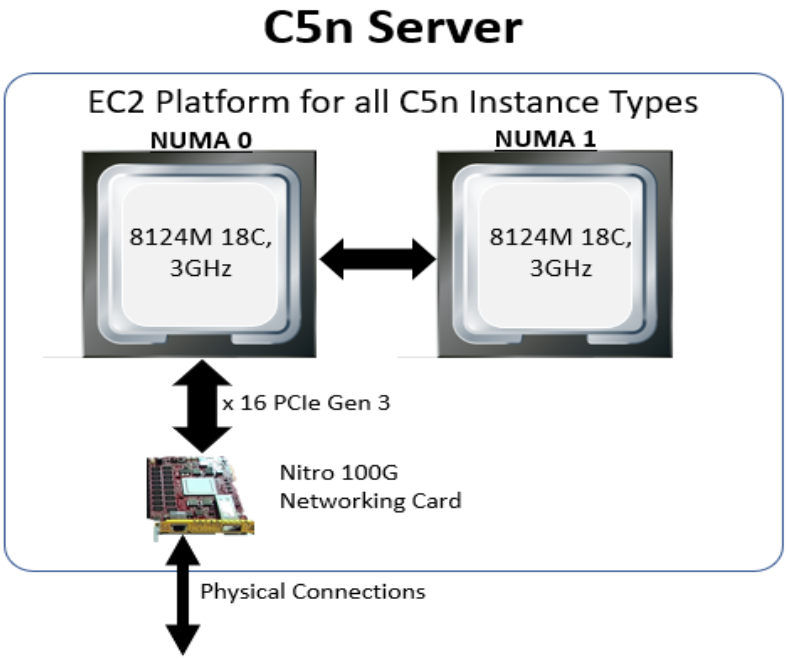


Figure 4 – AWS c5n server definition

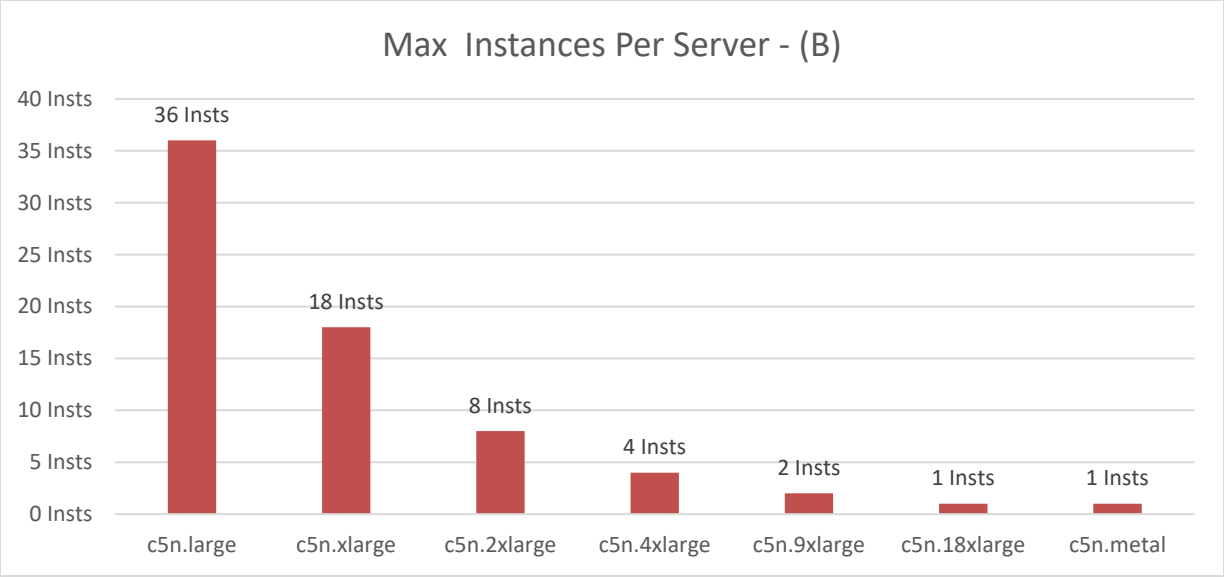


Figure 5 – Max number of instances per c5n server

Figure 3 is helpful to get a high-level understanding of the compute and memory resources per instance, however, what else can be said about the network bandwidth? This is key to set performance expectations correctly and perhaps even anticipate bottlenecks for a dataplane application. For example, is the defined bandwidth shared among other instances, are there limitations on packets per

second, are there any common packet offload features, how many ports/interfaces can be attached to an instance?

The network interfaces for the c5n family are provided through their own custom AWS Nitro System, which is a "combination of dedicated hardware and lightweight hypervisor[4]". This system includes a set of I/O acceleration cards for Virtual Private Cloud (VPC) and various storage options, a controller card that establishes a root of trust and interfaces to the control plane, and a few other elements designed for secure, accessible, and performant access to all cloud resources.

The main I/O card for network connectivity is called the Elastic Network Adapter (ENA), as shown in Figure 6. The ENA is a custom 100Gbps NIC supporting up to 15 physical interfaces (similar to VFs) per instance and up to 32 packet queues[5] per interface.

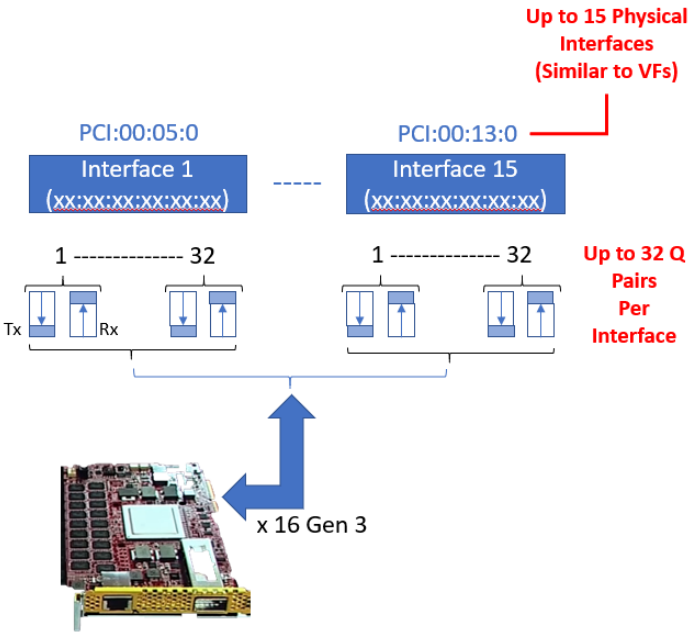


Figure 6 – Network connectivity via AWS Nitro ENA – c5n.18xlarge example

Even though all the c5n instance types make use of this 100Gbps ENA hardware, they do not have equal access to the bandwidth, the number of interfaces, the number of queues, or even if they need to share the bandwidth with other instances or not. Instances with less vCPUs and memory will have less networking bandwidth, less network interfaces (VFs), and less packet queues. This is important to note because of per-interface and per-queue performance limits that are discussed during later analysis.

Taking the public c5n instance listing and adding these key aspects for understanding the networking-related feature set yields the updated table shown as Figure 7.

[4] https://aws.amazon.com/ec2/nitro/
[5] https://aws.amazon.com/blogs/aws/new-c5n-instances-with-100-gbps-networking/

Model	vCPU	Memory (GiB)	Network Max Interfaces	Network Max Queues	Network Bandwidth (Gbps)	Network Shared or Exclusive	Number of Instances per Server
c5pn.large	2	5.25	3	96	Up to 25	Shared	36
c5n.xlarge	4	10.5	4	128	Up to 25	Shared	18
c5n.2xlarge	8	21	4	128	Up to 25	Shared	8
c5n.4xlarge	16	42	8	256	Up to 25	Shared	4
c5n.9xlarge	36	96	8	256	50	Shared	2
c5n.18xlarge	72	192	15	480	100	Exclusive	1
c5n.metal	72	192	15	480	100	Exclusive	1

Figure 7 – Adding ENA functionality to c5n instance definition

Finally, the software associated with the AWS Nitro system is generally lightweight since much of the control, management, and security of resources is done in the hardware of the various IO cards. That said, benchmarking shows notable differences in the performance of the standard ENA driver versus a standard DPDK-based PMD NIC driver and will be explored later.

Knowing how ENA resources are allocated to a given instance type along with ENA driver behavior will help explain the varied results across vCMTS benchmarking test configurations presented in subsequent sections. For example, is performance capped due to the compute capabilities of the instance, a bottleneck to and from memory, oversubscribing the ENA adapter in some way, or some other reason?

By definition, a networking application implies that data is being transported from one place to another, being "processed" (i.e., in this case, in the aforementioned cloud instances), and then, in most cases, sent somewhere else to provide some overall functionality. For this type of system, the topology of the network, complexity of the data transport layer, and location of compute resources, will all contribute to the expectations and limitations of deterministic behavior and a low end-to-end latency.

Early cloud options may have been limited, with servers and switches hosted in a far-off datacenter, but there are many more options today. As their real estate and infrastructure footprint has grown to accommodate a wide variety of workloads and markets, AWS offers compute, network, storage, and service products that can be located pretty much anywhere – including on a customer premise in various Enterprise and IOT form factors. Figure 8 shows several of the options relevant for Access network applications. The closer one gets the equipment and processing to end users, the lower possible latencies and higher potential for determinism for the application, but also the less possible abstraction and higher costs for hosting and management.

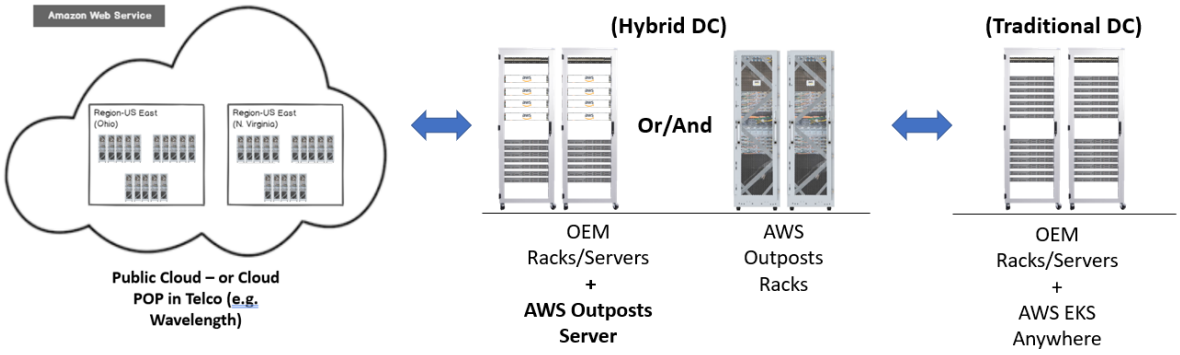


Figure 8 – AWS application continuum

AWS Regions define a physical location in the world where AWS clusters data centers[6]. Each Region is made up of so-called Availability Zones (AZs), which are one or more data centers logically connected for multiple layers of load balancing and redundancy. Some regions include ancillary Local Zones (LZs), which place AWS infrastructure and services even closer to population centers, enterprise customers, and IT centers to allow for single-digit millisecond response times. These concepts together allow for the scale, high availability, and relative localization CSP offerings are prized for and fit the needs of most applications.

But vCMTS or other Access dataplane workloads have more difficult-to-achieve requirements in terms of latency, jitter, and overall throughput than a typical Enterprise application. AWS is responding to this market by offering additional cloud deployment options. The first is called AWS Wavelength, which installs and maintains AWS equipment right inside the Edge Data Centers, Next Generation Central Offices, or Headends/Hubs owned by Communications Service Providers. This is a co-location network edge deployment model[7] and has "better together" benefits for the CSP and the network operator in being able to offer extremely low latency hosting or services for end users.

The second option, known as AWS Outposts, are packaged AWS hardware with AWS management infrastructure delivered onsite to a customer. The hardware could be a single server or a set of full racks, made for any instance type. From the technical perspective, AWS Wavelength is, perhaps, a scaled version of AWS Outposts that also includes an explicit agreement to resell or co-resell the use of this infrastructure to Enterprise customers as opposed to using to support the network directly. Given this paper focused on Access workloads (and vCMTS in particular) for an MSO, AWS Wavelength or AWS Outposts would provide approximately the same baseline performance in terms of latency and jitter.

The final option is to use non-AWS hardware running EKS, a Kubernetes offering from AWS, enabling a seamless workload orchestration across public, hybrid and private cloud. This provides an option to use

[6] https://aws.amazon.com/about-aws/global-infrastructure/regions_az/
[7] https://builders.intel.com/docs/networkbuilders/strategies-for-implementing-edge-services-in-the-10g-cable-network.pdf

familiar well proved and known quantity hardware where performance density and determinism are guaranteed while still availing of the flexibility and scalability offer by the cloud.

In short, AWS Regions, AZs, LZs, Wavelengths, and Outposts, as well as EKS Anywhere provide a gamut of options for hosting cloud instances and realizing a cloud deployment. The general interfaces and procedures to access and manage these resources is the same whether the associated server is local or (varying degrees of) remote.

In the context of the vCMTS dataplane and similar Access/Edge workloads, the main technical factors (i.e., ignoring cost) for choosing one over another are:

1. Meets or exceeds minimum required latency
2. Meets or exceeds maximum possible jitter
3. Maintains deterministic behavior for throughput/performance
4. Supports DOCSIS Time Protocol or underlying IEEE 1588 Precision Time Protocol (PTP)[8]
5. If going with onsite solution like Outpost, have the space, power, and skillset to host the infrastructure
6. Allows for required level of high availability/redundancy

Latencies and jitter are going to depend on distances, mediums, and network complexity (ex. switch hops, congestion, etc.) between the cloud site and end users. Deterministic behavior of the workload (i.e. beyond jitter on the network) will rely on having consistent access to compute, memory, and network resources with no major contention with other software applications. AWS provides several facilities for managing workload placement, which can constrain the variables for achieving solid and consistent performance. Two of these schemes are employed in this study.

The first scheme is Amazon ECS placement groups and specifically Cluster Placement Groups, which are logical groups of tasks or services[9] that can run on any type of AWS cloud location. In this way, a cluster can be defined to only run instances for the vCMTS dataplane and packet generators to achieve high and consistent throughput between deployed machine instances. Instances will be created within collocated Servers and Racks – maximizing communications efficiency. In this model is not possible to target specific servers within specific racks to host machine instances, AWS will manage the actual placement and hence actual location may vary as Instances are stopped and started.

The second scheme is the use of an Amazon EC2 Dedicated Host, which is a physical server with EC2 instance capacity fully dedicated for (one's) use. Dedicated Hosts support different configurations (physical cores, sockets and VCPUs) which allow (one) to select and run instances of different families and sizes depending on business need[10,11]. Though specifically designed for handling special Software licensing, in the context of this vCMTS dataplane analysis, it was used to run a number of c5n instances on the same physical server as shown in Figure 9.

[8] Technically DTP is part of vCMTS control plane and not strictly considered in this paper
[9] https://docs.aws.amazon.com/AmazonECS/latest/developerguide/clusters.html
[10] https://aws.amazon.com/ec2/dedicated-hosts/faqs/
[11] Note that there may be limited quantities and quotas for dedicated hosts in various regions.

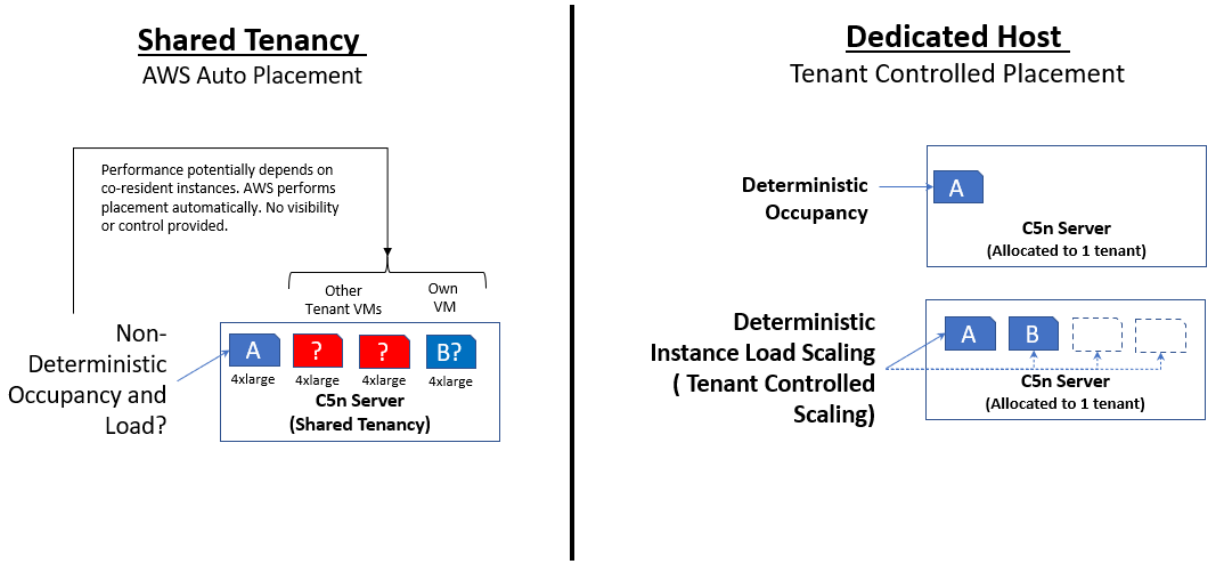


Figure 9 – AWS auto placement versus Dedicated Host

All that said, the cloud-based test setups described in the next section generally make use of AWS Regions as well as Dedicated Hosts to gather data and figure out the limits at each end of the spectrum of physical options. The other locations described in this section can be considered for future work but generally employ the exact same server hardware. As such, learnings in terms or throughput, scaling, and server capacity explored in this paper will translate to these locations too.

1.2. Application Porting and Test Setup

Intel had previously developed and released a reference vCMTS dataplane implementation to demonstrate the capability of various server hardware for the Cable Access network[12]. The reference package includes the vCMTS dataplane itself, packet generation and sink software, a cloud-native software infrastructure based on Kubernetes, and scripts for on-premise deployments. While there were some modifications that needed to be made for the cloud case (discussed later), all benchmarking was done using this vCMTS dataplane and existing packet generation software.

The on-premise tests were done on a server with the same CPU family that underlies the c5n instances – 1st Gen Intel® Xeon® Scalable Processors using the Skylake architecture. While this is an older platform than what is used for current vCMTS deployments[13], this apples-to-apples comparison makes it easier to identify underlying causes of unexpected behavior of the cloud infrastructure.

Figure 10 shows how the vCMTS dataplane and test software is deployed via Kubernetes across two different servers, one for the packet generation and sink facilities and one as the vCMTS dataplane device under test (DUT). These systems are connected back-to-back via Intel® 810 Series NICs and each SG is essentially assigned a dedicated 10GBE port (or equivalent VF). The traffic packet generation and

[12] https://01.org/access-network-dataplanes/overview
[13] https://networkbuilders.intel.com/solutionslibrary/maximizing-vcmts-data-plane-performance-with-3rd-gen-intel-xeon-scalable-processor-architecture

packet sink capabilities for a given vCMTS pod are provided by the same Traffic-Sim (DPDK pktgen) instantiation per SG.

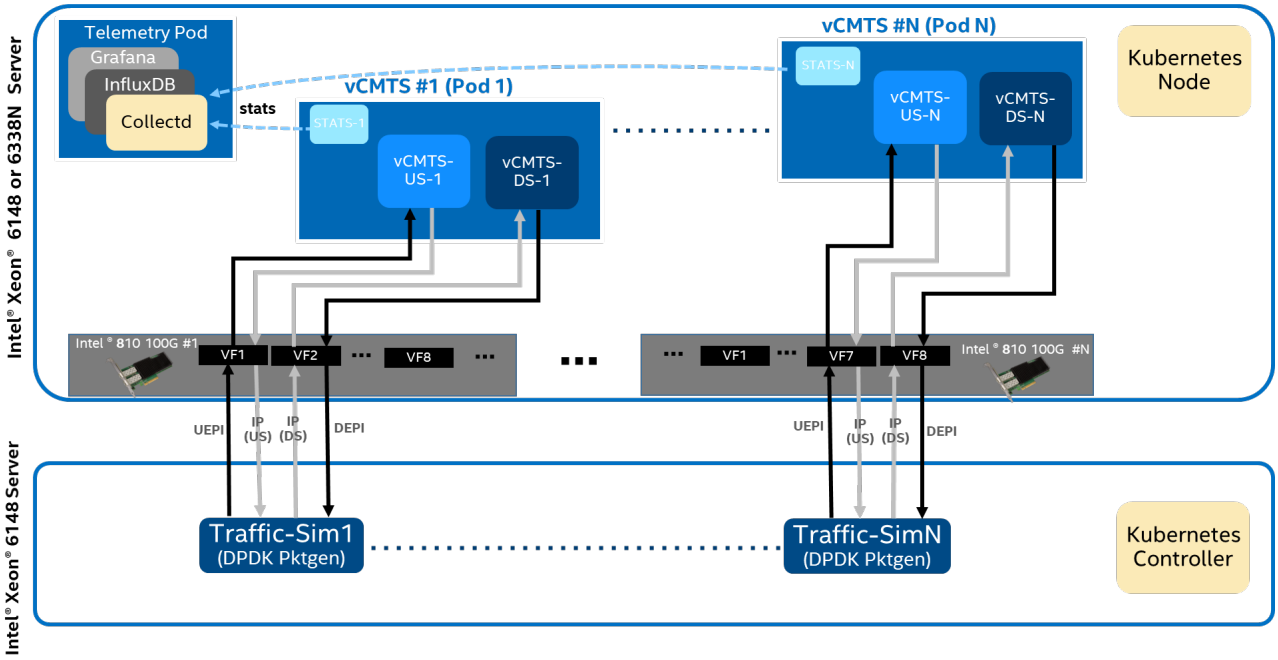


Figure 10 – On-premise vCMTS performance analysis setup

Figure 11 shows the test setup used for the AWS c5n performance analysis. At the high level, it is generally the same as the on-premise test setup insofar that the same vCMTS application and DPDK pktgen software is being used, however, there are a few key differences.

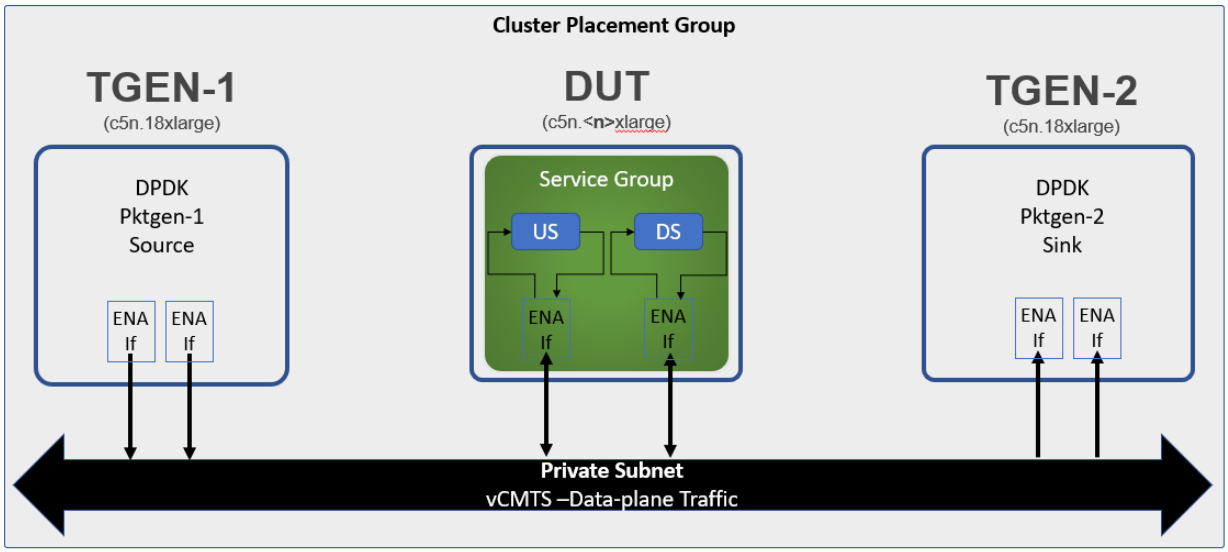


Figure 11 – vCMTS performance analysis set-up on AWS

From a system point of view, the packet generator and packet sink are running in different instances due to traffic throttling in the Nitro hardware observed during initial testing. That is, even though a single c5n.18xlarge instance has the raw compute capability to generate and receive the required number of packets, testing found there was a bottleneck through the networking sub-system that limited the bi-directional throughput. There will be more discussion of this finding later in the paper as it relates to scaling of the vCMTS dataplane itself.

On the infrastructure front, Kubernetes and the on-premise install scripts are not needed as the AWS Nitro system has its own hypervisor and application management facilities. In fact, the AWS cloud management tools were used to instantiate the desired instances and connect them via a Virtual Private Cloud (VPC) per the network diagram in Figure 12. Correspondingly, the existing reference testbed – which assumed that physical systems are connected back-to-back or through a simple switch – was modified so that L2TP tunnel and Cable-modem IP addresses were adapted to the sub-net properties of this VPC. At the network device level, the DPDK PMD for the Intel® NICs were switched to use the ENA PMD. This change also required the use of the DPDK out-of-tree IGB_UIO kernel module instead of the standard VFIO kernel module due to performance degradation with the ENA PMD.

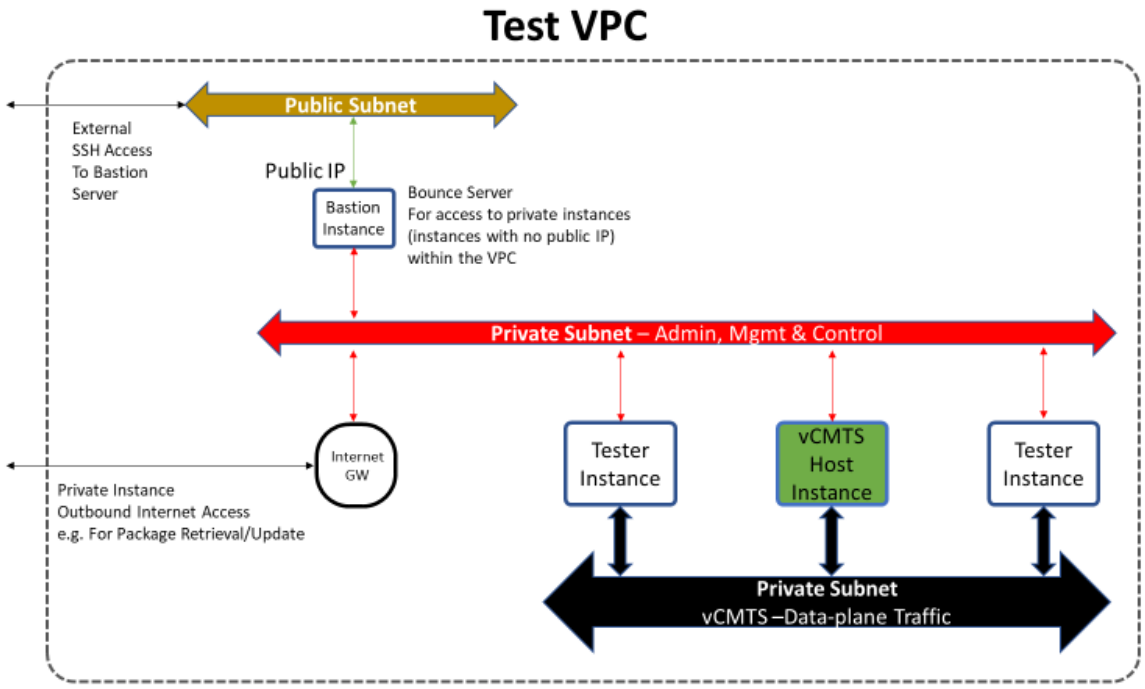


Figure 12 – VPC network architecture for vCMTS dataplane testing

All benchmarking, both on-premise and cloud setups, used the same general process as defined by RFC-2544[14], but limited to steady state load testing across a wide variety of packet sizes. In addition to the packet sizes recommended by RFC-2544, an iMix representative of MSO network traffic was also employed to replicate real-world circumstances. All tests were bi-directional in nature (i.e., sending traffic in both upstream and downstream directions simultaneously) and maximum throughputs were measured under zero packet loss conditions. Upstream traffic was configured to be a minimum of 10%

[14] https://datatracker.ietf.org/doc/html/rfc2544

of Downstream traffic for cloud performance tests. All tests employ the same number of (virtual) cable modems and use the same filter rules. In some situations, supplemental data, like per-pipeline-stage CPU cycle-counts, were captured later to the original test to help explain certain observations. Such data will be discussed in the relevant sections below.

2. Single-SG Throughput Analysis

Initial testing established single-SG performance across both the on-premise systems and the following AWS c5n instance types: c5n.metal, c5n.4xlarge, and c5n.2xlarge. The SG was chosen as the base unit of interest because a vCMTS dataplane would be scaled across vCPU cores based on the number of service groups required for a given network footprint.

Since all c5n instances make use of the same c5n server hardware it was expected that the single-SG performance would be similar across instances. It was also expected that this cloud performance would be similar to what is achievable on the on-premise platform. In both cases, there were surprises.

2.1. Comparing Instance Behavior to On-Premise

As mentioned in Section 1.2, bi-directional upstream and downstream throughput was measured for each test configuration with zero packet loss with traffic of several packet sizes and iMix types. Figure 13 plots the results for a c5n.metal instance.

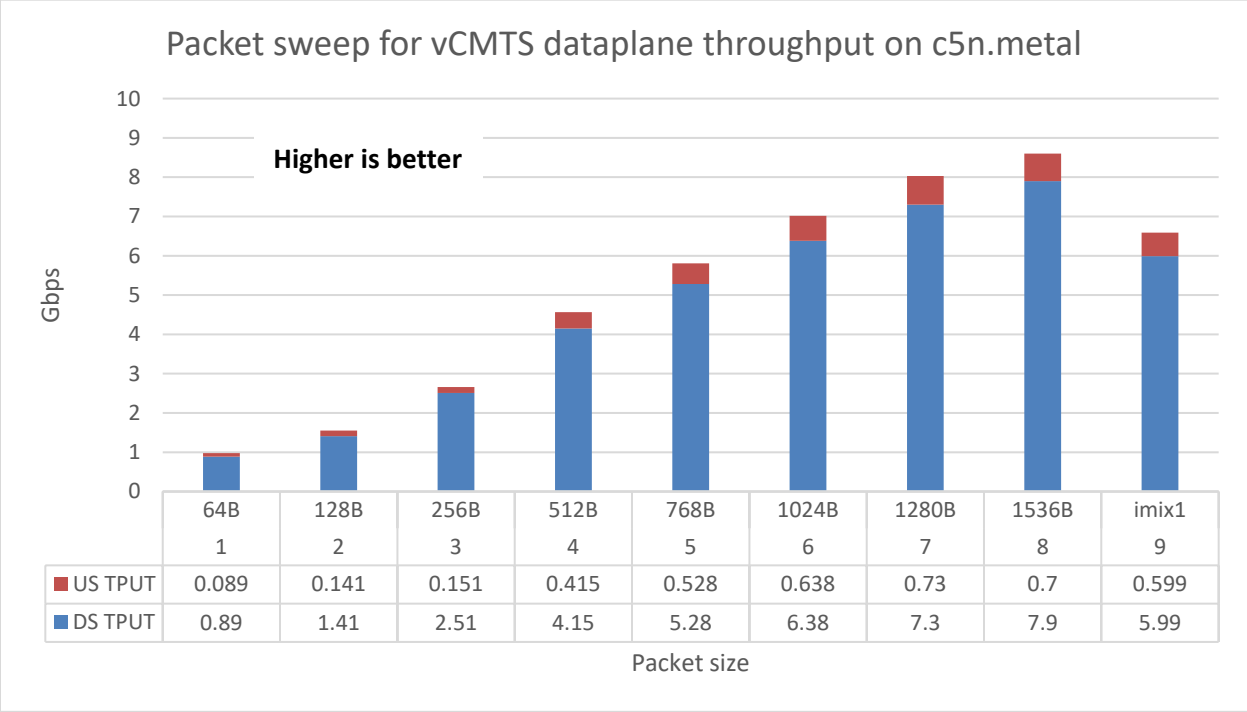


Figure 13 – Max bi-directional throughput for single vCMTS SG (c5n.metal)

The performance variance across packet sizes is expected and is a common phenomenon among any network function because smaller packet sizes at a certain network speed give less time to process that packet than larger ones. This behavior is seen in the other instances and both the on-premise platforms.

Since looking at any given packet size will yield the same themes, from here the analysis will use iMix data only. Figure 14 shows the iMix upstream and downstream aggregate throughput results across all tested instance types and on-premise platforms.

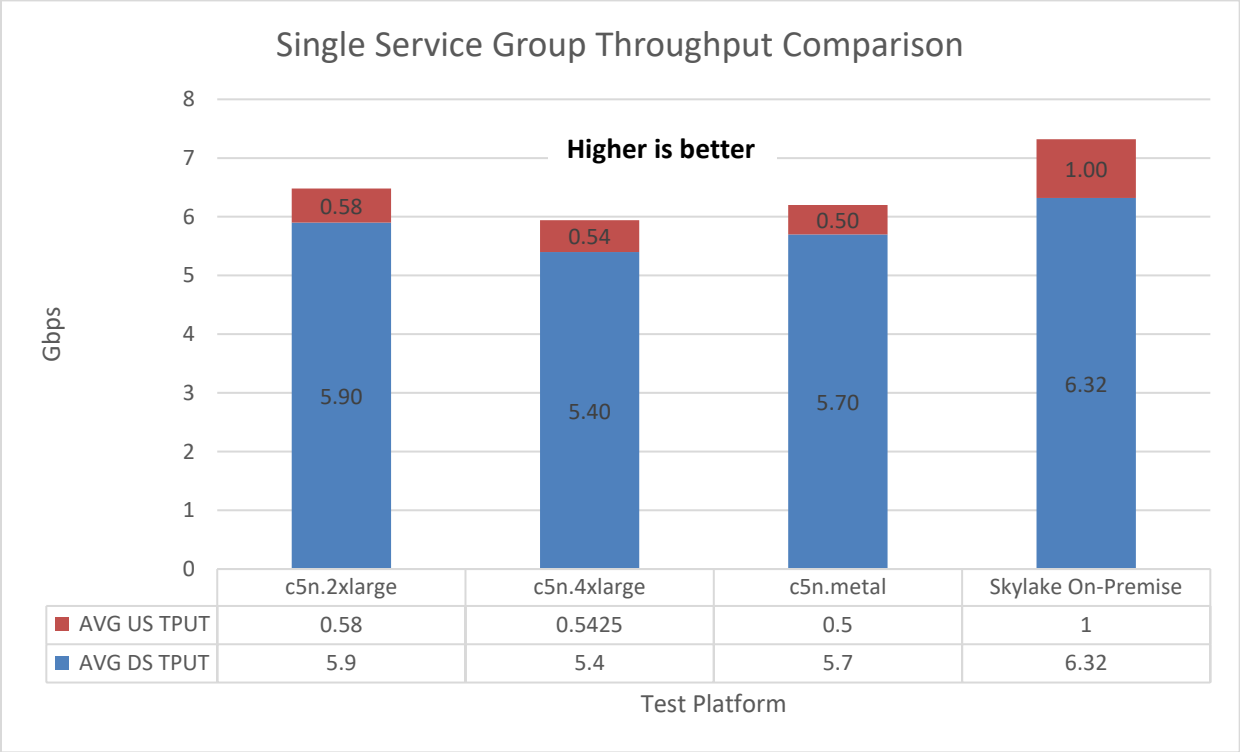


Figure 14 – Comparing per-SG vCMTS Dataplane Performance for iMix

The first observation is that the on-premise solution yielded the best performance at 7.32 Gbps of total upstream and downstream throughput. This was 13% better than the best cloud instance (c5n.2xlarge) at 6.48 Gbps. The result is surprising given that the cloud instances are running on CPUs with a higher base clock frequency than the CPU in the on-premise system, and the fact that the cloud instances may be able to take advantage of CPU Turbo operation for an even bigger boost (the on-premise servers disable Turbo in the name of determinism). Without mitigating factors in the memory or network, the performance of the vCMTS dataplane tracks linearly with the CPU speed. Thus, the cloud instances running at a minimum of 3.0GHz should get 25% better throughput than the 2.4GHz on-premise servers, but instead they are far in deficit.

How to explain this difference?

To answer this question, CPU cycle counts for each stage of the DOCSIS MAC pipeline were collected for both on-premise and cloud instance test systems. This was done in runs separate from the throughput measurements so one test did not affect the other. Figure 15 compares the results and sheds light on where each system is spending its CPU budget.

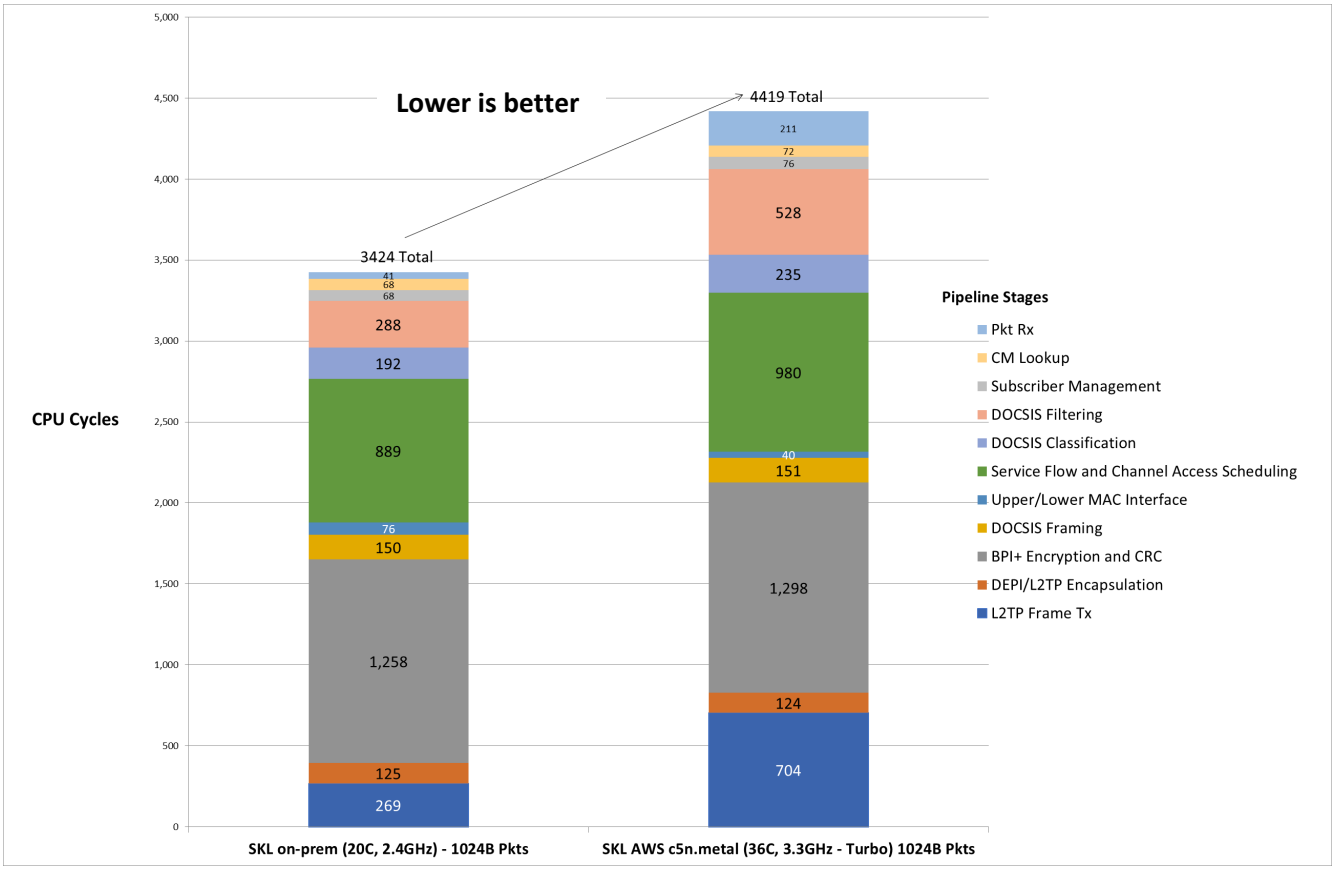


Figure 15 – DOCSIS MAC downstream CPU cycle count comparison – On-premise vs Cloud

Many of the pipeline stages have similar cycle counts. This means that both the cloud instances and the on-premise servers spent equal amounts of compute capability in those stages. However, the c5n instances took over 400%, 160%, and 150% longer in the packet receive, the packet transmit, and the DOCSIS filtering stages, respectively. Together this accounts for 29% more cycles and thus would process 29% less packets without even accounting for other possible bottlenecks.

The software part of the AWS Nitro system is generally lightweight; however, this data indicates that the ENA networking driver has quite a bit more overhead for basic receive and transmit functionality than a typical DPDK PMD in Linux. It is not clear what portion of this overhead can be won back with optimization to the ENA software driver or Nitro hypervisor (by AWS) and what portion is due to immutable behavior in the ENA/Nitro hardware itself.

The additional cycles for the DOCSIS filtering stage are not fully explained insofar that all tests use the same number of cable modems and the same filtering rules (i.e., parameters that are known to change the processing needs of this algorithm), but the current theory is that the significantly greater number of CPU cycles spent on packet receive is having some secondary effect on a shared resource in the CPU, for example, the L3 cache, instruction pipelines, or memory controller, and causing additional latency through the system.

The other main observation from Figure 14 is that the throughput was not consistent across the cloud instance types even though they are all running on the same underlying hardware and the throughput for a single service group should not be saturating the available network bandwidth. The c5n.2xlarge and c5n.metal data is approximately the same, within 5% variation, but the c5n.4xlarge performance was more than 10% worse.

2.2. Use of AWS Dedicated Hosts

Given that all c5n instances run on the same physical hardware and do not hit hard vCPU, memory, or network constraints to support a single SG, what could be the reason for the c5n.4xlarge performance difference? Looking back at Figure 5, four c5n.4xlarge instances can be instantiated on a given c5n server at any one time, and there are no guarantees where a given instance will be scheduled within an AWS Region (even if employing clusters). In other words, there could be other workloads from other AWS customers trying to make use of the same CPU L3 caches, memory, network interfaces, and other shared resources critical to vCMTS dataplane performance running on the same c5n server. These types of applications are known as noisy neighbors.

Per Section 1.1, AWS does provide a way to limit noisy neighbor risk via an AWS EC2 Dedicated Host. With a Dedicated Host, a customer reserves a whole server type (in this case a c5n) for themselves and therefore has full control on what instance types will be deployed there and what workloads will be running within those instances. Figure 16 shows how the initial single-SG test setup was modified from the original of Figure 11 to incorporate a Dedicated Host running the vCMTS dataplane in a c5n.4xlarge instance.

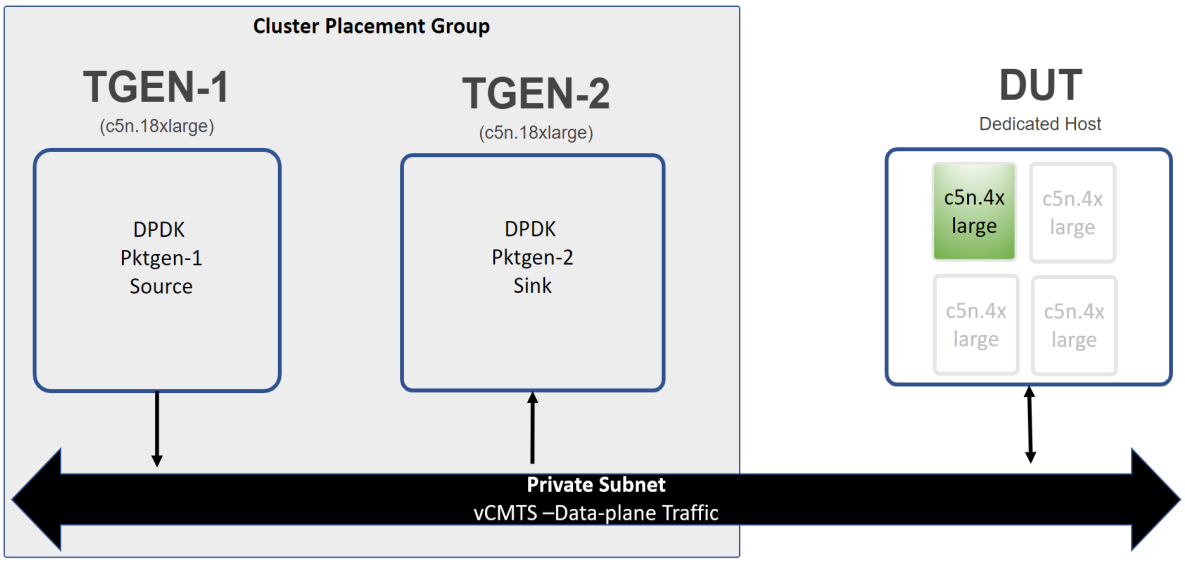


Figure 16 – vCMTS test setup with Dedicated c5n Host

Figure 17 shows the new c5n.4xlarge single-SG throughput test results alongside all the previous data.

Figure 17 – Affect of using a dedicated host on per-SG Performance

It is clear that using the Dedicated Host helped bring the c5n.4xlarge results in-line with the performance of the other c5n instance types, for example, within 1% of the c5n.metal bi-directional throughput. By definition a c5n.metal instance takes up a whole c5n server, so it makes sense that the behavior of a single known application in that environment would be consistent. However, in the c5n.4xlarge case, there are no guarantees on where and how AWS will schedule such instances in real time from any customer.

It is interesting to note, anecdotally, that during the initial single-SG cluster-based testing the c5n.4xlarge throughput would be in line with the performance of the other instances as if it was running on a Dedicated Host. However, it was uncommon and there was no discernable pattern to when the throughput would be higher and when it would be lower. A network operator cannot rely on a random best-case scenario when dimensioning cloud resource needs for critical Access infrastructure.

That said, why did the testing not show this same behavior in the c5n.2xlarge instance case? It, too, can share a c5n server with other instances; per Figure 5 a c5n server can host up to 8 c5n.2xlarge instances and chances are that at some point there will be a noisy neighbor application running in one of them. Without knowing the full details of the AWS instance scheduler, the main lesson is that without full control of the hardware performance cannot be guaranteed. With enough tests it is expected that the c5n.2xlarge instances would also see the same performance variability as measured in the c5n.4xlarge case.

Why wouldn't a user always employ a Dedicated Host then? While the benefits are clear, there are some trade-offs to using this type of cloud infrastructure that would need to be factored into a total cost of ownership calculation. First, AWS limits the number of dedicated hosts per account to two per AWS

Region. Perhaps exceeding that stated limit is due to the fact they their original intention is to support legacy software licensing models, or perhaps it puts the user into the realm of deploying their own AWS Outpost servers? Regardless, it is not an approach that scales on its own. In addition, for every dedicated host available a user will pay a higher per time unit fee versus equivalent EC2 product offerings (ex. c5n.metal) at the time of this writing[15].

While it was not originally designed for networking applications, the AWS Dedicated Host concept is very useful for workloads that prize deterministic behavior and/or have hard requirements on system resources to achieve key performance indicators (KPIs). However, their utility for large scale deployments may be limited due to higher cost and capacity constraints.

3. Per-Instance Throughput Analysis

The per-SG analysis uncovered functional differences between c5n instance types and comparable on-premise solutions for the vCMTS dataplane and subsequently, insights into the c5n hardware itself, how those servers are shared among multiple customers, and key insights for managing performance expectations.

The next phase of testing is designed to determine the following for each c5n instance type:

1. Scaling capability (i.e., throughput per SG)
2. Maximum throughput possible via multiple SGs
3. The limiting factor for performance (ex. vCPU, memory, networking)

This information will be used to deepen the technical understanding of each instance type and, ultimately, to plan effective vCMTS dataplane application scaling strategies.

3.1. Finding Performance Limitations

This scaling analysis uses the same testing set up that was used for the single-SG benchmarking. Data will be collected from c5n.metal, c5n.4xlarge, and c5n.2xlarge instance types, as well as the on-premise platform. In the ideal case, good scaling means that the throughput per SG stays as constant as possible as more SGs are added into the system. However, the reality is that the scaling is typically non-linear, in the negative direction, for any system with shared resources such as all of these under test.

The vCMTS dataplane reference application requires two network interfaces per SG (one for upstream and one for downstream), so the maximum number of SGs for a given instance type is the total number of network interfaces minus 1 for an administration port, then that number divided by two. For example, the c5n.4xlarge instance is allowed up to 8 total interfaces, so with 1 interface used up for administration, the maximum possible number of SGs for that setup is 3.

Figure 18 and Figure 19 below show maximum service-group scaling for c5n.metal and c5n.4xlarge instance types, respectively. The data for the c5n.4xlarge instance type was collected using the standard cluster setup versus using a Dedicated Host. Scaling data is not shown for the c5n.2xlarge instance since its 4 network interfaces limit it to supporting 1 SG at a time.

[15] https://aws.amazon.com/ec2/pricing/

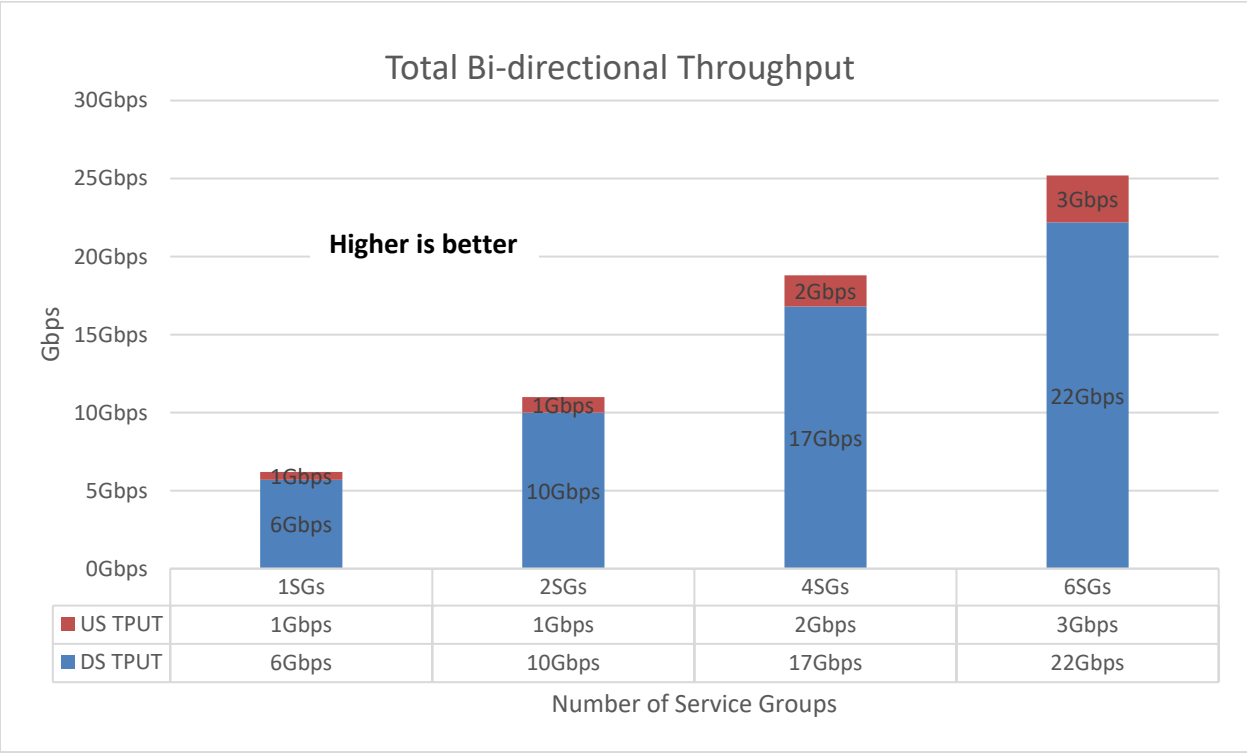


Figure 18 - Max bi-directional throughput for vCMTS service-group scaling (c5n.metal)

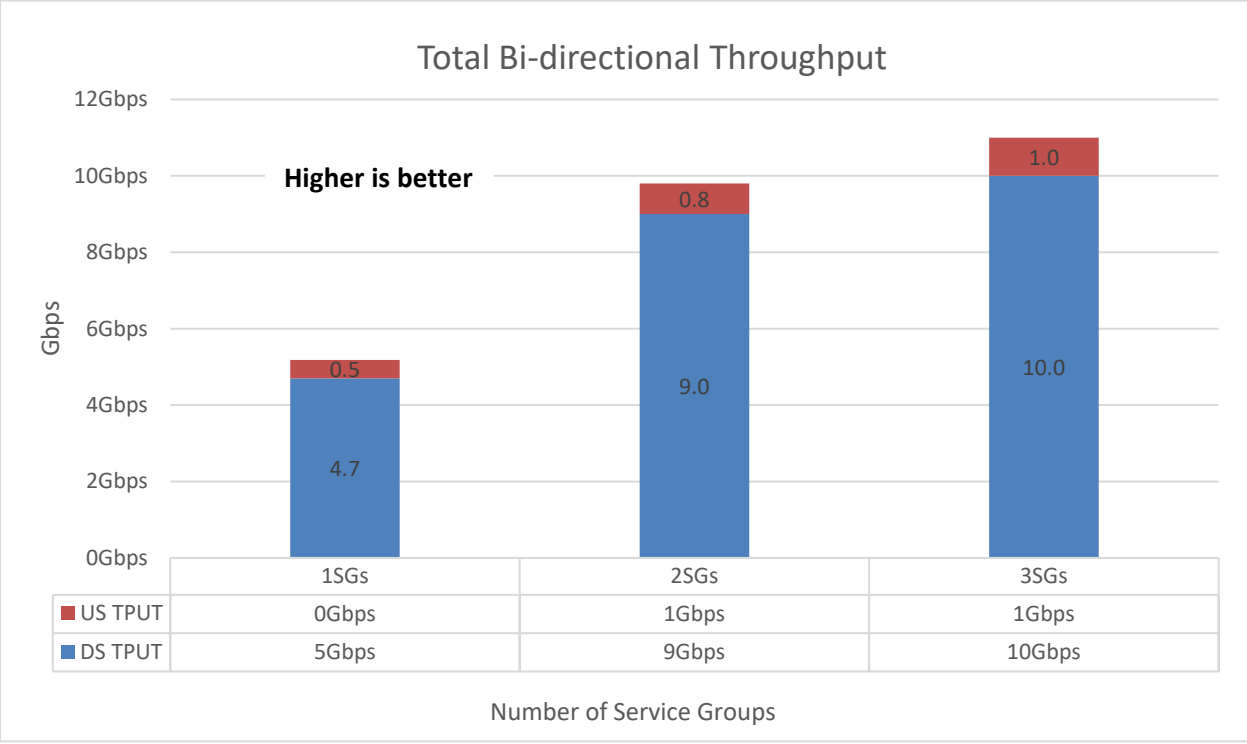


Figure 19 - Max bi-directional throughput for vCMTS service-group scaling (c5n.4xlarge)

The total bi-directional throughput of the c5n.metal instance shown in Figure 18 is not linear, but it tracks similarly to an on-premise solution insofar that it gets progressively worse as more SGs are added. The data points start at a single SG achieving 6.2 Gbps, then going to 2 SGs at 5.5 Gbps each and 4 SGs at 4.7 Gbps each, finally landing at 4.2 Gbps per SG in the 6 SG case. With 3 vCPUs supporting each SG, only 18 vCPUs were used out of the 72 available to achieve the total 25.2 Gbps bi-directional throughput for this instance type. Thus, the maximum throughput per instance is limited by the number of networking ports. Each additional two ports would allow the addition of another SG.

The story is different in the c5n.4xlarge case. There is decent scaling up to 2 SGs, with the throughput per SG at 5.2 Gbps and 4.9 Gbps in the 1 SG and 2 SG cases, respectively. However, the total throughput for the instance seems to be capped at about 11 Gbps, devastating the throughput per SG in the 3 SG case (bringing it down to 3.67 Gbps per SG). Prior testing showed that each network interface can only pass a maximum of ~4.0Mpps bidirectional or ~9.0Mpps unidirectional traffic, so total throughput for the system will be limited for small packet sizes. The iMix traffic in this testing averages ~1000 bytes in the downstream direction and ~300 bytes in the upstream direction and thus well outside of realm of possibility for saturating the theoretical 25 Gbps network bandwidth for the c5n.4xlarge instance type.

Bringing the on-premise platform into the picture, Figure 20 compares c5n.metal instance scaling to on-premise scaling and highlights the reasons for a huge difference in overall performance.

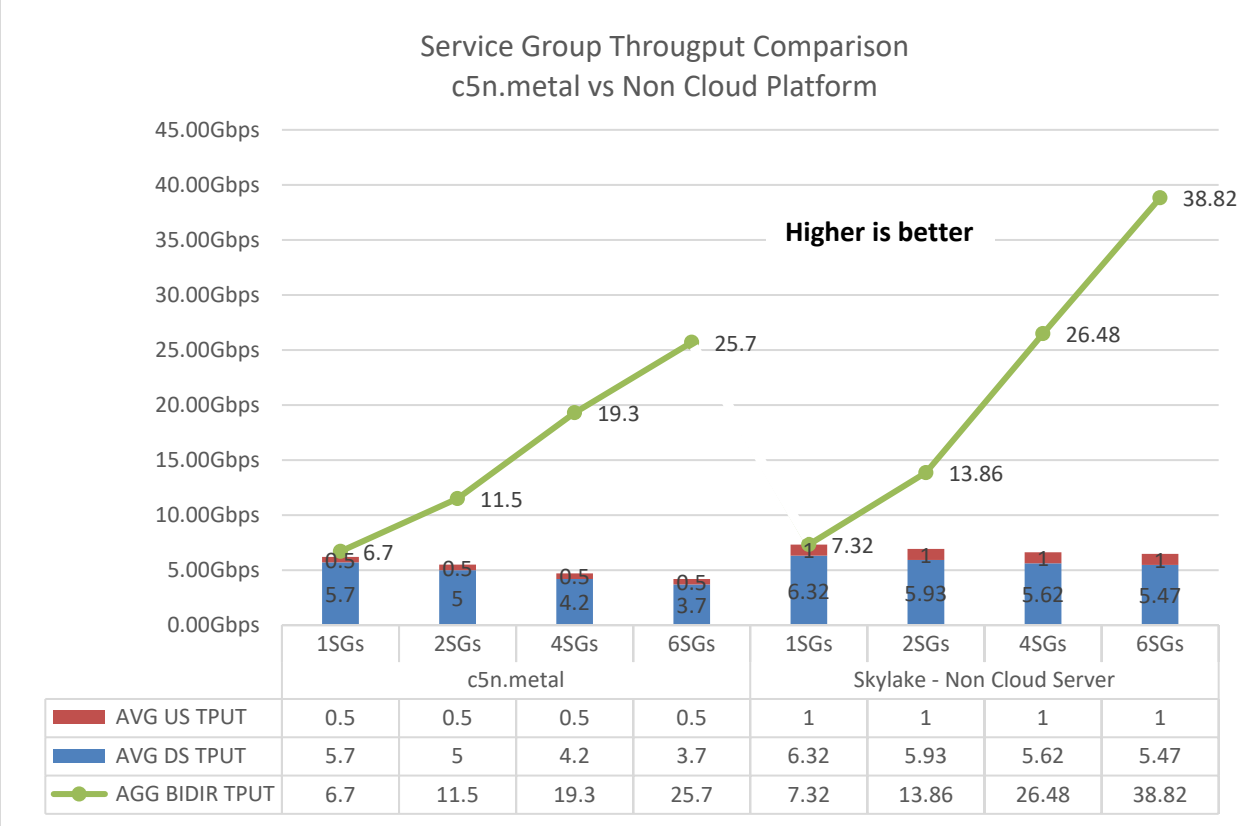


Figure 20 – Comparing on-premise to cloud scaling

There are three reasons why the on-premise platform scales better and achieves much higher overall performance than the c5n.metal instance. First is that the number of vCPU cycles needed per packet is

much higher for c5n instances due to large receive and transmit overhead in the ENA driver as described in Section 2.1. Second, the slope tracking the aggregate bi-directional throughput for the on-premise case in Figure 20 is much steeper versus the one tracking the c5n.metal performance, meaning that its average throughput per SG goes down at a slower rate (as seen in the per-SG bars). Finally, per prior analysis[16], the on-premise solution is only limited by the number of vCPUs and the amount of L3 cache; in the case of the Intel® Xeon® 6148 Scalable Processor with 20 physical cores (40 vCPUs) and 27.5 MB of L3 cache, the total achievable throughput per server with two sockets is ~100Gbps versus the ~25 Gbps for the c5n.metal case.

The main take-away is that the maximum vCMTS dataplane throughput for every c5n instance type is being limited by the networking sub-system. Specifically, the c5n.metal instance is limited by the number of networking ports and the c5n.4xlarge instance is limited by maximum packets per second in the ENA. Either way, this means that c5n instances, as defined today, have a mismatched compute-to-networking ratio and thus cloud resources that are being paid for (i.e., vCPUs) are not being used when the system is targeted at network-heavy applications.

3.2. Breaking Through Per-Instance Bottlenecks

Looking for clues in Figure 7, the compute to networking mismatch can be minimized with careful planning across multiple instances. In other words, can the c5n.metal bottleneck due to lack of network ports be solved by deploying multiple other c5n instance types with a better vCPU to port ratio (even if those instances have different per-instance limitations of their own)?

Figure 21 illustrates the example by comparing a single c5n.metal instance to four c5n.4xlarge instances, both of which fully occupy a single c5n server. These two system configurations are cost-equivalent in the financial sense, but also in terms of having the same amount of aggregate network bandwidth (100 Gbps) and a similar number of vCPUs (72 and 64). However, due to the greater total number of network ports available, the four c5n.4xlarge set up can support up to 12 SGs while the c5n.metal instance is limited to 7 SGs.

[16] https://networkbuilders.intel.com/solutionslibrary/maximizing-vcmts-data-plane-performance-with-3rd-gen-intel-xeon-scalable-processor-architecture

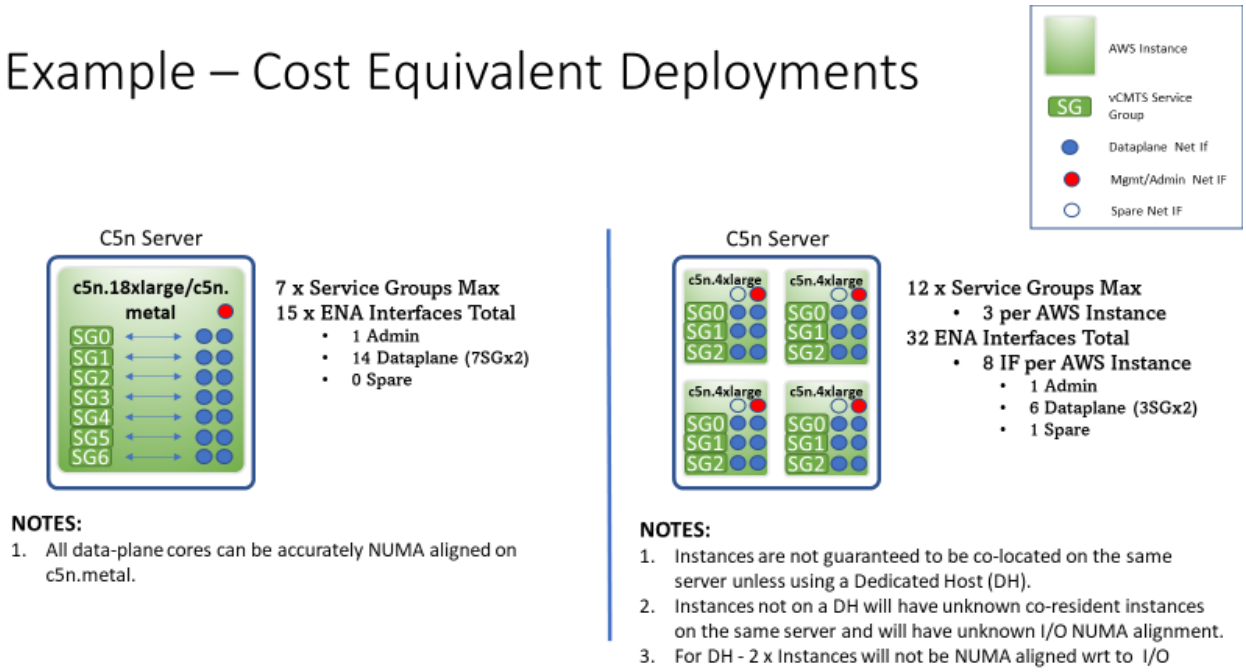
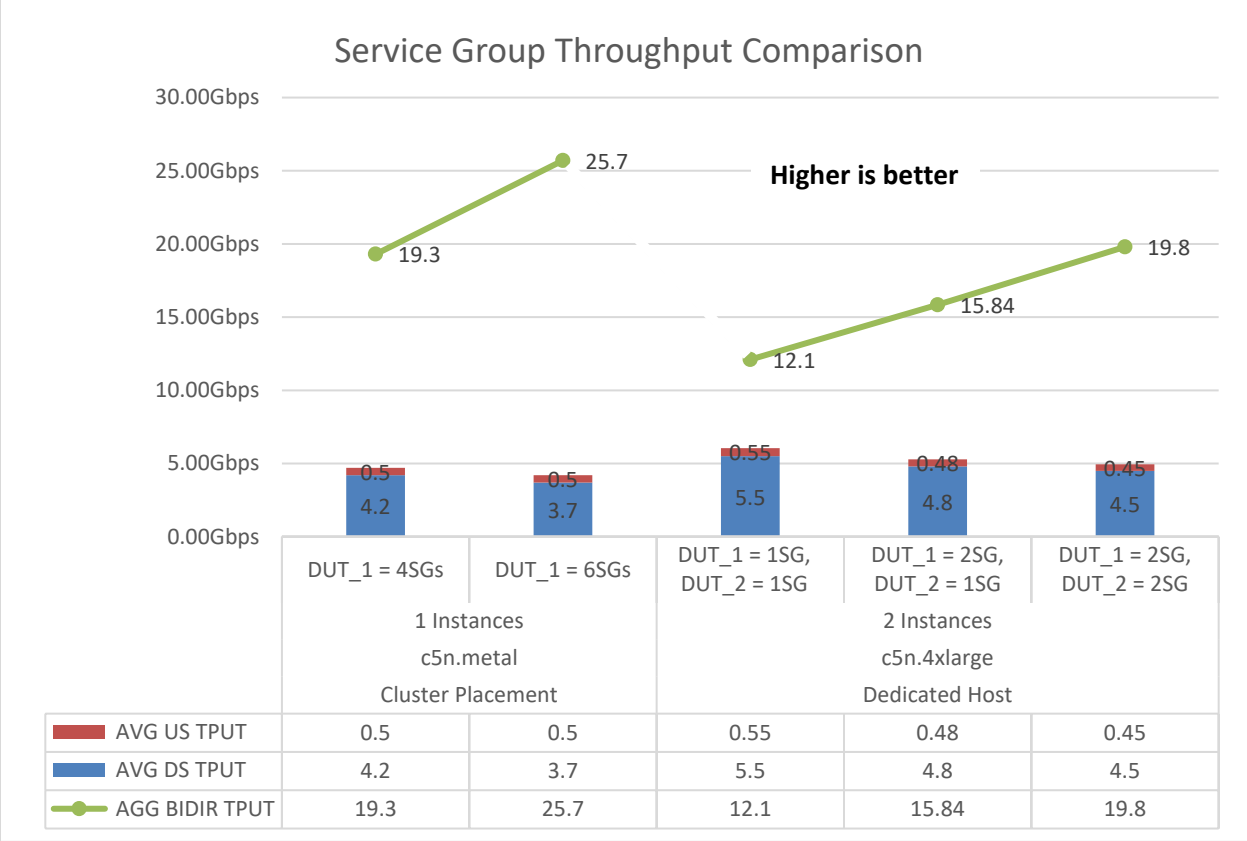


Figure 21 – Example of cost equivalent c5n instance configurations

Per the scaling data in the previous section, each c5n.4xlarge instance has a maximum possible throughput of 11 Gbps and therefore the maximum theoretical throughput for this system is 44 Gbps vs the 25.7 Gbps measured for the c5n.metal instance, a possible 71% improvement.

Taking the effect of real-life scaling factors into account, Figure 22 shows the total throughput measurements of two c5n.4xlarge instances. And while there is some degradation from the theoretical maximum, the data still demonstrates how compelling this approach can be. Note that due to organizational logistics (i.e., not any issue with the AWS infrastructure), benchmarking was completed with only two of the full four instance configuration.



	DUT_1 = 4SGs	DUT_1 = 6SGs	DUT_1 = 1SG, DUT_2 = 1SG	DUT_1 = 2SG, DUT_2 = 1SG	DUT_1 = 2SG, DUT_2 = 2SG
	1 Instances c5n.metal Cluster Placement		2 Instances c5n.4xlarge Dedicated Host		
AVG US TPUT	0.5	0.5	0.55	0.48	0.45
AVG DS TPUT	4.2	3.7	5.5	4.8	4.5
AGG BIDIR TPUT	19.3	25.7	12.1	15.84	19.8

Figure 22 – Throughput of two c5n.4xlarge instances to one c5n.metal

Starting with these measurements of 19.8 Gbps for two c5n.4xmetal instances supporting 2 SGs each, we see a 10% scaling penalty to the theoretical maximum of 22 Gbps. Conservatively assuming a further scaling penalty of 15% when going from two to four c5n.4xlarge instances, one could expect up to 33.6 Gbps of bi-directional traffic, which is over 76% of the theoretical maximum and still a 30% improvement over the c5n.metal capacity at the same cost.

Table 1 summarizes the vCMTS dataplane performance for each test setup alongside the resource usage for each scenario to give an idea of the constraint that is causing the bottleneck. It also calculates what percentage of the instance resources are being used and what is left unused – specifically for the vCMTS dataplane at least.

Instance Type	vCPU	Ethernet	Ports	Maximum vCMTS tput	Tput limitation	Used Resources	Unused Resources	% Unused Resources
Single c5n.metal	72	Shared 100G	15	25.7Gbps with 6 SGs	Number of ports	18 vCPUs, 25.7Gbps, 12 ports	54 vCPUs, 74.3 Gbps, 3 ports	75% vCPUs, ~74% network

Four c5n.4xlarge	64	Shared 100G	32	33.6 Gbps[17] with 12 SGs	pps per instance	36 vCPUs, 33.6 Gbps, 25 ports	28vCPUs, 66.4 Gbps, 7 ports	~43% vCPUs, 66% network
Two c5n.4xlarge	32	Shared 50G	16	19.8 Gbps with 6 SGs	pps per instance	18 vCPUs, 19.8 Gbps, 13 ports	14 vCPUs, 30.2 Gbps, 3 ports	~43% vCPUs, ~60 Gbps network

Table 1 - Scaling via multiple smaller instances

Clearly, aggregate performance of four c5n.4xlarge instances (based on the measurements from two c5n.4xlarge instances) is better than a single c5n.metal instance from both absolute throughput and percentage of resource usage perspectives. This configuration drives better utilization of vCPU cores by having a greater aggregate number of network interfaces to work with. That said, all c5n platforms used for vCMTS dataplane applications will not make use of all available compute and network capability.

So, what can be done with those unused resources? First, given that one type of bottleneck to scaling is the number of Ethernet ports per instance, a software vendor could modify the workload to use less ports overall. For example, the vCMTS dataplane application dedicates a port for US and a port for DS for every SG to optimize its transmit and receive logic, but this scheme could be modified to make better use of this constrained resource. Even if the new code to manage port usage adds more cycles per packet, the change would be well worth the effort in terms of maximizing overall throughput.

Second, unused vCPUs could be used for applications that do not require network access and do not otherwise interfere with the vCMTS dataplane resources, for example, to pre-process network or system telemetry. Alternatively, those unused vCPUs could be used in a high availability scheme, with backup vCMTS dataplane instantiations pointed to relevant databases and network ports but not actually processing packets until the primary applications fail.

No matter what is done with resources unused by the vCMTS dataplane, this testing shows that it is possible to use multiple smaller instances (in this example, c5n.4xlarge) to overcome sub-optimal compute-to-networking ratios and achieve better total bi-directional throughput for c5n-based infrastructure.

4. Conclusions

There is a lot of interest in moving applications found in the modern network into cloud infrastructure for all the typical reasons: offloading deployment and configuration of hardware, managing all infrastructure and over the top services though a consistent pane of glass, and having all needs for scale and multiple levels of redundancy taken care of by someone else. These are good and compelling reasons.

However, not all applications are created equal; each one must be evaluated to see if it makes financial, business, and technical sense to move from on-premise equipment to the cloud. This paper focused on only technical considerations related to making best use of cloud infrastructure for vCMTS dataplane

[17] Four c5n.4xlarge performance estimated based on the two c5n.4xlarge lab measurements

processing, with its challenging requirements for high throughput, cryptographic elements, low latency, and consistent performance, and leaves the rest to other venues. Only AWS product offerings were evaluated in this context, but the same methodology and criteria can be used to choose the best cloud infrastructure from any CSP.

The results were clear that even though the AWS c5n family of cloud instances was shown to be best among AWS offerings because of its high capability with Intel® Xeon® Scalable Processors paired with 100 Gbps network interfaces, there is a huge gap in their performance relative to a similar on-premise deployment. Per testing, there is too much overhead and a lack of performance guarantees that make it costly or operationally prohibitive to widely deploy the vCMTS dataplane to the cloud.

That said, AWS and other CSPs are continually innovating and introducing new cloud products; it is inevitable that they will address some of the observations described in this paper:

- Networking capability was not matched to the compute capability, limiting overall throughput
    - Networking ports per instance was generally too low
    - Compute cycles to support packet receive and transmit on the ENA was too high
    - Published bandwidth capacities had too many caveats to achieve in real-world tests (ex. packet per second limitations in the ENA capped iMix performance)
- Lack of resource isolation and/or control of workload placement leads to non-deterministic performance

A cloud instance (or rather, family of instance types) that addresses these issues would increase overall system performance, reduce custom software development efforts, assure customers with expected and consistent behavior, and generally give maximum value to the resources and services a network operator is purchasing to run an important part of their network.

In the meantime, even if not suitable for a large-scale vCMTS deployment, the convenience of cloud infrastructure still makes it interesting for other use cases, including the expansion of test capability, the enabling of remote collaboration, as a temporary solution for surge traffic, or to expand fail-over capacity.

Again, the most value from this work is in the education and process. The particulars (ex. test data for a c5n instance type) may have a limited lifespan, but the evaluation framework will help network operators make informed decisions on CSP offerings for any dataplane applications they are looking to move into the cloud now and into the future. On the flip side, it may help a CSP create the perfect cloud product optimized for the Access network.

4.1. Future Work

While cloud infrastructure in various forms has been around a long time and has proven its value in many business arenas, it is still early days for applying its underlying technologies and processes to the high demand needs of Access and Edge workloads like the vCMTS dataplane used in this study. There are many compelling topics in this domain demanding further research; some of these are: to expand this analysis to future AWS offerings (especially the next generation of the c5n family), to expand this analysis to similar products from other CSPs, to study questions around latency in the network by, for example, varying the location of the packet generator and sink relative to the dataplane function, and evaluate cloud offerings for vCMTS control plane needs.

Abbreviations

AWS	Amazon Web Services
AZ	AWS availability zone
bps	bits per second
CMTS	cable modem termination system
CoSP	communications service provider
COTS	commercial off the shelf
CSP	cloud service provider
DAA	distributed access architecture
DPDK	data plane development kit
DTP	DOCSIS time protocol
DUT	device under test
ECS	AWS elastic container service
ENA	AWS elastic network adaptor
KPI	key performance indicator
ML	machine learning
MSO	multi service operator
NIC	network interface card
PCIe	PCI Express
PMD	poll mode driver
pps	packets per second
PTP	IEEE 1588 precision time protocol
SCTE	Society of Cable Telecommunications Engineers
SG	service group
TCO	total cost of ownership
VF	virtual functions (in context of a PCIe device)
VPC	virtual private cloud

Bibliography & References

Amazon Web Services (2021). “Amazon EC2 Overview”, https://aws.amazon.com/ec2/ (Accessed: 25 August 2021).

Amazon Web Services (2021). “Amazon EC2 Instance Types”, https://aws.amazon.com/ec2/instance-types/ (Accessed: 25 August 2021).

Amazon Web Services (2021). “AWS Nitro System”, https://aws.amazon.com/ec2/nitro/ (Accessed: 25 August 2021).

Barr, J. (2018). “New C5n Instances with 100 Gbps Networking”, https://aws.amazon.com/blogs/aws/new-c5n-instances-with-100-gbps-networking/ (Accessed: 25 August 2021).

Intel Corporation (2021). “Intel vCMTS Reference Dataplane and NFV Stack”, https://01.org/access-network-dataplanes/overview (Accessed: 25 August 2021).

Ryan, B. et. al. (2021). “Maximizing vCMTS Data Plane Performance with 3rd Gen Intel® Xeon® Scalable Processor Architecture”, https://networkbuilders.intel.com/solutionslibrary/maximizing-vcmts-data-plane-performance-with-3rd-gen-intel-xeon-scalable-processor-architecture (Accessed: 25 August 2021).

Heaton, E. (2020). “Strategies for Implementing Edge Services in the 10G Cable Network”, https://builders.intel.com/docs/networkbuilders/strategies-for-implementing-edge-services-in-the-10g-cable-network.pdf (Accessed: 25 August 2021).

Bradner, S. and McQuaid, J. (1999). “Benchmarking Methodology for Network Interconnect Devices”, https://datatracker.ietf.org/doc/html/rfc2544 (Accessed: 25 August 2021)

Appendix A – On-Premise Test Configuration

Test Environment Configuration Information and Relevant Variables	
CM Lookup & Subscriber Mgmt	300 subscribers per service group, 4 IP addresses per subscriber
DOCSIS Filtering	6 filter groups, 2 filter groups associated with each cable-modem 16 filter rules per filter group 10% matched, 90% unmatched (default action - permit)
DOCSIS Classification	16 rules per subscriber, 10% matched - enqueue to one of 3 service-flow queues 90% unmatched - enqueue to default service-flow queue
Downstream Service-Flow Scheduling	8 service-flow queues per subscriber (4 active)
Downstream Channel Scheduling	6 x OFDM (1.89 Gbps) Channels, 2 x channel-bonding groups. Or 2 x OFDM (1.89 Gbps) and 32-SC-QAM (42.24 Mbps) Channels, 4 x channel-bonding groups NOTE: channel-bonding groups are distributed evenly across cable-modems
Upstream Bandwidth Scheduling	Upstream Scheduler not used. Upstream bandwidth pre-allocated in grants of 2KB per service ID. Bandwidth grants balanced evenly across 300 cable-modems.
Ethernet CRC	Downstream: 100% CRC re-generation Upstream: 0% CRC verification NOTE: CRC relates to inner frames
Encryption	100% AES, 0% DES
Packet IMIX Distribution	Upstream 65% : 84B, 18% : 256B, 17% : 1280B Downstream 15% : 84B, 10% : 256B, 75% : 1280B

vCMTS DUT	
Hardware	
Platform	Advantech Skylake SKY-8201L1

CPU	Intel® Xeon® Gold 6148, Dual Socket, 20C @ 2.4GHz Microcode : 0x2006a08
Memory	12 x 16GB DDR4-2667
Hard Drive	Intel® SSD (480G)
Network Interface Card	4 x Intel® Ethernet Converged Network Adapter 810 100GbE
Crypto Acceleration Card	4 x Intel® QuickAssist Technology Adapter 8970
Software	
Host OS	Ubuntu 20.04, Linux Kernel v5.4.x
DPDK	DPDK v20.08
vCMTS	Intel vCMTS Reference Data plane v20.10
Date Tested	July 9th, 2021

vCMTS Traffic Generator	
Hardware	
Platform	Intel® Wildcat Pass S2600WTTR
CPU	Intel® Xeon® E5-2699 v4, Dual Socket, 22C @ 2.2GHz Microcode: 0xb000036
Memory	6 x 16GB DDR4-2400
Hard Drive	Intel® SSD (480G)
Network Interface Card	4 x Intel® Ethernet Converged Network Adapter 810 100GbE
Software	
Host OS	Ubuntu 20.04, Linux Kernel v5.4.x
DPDK	DPDK v20.08
Traffic Generator	DPDK Pktgen v19.10

Appendix B – AWS c5n Test Configuration

Test Environment Configuration Information and Relevant Variables	
CM Lookup & Subscriber Mgmt	300 subscribers per service group, 4 IP addresses per subscriber
DOCSIS Filtering	6 filter groups, 2 filter groups associated with each cable-modem 16 filter rules per filter group 10% matched, 90% unmatched (default action – permit)
DOCSIS Classification	16 rules per subscriber, 10% matched - enqueue to one of 3 service-flow queues 90% unmatched - enqueue to default service-flow queue
Downstream Service-Flow Scheduling	8 service-flow queues per subscriber (4 active)
Downstream Channel Scheduling	6 x OFDM (1.89 Gbps) Channels, 2 x channel-bonding groups. Or 2 x OFDM (1.89 Gbps) and 32-SC-QAM (42.24 Mbps) Channels, 4 x channel-bonding groups NOTE: channel-bonding groups are distributed evenly across cable-modems
Upstream Bandwidth Scheduling	Upstream Scheduler not used. Upstream bandwidth pre-allocated in grants of 2KB per service ID. Bandwidth grants balanced evenly across 300 cable-modems.
Ethernet CRC	Downstream: 100% CRC re-generation Upstream: 0% CRC verification NOTE: CRC relates to inner frames
Encryption	100% AES, 0% DES
Packet IMIX Distribution	Upstream 65% : 84B, 18% : 256B, 17% : 1280B Downstream 15% : 84B, 10% : 256B, 75% : 1280B

vCMTS DUT	
Hardware	
Platform	AWS c5n family
CPU	Intel® Xeon® Platinum 8124M, Dual Socket

Memory	Up to 196GB, depends on instance type
Hard Drive	N/A
Network Interface	Up to 100GBE, depends on instance type
Software	
Host OS	Ubuntu 20.04, Linux Kernel v5.4.x
DPDK	DPDK v20.08
vCMTS	Intel vCMTS Reference Data plane v20.10
Date Tested	July 30th, 2021

vCMTS Traffic Generator	
Hardware	
Platform	AWS c5n family
CPU	Intel® Xeon® Platinum 8124M, Dual Socket
Memory	Up to 196GB, depends on instance type
Hard Drive	N/A
Network Interface	Up to 100GBE, depends on instance type
Software	
Host OS	Ubuntu 20.04, Linux Kernel v5.4.x
DPDK	DPDK v20.08
Traffic Generator	DPDK Pktgen v19.10

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Convergence of Services Using Network Slicing

A Practical Implementation

A Technical Paper prepared for SCTE by

Fernando X. Villarruel
Chief Architect, MSO Practice
Ciena Corporation
1185 Sanctuary Pkwy, Alpharretta
fvillarr@ciena.com

Geoff Eaton
Senior Solutions Architect
Ciena Corporation
2351 Alfred-Nobel Blvd, Saint Laurent, Quebec
geaton@ciena.com

Marco Naveda
Senior Director, Architecture
Ciena Corporation
3500 Carling Ave, Ottawa, Ontario
mnaveda@ciena.com

1. Introduction

By now the convergence of service signals over one digital MSO network needs no further motivation. Over the last some years there have been resources to describe the financial and engineering benefits of this transition (Chamberlain, 2017), (BAUMGARTNER, 2019), (Chamberlain, 2017).

In this paper we turn to the mechanism for creating the right form of infrastructure to make convergence an implementable reality. According to (Chamberlain, 2017) there are three requirements to make convergence successful: powering, backhaul and site availability. We find this a good generalization where power refers to the active power availability in the MSO outside plant, and site availability implies the fiber, coax, and pole access real estate to work with. In this paper we focus particularly on the backhaul section for successful convergence, which includes the connectivity of hubs and customer endpoints with the recent differentiator of digital fiber, see Figure 1 below.

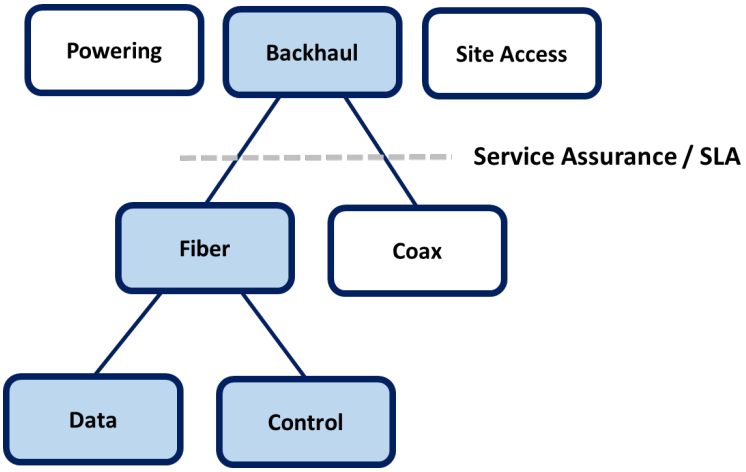


Figure 1, Infrastructure Needs for Convergence

This paper more specifically focuses on the transmission and management of multiple services over the fiber infrastructure, otherwise known as the converged interconnect network (CIN), spanning the fiber access endpoints or aggregation points to the signal packet cores at hubs or headends (Villarruel, 2020). While our focus is on the fiber networking, we go beyond to point out general implications and expectations for service assurance—code word for service layer agreements (SLA). The combination of transmission, management and sensitivities to service assurance is what drives the need for network slicing, thus the need for a paper that covers this topic.

Network slicing as a principle to streamline convergence of services for the cable industry has been described in useful detail in the reference paper: “Framework For Convergence Of Services On The MSO Network Using the Principles of Network Slicing” (Villarruel, 2020). This paper uses this reference paper as a launching point and further shows a practical implementation of network slicing on multiple services implemented in a laboratory setting.

We point out two disclaimers. First, we refer to network slicing as a general term, but the term network slicing also has a particular meaning within the 5G mobile space, as standard bodies like 3GPP have specific definitions for qualities of network slices for specific subservices. We use the generalized definition of network slicing because we can apply similar principles to a new space, in this case the MSO industry. Secondly, the work we show here is descriptive and not prescriptive in that there can be other ways to organize and share network resources and enforce service assurance.

2. Network Overview

The purpose of this paper is to position an end-to-end digital network that has the right sensitivities for convergence. At a high level we partition the converged network in two sections, access aggregation and core network transmission. See Figure 2 below. We position Flexible Ethernet (Flexed) technology for access aggregation and segment routing (SR) for the core network transmission. We also describe the principle of a Transition Map between FlexE and SR. This paper aims to describe the methodology for network slicing in the fiber backhaul and finally present a proof of concept from a lab buildout of the described principles.

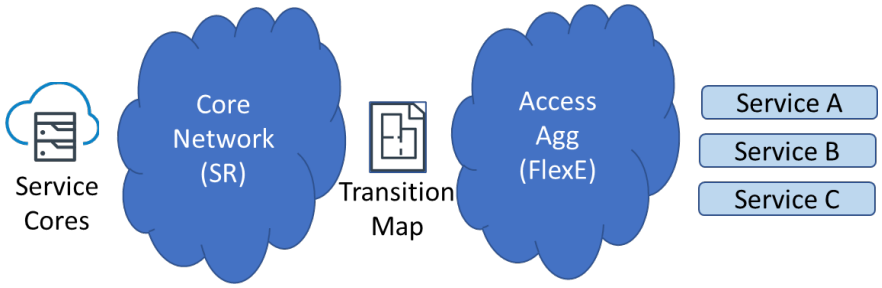


Figure 2, Network Overview

3. Convergence and Network Slicing

It is useful first review what we mean by service convergence and its relation to network slicing. Network convergence has been shown to include several access technologies with several services and subservices embedded, and span outside plant access and core networks, as seen in Figure 3 below (Villarruel, 2020).

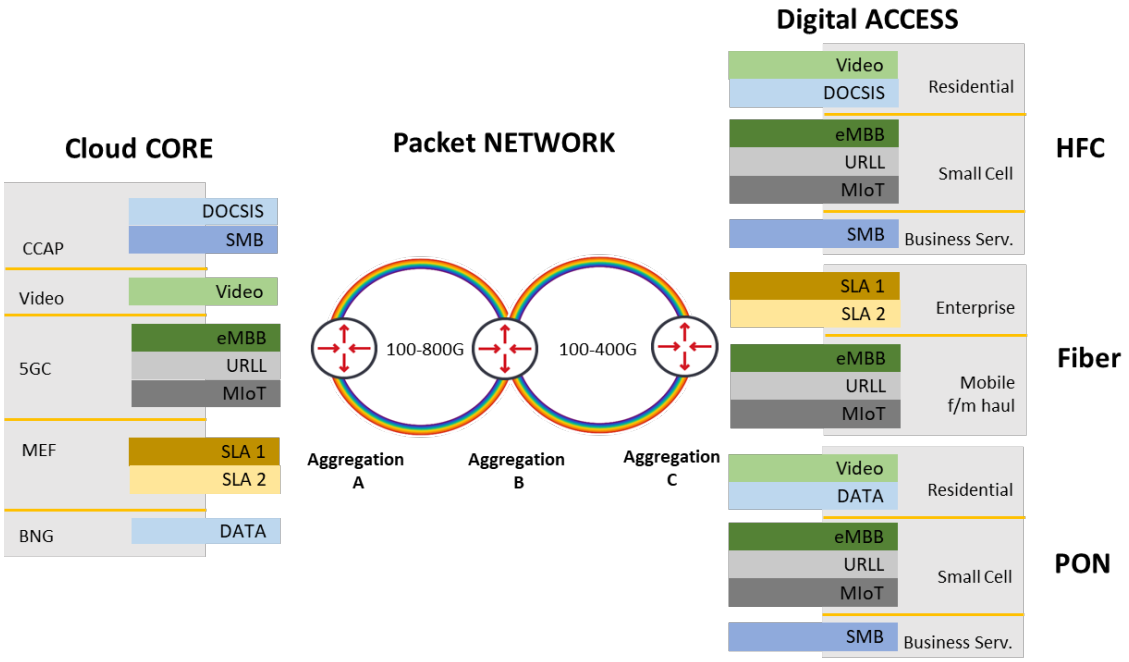


Figure 3, Convergence and Network Slicing

Notice that while Figure 3 is the view of an end-to-end system that works together, there is a delineation of domains, as each would have its own functions and expectations. See Figure 4. In the context of network slicing we call these resources the access would have its own resources, the network domain would have its resources and the core would have its resources. Connecting them might be a global orchestrator, but it is worthwhile to evaluate principles for each section in its own light. As alluded to in the introduction, the focus of this paper is on the network domain.

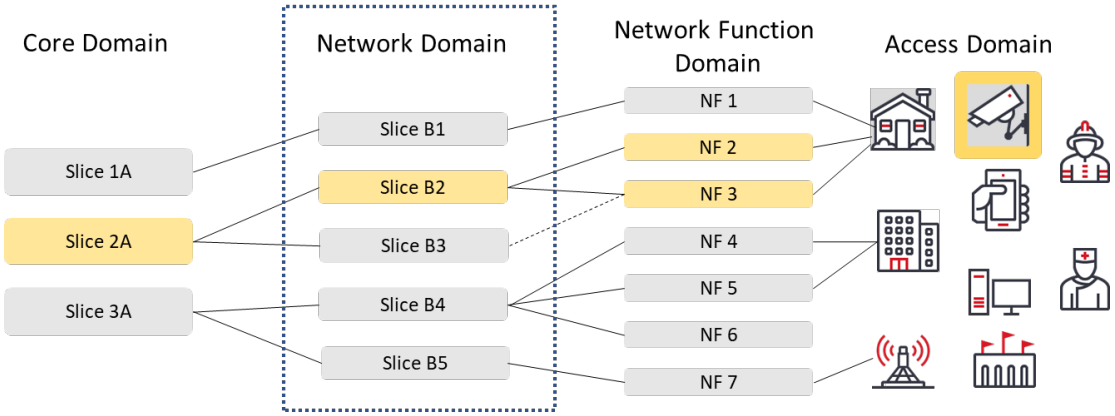


Figure 4, Slice Domains and Functions

Formally network slicing is: "a method to run multiple end-to-end logical networks on a common set of resources." But within and end to end framework there are domains which also have their own slices, so end-to-end network slicing is really a patch of slices from each domain to make a particular service. Using the example shown in (Villarruel, 2020), Figure 4, where each domain has a partition of resources that create domain slices, a subservice would use a unique collection of slices from each domain for end-to-end transmission of a subservice, otherwise known as a service blueprint. Consider the example

shown of a residential camera subservice, which might need a network function of face recognition and encryption, a tunnel from the network and a unique profile from a service core. The emphasis of this paper, if we take Figure 4 as reference, are the slice types available and created in the network domain. In this case the network domain is specifically the Ethernet and IP resources available to MSO systems that have evolved to a digital fiber plant, with the advent of distributed access architectures.

For a converged system it takes some effort to create the slice profiles and instantiate services as a collection of slices and manage their evolution through a lifecycle, but the primordial challenge is how to share sources in a way that allows legacy networks the same service assurance that they've had as independent networks. This is the fundamental challenge and so it must be addressed first, before any high-level orchestration or management is determined. The question is then, how does one converge services in a digital network such that they remain autonomous? The methodology for network slicing answers this question and is thus the kernel of this paper. We particularly revolve our work around utilizing concepts of soft and hard slicing, where hard slicing describes mechanisms that are rigid in how resources are dedicated, and soft slicing refers to mechanism that are made for flexibility in the sharing of resources. For the application of these hard and soft slicing concepts we leverage Flex Ethernet (Flex-E) and managed segment routing (SR) respectively. We now take a focused view of Flex Ethernet and Segment Routing.

4. Flexible Ethernet

Flexible Ethernet, otherwise known as Flex-E, is a mechanism that was created to provide transmission rate flexibility between client and line signals. This capability addresses the variety of client rates new and installed, along with line side transmissions and their ability to work together in a network. Flex-E dissociates the client rate from the physical interface. It does this by adding a shim layer to the 802.3, 100GBASE-R protocol stack and the 802.3bs 200/400GBASE-R protocol stack, separating the MAC from the PCS and PMA layers. See Figure 5.

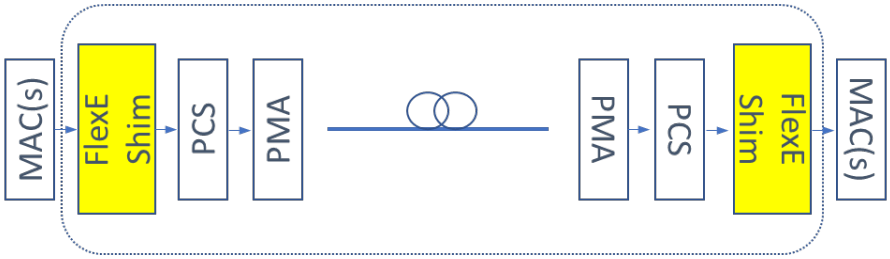


Figure 5, Flexible Ethernet Protocol

More generally, the Flex-E shim layer allows for not just one client MAC, but various, such that a collection of various clients at different rates can be transmitted over one or more Flex-E channels, see Figure 6. Specifically, FlexE clients have their own separate MAC, reconciliation layer and MII above the FlexE Shim layer, while layers below the PCS keep the intact rates as specified for their own Ethernet implementation (Optical Internet Working Forum, 22).

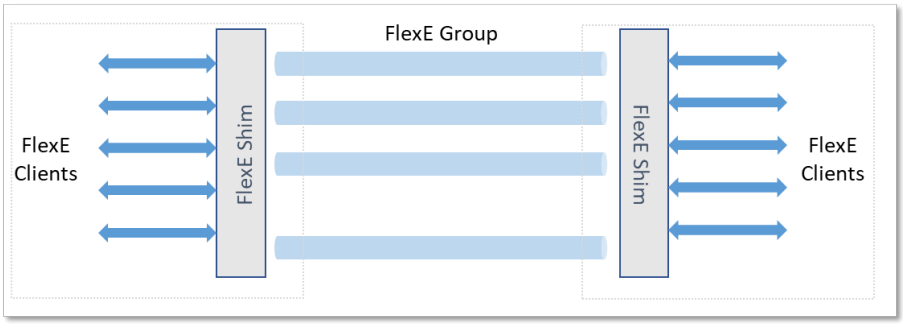


Figure 6, Clients and FlexE Group

The shim layer is shown in more detail in Figure 7; it shows the muxing function of a shim layer, where a collection of clients present themselves in a serial 66 bit stream, created from an ethernet MAC, operating at 10, 40 or mx25 Gb/s. Entering the calendar function, the client signals are adapted to the FlexE channel rate by a clock transition and an insertion or deletion of idle blocks. The calendar consist of 66 bit based time slots, and the control function controls the sequential insertion of client blocks to the calendar, along with insertion of necessary overhead. The subcalendars are logical groupings of calendar slots that prepare for interleaving that puts together signaling for the expected rate of the FlexE channel. Show in the Figure 7 is the description of the muxing function. There is also a demuxing function which is effectively the inverse of the muxing function.

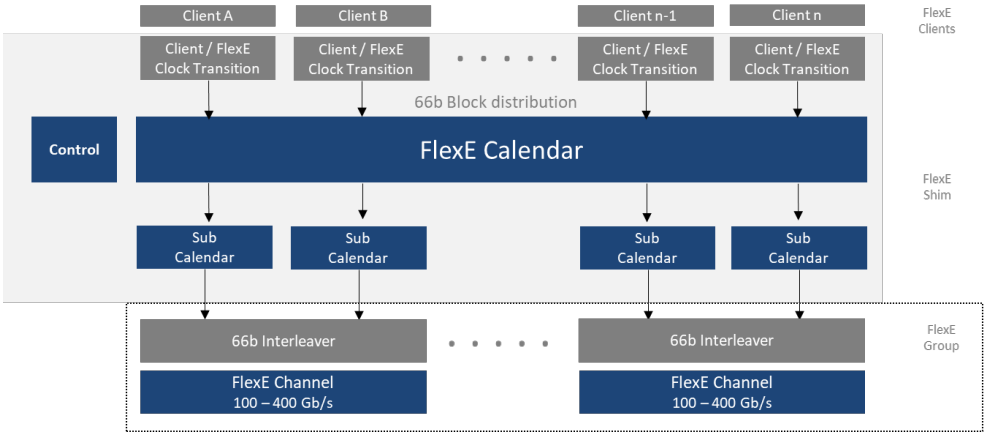


Figure 7, FlexE Shim Layer

The functionality of the FlexE calendar is interesting and a full treatment is beyond the focus of this paper, but as an example Figure 8 shows the calendar function in more detail for a 100Gb/s FlexE channel. Here the calendar has a granularity of 5G slots and every 20n slots create the sufficient 100Gb/s signal for a subcalendar. This calendar shows the serially inserted bitstream of mixed PHY clients at 10Bb/s and 25Gb/s. Similarly, the calendar can have up to 25G slots, a quality more natural for 25 G clients and greater speed FlexE channels at 200 or 400 Gb/s.

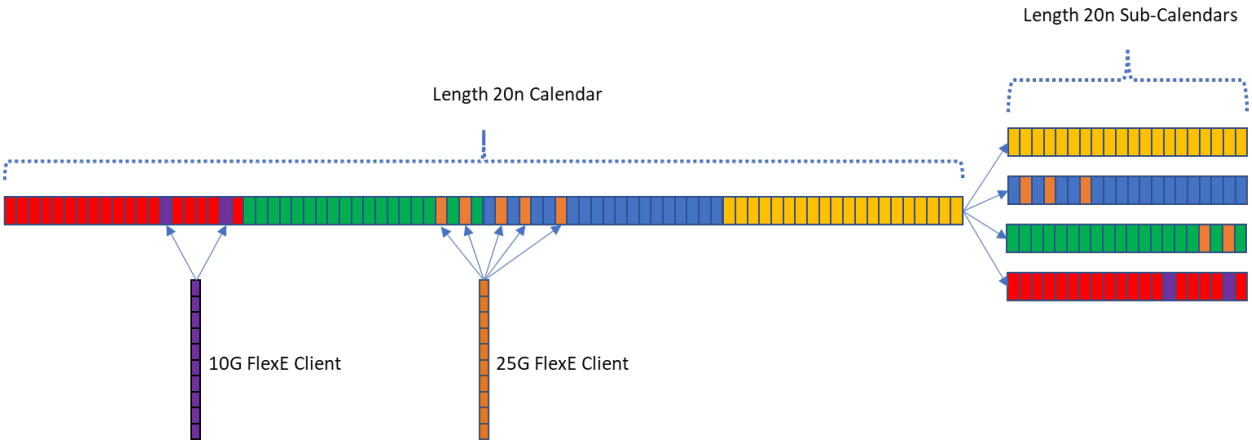


Figure 8, FlexE Calendar

With the presented characteristics FlexE has the capability of three basic functions, as shown in the Figure 9. Bonding allows the user to create a higher rate service, like 400G, with a collection of lower rate interfaces. The example shows 2, 100G Ethernet PHY signals carrying a 200Gb/s service. This is sometimes applied as an alternative to LAG without the extra layer of hashing algorithms. Subrating, service rates that don't match a greater service to be carried over on interface of a larger channel. More specifically the transmission rate is downsized in order to match the necessary client rate. In the example we show a 75G service utilizing a 100GB/s ethernet interface. And finally, there is the capability of channelization. which aggregates lower rate channels onto a higher rate interface. In the example we have clients of various native rates that manage to use one 100G FlexE channel. It is the case that these capacities can be mixed and matched depending on the circumstance.

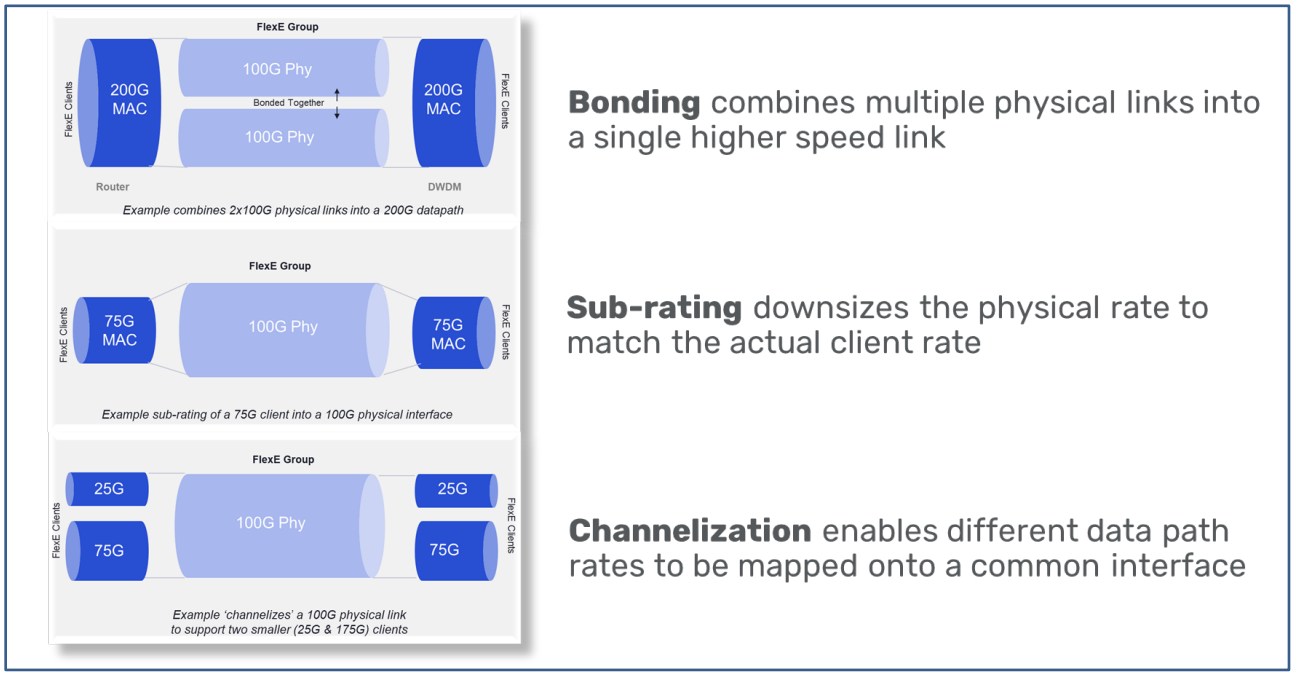


Figure 9, FlexE Functions

For our MSO convergence discussion channelization allows for the transmission of 10 and 25Gbs signals to leverage naturally higher rate backhauls, without having to go through an extensive forwarding plane, which is particularly useful in low footprint low energy platforms, such as outdoor shelves and nodes. Thus, FlexE is a good resource for the access outside plant, creating a robust combining mechanism that resembles straight forwarding muxing functions, like OTN, but with a more streamlined overhead.

But if FlexE is a good option for the outside plant access, what about the core network? This brings us to our next portion of end-to-end network slicing—segment routing. SR is an IP function which allows us to present a solution for soft network slicing.

5. Segment Routing – Data Plane

Segment routing is a topic that has been well described in other resources so here is a brief treatment reviewing basic principles (Bonica, 2017). SR has also started to show merit specifically in MSO networks as a way to enhance operational efficiencies for business services delivery (Yeo, 2019). SR per the name focuses on breaking down network topologies into segments, where a segment is a path between two routers. As seen in Figure 10, the router segments can be between two nearest neighbors like (AB) and (BC), and only need local awareness between nearest neighbors. Or segments can also describe routes that span across routers like (AC), and these segments need global awareness beyond just neighboring routers—these segments identifiers (SID)are described as adjacent SIDs or binding SID's respectively. Routing protocols can then be applied to segments as is done with regular routing that relies on unique addressing of source and destination. The routers that participate with capability to recognize SIDs and routes accordingly is called the SR domain.

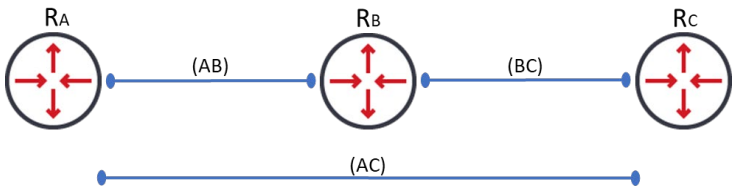


Figure 10, Segment Routing IDs

There are two methods for encapsulating SID's, one is to use labeling space in the MPLS header, and the other is for to use an IPv6 header with an SR extension. In either case the principles for inclusion of labels in headers throughout the life of a packet is the same. SR aims to keep the route of a packet known to the packet itself and reduce the calculation of routes as is done in legacy routing. This occurs by encoding the path of a packet with a stack of labels that describe the route to take and for the most part avoids large route calculations at every hop, as is done in regular routing by analyzing IP addresses and deriving paths from large look up tables.

In Figure 11 we show a very simple example of a six routers SR domain with a collection of local SID's, advertised and known only to individual nearest neighbors, and binding (global) SID's advertised and known to all (which allows packets to skip treatment at certain hops.) In Figure 12 and Figure 13 we show SR label stacks and accompanying payload progress through a route. In Figure 12 an instance using only local SID's shows a path from router Ra to router Rf. Note how the ingress packet from Ra has the whole path encoded on its label stack. At Rb, the router pops off the incoming label and forwards

the packet onto its next location, at Rc the router pops off the incoming label and forwards the packet onto its next location, and so on until it reaches Rf, where only the payload is left as was desired. In Figure 13 we note that the labels used include global SID's, where label processing skips over Rb and Re, with a local label between Rc and Rd. Here Rb and Re pass the packet onto the destination known globally without any manipulation of labels. At Rc and Rd however the label stack is depleted on route to Rf where it terminates as desired. Because of this simple approach, SR relies heavily on Interior Gateway Routing protocols, which simplify the necessity to know routes beyond local networks. There are other nuances and complications to label stack building and recognition, but this example highlights the general interworking of SR.

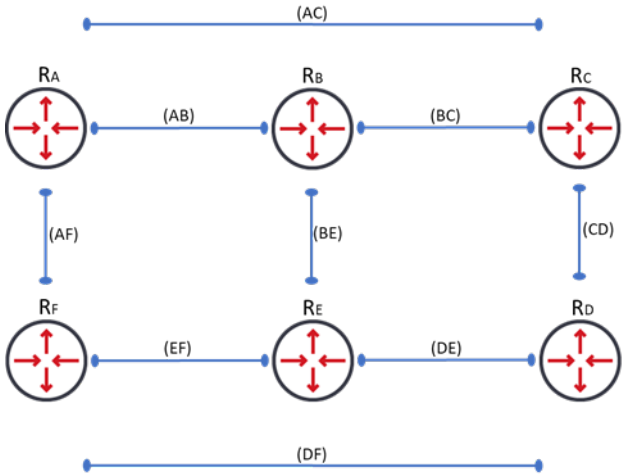


Figure 11, Routing Mesh and SID assignments

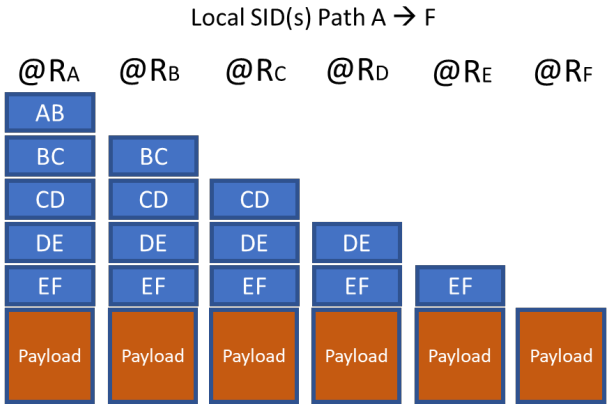


Figure 12, Local SID Path Progression

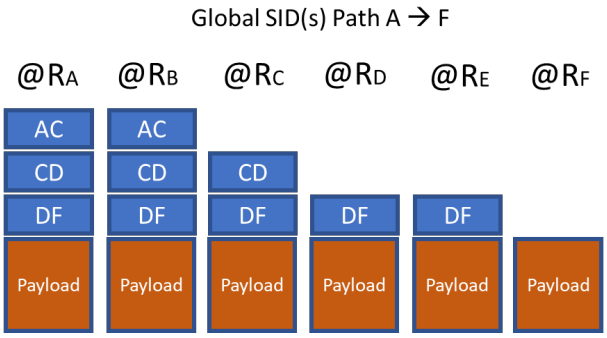


Figure 13, Global SID Path Progression

6. Segment Routing – Control Plane

In parallel to the data plane in SR there is also a control plane that calculates paths and decides what segments to visit, then communicates to routers instructions on how to build label stacks. The methodology used to control SR could allow us the ability to create an umbrella of reliable service assurance, which manifests itself as sellable SLA's, thus its importance. The decisioning of paths is done in path compute engines (PCE) where the cost of segments is evaluated according to variable state qualities such as latency or congestion, or other metrics. As shown in Figure 14, there are two schools of thought for location of path computation engines. One is to have path computation local, where each router participates in computing route characteristics and contributes cost of segments to a database that is available to all IGP participants. Each router then has the ability to influence SR labeling accordingly. The second school of thought is to have a centralized path compute engine which peers into networks and collates a global view of segment constraints based on telemetry and external influence and thus can create a wholistic cost of routes. The centralized PCE then has responsibility, not the local routers, for how label stacks are built and so directly controls the flow of information throughout the network.

For network slicing a centralized PCE, or "off-board" PCE is preferred because it can execute policy structures that could be missed if left to generic SR support mechanisms like Topology Independent Loop-Free Alternate (TI-LFA). A centralized PCE then takes control of creating SR stacks and distributing to origination points without burdening any path responsibility to middle hop routers. This is a particularly efficient way to meet service SLAs, particularly as SLA contracts can include certain non-typical networking constraints. For example, consider an example that has been considered recently, where a sensitive government service cannot cross country borders. If left to on board PCE it would be difficult to maintain data from crossing borders as pure networking mechanisms would not naturally digest that policy. With a centralized view of routes and costs expectations outside of typical network activity this important policy can be met.

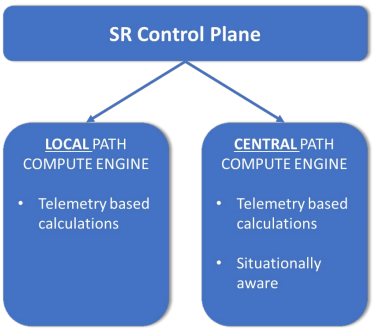


Figure 14, SR Control Plane Options

7. Transition Mapping

We have shown the operating principles of FlexE and SR, and while these are powerful independently in the case of a converged end-to-end network there needs to be an elegant and efficient, service aware, transition between the FlexE and SR domains, as shown in Figure 2. Figure 15 shows in more detail and how the Transition Map works in the context of an end-to-end network as it converges services. For Figure 15 this example we borrow some of the typical services and subservices listed in Figure 3. The transition function is a mapping between established FlexE channels and segment routed flows, where channels and flows have specific meaning and relationship to services and subservices carried by them from customer endpoints and handed off to service packet cores, then vice-versa.

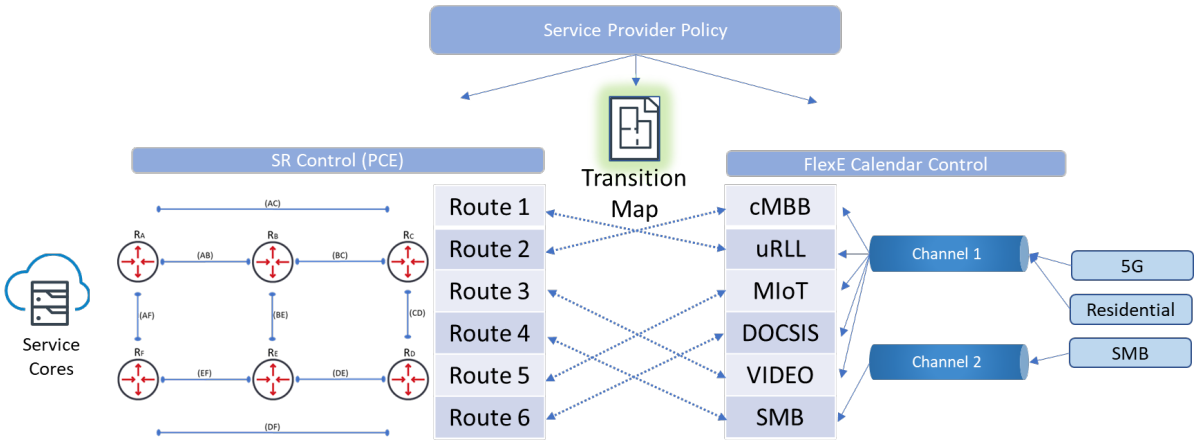


Figure 15, Transmission Mapping and Context

The transition function map executes a policy determined by the service provider which describes the relationship of subservices to SR labels and services to FlexE channels, in accordance to known SLAs. Note that the function map is intimately connected to the PCE and Calendar control functions, as they shed light to policy on how the map should be built.

In practice the transition function can be statically provisioned or dynamic. In a static set up very specific FlexE channels map directly to particular IP flows with variations over time executed by hand. On the other hand, in a dynamic state the mapping between Flex channels and SR labels could change depending on situations that allow services to best meet their SLA—again this points to the relationship of PCE and Calendar control with the policy mechanism and the transition map.

Note that SR and FlexE are well defined processes and mechanisms while the transition map is not. This implies that vendor differentiation is created around management of policy via PCE and Calendar control, and importantly the robustness of the Transition Map, how it is built and how it evolves as the system changes.

As an example, in the world of 5G, where the original work for formally defining network slicing, there is a standard list of attributes that define a slice, and that list is populated according to the particular needs of a service thus creating particular network slices, which then get assigned to subservices (GSMA, 2016). The service provider policy shown in Figure 15 would have an understanding of these slice definitions and their particular attributes and express this to the centralized PCE, the FlexE Calendar control and the transition Map where the policy is ultimately executed. In the MSO world, this is ripe definition work for CableLabs, where MSOs could benefit from Network Slicing attribute tables which could readily be assigned to subservice types.

8. Laboratory Set Up

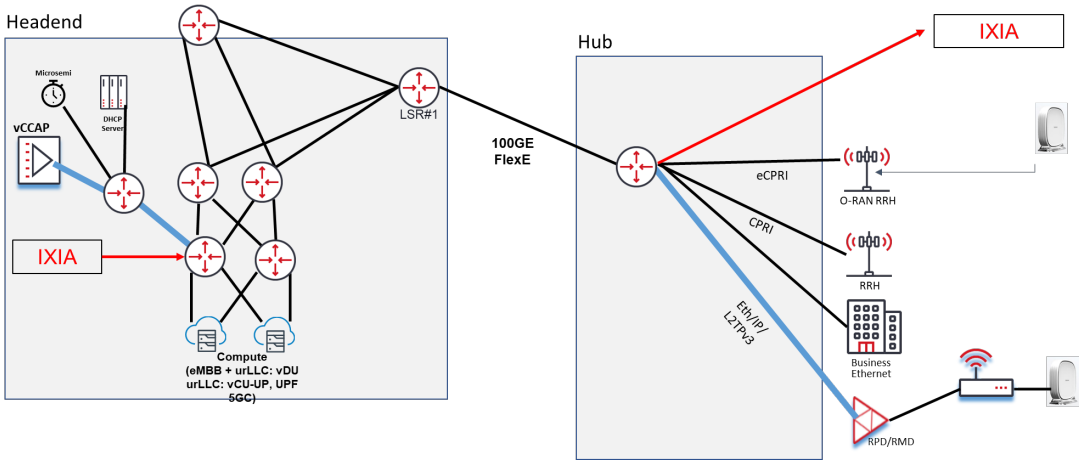


Figure 16, Planned Laboratory Setup

Figure 16 shows the planned laboratory setup with a convergence of mobile, residential and business services. Note that while this set up is built for this paper, it is also part of a more general laboratory set up for convergence in general, so some of the portions go beyond what would be needed for our testing, but useful in to verify other architecture questions.

The setup represents and-end-to end network with equivalents of outside plant, Hub and Headend. At the Headend there are the packet cores and supporting applications. There is cloud implementation of a 5G core and accompanying fabric, not in focus here. There is a virtual CCAP from a leading provider and supporting IEEE 1588 Master clock running ITU-T G.8275.2, and a DHCP server. There is also two entry routers to the Headend, an edge router located at the top, and the LSR#1 which is of our focus.

At the outside plant equivalent there is mobile signaling, from legacy backhaul in CPRI and next generation fronthaul in eCPRI supporting macro and microcells respectively. There is a point-to-point ethernet connection representative of business services. There is also a cable modem feeding into a Remote PHY Device, which ultimatately terminates at the CCAP. There is also an Ixia signal generator present to add rogue traffic which assists in testing the robustness of the FlexE and SR tunneling mechanisms.

We show that the service signals coincide on an aggregation switch, at the hub, which handles the execution of FlexE channels for all services. The aggregated signals are backhauled from the Hub to the Headend on a 100G signal, which is spent according to the FlexE channelization. The router at the headend LSR#1 terminated the FlexE channels and using a transition map executes service or subservice signals onto specific SR labels and flows, which have their own quality of service COS assignments. The flow prioritization is kept throughout the rest of the SR domain. Note that the signals are taking a non-direct path after LSR#1, as would be determined by typical routing protocols. This is to show the deliberate nature of the SR policy.

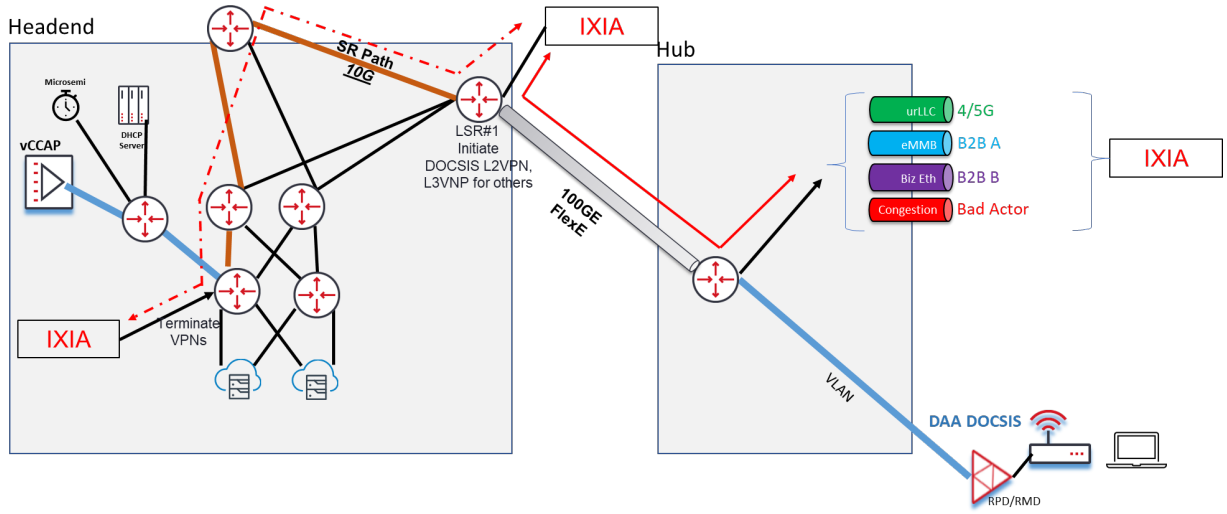


Figure 17, Actual Laboratory Setup

In Figure 17 we show the actual laboratory set up, which is a functionally networking equivalent to what is shown in Figure 16. Figure 17 shows the residential cable signaling is the only service that is running specifically on its native endpoints and cores. The system on the other hand mimics mobile and Business Ethernet flows via an Ixia traffic generator at the Hub. There is a 4/5G flow to mimic a signal from a wireless network source. The B2B-A and B2B-B signals represent enterprise grade circuits and the Ixia also provides a congestion flow to stress the FlexE channel—shown in the solid red line. In the Headend, at LSR#1 there is a breakout to another Ixia. This Ixia temporarily evaluates the FlexE signals and is the permanent stop for the congestion signal added to test the FlexE group. It is also the source for another "bad actor" signal into the SR domain in order to stress its class of service (CoS) mechanism—shown in the dot and slash red line. The LSR#1 is also where VPNs are originated, L2VNP for the vCCAP and RPD circuit, along with an L3VPN for the other signals. This is done to mimic typical business grade scenario in the case of L3VPNS, and in the case of DOCSIS signaling because we did not want to disturb same network settings for the vCCAP and the RPD. Note, that the FlexE segment relies on a 100G coherent direct connect signal, while the SR segment first path uses a 10G signal, for reasons that will be apparent in the next section. The VPNs are terminated ate the router before the last Ixia, through the representative cloud fabric.

The reason for differences between Figure 16 and Figure 17 was shortened set up and test time, a result of product transit, availability, and laboratory access in our current state with a pandemic.

9. Signal Settings and Test Results

	FlexE Settings	
Service Type	FlexE Group (Gbps)	FlexE Channel (Gbps)
B2B - A	100	5
B2B - B		5
4G/5G		5
DAA - DOCSIS		5
Bad Actor – FlexE*		75

Figure 18, FlexE Channel Settings

Figure 18 describes the FlexE bandwidth settings for each service type: Each service is assigned 5Gbps hard slices, and the "bad actor" congestion circuit is assigned 75Gbps, for a total composite of 95Gbps of client traffic and one backhaul signal and interface of 100Gbps.

Traffic Type	Tx (Mbps)	Rx (Mbps)
B2B - A	2G	2G
B2B - B	3G	3G
4G/5G	4G	4G
DAA - DOCSIS	25Mbps	25Mbps
Bad Actor	90G	75G

Figure 19, FlexE Throughput Results

Figure 19 shows the transmit and receive results for the Flex Ethernet group signal. We note three items. First when a bad actor that has a hard slice limit of 75G, as shown in Figure 18, tries to use more than its assigned SLA its signaling is curtailed back down to 75G, as determined by the FlexE channel setting, and second in this exercise of limiting the bad actor, none of the other service signals are affected. This is as expected. Third we note that the DAA DOCSIS signal is very low. This is per the one modem setting we had on the vCCAP. With more time we would have built a more extensive system. As it stand, with regards to bandwidth scale, the DDA DOCSIS signal is noise.

Service Type	DSCP	MPLS EXP	SR CIR	SR EIR (10G)
B2B - A	0 (Routine)	0	2G	8G
B2B - B	24 (Flash)	3	3G	7G
4G/5G	46 (Critical)	5	4G	0G (no burst)
DAA - DOCSIS	32 (Flash Override)	4	200Mbps	9.8G
Bad Actor – SR	1 (Priority)	1	200Mbps	9.8G

Figure 20, CoS Settings in SR Domain

Figure 20 shows the CoS settings for the service signals in the SR domain. The priority settings are expressed in the IP domain following the Differentiated Services Code Point (DSCP) structure. The MPSL priority settings are expressed in the MPLS Experimental bits mechanism (EXP), from lowest to highest. Basically, the MPLS priority structure can reflect the IP priorities to expedite any packet processing without digesting any IP headers. Note, the DCSP uses six bits and 64 possible values. The MPLS EXP on the other hand uses only three bits, and generally reflects the first three bits setting of the DCSP setting. Note, this allows for a further DSCP structure of priorities for subservices at the customer domain, to use as necessary.

Figure 20 also shows the bandwidth assignments for the different signals. The Committed Information Rate (CIR) shows the minimum bandwidth that the given signal will receive, this is typically a contractual SLA value. The Excess Information rate is the burstable rate possible, and it is the difference between the capable signal interface and the CIR. Note that the SR segment in egress from SLR#1 is 10G, thus the EIR values shows are calculated at 10G minus CIR. This is the nature of a flexible network slice as discussed earlier in the paper.

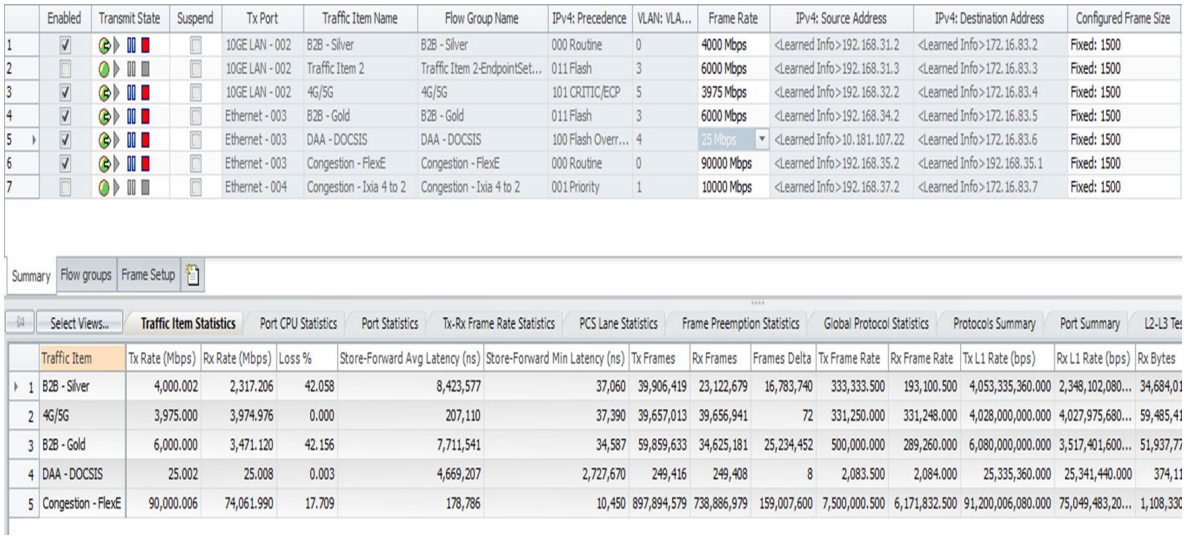
Traffic Type	Tx (Mbps)	Rx (Mbps)
B2B - A	2G	2G
B2B - B	6G	3G
4G/5G	4G	4G
DAA - DOCSIS	25Mbps	25Mbps
Bad Actor	10G	800Mbps

Figure 21, Transmission Results with SR CoS and FlexE

Figure 21 shows the end-to-end transmission results for the SR and FlexE segments. The focus here however is the priority structure set of the SR domain. Note the “Bad Actor” signal, which was assigned 200 Mbps, is trying to push 10G through the system which would take up all the line rate value. The output however shows the limiting nature of the soft slice as the system determined only 800 Mbps worth

of capability. Also note the B2B-B signal, which has a limitation of 5G in the FlexE domain, and a limitation of 3G in the SR domain. When it tries to put through its 6G signal, it gets throttled to 3G. Also note, that the DOCSIS signal is unaffected. Ultimately the 10G SR segment is all used up based on the decision structure set by CoS.

Figure 22 below shows a snapshot of the raw data presented above. The one item to note in the tests was latency nature of the DOCSIS signal. We conjecture that this loss value and the high latency is due to the very small nature of the DOCIS signal in this case. We describe it as being in the noise and thus prone to calculation error, which is not the case of all the other signals which show miniscule latency.



	Enabled	Transmit State	Suspend	Tx Port	Traffic Item Name	Flow Group Name	IPv4: Precedence	VLAN: VLA...	Frame Rate	IPv4: Source Address	IPv4: Destination Address	Configured Frame Size
1	✓			10GE LAN - 002	B2B - Silver	B2B - Silver	000 Routine	0	4000 Mbps	<Learned Info>192.168.31.2	<Learned Info>172.16.83.2	Fixed: 1500
2				10GE LAN - 002	Traffic Item 2	Traffic Item 2-EndpointSet...	011 Flash	3	6000 Mbps	<Learned Info>192.168.31.3	<Learned Info>172.16.83.3	Fixed: 1500
3	✓			10GE LAN - 002	4G/5G	4G/5G	101 CRITIC/ECP	5	3975 Mbps	<Learned Info>192.168.32.2	<Learned Info>172.16.83.4	Fixed: 1500
4	✓			Ethernet - 003	B2B - Gold	B2B - Gold	011 Flash	3	6000 Mbps	<Learned Info>192.168.34.2	<Learned Info>172.16.83.5	Fixed: 1500
5	✓			Ethernet - 003	DAA - DOCSIS	DAA - DOCSIS	100 Flash Overr...	4	25 Mbps	<Learned Info>10.181.107.22	<Learned Info>172.16.83.6	Fixed: 1500
6	✓			Ethernet - 003	Congestion - FlexE	Congestion - FlexE	000 Routine	0	90000 Mbps	<Learned Info>192.168.35.2	<Learned Info>192.168.35.1	Fixed: 1500
7				Ethernet - 004	Congestion - Ixia 4 to 2	Congestion - Ixia 4 to 2	001 Priority	1	10000 Mbps	<Learned Info>192.168.37.2	<Learned Info>172.16.83.7	Fixed: 1500

Summary | Flow groups | Frame Setup

Select Views... | Traffic Item Statistics | Port CPU Statistics | Port Statistics | Tx-Rx Frame Rate Statistics | PCS Lane Statistics | Frame Preemption Statistics | Global Protocol Statistics | Protocols Summary | Port Summary | L2-L3 Tes

	Traffic Item	Tx Rate (Mbps)	Rx Rate (Mbps)	Loss %	Store-Forward Avg Latency (ns)	Store-Forward Min Latency (ns)	Tx Frames	Rx Frames	Frames Delta	Tx Frame Rate	Rx Frame Rate	Tx L1 Rate (bps)	Rx L1 Rate (bps)	Rx Bytes
1	B2B - Silver	4,000.002	2,317.206	42.058	8,423,577	37,060	39,906,419	23,122,679	16,783,740	333,333.500	193,100.500	4,053,335,360.000	2,348,102,080...	34,684,01
2	4G/5G	3,975.000	3,974.976	0.000	207,110	37,390	39,657,013	39,656,941	72	331,250.000	331,248.000	4,028,000,000.000	4,027,975,680...	59,485,41
3	B2B - Gold	6,000.000	3,471.120	42.156	7,711,541	34,587	59,859,633	34,625,181	25,234,452	500,000.000	289,260.000	6,080,000,000.000	3,517,401,600...	51,937,77
4	DAA - DOCSIS	25.002	25.008	0.003	4,669,207	2,727,670	249,416	249,408	8	2,083.500	2,084.000	25,335,360.000	25,341,440.000	374,11
5	Congestion - FlexE	90,000.006	74,061.990	17.709	178,786	10,450	897,894,579	738,886,979	159,007,600	7,500,000.500	6,171,832.500	91,200,006,080.000	75,049,483,20...	1,108,330

Figure 22, Raw Data Sample From Ixia.

10. Conclusion

This paper motivates the principles for hard slicing in an Ethernet domain via Flexible Ethernet, and soft slicing in a Segment Routing domain via the CoS structure. The paper then successfully shows these principles exercised on a laboratory system with multiple types of services. The systems shown is highly representative of what MSOs are trying to achieve with one digitized access and metro network. Please contact the authors for more information and context.

The authors would like to thank Regis Mounkala, Kaylan Kanchumarthy and Tina Wu, Solution Engineers at Ciena for their lab support and dedication to excellence.

Abbreviations

3GPP	Service Layer Agreement
CIN	Converged Interconnect Network
CIR	Committed Information Rate
CoS	Class of Service
DSCP	Differentiated Services Code Point
EIR	Excess Information Rate

FlexE	Flexible Ethernet
GSMA	Groupe Speciale Mobile Association
IGP	Interior Gateway Protocol
IP	Internet Protocol
MAC	Media Access Controller
MII	Media Indipendent Interface
MPLS	Multiprotocol Label Switching
MPLS EXP	MPLS Experimental Bits
MSO	Multiple Service Operator
OIF	Optical Internetworking Forum
OTN	Optical Transport Network
PCE	Path Computation Engine
PCS	Physical Coding Sublayer
PHY	Physical Layer Implementation
PMA	Physical Medium Attachment (sublayer)
SLA	Service Layer Agreement
SR	Segment Routing
TI-LFA	Topology Independent Loop Free Alternate

Bibliography & References

BAUMGARTNER, J. (2019, April 16). *Comcast, Charter MVNO Deals Are Bad for Everyone – Analyst.* Retrieved from LightReading.com: https://www.lightreading.com/mobile/comcast-charter-mvno-deals-are-bad-for-everyone---analyst-/a/d-id/750837

Bonica, R. (2017, June 6). *A Segment Routing (SR)*. Retrieved from NANOG Archives: https://archive.nanog.org/sites/default/files/1_Bonica_Tutorial_Segment_Routing.pdf

Chamberlain, J. (2017). Competitive Advantages Of HFC Networks for. *SCTE Tech Expo 2017.* Denver: Society Of Cable Television Engineers.

Chamberlain, J. (2017, April 11). *Reality Check: Competitive advantages of HFC networks for wireless convergence.* Retrieved from RCR Wireless News: https://www.rcrwireless.com/20170411/opinion/reality-check-competitive-advantages-hfc-networks-wireless-convergence

GSMA. (2016). *Generic Network Slice Template Version 2.0.* GSM Association.

Optical Internet Working Forum. (22, June 2018). *Flex Ethernet 2.0.* Retrieved from oiffourm.com: https://www.oiforum.com/wp-content/uploads/2019/01/OIF-FLEXE-02.0-1.pdf

Villarruel, F. (2020). Framework For Convergence Of Services On The MSO Network Using The Principles of Network Slicing. *Cable Tech Expo.* Virtual: Society Of Cable Television Engineers.

Yeo, E. (2019). Segment Routing Proof of Concept for Business. *SCTE Cable Tech Expo.* New Orleans: Society of Cable Television Engineers.

Creating Confidence Among Subscribers Faced with Growing Cyberthreats

A Technical Paper prepared for SCTE by:

Bruce Van Nice
Senior Product Marketing Manager
Akamai
Santa Clara, CA
650-575-4008
hvannice@akamai.com

1. Introduction

Today people depend on the Internet for just about everything: transacting, interacting, learning, traveling, and a multitude of other activities. The Internet also allows businesses of all sizes to reach new and larger markets and provides opportunities to work more efficiently by using digital tools that are increasingly economical and targeted at organizations with limited technical expertise. One of the downsides is theft of digital assets has become commonplace and for many organizations it represents a greater threat than physical theft.

Remote work challenges that began in 2020 continue as cybercriminals take advantage of elevated stress and unprecedented focus on topics related to the pandemic to maximize the value of their exploits. They saw a good opportunity to launch attacks and profit from dependence on the internet and technology. They used phishing and other kinds of internet-enabled fraud to target everyone from workers searching for personal protective equipment to families looking for information about help paying bills, and many others.

For ISPs there are opportunities to help subscribers navigate the always changing security landscape with new services that deter internet threats and are easy to use. This paper will talk about the product development process behind a security service targeted at Small Midsize Businesses (SMB) that was launched across a major North American cable network. It will:

Summarize the threats internet users face which motivated product teams to evaluate potential services

Outline a service strategy developed to help small and midsize businesses deter internet threats

Offer perspectives on the threat intelligence used to support the new security service

Discuss go to market to build awareness of security services and extend the value proposition of reliable high speed internet access

2. Internet Users Face a Diverse Threat Landscape

Cybersecurity is a visible topic today, news reports in mainstream media highlighting internet related crimes have become commonplace. The news is fueled by a multitude of organizations that track and analyze malicious activity - there's lots of data about cyber threats.

The US Federal Bureau of Investigation is a widely respected organization that runs the Internet Crime Complaint Center (IC3) whose mission is "to provide the public with a reliable and convenient reporting mechanism to submit information to the FBI concerning suspected Internet-facilitated criminal activity, and to develop effective alliances with industry partners."

The IC3 tabulates the data they receive and publishes it in the Internet Crime Report every year. The chart on the following page, copied from the 2020 report[1], summarizes the year over year trends from 2016 to 2020.

[1] https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf

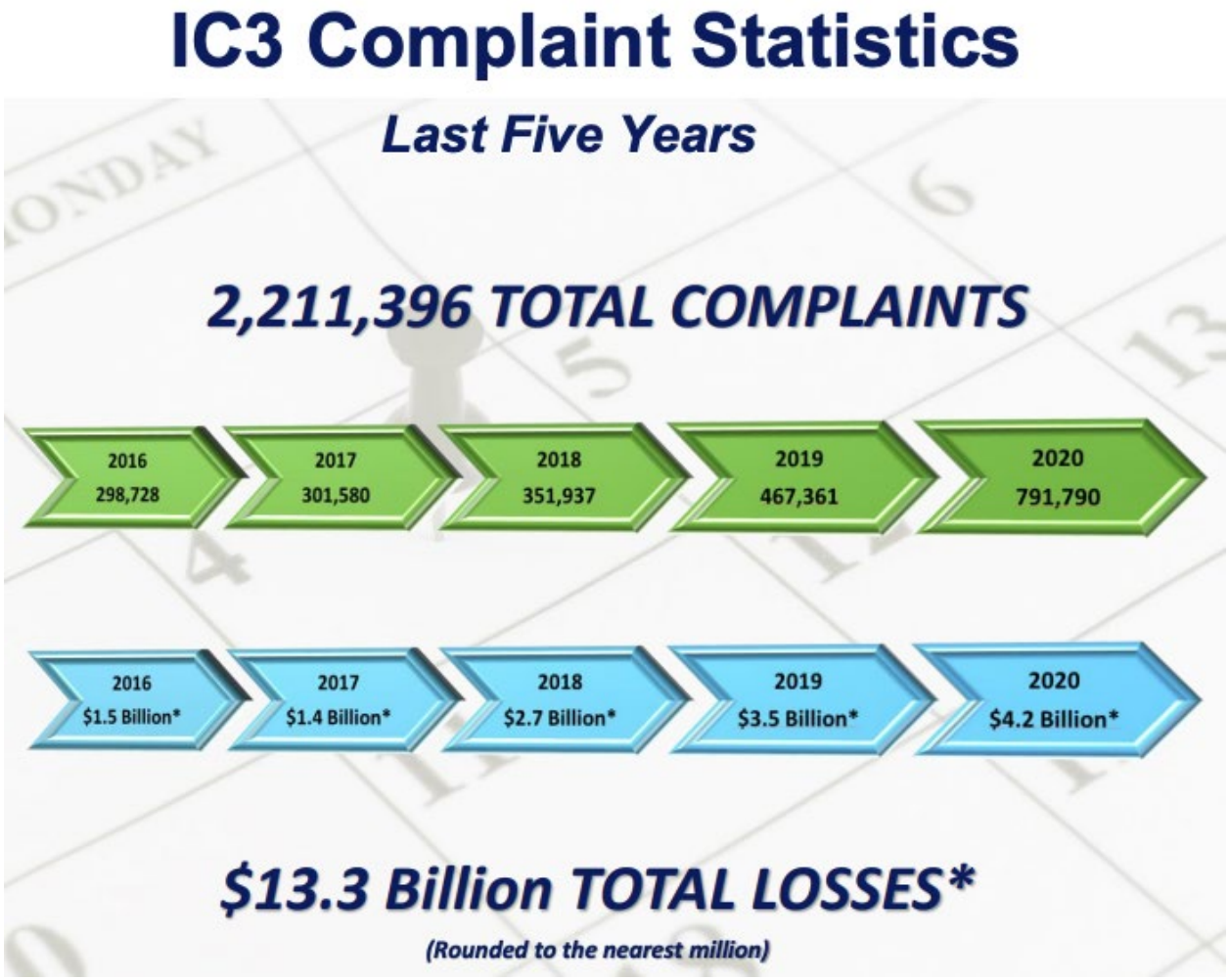


Figure 1 - Internet Crime Complaint Center (IC3) Internet Crime Report 2020 Summary of Complaints Received

The 791,790 complaints received from the American public in 2020 were a new record and represented a 69% increase over 2019. Reported losses exceeded $4.1 billion, a 20% increase over the previous year. The report also states: “These criminals used phishing, spoofing, extortion, and various types of Internet-enabled fraud to target the most vulnerable in our society - medical workers searching for personal protective equipment, families looking for information about stimulus checks to help pay bills, and many others. “

Another chart from the report summarizing threat vectors year over year is below. The 241,342 complaints tabulated for phishing scams more than doubled from 2019. To add color to the data and show how cyber criminals observe trends and tailor their exploits accordingly. The IC3 also received over 28,500 complaints about scams related to COVID-19.

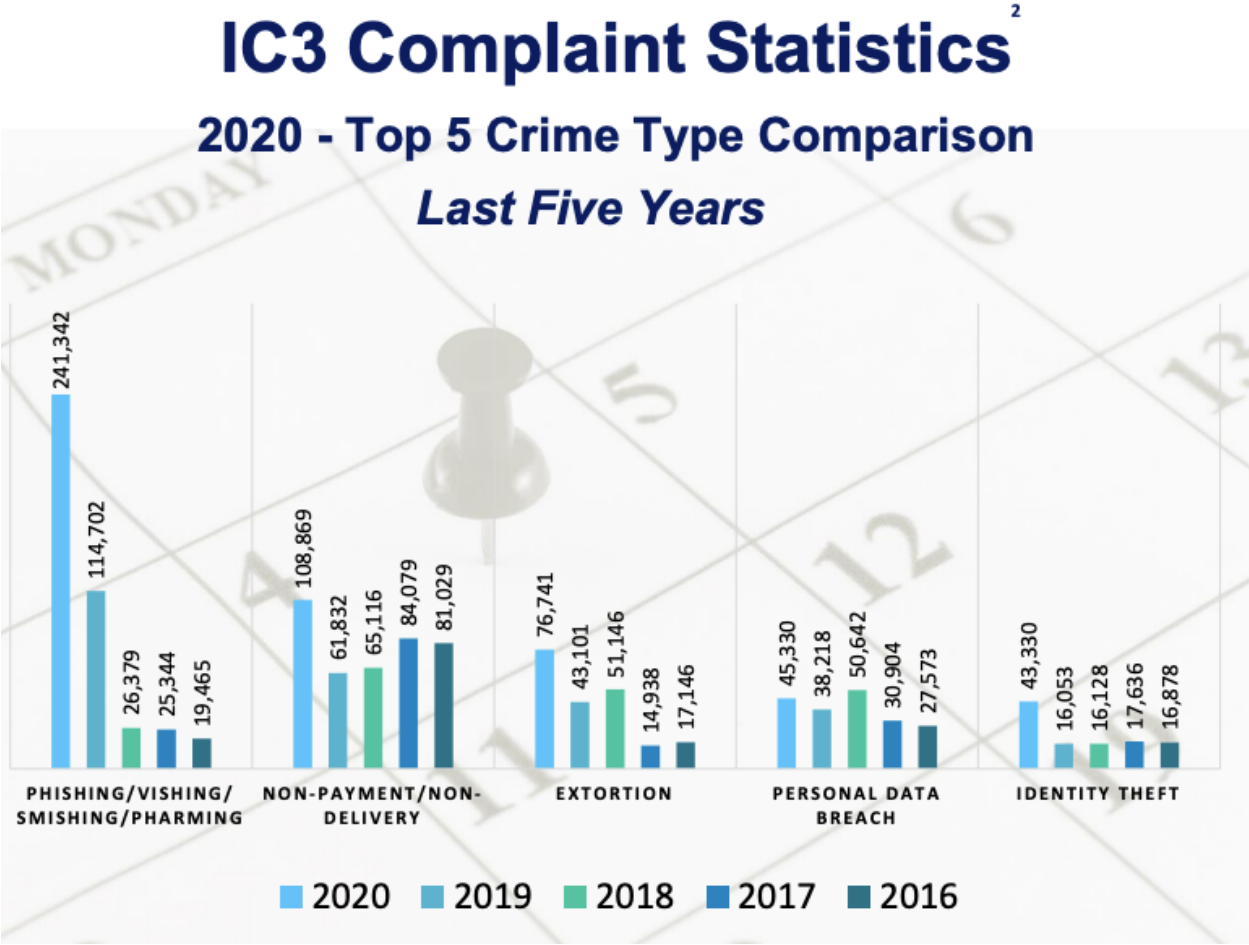


Figure 2 - Internet Crime Complaint Center Internet Crime Report 2020 Summary of Types of Internet Crime Reported

To add even more detail to this point, early in 2020 Akamai, a security company who participated in the development of the service covered in this paper, also showed a large amount of phishing activity focused on anything related to the pandemic[2]. The research showed hundreds of new COVID-related domains each day, most of which disappeared within 24 hours, which is consistent with longevity statistics they've historically gathered for names used for phishing.

Pandemic related scams are diminishing but hackers always innovate and find new themes to mine. Additional published research from Akamai shows a large increase in phishing prior to the holiday season last year[3]. Still more work discussed phishing campaigns modeled around cryptocurrencies and even Elon Musk[4]!

[2] https://blogs.akamai.com/2020/04/covid-19-phishing-exploiting-a-global-pandemic.html
[3] https://blogs.akamai.com/2021/02/phishing-holiday-season-attacks-on-the-rise.html
[4] https://blogs.akamai.com/2021/06/crypto-threats-surge-by-500-and-its-all-about-the-money.html

The point of these examples is not to offer a comprehensive overview of the threat landscape but to illustrate how hackers change their stripes to attract the attention of internet users and more importantly, get them to click on links and give up valuable data so they can make money. Diverse, fast changing exploits can make it hard for expert internet users to avoid being tricked, and it's even harder for average people.

3. The Road to Service Success

Like every company offering internet access services this organization was always looking for ways to improve their services. They were well known for performance and reliability and saw an opportunity to continue to contribute to online safety and help instill customer confidence with new security offerings.

The idea was to build on existing products which had embedded security features, like CPE for business customers with an integrated firewall and selectable levels of protection to control which kinds of protocol traffic are allowed or denied. Users can configure options to manage which protocol traffic will be permitted and blocked, like P2P applications.

More recently the team defined an optional DNS-based security service to defend workplaces against threats like malware, ransomware, phishing, botnets etc. The goal was to provide a foundational layer of defenses that take advantage of embedded network and operational strengths - such as scale, reach, reliability. This offering will be the focus of this paper.

In formulating product plans for this new service, initially targeted at Small and Midsize Businesses (SMB), several essential issues were considered to ensure success.

- Simplicity was paramount, surveys confirmed high awareness of cyberthreats but limited understanding of how to deter them - a service had to be simple to procure and operate.
- Complementing other security services was a critical consideration, most customers had security solutions and wanted to understand how anything new improved their posture
- Transparency was another requirement, subscribers had to be aware the service was operational, and what kinds of activity it was blocking.
- Customers had to have control and over time would expect even more "knobs and dials" to tune the experience

4. Keep it Simple

Large businesses are challenged dealing with internet security exposure but SMBs are even worse off because they may lack IT resources with specialized security expertise. A key product goal was to minimize the need for deep security expertise - with a model where the service "just works" after the service is activated (with some value add options discussed below) and covers all of the devices typically found in businesses today - PCs, phones, and even smart connected "things" like POS terminals, cameras etc. Other goals for the service included reducing installation overhead and ongoing maintenance - by eliminating software and hardware upgrades - as well as minimizing the expertise for configuring any optional features.

We provide every business subscriber with a graphical web portal that can be accessed anytime with a browser. The portal show's what's happening on their network and can be used to configure optional filters to manage what kinds of content is available on a workplace network, and what times certain websites are available to support Acceptable Use Policies.

5. Complementing Other Security Services

It was expected most business customers would already have security solutions and of course want to understand how anything new would improve their security posture. Examples of other services likely to be encountered include endpoint protection, specialized appliances, and “over the top” DNS-based filtering services. They considered how these options would shake out with respect to the planned service.

Endpoint software must be loaded on devices and requires ongoing maintenance, and it’s not available for smart connected devices. Appliances offer protection but web traffic on the internet is increasingly encrypted and malicious activity can be completely invisible to security appliances, especially lower end models that don’t have extremely sophisticated features to decrypt traffic (and experts to manage them). They also usually require periodic security expertise for ongoing care and feeding to be sure they operate properly. The user interface for security appliances is also typically aimed at someone with security knowledge, which an SMB may or may not have.

Security is about examining network traffic to determine whether it is malicious or legitimate. Approaches that operate in the data plane implement inline packet inspection to evaluate traffic. In the past it was possible to get visibility into most of the traffic on a network but the predominance of encryption has introduced significant limitations or complexity decrypting traffic. Inspecting packet traffic at line speeds has always been a costly operation, and inspecting encrypted traffic adds even more costs and operational overhead.

DNS filtering operates in the control plane. Incoming subscriber queries are matched against dynamic threat intelligence provisioned in resolvers, and policies can be applied to manage unwanted traffic. The team recognized attractive scaling capabilities since there is no need for pervasive monitoring and filtering of network traffic (and possibly decryption) and it can be layered on infrastructure already deployed and managed. Traffic from any device that makes DNS queries – the overwhelming majority – can be observed without the need for any client software. Threat coverage can be expanded by selectively forwarding suspicious traffic (typically domain names that point to both malicious and legitimate web resources) to proxies for further inspection. Experience has shown this is a small percentage of traffic, usually around 2%.

Based on these considerations DNS filtering was selected. The service is integrated with the subscriber internet access service (versus and off network “over the top” alternatives) to ensure performance and reliability.

6. Ensuring Transparency and Customer Control

Subscribers need a window into their service to understand how it is functioning. A portal user interface was specified to inform subscribers about their security posture and to allow them to configure other filtering features. Requirements included:

- Threat activity graphs showing threats blocked
- Configuration user interface for specifying web filters to support Acceptable Use Policies (AUP)
- Web filter graphs showing websites blocked by subscriber configured web filters
- Automated discovery of subscriber devices to support configuration of device-specific filters
- Subscriber-specified reports for archiving or offline analysis

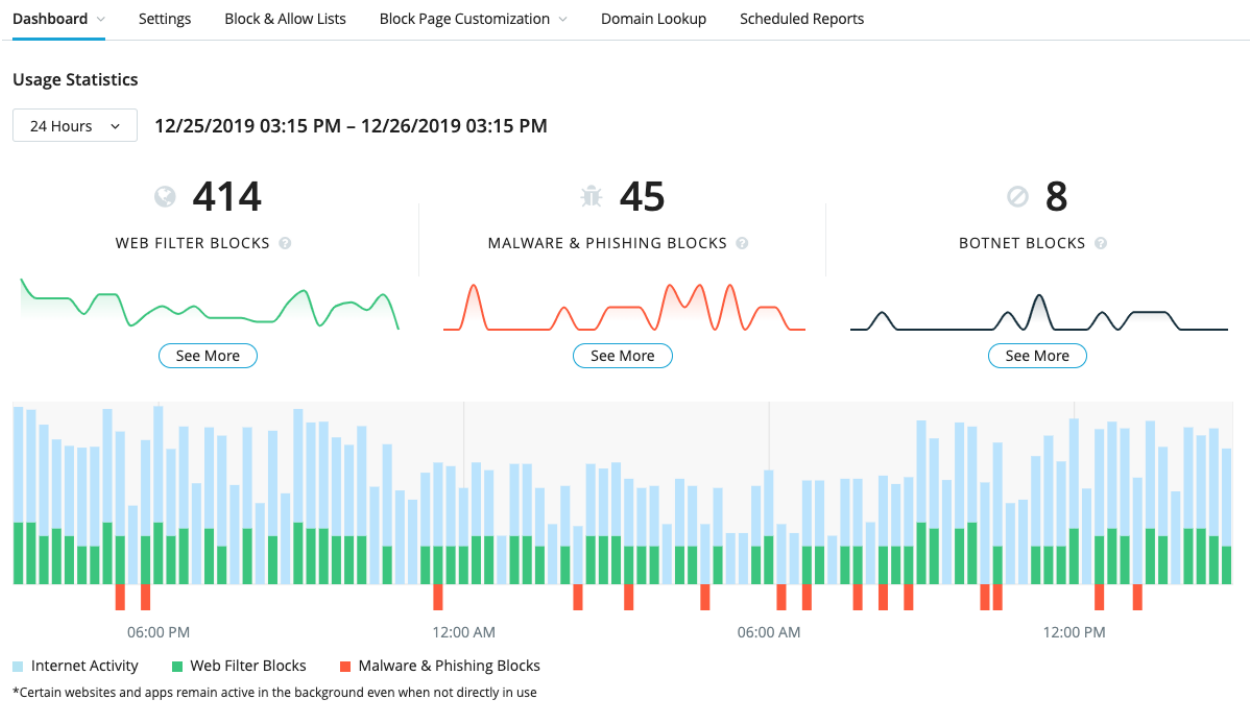


Figure 3 - Graphical User Interface Showing Internet Activity and Impact of Web Filters and Various Kinds of Malicious Activity

7. Sourcing Threat Intelligence

The value of security services is heavily influenced by the quality of the threat intelligence they employ. The threat landscape is incredibly diverse and it's simple for attackers to expose internet users everywhere to maximize the value of their efforts. Exploits also change quickly so attackers can stay ahead of security defenses that block their handiwork. For service providers in particular false positives can be extremely costly if subscriber access to favorite obscure websites is inadvertently blocked.

High level requirements for reconciling these objectives include:

- Large amounts of unencrypted raw data to maximize the coverage of increasingly sophisticated threats such as phishing, botnets and other kinds of malicious web resources
- High performance infrastructure to support near real-time processing of data to identify threats quickly and accurately to match the speed of change of today's exploits
- Machine learning algorithms that extend threat coverage by identifying subtle patterns and linkages in raw data that statistically match known threats used to train the algorithms
- Additional functions to validate threat entries are malicious to avoid blocking of legitimate web resources
- Software infrastructure to interconnect different processing systems so data flows among the "layers" in a structured way that allows inferences from one layer to the next

- Specialized processes to cull stale entries to keep threat lists from growing exponentially
- An iterative process to continuously refine, specialized algorithms that work together

Creating robust threat intelligence is complicated by the fact that some kinds of data is less available than in the past due to privacy regulations. More and more data is encrypted so it's opaque to analytical tools, and hackers increasingly use encryption to cover their tracks. Another advantage of DNS filtering is it overcomes these limitations. Unencrypted resolution data can be sourced from production resolvers or various kinds of network taps and anonymized to eliminate Personally Identifiable Information (PII).

DNS data has other useful characteristics for security research. Domain names and the Domain Name System (DNS) authorities and resolvers that support them, are fundamental to most security exploits. The DNS is widely used by malware developers because it connects everything on the internet from anywhere and virtually every network and device where an exploit might be activated will have access to the DNS. This means any device that emits a DNS query known to be associated with malicious activity can be identified as infected with the associated malware.

From a practical standpoint DNS queries are usually the first threat "signal" that's visible on a network where it can be detected remotely. Identifying activity at this stage is extremely useful as an exploit can potentially be disrupted before it does any real damage.

More details about how threat intelligence is developed can be found in a paper from SCTE 2018 "When Security and Privacy Collide New Approaches are Needed"

8. Marketing to Ensure Success

This is primarily a technical venue, but marketing is a critical part of any product launch process. Nothing sells itself anymore and building brand equity takes work - marketing investment helps ensure success. Product teams worked cooperatively with marketing to develop messaging that's aligned with other product and corporate marketing objectives.

For this service the marketing teams executed on a wide variety of initiatives. They built out a dedicated web presence in the commercial section of their website that provided all the information a prospective customer might be interested in. It highlighted all the major features of the service and offered complete descriptions of the value they offered to the intended customer base. Recognizing many potential customers were unlikely to be immersed in security jargon they used approachable language that was easy for typical SMB subscribers to understand. Other web pages present background information on security exposure so prospective customers can understand why protections are needed.

The marketing team also launched television advertisements across their US territory. Additional marketing campaigns reached customers directly with other media like email and marketing inserts as part of billing or other promotional packages.

There was additional effort for sales enablement. Teams were prepped on internet security basics so they could engage comfortably with prospective customers. They were equipped with simple product benefits they could convey and learned how to respond to questions about the service and how it compared with other security products prospects might already have. Business results

backed up the product development and marketing effort, with steady continuous growth in uptake.

9. Summary

Internet service providers everywhere are always looking for opportunities to improve their services. Speed and reliability aren't sufficient anymore. Awareness of security exposure on the internet creates an opportunity to contribute to online safety and help instill customer confidence with new security offerings. This DNS-based security service defends workplaces against threats like malware, ransomware, phishing, botnets etc., providing a foundational layer of defenses that take advantage of the power of network and operational strengths - scale, reach, reliability.

The product strategy for the new service considered several essential issues to ensure success:

- Simplicity was paramount because customers in the target segment often lack security expertise and specialized resources in general to oversee anything technically oriented
- It needed to complement other security services since most customers had security solutions and wanted to understand how anything new improved their posture
- Transparency was another requirement, subscribers had to be aware the service was operational, and what kinds of activity it was blocking.
- Customers had to have control and be able to tune the experience to match their business needs, preferences, and values

The new service was designed to complement other Small Medium Business (SMB) cybersecurity solutions such as firewalls and anti-virus. It adds threat defenses backed by machine learning algorithms that process live streamed DNS data in near real time to uncover new malicious activity quickly and accurately. Customized web content filters allow business managers to enable "Acceptable Use Policies" to manage the kinds of web content that are accessible on workplace networks. Investments in marketing and sales enablement were made to ensure good penetration of the service and attainment of business goals.

Abbreviations

CPE	Customer Premises Equipment
DNS	Domain Name System
FBI	Federal Bureau of Investigation
IC3	Internet Crime Complaint Center
ISP	Internet Service Provider
P2P	Peer to Peer
PII	Personally Identifiable Information
SMB	Small and Midsize Business

Bibliography & References

1. Federal Bureau of Investigation Internet Crime Complaint Center Internet Crime Report 2020 https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf
2. The Akamai Blog *Exploiting a Global Pandemic* Bruce Van Nice https://blogs.akamai.com/2020/04/covid-19-phishing-exploiting-a-global-pandemic.html
3. The Akamai Blog *Phishing Holiday Season Attacks on the Rise* Or Katz https://blogs.akamai.com/2021/02/phishing-holiday-season-attacks-on-the-rise.html
4. The Akamai Blog *Crypto Threats Surge by 500% and it's All About The Money* Or Katz https://blogs.akamai.com/2021/06/crypto-threats-surge-by-500-and-its-all-about-the-money.html

Delivering Access Beyond 10G

Coherent Subcarrier Aggregation as Backhaul for Next-Generation R-OLT, RMD, and Wireless

A Technical Paper prepared for SCTE by

Colin Howlett
Chief Technology Officer
Vecima
771 Vanalman Ave, Victoria, BC, Canada V8Z 3B8
+1 (250) 881-6235
colin.howlett@vecima.com

Aaron Chase, Infinera
achase@infinera.com

Antonio Napoli, Infinera
anapoli@infinera.com

Kevin A. Noll, Vecima
kevin.noll@vecima.com

Jay Rolls, Pacband
jrolls@pacband.com

1. Introduction

Over the last decade, the demand for network capacity has been continuously growing. This is particularly true in the last mile of the access and metro networks. This has been fueled by an unstoppable sequence of web-based applications such as video streaming, cloud services, mobile transport, and machine-to-machine communication. Recently, the traffic demand further increased driven by the ongoing pandemic though the impact has been, so far, manageable.

This growth was enabled by the wide deployment of broadband connectivity. This is an evolving landscape, which is now being addressed by the realization of newer technologies such as 10 gigabit per second (10G) capable Passive Optical Network (PON), DOCSIS 3.1, and fifth generation (5G) wireless, each offering access to subscribers at gigabit speeds. At the same time, the consolidation of hub sites and plant extensions are driving deployment of Distributed Access Architecture (DAA). As a result, the Converged Interconnect Network (CIN), the glue that connects these diverse access networks to the operator's headends, is being asked to provide more capacity to an increasing number of devices that are spread over an ever-widening geographic area.

Modern telecommunication networks transport Internet Protocol (IP) traffic utilizing primarily hub & spoke topologies. Here low-speed transceivers (spokes at the end user's location) are connected to high-speed transceivers (hubs). The optimal network topology for this would be point-to-multipoint (P2MP) and, usually, the physical plant is built as a P2MP architecture. However, the traffic is usually transported via point-to-point (P2P) architectures, such as those typically deployed for traditional telephony services, with the exception of PONs, where P2MP is realized using time-division multiplexing (TDM).

These network segments are the closest to the end users, and because of the large number of devices, cost and power consumption play a crucial role. At the moment, the preferred transmission method in access and metro networks is Intensity-Modulated Direct-Detection (IM-DD) systems, which employ modulation formats such as non-return-to-zero (NRZ) or pulse-amplitude-modulation (PAM)-4.

This technical solution has significantly contributed to the boom of the Internet, but initial doubts about its long-term sustainability – in terms of capacity – are arising as next-generation optical networks are being called upon to support even larger amounts of data. In fact, IM-DD is limited in spectral efficiency compared to the advanced modulation formats employed in coherent systems. Another strong limitation of the technology used in the access network is the utilization of TDM architectures, which significantly reduce the maximum throughput of the network and does not allow the usage of advanced digital signal processing (DSP) algorithms. These boundaries, and the underlying P2P network architecture, have resulted in a growing awareness in the industry that P2P 10G IM-DD systems cannot meet the requirements for backhaul while simultaneously making efficient use of network infrastructure, e.g., in the case of fiber scarcity.

In this context, coherent optics, as has happened in other segments, might come to the rescue where capacity and reach become the core concerns, but they still fall short where high device counts or simply geographic distribution are the key issues. A first important step towards a low-cost coherent marketplace is represented by the Optical Internetworking Forum (OIF) implementation agreement of 400ZR [OIFIA2020]. Coherent offers a wide array of advantages ranging from the ability to employ advanced DSP technologies – that can compensate for fiber propagation effects such as accumulated dispersion – to the enabling of wavelength division multiplexing (WDM).

This evolving storyline requires new solutions to address the given requirements. This paper provides operators with an evaluation of possible approaches to this issue ranging from higher speed IM-DD

systems to coherent optics with a special focus on the emerging concept of coherent P2MP transport using digital subcarrier multiplexing (DSCM) [Sun2020] for aggregation [Welch2021].

Fiber access links have traditionally been arranged in a P2P fashion, with paired electronics on each end of the fiber. If electronics are changed on one end, they must be changed on both ends to remain interoperable. Consequently, to combine low-speed interfaces to higher-speed ones, an intermediary aggregation device is required. In [Welch2021], a paradigm shift has been proposed that targets a universal approach to connectivity in telecommunication systems. This exploits DSCM to simplify the network architecture by removing the electrical aggregation layer and the bookended transceivers and replacing them with a simple passive optical combiner. Moreover, DSCM allows the realization of a truly P2MP network architecture [Welch2021].

This is relevant because most current-day last-mile architectures are broadcast in nature – meaning that transmitted downstream signals are "heard" by all endpoints. This is true for PON, DOCSIS, and Mobile networks. Wired networks (PON/DOCSIS) might be single or dual fiber. Whether one considers ring topologies, star topologies, passive tree architectures, PON overlay architectures, or enterprise/wavelength extensions, the deployment approaches are all similar. "Last mile" networks are predominantly P2MP in nature, but still served by P2P optics. Such designs tend to drive high optical port counts, while the solution based on DSCM might lead to significant savings [Bäck2020].

At the same time, that network capacity continues to grow, marketed network speeds are also growing, with 1 gigabit per second (Gbps) service now commonly available, and plenty of industry discussion around emerging multi-Gbps speeds on the horizon. We see this from a few perspectives, including residential broadband rates, demand from business customers, wireless capabilities now made possible via 5G, as well as an accelerating build rate of fiber networks. In addition, the cable industry has put forward its strong 10G mantra over the past few years. The discussion will address critical operator needs such as capacity, scalability, deployment lifespan, network reach, CIN aggregation architecture impacts, the convergence of DAA/wireless/business services, and operational simplification. A road map toward a new network paradigm will be also provided.

The remainder of this paper is organized as follows. Section 2 describes the state-of-the-art in terms of traffic, networks, integration with cable access technologies, and the reasons to move beyond 10G. Section 3 describes existing optical transmission technologies and how those can move beyond 10G. Section 4 presents a proposal that aims to cope with existing traffic patterns and high-capacity networks. Here, the concept of DSCM is introduced and we explain how this helps to realize fully capable point-to-multipoint optical networks. Section 5 provides a high-level qualitative comparison of technology options in moving beyond 10G before wrapping up.

2. Evolving Traffic and Deployment Requirements

In this section, we describe the evolving traffic and deployment requirements to serve various cable operator access solutions for HFC, PON, and wireless.

2.1. Network segments and traffic patterns

Figure 1 illustrates a generic tiered telecommunication infrastructure typical of systems deployed by cable operators today, spanning all the way from subscribers back to a wide area core network, with all links except for most subscribers, consisting of digital optical transport.

The access segment includes equipment such as Converged Cable Access Platform (CCAP) in DOCSIS hybrid-fiber coax (HFC), optical line terminal (OLT) in PON, and centralized unit (CU), distributed unit

(DU), radio unit (RU) in 5G wireless which provide last mile connections to users. Connections are generally tens of kilometers and limited to 80-120 km maximum.

The metro segment connects multiple access segments together across a larger, generally metropolitan area with links <200 km.

The core segment provides multi-city transport between different metropolitan areas covering hundreds or thousands of kms.

In this paper, we focus primarily on the access segment, but the solution described is also applicable to the metro segment.

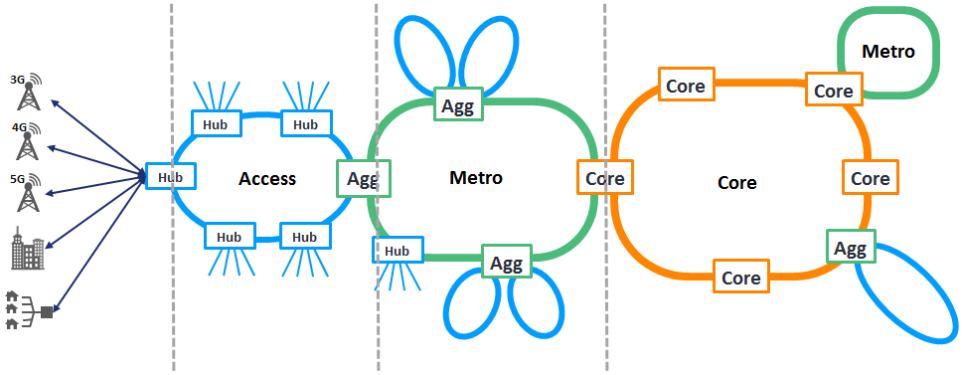


Figure 1 - Generic Cable Operator Network Architecture

A significant portion of the traffic within a cable operator's network is used for providing access from a large number of subscribers to a smaller number of service access points delivering services such as operator-delivered video content, streaming video providers, or general Internet services. This traffic is inherently organized in a hub and spoke pattern. Relatively low-speed (compared to optical transport) subscriber last mile interfaces connect to the Internet and other services via a high-speed hub, and traffic patterns resemble a point-to-multipoint (P2MP) network.

Today though, optical networks are still designed utilizing a point-to-point (P2P) approach, with the exception of access via passive optical network (PON). Figure 2 illustrates the differences between these traffic patterns.

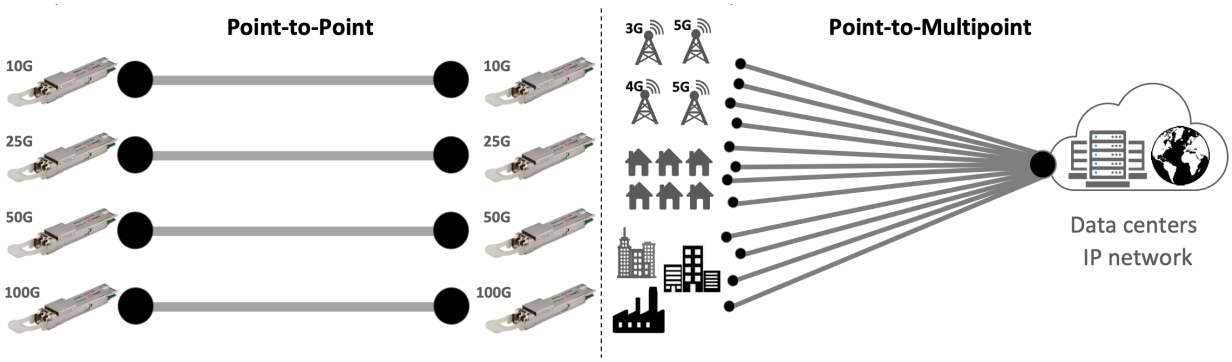


Figure 2 - Point-Point and Point-to-Multipoint Network Traffic Patterns

2.2. Evolving capacity needs for cable access segments

Over the last decade, the demand for capacity in the last mile access network has been growing continuously. At the same time, the access topology has also changed, with the "last mile" links becoming shorter, and active electronics being pushed deeper. Whereas fiber, coaxial, and twisted pair copper north bound endpoints previously terminated in facilities (headends, central offices, huts, etc…), it has become far more common for these termination points to be placed somewhere in the outside plant infrastructure.

Examples include:

- Cable operators' hybrid fiber coax networks have typically been fed by a combination of video QAMs (signal modulators) and Converged Cable Access Platform (CCAP) systems for broadband. These were typically rack mounted in an environmentally controlled facility. With the advent of DAA using Remote physical layer (PHY) or Remote media access control + physical layer (MACPHY) devices, those electronics take on a smaller more modular format that gets pushed out of the rack and into the fiber node installed in the neighborhood that is being served.

- Since the early 2000's, TelCo's have deployed remote digital subscriber line access multiplexers (DSLAMs) to deliver broadband over twisted pair copper.

- Mobile network operators (MNOs) have experienced a fast evolution of cellular radios and the radio access network (RAN) with electronics moving in multiple directions within the network topology. Increased use of virtualization and evolution of the location at which processing happens within the RAN has created increased need for a variety of optical transport solutions ultimately feeding radios located deep in the network.

- And for operators deploying fiber to businesses and residences, PON fiber terminations now frequently occur in remote OLTs that are pedestal or strand mounted, or within MDUs and office buildings.

The result is that the number of locations requiring fiber terminations is growing exponentially, while requiring higher and higher speed interfaces at the same time. Such network demands require a well thought out approach to connectivity.

Figure 3 below shows common scenarios for cable operators as they consider the required capacity in the access segment of their network.

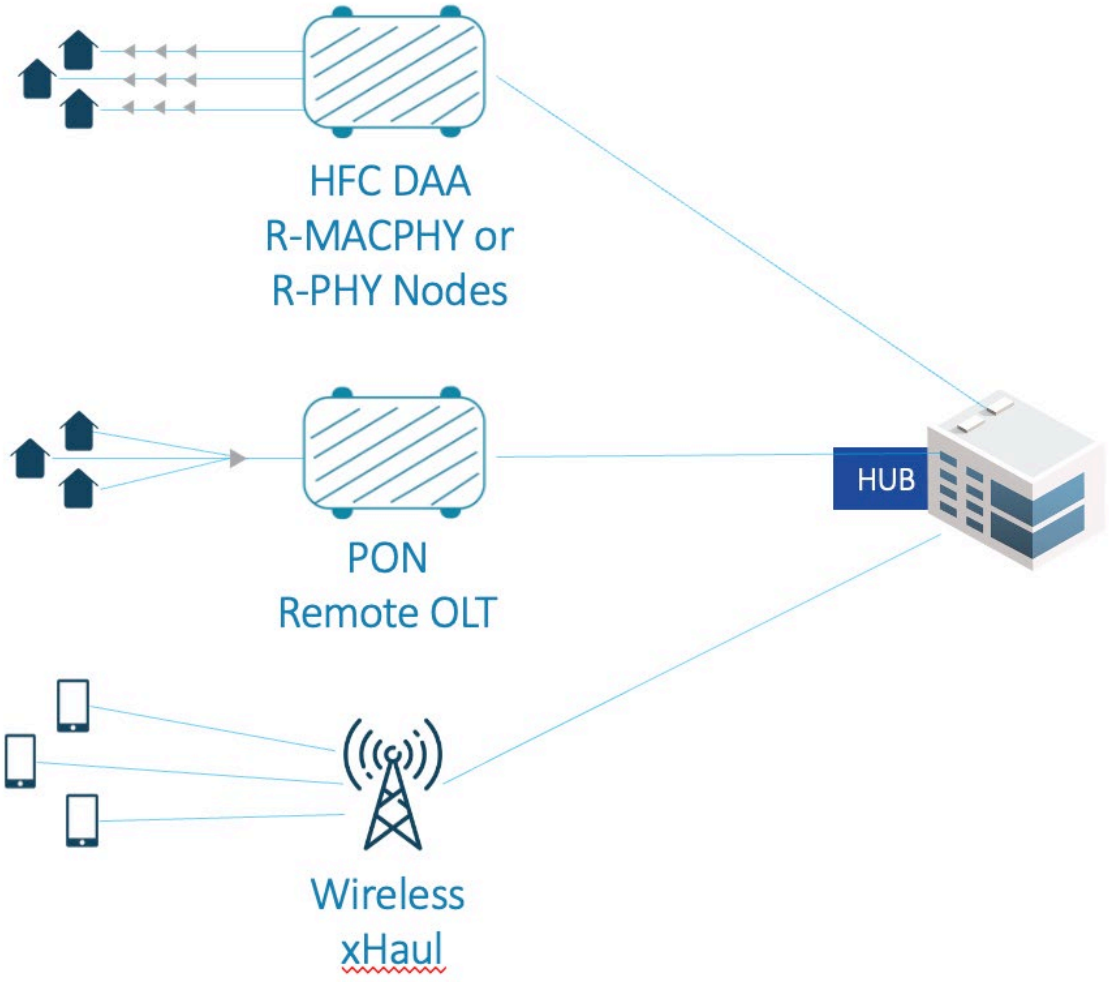


Figure 3 - Access Segment Transport Applications

2.2.1. HFC DAA

Many operators are evolving their DOCSIS and QAM video delivery infrastructure to DAA to:

- Save space and power in hub facilities as the number of HFC service groups expands with ongoing capacity growth
- Improve signal fidelity and resulting capacity by moving RF processing to the edge
- Converge HFC access backhaul with other forms of access backhaul by removing analog optics and replacing them with IP/Ethernet

DAA deployments today typically consist of 1 downstream (DS) service group with a mix of DOCSIS and QAM video. The total throughput required to feed a single DS service group with 1218 MHz of spectrum can theoretically exceed 11 Gbps but operators have generally coalesced around using 10G small form factor plus (SFP+) with 10GBASE-LR (long reach), ER (extended reach), or ZR/dense

wavelength division multiplexing (DWDM) optics based on IM-DD as a pragmatic, cost-effective solution.

The expense of updating the nodes once deployed leads operators to build outside plant (OSP) and the converged interconnect network (CIN) feeding DAA nodes for long term (5-10 year) capacity. Capacity needs are soon expected to exceed 10 Gbps in DAA nodes due to:

- Increasing desire to deploy segment-able nodes which provide an option to support 2 (or more) DS service groups in a single DAA node at 10 Gbps each (20 Gbps for 2 DS SG)
- DOCSIS 4.0 (D4.0) frequency division duplex (FDD) extension to 1794 MHz can support DS capacity in an individual service group beyond 13 Gbps; this allows D4.0 to exceed the capacity of 10 Gbps PON and offer true 10 Gbps service which may be an important marketing difference in years to come

As a result, 25G links to the DAA nodes is becoming a desirable target for the combination of D4.0 and multiple-downstream service groups in next-generation DAA nodes.

HFC DAA nodes are also generally power constrained, with increasing pressure to fit more and more processing or other features in a housing (in North America) which is thermally limited to 160-180 Watts in harsh outdoor environments. In many DAA node designs, pluggable uplink Ethernet optics are generally the constraint for maximum operating temperature. Larger and more thermally friendly pluggable form factors such as C form-factor pluggable half-size (CFP2) are favored in this environment over small form factor pluggable (SFP) or quad small form factor pluggable (QSFP) formats.

The upcoming standardization of a generic node housing with the SCTE Generic Access Platform (GAP) standard will drive this further by allowing other modules for edge compute or Ethernet switching in the same node. Power efficiency and power consumption are critical in DAA applications.

2.2.2. Remote PON

With advancements in IP video delivery and the ability for operators to deploy video service without traditional QAMs, cable operators are increasingly deploying more fiber to the home (FTTH) in the network. This is especially true in greenfield builds such as housing subdivisions, and also in rural areas fueled by government broadband funding. Common technologies for deployments today are 10G-EPON or XGS-PON.

Due to the nature of the cable network with long distance links to hubs and the desire for high split ratios of 64:1 or 128:1, many operators cannot easily support hub-based OLT. As a result, the emerging dominant deployment model is the Remote OLT (R-OLT), containing 4 PON segments within a single outside plant housing powered from the coax cable.

In terms of capacity, each PON segment is generally limited to ~9 Gbps of maximum throughput due to the nature of the PON system overhead. An R-OLT is typically fed using the same 10G SFP+ that would be used for HFC DAA including 10GBASE-LR/ER/ZR optics as a pragmatic solution. A one-to-one arrangement of SFP+ to PON segment is common, but some operators optimize cost using R-OLT platforms that support aggregation of the uplink interfaces to multiple PON segments.

R-OLT implementations could already use a cost-effective solution beyond 10G to support aggregation over a single interface for fiber/wavelength savings. Year over year capacity growth and evolution in the R-OLT is also expected to drive further need for capacity beyond 10G due to:

- Support for 25G-PON (25GS-PON or 25G-EPON)
- Potential increases in PON segments beyond 4 per R-OLT
- In some countries, cabinet-based deployments with up to 16 PON segments in a single chassis are also common

R-OLT are not as thermally or space constrained as HFC DAA nodes due to the lack of RF amplifiers, but pluggable uplink optics are still a significant design consideration.

2.2.3. Wireless xHaul

Many cable operators operate wireless networks directly or sell services to provide transport for wireless operators. With convergence of the access segment of the network to serve HFC and PON, it's natural to also consider wireless xHaul within that same network.

Figure 4 illustrates the split options in various 5G deployment architectures which differ in where and how functions are positioned between elements containing some or all RAN components consisting of CU, DU, and RU. Cable operators may deploy a mix of these architectures depending on the siting and capacity needs within a particular market.

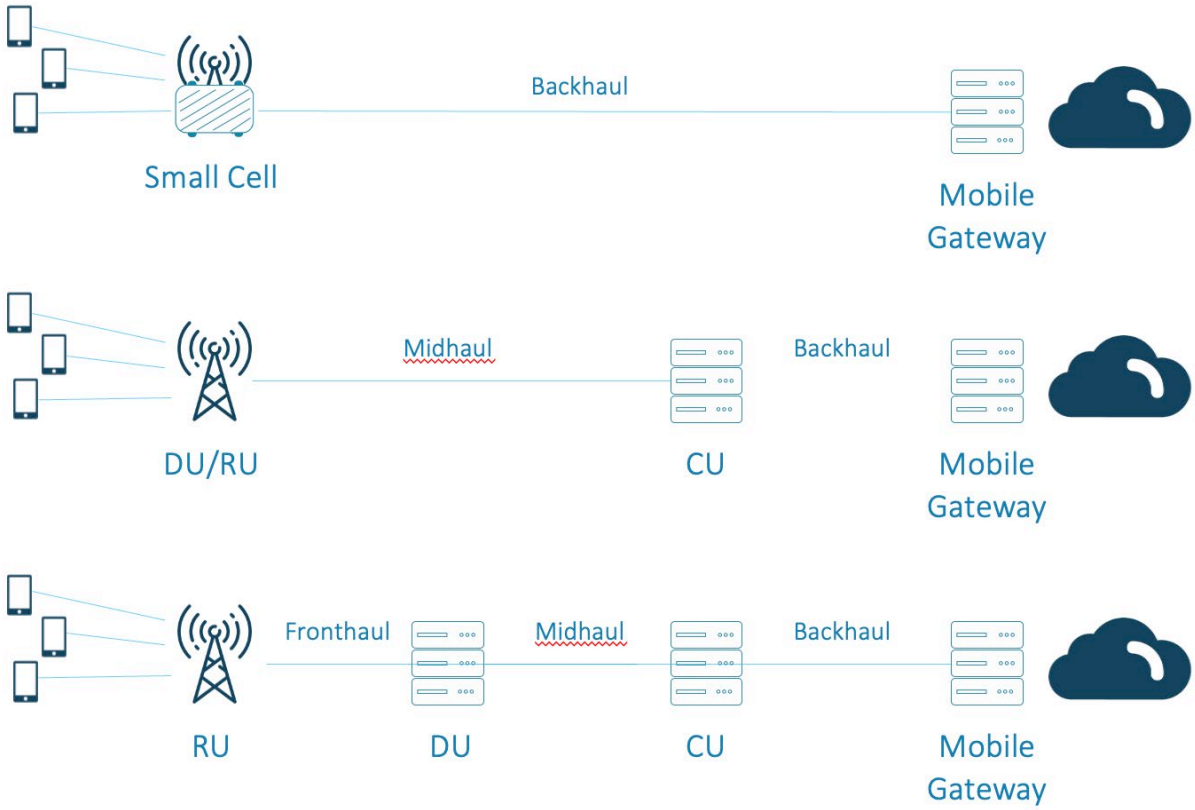


Figure 4 - 5G xHaul Transport Splits

Architectures feeding backhaul and common mid-haul options (F1 split, option 2) through the access segment will require overall capacity approximately equal to user bandwidth. Except in the case of massive multiple-input, multiple-output (MIMO) mm-wave bands, this capacity will generally not consume > 10 Gbps.

The increased use of virtualized RAN, with significant benefits in using elasticity to match capacity to where users move during the day, leads to growth in a fully disaggregated RAN with significant fronthaul usage possible through the access segment. While it is relatively common today to use dark fiber and gray optics with the need for fiber construction, an access segment which is naturally increasing capacity could allow for convergence and increased reuse of the existing fiber to boost wireless capacity.

Fronthaul bandwidth requirements, as indicated in [Cisco5G] can easily exceed 10G or even 25G as cell site capacity and the use of massive MIMO expands. An access segment architecture which supports endpoint capacities up to 100G could be well suited to serve these wireless needs.

Table 1 - 5G xHaul Transport Capacity Requirements

Band	Bandwidth	MIMO Layers	Fronthaul Data Rate for 3 Sectors No Compression	Midhaul/Backhaul Data Rate for 3 Sectors
850 MHz	10 MHz	4T/4R	7.35 Gbps	330 Mbps
1.8 GHz	20 MHz	4T/4R	14.7 Gbps	660 Mbps
3.5 GHz	100 MHz	64T/64R, 8 layers	57.87 Gbps	13.5 Gbps
28 GHz	400 MHz	64T/64R, 4 layers	82.32 Gbps	19.5 Gbps

2.3. Pushing beyond 10G

In the access space, service providers are faced with a challenging question – once an operator needs to upgrade or support capacities beyond 10G, should they upgrade to 25G or to 100G? If an operator upgrades to 25G, that increased capacity, as is often the case, will have a limited lifetime because there will be a need to upgrade soon again, assuming today's 30% yearly traffic growth persists. However, deploying 100G would likely be cost prohibitive today.

A quandary operators face in the "last mile" access network is that the number of network touch points far eclipses those of all other parts of the network. This leads to a situation where the operational expenses encountered for equipment upgrades drive costs that are often vastly more than for the equipment itself. Operators are well served to think this through and work to deftly address these ever-growing needs, going beyond just a repetitive process of constantly adding capacity in an incremental fashion. Minimally, such changes should have a lifetime of at least five years, and more ideally upwards of ten years.

Many choices have a ripple effect beyond just the end point solution. Commonly end points require an aggregation router at a hub/headend location. Aggregation devices with SFP/SFP+ interfaces usually have two or more 100G interfaces and those are used to communicate with two or more spine switches. When a network operator is considering bandwidths above 10G, each leaf switch is only aggregating four to seven 25G optics, or in other words, the cost of that leaf switch is applied to only six 25G end points on average. The hub facility at the edge of the network may also be experiencing space and power constraints so density of connections is increasingly important, especially when considering operator intention to add new functions like mobile edge compute (MEC). If an operator chooses to upgrade to 100G optics, then they might only be using 15%-20% of the capacity, so again, an aggregation device is needed higher up to drive more efficient port utilization on routers.

In HFC networks, it is becoming increasingly observed that fiber termination points at the fiber node demarcation points (where fiber transitions to coax occurs), are becoming an increasingly strategic point in the network that can be leveraged for a multitude of services. Fiber is such an important resource, as it represents an asset that was constructed at great expense. Furthermore, it is also a medium which is incredibly flexible and expandable at modest cost. These fiber end points represent the "new network edge" for MSOs that can be used to serve not only coax-fed residential and SMB customers, but also enterprise customers, cell backhaul services, as well as an operator's own RAN.

3. Existing optical transmission technologies

Direct detection (DD) was one of the first solutions to detect optical signals at the receiver. It is a technique where only the amplitude information is preserved after the photodiode optical signals are converted into electrical currents. The first commercial optical systems were realized with this technique, and DD dominated all market segments until approximately 2010. Up to that time, optical systems could transport 10 Gbps per channel over a total of 80 wavelength division multiplexed (WDM) channels on a single fiber.

Since 2010, there have been rapid advancements in coherent optical transmission technology to extract more and more capacity from a single channel. As a result, transmission systems with 800 Gbps or higher capacity are now commercially available. Coherent optics, in a similar way to the classical concept of coherent radio transmission, uses polarization along with amplitude and phase modulation to realize higher order quadrature modulations and provide higher capacity from a specific line rate or channel bandwidth.

3.1. 10G

10 Gbps per channel is widely used today. 10 Gbps Ethernet, first standardized as IEEE 802.3ae in 2002, is used widely in enterprise networks and access network backhaul especially for DAA in DOCSIS and fixed wireless networks. 10 Gbps was standardized as an option for FTTH in 2009 as IEEE 802.3av (10G-EPON) and in 2010 as ITU-T G.987 and is now the fastest growing access network technology in the world. 10 Gbps isn't just for fiber, either – DOCSIS support for DS speed approaching or exceeding 10 Gbps is becoming possible with DOCSIS 4.0 technology.

3.2. 25G (and beyond)

IM-DD has continued to evolve. 40 Gbps was the next step, with IEEE 802.3bg in 2011 creating the standard for single-lane 40 Gbps Ethernet (4 × 10 Gbps was standardized in 2010), but it is limited to a distance of 2km and is all but defunct in the market.

25 Gbps, as a by-product of 100 Gbps (4 × 25 Gbps) was the next step beyond 10 Gbps that was attractive to the market. IEEE 802.3by and IEEE 802.3cc standardized 25 Gbps in a single lane, and 32G Fiber channel (really 28 Gbps) was introduced at the same time.

25 Gbps has taken on a life of its own, independent of 100 Gbps. Economically 25 Gbps fits a sweet spot in network designs – fulfilling a need for more bandwidth but at a much lower cost than a jump to 100 Gbps. 25 Gbps per channel is now being applied to passive optical networking in the 25/50G EPON standard, IEEE 802.3ca. The limits of today's IM-DD technology began to be exposed in development of 50G PON at the ITU-T where 50Gbps over a single channel in the downstream was adopted for the ITU-T G.9804 series. Nonetheless, specification of 50 Gbps in the upstream was postponed in favor of 10 Gbps and 25 Gbps in the upstream due to the difficulties of burst reception at 50 Gbps.

Achieving 100 Gbps on a single channel with IM-DD was no small feat. In 2021, IEEE 802.3cu was able to achieve a standardized 100 Gbps on a single channel at a distance of 10 km. There are proprietary versions of single-lane 100Gbps that are able to achieve up to 40 km, but greater distances are relying on advanced modulations.

4. Future coherent transmission technologies

A race is on to establish the next-generation technology that will replace current P2P 10G optics based on IM-DD to support increasing capacity requirements for access and metro to serve DOCSIS, PON, and 5G wireless services. Among the most relevant challenges, one could single out bandwidth demands, compensation of the physical layer impairments, and continuous and dynamic network upgrades. Next, we will describe coherent systems in the form of single-carrier systems as well as DSCM.

4.1. Single-carrier coherent systems

Coherent detection has two key advantages over IM-DD. The first is that coherent detection preserves the phase information, and DSP algorithms can be applied within the transceiver to decode information embedded in the phase. Secondly, the local oscillator allows excellent channel selection. The first property was not fully exploitable until complementary metal–oxide–semiconductor (CMOS) and digital-to-analog (DAC) and analog-to-digital converter (ADC) technologies became mature. The second became less important after the introduction of the first commercial optical amplifiers. Because of this, coherent was not used during the first optical revolution that started with WDM systems and IM-DD. After the dot-com era and explosion of IP data traffic, the throughput of IM-DD systems became insufficient to cope with the exponential bandwidth growth.

This led to the rapid development of 100G thanks also to the implementation of the DSP algorithms and advanced forward error correction (FEC) code within the first generation of application-specific integrated circuit (ASIC) [Sun2008, Roberts2009]. These achievements led to the spreading of coherent technology from subsea to core and regional networks, where it replaced the IM-DD. Now, coherent transponders can achieve data rates beyond 1 Tb/s [Sun2020, Buchali2016], employing high-order modulation formats enabling a multitude of modes by varying the symbol rate, FEC overhead, and probabilistic constellation shaping (PCS). Although coherent has been so far confined to core and submarine networks, thanks to recent initiatives such as 400ZR, it is set to enter new markets such as data center interconnects.

IEEE 802.3ct has released a specification for 100 Gbps over a single channel for 80 km using dual polarization-differential quadrature phase shift keying (DP-DQPSK) (not direct detection). CableLabs® released a coherent optics specification in 2019 that describes the architecture and requirements for a 100 Gbps and 200 Gbps capable of up transmitting data up to 120km. As of today, though, only a handful of vendors are offering products that meet the 100Gbps requirements and fewer are offering a 200 Gbps option. The CableLabs coherent optics specification requires only a single fiber for transmission which is a big advantage over most other options that require 2 or more fibers. One potential disadvantage of coherent optical modules is that they incorporate a DSP into the housing which requires more power and a larger format due to the component size and required heat dissipation.

Type	Launch Power (dBm)	Loss Budget (dB)	Distance (km)	Speed (Gbps)	Fiber Count
10GBASE-ZR	0	25	80	10	2

Type	Launch Power (dBm)	Loss Budget (dB)	Distance (km)	Speed (Gbps)	Fiber Count
25GBASE-ER	-1	20	40	25	2
25GBASE-LR	-5	18	10	25	2
10GBASE-ER	-1	14	40	10	2
10GBASE-LR	-11	8	10	10	2
50GBASE-LR	-4.5	12	10	50	2
50GBASE-ER	-0.5	15	40	50	2
100GBASE-FR1	-2	4	2	100	2
100GBASE-LR1	-1	8	10	100	2
CableLabs P2PCO 100G	-6	21	80	100	1
CableLabs P2PCO 200G	-7.5	19	80	200	1
10GBASE-PR	2	30	20	10	1

Table 2 - Comparison of Single Channel Optical Systems

Coherent is also gaining momentum in the PON space with CableLabs recent kickoff of Coherent PON which is targeting 100Gbps PON using coherent optics as the underlying physical layer.

4.2. Digital subcarrier multiplexing based coherent systems

An alternative and more flexible solution to standard coherent is digital subcarrier multiplexing (DSCM), which creates subcarriers (SC) in the digital domain using only one laser. From a networking perspective, it is similar to the sliceable bandwidth variable transceiver (S-BVT) introduced by [Sambo2015]. The key difference between the single carrier of Section 4.1 and DSCM is that now the DSP operates at a symbol rate (R_s) of R_s/N, where N is the number of digital SCs. Clearly, in this case. the transceiver performance might be limited by the value of the laser linewidth Δf [Dris2013]. In fact, given the larger symbol period T_s, the $\Delta f \bullet T_s$ product for a DSCM signal with N SCs is N times larger than that of a single wavelength [Welch2021].

DSCM is used in P2P links [Sun2020], and it provides different benefits. First, it enables, by parallelizing with respect to the individual SCs, simplifcation of the DSP algorithms. For example, the compensation of accumulated dispersion. Second, it significantly reduces the impact of equalization enhanced phase noise (EEPN), which is relevant in the case of long-haul transmission. Third, by optimizing the modulation format of the SCs, it helps to increase the tolerance against filter cascade [Rahman2016]. Last, thanks to its intrinsic flexibility, it enables transmission to be adapted to the current traffic, thus reducing the power consumption [Velasco2021].

A pictorial representation of the spectrum of a 400G 16QAM DSCM signal – with 16 SCs at 4 gigabaud (GBd) each, and its equivalent single-carrier are depicted in Figure 5, where both channels occupy the same 64 GHz bandwidth. Only in case of ideal Nyquist shaping (i.e., roll-off = 0) there is no need for a guard band as there is no overlap. Roll-off = 0 is not possible, and therefore a guard band is needed [Welch2021].

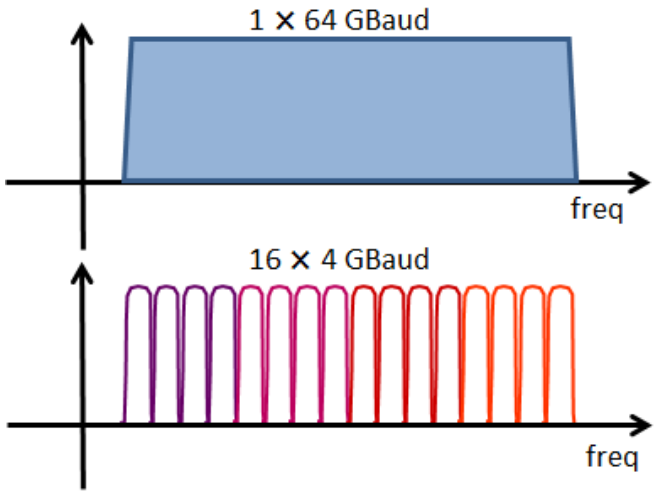


Figure 5 - Single Carrier and DSCM Spectrum

4.3. Enabling point-to-multipoint optical aggregation networks with DSCM

Beside the benefits described in Section 4.2, DSCM and coherent can enable the re-architecting and simplification of future optical networks. Figure 6 illustrates a simplified metro/access aggregation network realized with P2P transceivers (a) and with P2MP ones enabled by DSCM (b).

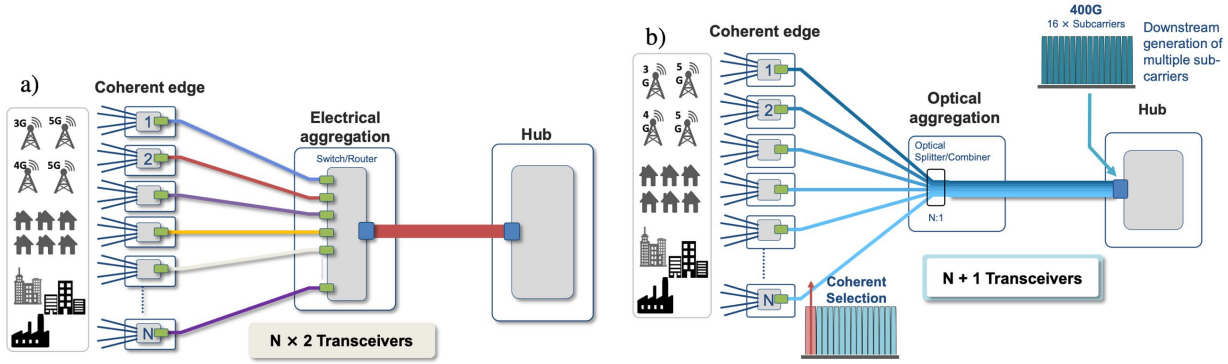


Figure 6 - Point-to-Point vs. Point-to-Multipoint Aggregation

In Figure 6-a, the N end users are equipped with N low-speed transceivers connected with their N pairs at the electrical aggregation stage. Next, a high-speed transceiver communicates to the hub, for a total of 2N

(low-speed) + 2 (high-speed) transceivers. The architecture shown in Figure 6-a can be realized by employing IP routers for aggregation and multiplexing of the low-speed into the high-speed transceivers, where each link uses the optimal rate P2P transceivers.

This approach is sub-optimal for hub and spoke traffic and is a major drawback of this solution. As optical spectrum in fiber is not scarce outside core networks, to achieve the lowest cost solution, we minimize the number of regeneration points. This approach hides a waste as there is no value to the end-users in regenerating traffic unnecessarily. In fact, this approach is constrained by decisions that are based on the peak traffic value. Consequently, operators will decide (1) to minimize the number of physical site visits, thus deploying higher-capacity transceivers – i.e., higher CAPEX or (2) to optimize the capacity of the transceivers; thus, increasing the number of truck-rolls, i.e., higher OPEX. Neither solution is economically optimal. With this architecture if an endpoint requires more capacity, the two transceivers in that link must be replaced. In the situation where the electrical aggregation devices utilize common port capacities, as is typical in most commercial high density devices, the entire aggregation device and all attached transceivers may need to be replaced. A typical example of this could include a DAA CIN built to support 10G with high density 10G switches for electrical aggregation; a need to change a link to the next step higher (25G) cannot be easily accommodated in the existing aggregation device.

Figure 6-b shows the same aggregation network realized with DSCM. The electrical aggregation stage is now replaced by a simple 1:N passive optical combiner, which greatly reduces the number of devices and stages needed to aggregate and up-speed the traffic to be transported to the next hub. In Figure 6-b, only N low-speed transceivers and 1 high-speed are required, i.e., 50% less than in Figure 6-a. This P2MP architecture connects multiple low-speed transceivers (spokes) to high-speed ones (hubs), breaking the bookended transceiver paradigm. These benefits directly translate into lower power consumption, smaller footprint, less sparing of parts and grooming equipment.

Furthermore, P2MP with DSCM enables cross-layer savings related to an efficient utilization of the optical transceivers, which can flexibly provide the required amount of capacity. For example, an operator deploys 100G (4×25G) P2MP pluggable optics in 12 leaf nodes. On day 1, only 1 SC per node is active, and all 12 SCs are terminated in a single 400G hub router port, which can receive up to 16 SC at 25G each. This is equivalent to approximately one-half (13 vs. 24) of the devices compared to what would be required with P2P optics on day 1. If the capacity grows in the edge nodes, then more SCs will be needed over time, and they can be activated via software. This operation is possible if the definition of SCs is maintained over more generations of P2MP optics, so that different versions of pluggable can interoperate, and operators can thus define multi-generational network architectures. By enabling this, routers and hubs can be independently upgraded, thus decoupling nodal upgrades from network-wide ones. Network operators can maximize the return on investment and ensure a smooth and cost-effective capacity upgrade of the network.

Nevertheless, there are challenges for P2MP architectures. First, it requires time to remove routers from the network, as these elements are rich in features, and operators use them for several network functions. Second, having fewer transceivers leads to network simplification, but if one device fails, the amount of traffic that is affected will be larger. A third challenge is the transmission of SCs over different light paths. If different SCs are propagated over different links, at the hub, they might suffer power imbalance. The pluggable will have an adjustable range to transmit and receive, and this will be monitored via a power control loop between the hub and each edge transceiver. Nevertheless, if the path loss difference exceeds the transmit output adjustment range, an external attenuator might be required.

As bandwidths increase and the number of devices in the geographic area expands, there may be a desire for protection or redundancy. The access network is typically unprotected over the last mile, but

redundant optics can be used at the hub location to protect against module failure; where there is an active/standby architecture deployed. If diverse fiber paths are available, then current protection mechanisms can be used to protect against fiber cuts. The cost of redundant equipment at the access edge of the network is typically prohibitive but for those applications that can justify the incremental expense that is also an option.

Overall, coherent with DSCM is a combination of mature technologies and it has been successful in high-end products, but when it comes to mass production, the required components, e.g., the laser, might have higher costs relative to traditional IM-DD or single carrier coherent solutions. On the other hand, coherent with DSCM significantly reduces the number of devices, eliminates the electrical aggregation stage and reduces the number of truck-rolls. It has been shown that, if the coherent pluggable price is below 50% of the existing pluggable, the solution presented in [Welch2021] brings cost benefits, when we consider the entire network [Marino2021].

4.3.1. Time-division versus frequency-division multiplexing

Time division multiplexing (TDM) and frequency division multiplexing (FDM) serve different scenarios with their advantages and disadvantages. In access, TDM – and its variants, e.g., statistical TDM – is used in PON networks before aggregation, where the multiplexing is carried out in time. This is a complex operation in the case of high data rates.

FDM – with DSCM – has been deployed in core networks [Sun2020] and enables Tb/s channels after aggregation via high-speed coherent transponders. In transport, the FDM-based channels are seen as one carrier by the remaining elements of the network. Hereafter, we compare TDM versus FDM, based on network application scenarios.

In PONs thanks to the application of statistical TDM, end users can receive streams down to Mb/s speed. This differs from the proposed coherent solution with DSCM approach, as the laser instability does not allow it to operate at symbol rate well below 1 GBd. From a transmission point of view, PONs today use the highly cost-effective IM-DD, whose transceivers are manufactured in volumes exceeding millions of units per year [Nesset2017].

On the other hand, in the case of data rate upgrade in a PON network based solely on TDM, all devices (hub + leaf modules) need to be replaced because the same peak communication rate must be supported. With FDM, the technology used and power consumption applicable to lower data rate locations can be optimized for that lower data rate. Transmission and reception at the peak rate is not needed. Maintaining high peak rates in all locations may increase costs and integration challenges for those optoelectronics in applications that do not need the full channel rate. .

Next, PON TDM transceivers cannot exploit the capability of advanced DSP techniques. DSP algorithms have difficulties in the learning phase required for burst mode transmission that is used in today's PON. For this reason, data rate upgrade might be limited for what concerns reach and synchronization. The last is particularly critical as the leaf modules will be placed at different locations before the aggregation stage. Because of these limitations, PON standards advance at a slower pace compared to other technologies.

In opposition to PON based solely on TDM, FDM simplifies data rate upgrades and enhances the network flexibility. By using DSP and continuous mode transmission/reception in both directions, it can extend the reach, and it supports DWDM. An FDM network enables different costs for users versus hub, the upgrades are independent, and they can be performed per individual end user (at a given coarse granularity). Thanks to these characteristics, P2MP based on DSCM and coherent can achieve a capacity

of 400G. P2MP optics based on coherent transmission with DSCM offers a roadmap to higher end user capacities than IM-DD implementations can achieve, and as traffic demands in access networks grow, the proposed implementation provides a number of other advantages, as discussed above.

5. Beyond 10G Solution Comparison

A flexible access segment optical network topology should provide:

- Cost effective technology
- Scalability as well as ease of upgrading
- Sufficient distance reach
- Integration into outside plant node platforms
- Service convergence to allow for a unified access network segment
- Efficient use of hub space and power
- Long deployment lifespan
- Operational simplicity

Table 2 provides the authors' view of a summary comparison of the available technologies with these criteria in mind . A "+" score indicates a technology which excel in a particular area while "–" indicates the poor ability to support the desired criteria. For single-carrier coherent, some criteria are greatly impacted by the use of electrical aggregation (EA) and the "with EA" variation covers the case where electrical aggregation in the outside plant is added.

Table 3 - Beyond 10G Solution Comparison

Criteria	IM-DD 25G	Single-Carrier Coherent 100G/200G+	DSCM 25G-400G
Aggregate Module Capex	+	-- - with EA	-
Scalability	-	Neutral + with EA	++
Reach	-	+	+
Node Integration	+	Neutral	Neutral
Service Convergence	Neutral	+	+
Hub Space/Power	Neutral	Neutral	+
HFC DAA Suitability	+	- Neutral with EA	Neutral
PON R-OLT Suitability	Neutral	Neutral	+
Wireless xHaul Suitability	Neutral	Neutral	+
Deployment Lifespan	-	+	+
Operational Simplification	Neutral	-	+

Aggregate Module Capex

- 25G IM-DD solutions provide very reasonable module costs while coherent solutions are expected to be significantly higher than IM-DD for some time

- DSCM allows for a significant reduction in the number of transceivers due to the P2MP operation and optical splitting/combining

Scalability

- The selective use of subcarriers in DSCM allows for a single module to support a range of total bandwidths while the capacity is fixed in IM-DD; this could allow for specific variants of the modules which only serve small numbers of subcarriers

- Single carrier coherent can scale in capacity through modulation, but does not generally support scaling below 100G per link without the use of electrical aggregation

Reach

- IM-DD solutions in the 25G and above range (and especially in DWDM configurations) are not readily available beyond 40km while coherent solutions can provide reach out to 120 km in point-point operation. P2MP operation with DSCM can be achieved beyond 40 km at rates up to 100G with split ratios of 32:1 or less

Node Integration

- IM-DD 25G solutions using SFP28 are nearly identical in space and thermal footprint compared to 10G SFP+ used for DAA and R-OLT nodes today

- The extra power in the optics and DSPs in coherent solutions drive increased power beyond 7W and require larger form factors for node integration such as CFP2 to support effective passive cooling

Service Convergence

- Moving past 10G generally allows service convergence but the maximum limit may prevent higher rate wireless fronthaul or business services from being offered over the same access segment optical network

- Coherent solutions allow for greater data rates to more easily aggregate multiple services and support higher xhaul capacity needs

Hub Space/Power

- The additional power consumed by single-carrier coherent solutions is offset by the reduction in total switch ports if aggregation of multiple services is possible externally (outdoor coherent termination device for example)

- DSCM provides significant benefits for hub space and power due to the P2MP operation which allows for aggregation of multiple end points into a single high-rate hub module along with simple optical passive combining/splitting

HFC DAA Suitability

- Single-carrier coherent, because it cannot scale down without separate electrical aggregation, is not well suited to the 20-25G requirements of segment-able D3.1 and D4.0 DAA nodes

- DSCM can scale down to a single subcarrier to suit 25G needs but does come with tradeoffs on power and space in HFC DAA nodes already constrained in those areas

PON R-OLT Suitability

- 25G IM-DD supports 2:1 aggregation relative to 10G PON technologies (peak capacity design targets)

- DSCM scales well through the range of possible uplinks in a typical 4 port R-OLT (40-100 Gbps), even considering some oversubscription while single-carrier coherent overserves capacity in today's 10G PON environments

Wireless xHaul Suitability

- DSCM supports lower scale than single-carrier and the dynamic P2MP operation is very well suited to the elasticity and number of radio endpoints that are served

Deployment Lifespan

- IM-DD will eventually bump into limits as 25G is no longer sufficient for transport to last mile access devices

- Coherent solutions provide significant capacity to cover near and long-term requirements

Operational Simplification

- IM-DD 25G operations are just like DAA and R-OLT today

- Single carrier coherent will generally require a coherent termination device to fit the wide range of devices driving up transceiver count and adding another layer of electrical aggregation devices into the network

- DSCM use of P2MP operation ideally suits the needs of the network and allows passive optical splitting/combining instead of electrical aggregation

6. Conclusions

This article presents a comprehensive discussion on why access segment optical transport Beyond 10G is required while reviewing the solutions available using IM-DD and coherent transmission technologies. A new concept of coherent digital subcarrier multiplexing (DSCM) provides benefits in scalability using individual subcarriers while supporting P2MP operation to allow for transceiver reduction and replacement of electrical aggregation from single-carrier coherent with passive optical splitting and combining.

The near future is sure to be very active in this area with interoperability and standardization happening on multiple levels including the newly created OpenXR Forum and just launched CableLabs 100G Coherent PON projects.

Abbreviations

5G	Fifth Generation
10G	10 gigabits per second
25G	25 gigabits per second
ADC	Analog-to-Digital Converter
ASIC	Application-Specific Integrated Circuit
CCAP	Converged Cable Access Platform
CFP2	C Form-Factor Pluggable Half-Size
CIN	Converged Interconnect Network
CMOS	Complementary Metal-Oxide Semiconductor
CU	Centralized Unit
D4.0	DOCSIS Version 4.0
DAA	Distributed Access Architecture
DAC	Digital-to-Analog Converter
DP-DQPSK	Dual Polarization Differential Quadrature Phase Shift Keying
DSCM	Digital Subcarrier Multiplexing
DSLAM	Digital Subscriber Line Access Multiplexer
DSP	Digital Signal Processing
DU	Distributed Unit
DWDM	Dense Wavelength Division Multiplexing
EA	Electrical Aggregation
EEPN	Equalization Enhanced Phase Noise
FDM	Frequency Division Multiplexing
FEC	Forward Error Correction
FTTH	Fiber To The Home
GAP	Generic Access Platform
GBd	Gigabaud
Gbps	Gigabits per second
HFC	Hybrid-Fiber Coax
IM-DD	Intensity-Modulated Direct-Detection
IP	Internet Protocol
MACPHY	Media Access Control + Physical Layer
MEC	Mobile Edge Compute
MIMO	Multiple-Input, Multiple-Output
NRZ	Non-Return-to-Zero
OIF	Optical Internetworking Forum
OLT	Optical Line Terminal
P2MP	Point-to-multipoint
P2P	Point-to-point
PAM	Pulse-Amplitude-Modulation
PCS	Probabilistic Constellation Shaping
PON	Passive Optical Network
QSFP	Quad Small Form Factor Pluggable
R_s	Symbol Rate
R-OLT	Remote OLT
RAN	Radio Access Network
RU	Radio Unit

S-BVT	Sliceable Bandwidth Variable Transceiver
SC	Subcarrier
SFP+	Small Form Factor Plus
T_s	Symbol Period
TDM	Time-Division Multiplexing
WDM	Wavelength Division Multiplexing

Bibliography & References

[OIFIA2020] OIF-400ZR-01.0: *Implementation Agreement 400ZR*; Optical Internetworking Forum (OIF); https://www.oiforum.com/wp-content/uploads/OIF-400ZR-01.0_reduced2.pdf

[Sun2020] H. Sun *et al*., "800G DSP ASIC Design Using Probabilistic Shaping and Digital Sub-Carrier Multiplexing," in *Journal of Lightwave Technology*, vol. 38, no. 17, pp. 4744-4756, 1 Sept.1, 2020, doi: 10.1109/JLT.2020.2996188.

[Nesset2017] D. Nesset, "PON roadmap," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. A71–A76, Jan 2017.

[Welch2021] D. F. Welch *et al*., "Point-to-Multipoint Optical Networks Using Coherent Digital Subcarriers," in *Journal of Lightwave Technology*, doi: 10.1109/JLT.2021.3097163.

[Bäck2020] J. Bäck *et al*., "CAPEX Savings Enabled by Point-to-Multipoint Coherent Pluggable Optics Using Digital Subcarrier Multiplexing in Metro Aggregation Networks," *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1-4, doi: 10.1109/ECOC48923.2020.9333233.

[Sun2008] H. Sun *et al*., "Real-time measurements of a 40 Gb/s coherent system," *Optics Express*, vol. 16, no. 2, pp. 873–879, 2008.

[Roberts2009] K. Roberts *et al*., "Performance of dual-polarization QPSK for optical transport systems," *Journal of lightwave technology*, vol. 27, no. 16, pp. 3546–3559, 2009.

[Buchali2016] F.Buchali *et al.*, "Rate adaptation and reach increase by probabilistically shaped 64- QAM: An experimental demonstration," *Journal of Lightwave Technol- ogy*, vol. 34, no. 7, pp. 1599–1609, 2016.

[Sambo2015] N. Sambo *et al.*, "Next generation sliceable bandwidth variable transponders," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 163–171, 2015.

[Dris2013] S. Dris *et al.*, "M-QAM carrier phase recovery using the Viterbi-Viterbi monomial-based and maximum likelihood estimators," in *2013 OFC*. IEEE, 2013, pp. 1–3.

[Rahman2016] T. Rahman *et al.*, "Digital subcarrier multiplexed hybrid QAM for data-rate flexibility and roadm filtering tolerance," in *Optical Fiber Communication Conference*, 2016, pp. Tu3K–5.

[Velasco2021] L. Velasco *et al.*, "Autonomous and energy efficient lightpath operation based on digital subcarrier multiplexing," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2021

[Marino2021] P. P. Marino *et al.*, "Point-to-Multipoint Coherent Optics for Re-thinking the Optical Transport: Case Study in 5G Optical Metro Networks," *2021 International Conference on Optical Network Design and Modeling (ONDM)*, 2021, pp. 1-4, doi: 10.23919/ONDM51796.2021.9492393

[Cisco5G] Bringing it All Together in 5G, S. Ajmeri, BRKSPG-3477, https://www.ciscolive.com/c/dam/r/ciscolive/apjc/docs/2020/pdf/BRKSPG-3477.pdf, 2020

Designing a Cloud-Based DOCSIS Time Protocol Calibration Database

A Technical Paper prepared for SCTE by

Roy Sun
Ph.D., Lead Architect
CableLabs
858 Coal Creek Cir, Louisville CO, 80027
303-661-6789
r.sun@cablelabs.com

Rahil Gandotra, Ph.D., and Mark Poletti, CableLabs, Inc.

Jennifer Andreoli-Fang, Ph.D., Amazon Web Services (AWS)

Elias Chavarria Reyes, Ph.D., Hitron Technologies, Inc.

John Chapman, Cisco Systems, Inc.

1. Introduction

This paper will detail the design and implementation of a cloud-based application to enable DOCSIS Time Protocol (DTP) deployment. This application is unique to the cable environment and its implementation can serve as a reference design for future cable cloud applications.

The key components of the app will be a database and an application programming interface (API). The application can be implemented on multiple cloud platforms. We chose to prototype our application on Amazon Web Services (AWS). The architecture design will follow the AWS Well-Architected Framework to leverage the benefit of the cloud and will provide horizontal scalability, pay-as-you-go cost structure, and high availability. The application will also be designed for service assurance, which is critical in field deployments. All cloud-based components and a CMTS emulator with the API client will be implemented on AWS.

DTP was invented in 2011 by John Chapman of Cisco [1], in anticipation of using the DOCSIS network to provide timing as a service (TaaS). In 2012, DTP was standardized as part of DOCSIS 3.1 [2]. This version of DTP defined the core algorithm and functionality. Since the primary use case for DTP is mobile backhaul over DOCSIS, CableLabs introduced the SYNC specification [3] in 2020 to address this use case. As part of the SYNC specification initiative, Elias Chavarria and John Chapman from Cisco redesigned the DTP algorithm to bypass a set of limiting assumptions in the original DTP design [4]. Also, in 2020, a group of companies, including Cisco, Hitron, Charter, and CableLabs, did a proof of concept (PoC) to validate the performance of DTP. The test results of the proof of concept are reported in a separate Society of Cable Telecommunications Engineers (SCTE) paper this year [5] and a CableLabs technical report [6].

At its core, DTP allows a DOCSIS network to interface its native timing and frequency to external timing protocols. To do this, DTP establishes a set of techniques and DOCSIS signaling messages between the cable modem (CM) and the cable modem termination system (CMTS).

In the original DTP design, the DTP messages contained CM and CMTS timing parameters. The underlying assumption was that the timing parameters for the CM could be measured separately from the CMTS ones before deployment. The timing parameters in the CMTS and CM would have been scalable in this distributed measurement model. For the assumption to be valid, a testing device capable of measuring those timing parameters needs to exist. However, no such testing device exists. Therefore, the usability of the original DTP design was limited. For this reason, DTP was redesigned as part of the SYNC specification effort [2].

In the redesigned DTP, the timing parameters are no longer measured separately for the CM and the CMTS, they are instead measured jointly. With this change, existing testing devices in the market can be used. A consequence of this change is that the number of measurements grows from linear to exponential. If n number of CMTS products and m number of CM products support DTP, the original DTP design required n+m measurements of timing parameters. The redesigned DTP calls for $n * m$ measurements of timing parameters.

The values of the DTP timing parameters that a DOCSIS network should use also depend on the CMTS and CM configuration, e.g. the interleaver configuration. For example, if there are w number of different interleaver configurations, the total number of measurements of timing parameters grows to $n * m * w$. In Phase 2 of the DTP PoC that Cisco, Hitron, Charter, and CableLabs are conducting [5], they will assess the impact of other configuration parameters, e.g., modulation, cyclic prefix, and frame size. If all these configuration parameters impact the DTP timing parameters, the total number of measurements will

continue growing exponentially. Assuming x different modulations, y different cyclic prefixes, and z different frame size, a total of $n * m * w * x * y * z$ timing measurements would be required.

The cable industry has two options to handle the measurement, storage, update, and accessibility of all the DTP timing measurements. One option is for every CM and CMTS vendor to do everything by itself, which would lead to a replication of effort and resource investments with little added value for the vendors. The second option is for a common entity, such as CableLabs, to lead the measurement, storage, update, and accessibility of the DTP timing parameters. This second option allows the cable ecosystem – both vendors and operators – to leverage a shared pool of resources. The opinion of the authors is that the second option is more feasible for the cable industry. To make this second option a reality, the authors propose a cloud application, as discussed in the following sections.

The DTP calibration includes three major steps: 1) collect the calibration data; 2) build a cloud app that distributes the data; and 3) a CMTS to access and apply the calibration data. The DTP calibration test could be conducted in many test labs. For example, CableLabs/Kyrio established a Network Timing Lab that could evaluate the DTP performance and collect the calibration data. One of the key contributions of this paper is to present the design of API and the AWS cloud server that distributes the calibration data. The CMTS will access the database to obtain calibration values in real-time via the API. Using these values, the CMTS will calculate the timing offsets for the CM and the 5G radios. CableLabs expects to continuously sponsor the application to enable the commercial deployment of DTP.

2. DTP Calibration Method

In this section, we further explain why DTP needs calibration and how to do the calibration. The DTP timing diagram is shown in Figure 1, where DS-T and US-T denote downstream (DS) and upstream (US) delays inside the CMTS, DS-H and US-H are DS and US delays in the hybrid fiber-coaxial (HFC) plant, and DS-C and US-C are DS and US delays caused by the CM. The CMTS sends the DOCSIS 3.1 timestamp. The timestamp is delayed when it arrives at the CM. In other words, the timestamp that arrived at the CM represents an early version of the CMTS timestamp. The time error (TE) is in the DS only. Ideally, DS-T, DS-H, and DS-C should be measured separately and used by DTP. However, DOCSIS does not provide the reference points to measure these delays. DTP provides a practical way to calibrate the DOCSIS 3.1 timestamp using the true ranging offset (TRO).

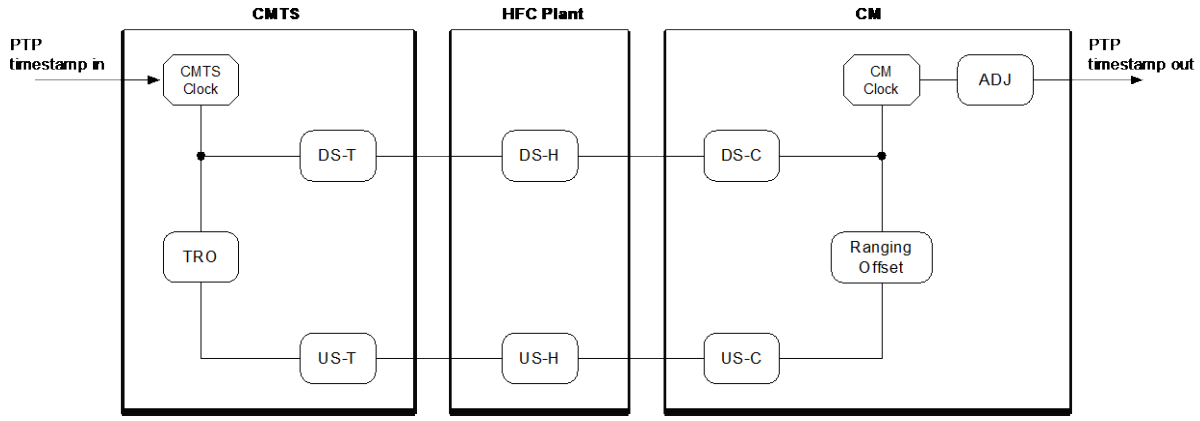


Figure 1 - DTP Timing [3]

DTP messages are exchanged between CMTS and CM, see Figure 2 for the case when the CMTS is the DTP master and the CM is the DTP slave. The CM measures the round-trip delay as the TRO. The CM

reports the TRO to the CMTS in the "DTP-Response" message. The CMTS sends the time adjustment *t-adj*, or *t-cm-adj*, to the CM using the "DTP-Info" message. The *t-cm-adj* is approximately equal to half of the TRO. The CM timestamp over its CM to CPE interface (CMCI) ports is equal to the DOCSIS 3.1 timestamp plus *t-cm-adj*. The *t-cm-adj* corrects the time error in the network. Note that the propagation delay through coaxial cable and fiber is theoretically symmetrical.

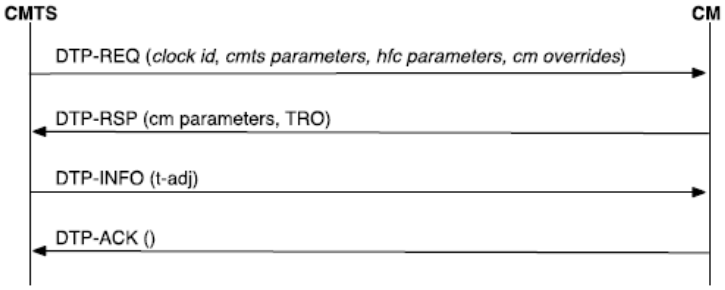


Figure 2 - DTP Message Flow [3]

If *t-cm-adj* is set to be half of the TRO, *t-cm-adj* cannot correct asymmetrical delay (different in DS and US). The asymmetrical delay is introduced by devices like CMTS, CM, remote physical RF layer (R-PHY) and remote physical and MAC layers (R-MACPHY), and any HFC elements. This asymmetrical delay needs to be measured in lab for each pair of CMTS and CM combos. The measured asymmetrical delay could be computed at the CMTS. For example, the following method is supported by the Cisco integrated CMTS (I-CMTS) cBR-8:

$$t\text{-}cm\text{-}adj = t\text{-}tro/2 + y, \quad (1)$$

where *y* is an additional time adjustment that applies in the CMTS to calibrate the DS delay in the DOCSIS 3.1 timestamp. This approach is described as method 2 in Section 5.4 in [6]. CM location, plant length, and other symmetrical TE are taken care of by the TRO. The asymmetrical TE is addressed by the additional time adjustment *y*. This additional time adjustment *y* can be mapped to formula (18) in the SYNC spec [3], which we copy here:

$$t\text{-}cm\text{-}adj = t\text{-}cm\text{-}adj\text{-}R + [t\text{-}tro + t\text{-}hfc\text{-}ds\text{-}o - t\text{-}hfc\text{-}us\text{-}o - t\text{-}tro\text{-}R]/2, \quad (2)$$

where *t-cm-adj* is the live time adjustment that the CMTS sends to the CM, while *t-cm-adj-R* is the value of the DTP time adjustment used in the calibration test that brings the average PTP two-way time error to zero [3]. Similarly, *t-tro* is the live TRO that the CM measures and sends to the CMTS. *t-tro-R* is the TRO reported by the CM in the lab calibration test. *t-hfc-ds-o* and *t-hfc-ds-o* represent any fixed delay elements in the HFC plant that contribute to delay [2], in DS and US, respectively. *t-hfc-ds-o* and *t-hfc-ds-o* are provided by the CMTS to the CM.

The above formula can be rearranged as:

$$t\text{-}cm\text{-}adj = t\text{-}tro/2 + [t\text{-}cm\text{-}adj\text{-}R + t\text{-}hfc\text{-}ds\text{-}o/2 - t\text{-}hfc\text{-}us\text{-}o/2 - t\text{-}tro\text{-}R/2], \quad (3)$$

Comparing Eq. (3) to Eq. (1), we get that:

$$y = t\text{-}cm\text{-}adj\text{-}R + t\text{-}hfc\text{-}ds\text{-}o/2 - t\text{-}hfc\text{-}us\text{-}o/2 - t\text{-}tro\text{-}R/2. \quad (4)$$

Note that if the HFC plant is assumed to introduce no asymmetry, i.e., *t-hfc-ds-o/2* = *t-hfc-us-o/2*, then Eq. (4) is further simplified as:

$$y = t\text{-}cm\text{-}adj\text{-}R - t\text{-}tro\text{-}R/2. \quad (5)$$

Eq. (5) captures the asymmetry in the reference-length plant (see Section 6.4.1.1 in [3]) introduced jointly by the I-CMTS and CM (in an I-CMTS architecture), or jointly by the RPD and the CM (in an R-PHY architecture). In DTP calibration, the values for *t-cm-adj-R* and *t-tro-R* that make the average PTP two-way time error to zero will be measured in the lab and the additional time adjustment *y* will be distributed by the AWS database. *t-tro* and *t-cm-adj* in Eq. (1) will be calculated lively in the field.

In summary, as shown in Figure 3, DOCSIS 3.1 timestamp is used in DOCSIS 3.1 networks that include a delay in DS. DTP uses TRO to correct the symmetrical part of the DS delay. DTP calibration further corrects the asymmetrical part of the DS delay. For example, DTP calibration reduced the time error from 3,223,800 ns to 13-31 ns as reported in Section 8.4 in [6].

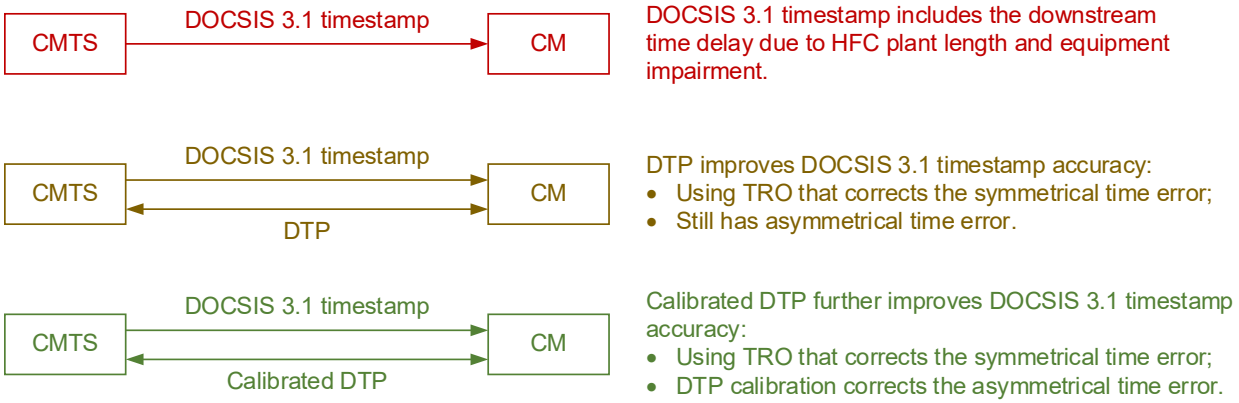


Figure 3 - DOCSIS 3.1 Timestamp, DTP and DTP Calibration

3. DTP Calibration Cloud Database Design

On a high-level, the primary components of the DTP calibration cloud database are – a database, a web-based graphical user interface (GUI) to provide a human interface for the lab engineer to add, read, and delete the DTP calibration entries, and an API framework to provide a machine interface for the CMTS to fetch the DTP calibration entries. Since the components to be implemented were functionally fairly uncomplicated, the decision to select a cloud provider was also relatively uncomplicated. AWS was selected as the cloud provider for DTP calibration since it is the most mature and enterprise-ready provider [7].

AWS provides multiple different methods to implement the abovementioned three components depending on the amount of modularity required in the system design. Figure 4 shows three possible system designs to implement a database, a web GUI, and an API in AWS with varying levels of modularity. Figure 4(a) illustrates implementing the three components within a single compute instance. While this approach provides vertical scaling capabilities, it offers limited flexibility and availability benefits. In Figure 4(b), the database is transitioned out to its own module while the web GUI and API reside on a single compute instance. Disaggregating the database offers added benefits of leveraging a database management system to abstract away the setup, operation, and scaling tasks. In Figure 4(c), all three components are implements in their own module. This approach provides the most modularity in terms of allowing

independent management and innovation of each component while hiding the complexity of each part behind an abstraction and interface.

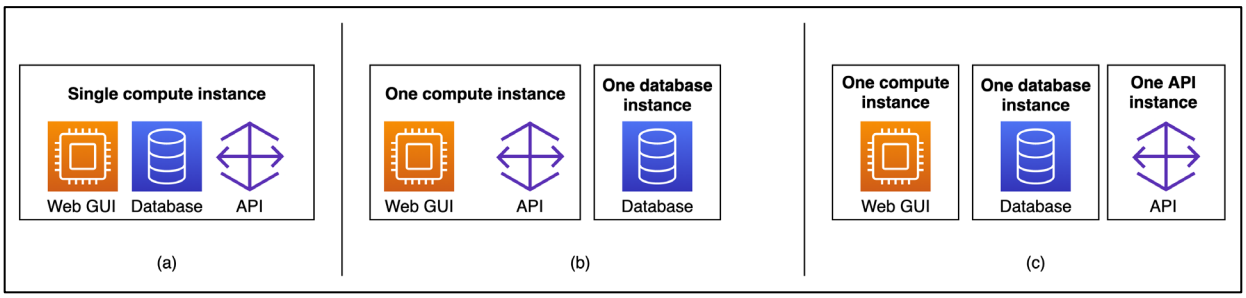


Figure 4 - Possible System Designs of the DTP Calibration Cloud Database

The complete framework of the DTP calibration cloud database is shown in Figure 5. The database is hosted using Amazon Relational Database Service (RDS), the web GUI is running on an Elastic Compute Cloud (EC2) instance, while the API framework includes Amazon API Gateway as the frontend and AWS Lambda service as the backend. To validate the end-to-end functionality of the cloud application, a CMTS emulator was developed to test sending requests and receiving responses from the cloud application.

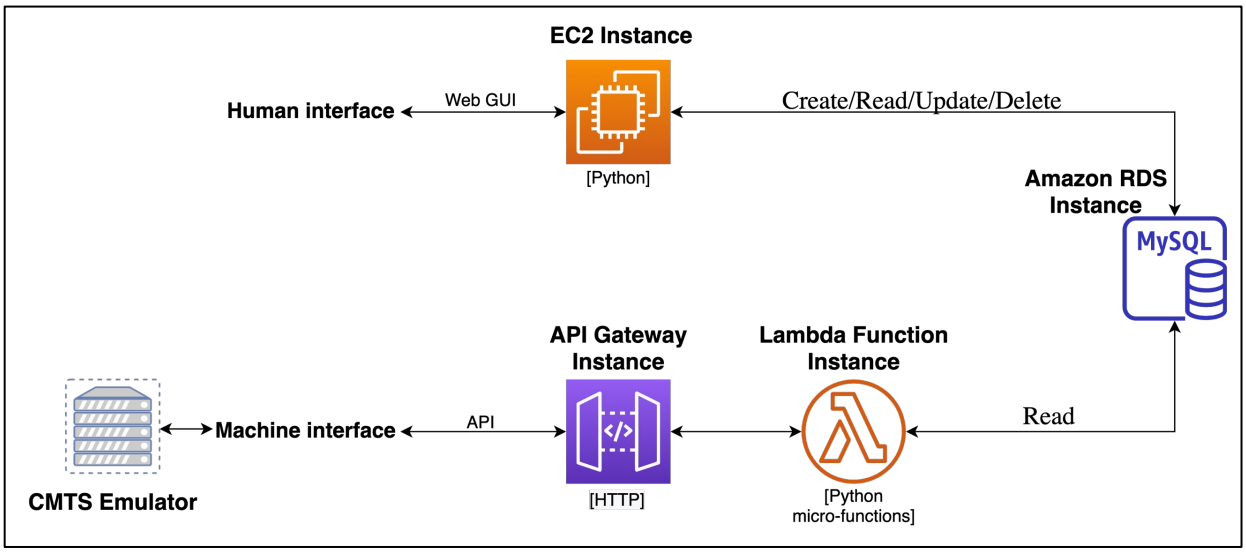


Figure 5 - Functional Design Diagram of DTP Cloud Calibration Database

Since the structure of the data to be stored in the database is not expected to change frequently, a relational database was selected as the database type. Amongst the different relational database engines available, MySQL was chosen as it is open-source and provides sufficient flexibility to run on any operating system. A Python script was developed to instantiate and create the database schema allowing for any possible changes in the future. Database reliability could be enhanced by utilizing the RDS Multi-AZ (Availability Zone) functionality provided by AWS wherein a standby database instance is automatically created in another AZ and data is synchronously replicated between the two instances.

The web GUI is implemented in Python using Flask [8]. Flask was selected as the web framework as it allows for easy addition of libraries or plugins for an extension and comes with a built-in development server and fast debugger. Additional modules for handling forms and enabling login using username and

passwords were implemented using Flask extensions. The web GUI provides the user with the capability to add new calibration entries to the database, read existing entries from the database, update existing entries in the database, and delete any existing entry from the database.

AWS offers two types of API Gateways – HTTP-based and websocket-based. HTTP APIs were selected as the API Gateway type since the communication between the CMTS and the cloud application is expected to be stateless and not based on stateful real-time two-way communication applications such as chat apps or streaming dashboards which require websocket APIs. Amazon's HTTP API Gateway provides API proxy functionality and low-latency, cost-effective integrations with other AWS services. Both HTTP GET and POST methods are implemented to enable sending query parameters via the URL of the GET request and the request body of the POST message. Since the connection between the CMTS and the cloud application is not expected to be persistent and would be infrequent (only when there is a new configuration pairing of CMTS-CM), the API backend also does not require persistent compute. Therefore, serverless Lambda functions are used in the API backend to consume compute resources only when needed – in case of an incoming request from a CMTS. The API Gateway is configured to pass the query/payload received from the client to the Lambda function and return the function's response to the client. A Lambda function is developed to run a database read query based on the received query/payload parameters (CMTS-CM Hardware-Firmware versions) and return the result (timing parameters) along with a valid HTTP status code.

The CMTS emulator is essentially an API client developed to test the responses of the cloud application by sending HTTP requests to the API Gateway with different Hardware-Firmware combinations. The content-type is set to application/json for these requests and responses. JSON is used as the payload format as it is lightweight and is suitable for both human reading and machine parsing. The motivation behind developing this emulator is to demonstrate the work needed by CMTS vendors to enable the remote collection of timing parameters from the cloud application.

The current application design also supports adding calibration entries, in addition to reading them, using the API Gateway and Lambda function framework. This would allow automation in the calibration process wherein a large number of timing entries could be added in the database without requiring manual work. Application security is considered at two levels: (i) API level: Access to the API can be restricted by either using HTTP request parameters-based authorization (such as username/password) or by using token-based authorization (such as JSON Web Token, JWT), and (ii) Network level: Access to the virtual network where the API framework is hosted can be restricted to known CMTS IP addresses only.

4. DTP Calibration Cloud Database API Design

4.1. Cloud App Message Flow

The message flow between the AWS server and CMTS is HTTP-based, as shown in Figure 6. HTTP uses TCP as transport layer to provide reliable network transmission using acknowledgments. Our DTP API uses "HTTP request" and "HTTP response" to exchange information between the CMTS and AWS server. The CMTS (client) sends Cal-data Request to the server that is contained in the HTTP Request. The Cal-data Request message includes network architecture and hardware/firmware combinations, see Section 4.2. The server sends Cal-data Response to the client that is contained in the HTTP Response. The data elements contained in the Cal-data Response message is described in Section 4.2.

The client sets a timer (i.e., 500 ms) using a status code when sending the Cal-data Request message. The client checks the status code and the received calibration data. If failure, the client resends the Cal-data Request message. The AWS server may provide multiple IP addresses for redundancy, and the client may

try another server IP address if failure over five times. There could be three failure reasons: 1) timer expired; 2) network architecture and Hardware/Firmware combinations do not match requested data; and 3) calibration data (test configurations or measured time error results) incomplete.

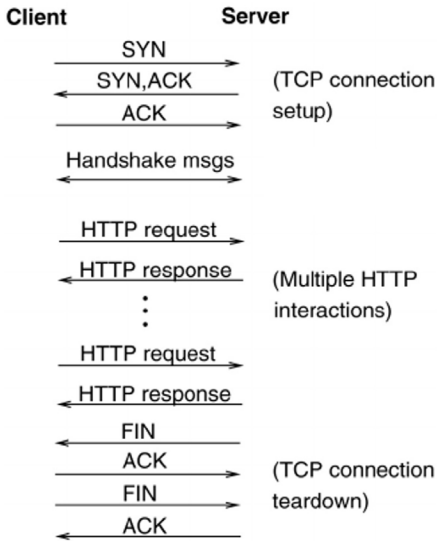


Figure 6 - HTTP and DTP Calibration Message Flow

4.2. Cloud App Data Structure

The data structure is illustrated in Table 1. Cal-data Request message only includes the first two data elements: network architecture and hardware/firmware combinations. Cal-data Response message includes all elements listed in Table 1. DTP could be deployed on the traditional network with I-CMTS and distributed access architecture (DAA). DAA also includes remote-PHY and remote -MACPHY. Network architecture is represented by two bits. Hardware and firmware combinations includes make, hardware and firmware version of I-CMTS chassis, I-CMTS line card, RPD (for DAA only), and CM.

Test configurations considered in the initial version of the cloud app include DS modulation scheme, DS interleaver, DS cyclic prefix, US modulation scheme, US cyclic prefix, and US frame size. DTP PoC [5] phase 2 testing is evaluating the impact of each parameter. Parameters that do not strongly impact DTP performance will be removed from Cal-data Response message. Other parameters that strongly influence DTP performance but are not listed above will be added to the Cal-data Response message.

Table 1 - Example Data Structure.

Network Architecture	Hardware & firmware combinations	Testing lab	Testing date	Test configurations	Additional time adjustment *y* (ns)	Constant Time Error (ns)
I-CMTS	Combo 1	CableLabs	11/1/2020	Config 1	200,000,000	-50
			11/1/2020		300,000,000	…
			11/2/2020	Config N	400,000,000	100
DAA	Combo 2	CableLabs	12/1/2020	Config 1	200,000,000	50
			12/1/2020		300,000,000	…
			12/1/2020	Config N	400,000,000	-100

5. Conclusion

4G/5G mobile networks require a high-accuracy synchronization source in the backhaul where the GPS signals are unavailable. DTP provides such sync signals in the backhaul over HFC networks. DTP needs automated calibration in the field to guarantee time accuracy.

This paper presented the DTP calibration method and design of a cloud app that distributes the calibration data. The cloud app is prototyped on AWS. A modular application design is developed allowing to leverage various cloud benefits such as abstraction, automation, and high-availability. The web GUI is implemented in Python using Flask allowing an engineer to add, read, and delete the DTP calibration data entries. The API uses HTTP protocol with JSON as the data format, and calibration data message flow was designed. Security and reliability enhancement features are considered and will be added based on costumers' requirements. Future work of automated DTP calibration includes collecting calibration data in test labs and adding them into the AWS database via the Web GUI. CMTS will need to add the corresponding feature to inquiry and apply the calibration data automatically. Proof-of-concept test for the AWS cloud app and automated DTP calibration is planned in the near future [5].

Abbreviations

API	application programming interface
AWS	Amazon Web Services
AZ	availability zone
CM	cable modem
CMTS	cable modem termination system
cTE	constant time error
DAA	distributed access architecture
DOCSIS	Data-Over-Cable Service Interface Specification
DS	downstream
DTP	DOCSIS Time Protocol
EC2	elastic compute cloud
GUI	graphical user interface
HFC	hybrid fiber-coaxial
I-CMTS	integrated cable modem termination system
PoC	proof of concept
RDS	Relational Database Service
RPD	remote physical layer device
R-PHY	remote physical RF layer
SCTE	Society of Cable Telecommunications Engineers
TaaS	timing as a service
TE	time error
TRO	true ranging offset
US	upstream

Bibliography & References

[1] John T. Chapman, Rakesh Chopra, Laurent Montini., "The DOCSIS® Timing Protocol (DTP), Generating Precision Timing Services from a DOCSIS System," *INTX/SCTE Spring Technical Forum*, 2011. [link]

[2] Cable Television Laboratories, Inc., "DOCSIS® MAC and Upper Layer Protocols Interface Specification", CM-SP-MULPI, December 2020. [link]

[3] Cable Television Laboratories, Inc., "Synchronization Techniques for DOCSIS® Technology Specification," CM-SP-SYNC, April 2021. [link]

[4] Elias Chavarria Reyes, John T. Chapman, "How the DOCSIS Time Protocol makes the SYNC Specification Tick," SCTE Cable-Tec Expo Fall Technical Forum, Denver, October 2020. [link]

[5] Ruoyu Sun, Jennifer Andreoli-Fang, Elias Chavarria Reyes, John T. Chapman, et al., "DOCSIS Time Protocol Proof of Concept," in SCTE-Expo 2021, Atlanta, GA, October 11-14, 2021.

[6] Cable Television Laboratories, Inc., "DOCSIS Time Protocol Proof of Concept Phase I Technical Report CM-TR-DTP-V01-210915," September, 2021. [link]

[7] M. A. Kamal, H. W. Raza, M. M. Alam, and M. M. Su'ud, "Highlight the features of AWS, GCP and Microsoft Azure that have an impact when choosing a cloud service provider," International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 5, Jan. 2020.

[8] "Welcome to Flask." Accessed on: Jul. 12, 2021. [Online]. Available: https://flask-doc.readthedocs.io/en/latest/.

Detecting and Mitigating Distributed Denial of Service Attack with Transparent Security

A Technical Paper prepared for SCTE by

Randy Levensalor
Principal Architect
CableLabs
858 Coal Creek Circle Louisville, CO 80027
(303) 661-3455
r.levensalor@cablelabs.com

Chris Sibley
Senior Engineer
Cox Communications
6305 Peachtree Dunwoody Rd. Atlanta, GA 3034
404-269-6701
chris.sibley@cox.com

1. Introduction

Transparent Security is an open-source solution for identifying and mitigating distributed denial of service (DDoS) attacks and the devices (e.g., Internet of Things [IoT] sensors) that are the source of those attacks. Transparent Security is enabled through a programmable data plane (e.g., "P4"-based) and uses in-band network telemetry (INT) technology for device identification and mitigation, blocking attack traffic where it originates on the operator's network.

Cox Communications and CableLabs conducted a proof-of-concept test of the Transparent Security solution in the Cox lab in late 2020. Testing was primarily focused on the following major objectives:

- Compare and contrast performance of the Transparent Security solution against that of a leading commercially available DDoS mitigation solution.
- Validate that INT-encapsulated packets can be transported across an IPv4/IPv6/Multiprotocol Label Switching (MPLS) network without any adverse impact to network performance.
- Validate that the Transparent Security solution can be readily implemented on commercially available programmable switches.

CableLabs and Cox completed the lab trial in conjunction with Intel® and Arista Networks. Transparent Security was able to identify and mitigate attacks in less than one second as compared to greater than one minute for the leading vendor. We also validated that inserting and removing the INT header had no observable impact on throughput or latency.

Distributed denial of service (DDoS) attacks and other cyberattacks cost operators billions of dollars, and the impact of these attacks continues to grow in size and scale, with some exceeding 1 Tbps. The number of Internet of things (IoT) devices continues to grow rapidly, many have poor security, and upstream bandwidth is ever increasing—this perfect storm has led to exponential increases in IoT attacks, by over 600% between 2016 and 2017 alone. With an estimated increase in the number of IoT devices from 5 billion in 2016 to over 20 billion in 2020, we can expect the number of attacks and the size of attacks to continue this upward trend.

Detecting attacks is difficult, but mitigating them is even harder, and several solutions have been proposed with varying degrees of success. Typically, these solutions focus on blocking attacks that originate offnet and which targets an on-net resource. Even solutions that do identify attacks that originate on-net are limited in that they cannot block the attack traffic until after it has already traversed the access network. As such an operator's access networks can still be seriously affected, resulting in connectivity loss and quality of service (QoS) issues for customers.

Enhanced device visibility and packet processing can be used to identify the sources of such attacks. Leveraging programmable ASICs and P4 applications provides support for these enhancements by making the behavior of the data plane expressible in software and customizable without affecting performance. Specifically, this project will pair a DOCSIS modem with a series of P4-enabled devices connecting back to the operator headend via the P4 Runtime or Barefoot Runtime Interface (BRI). Both the P4 Runtime and BRI leverage GRPC as the underlying protocol. This architecture allows for visibility throughout the access network.

Transparent Security can leverage a machine-learning controller that has been trained with patterns to identify conditions and to perform dynamic operations by deploying new packet processing behaviors in the network (e.g., DDoS mitigation, virtual firewall, QoS detection/enforcement, and DOCSIS data plane functions). All operations will be performed at line rate while leveraging P4 in-band network telemetry

(INT), which allows data to be collected for reporting and analysis without control plane intervention. By inserting telemetry into the packet header, telemetry can be added to all packets rather than simply a sampling, which significantly reduces the time required to identify and mitigate the attack.

2. Motivation

As the proliferation of IoT devices continues to increase, the number of devices that can be compromised and used to participate in DDoS attacks also increases. At the same time, the frequency of DDoS attacks continues to grow because of the widespread availability of DDoS for-hire sites that allow individuals to launch DDoS attacks for relatively little cost. These factors contribute to a trend of malicious traffic increasingly using upstream bandwidth on the access network.

Typical DDoS mitigation solutions use techniques such as BGP diversion and Flowspec to drop traffic at certain parts of the network. However, mitigating outbound attacks using these techniques isn't entirely effective because the malicious traffic will have already traversed the access network, where it has the greatest negative impact before the traffic can be diverted to a scrubber or dropped by a Flowspec rule. Additional information on Flowspec can be found in section 4.2.2.2.

Transparent Security offers the promise of near-instantaneous detection of outbound attacks, as well as the ability to mitigate that attack at the source, on the customer premises equipment (CPE), thereby preventing that traffic from using upstream access network resources.

In addition to Transparent Security's DDoS mitigation capabilities, there are additional benefits to network performance/visibility in general. Implementation of Transparent Security on the CPE means that network operators can derive the specific device type associated with a given flow. This allows the operator to determine the type of IoT devices being leveraged in the attack.

This also opens myriad other possibilities—for example, reducing truck rolls by enabling customer service personnel to determine that a customer's issue is with one specific device versus all the devices on the internal network. Another example would be the capability to track the path a given packet followed through the network by examining the INT metadata.

Consumers will see a direct benefit from Transparent Security. Once compromised devices are identified, the consumer can be notified to resolve the issue or, alternatively, rules can be pushed to the CPE to isolate that device from the internet while allowing the consumer's other devices continued access. Such isolation mitigates the additional harm coming from compromised devices. This additional harm can take the form of degraded performance, exfiltration of private data, breaks in presumed confidentiality in communications, as well as the access network bandwidth consumed through DDoS. Less malicious traffic on the network provides for a better overall customer experience.

3. The History and Updates of Transparent Security

We initially released the Transparent Security architecture and open-source reference implementation in October 2019. Since then, we've achieved several milestones:

- Added source-only metadata to the P4 in-band telemetry specification, along with Transparent Security as an example implementation.
- Added support in the Telemetry Report 2.0 specification to collate multiple packet headers in a single telemetry report.
- Released a document titled "Transparent Security: Personal Data Privacy Considerations."

- Created a Transparent Security landing page.

Most proposed DDoS solutions fall into one of two categories, detection, or mitigation. This section discusses additional possible solutions in those categories and examines their advantages and disadvantages.

4. Existing Solutions

4.1. DDoS attack detection

DDoS attack detection is typically identified by an analytics engine identifying network traffic trends. These trends are based on routers sampling a summary of the packets as they enter the network. These samples can use IPFIX, NetFlow and SFlow to export this data.

The sampling technique cannot offer an operator a complete view of the network. Sampling is a statistics-based method for measuring network traffic by collecting, storing, and analyzing a sample of traffic data. This method has the advantage of allowing many interfaces to be monitored without significantly affecting network traffic.

IPFIX, SFlow and NetFlow are protocols used for sampling data. IPFIX and NetFlow are very similar, with IPFIX extending NetFlow v9 (https://www.oreilly.com/library/view/practical-network-scanning/9781788839235/1d6b69c7-62c3-40d0-be58-1ad82b22c115.xhtml). All these methods are limited to the data in the packet and do not contain any information about the network path used by the packet, validation for the source IP address or visibility to the source IP/device on a network behind a NAT.

Although the sampling method can effectively identify larger trends, it can miss anomalies that occur in a smaller portion of the traffic. That is, it can effectively detect a DDoS attack directed at a target, but it can miss or take longer to detect a DDoS attack originating at a given source. With source-based DDoS attacks, the system is trying to identify smaller anomalies, which are less likely to be captured in the sample.

4.2. DDoS Attack Mitigation

There are a wide variety of methods for mitigating DDoS attacks. A few of the common methods and the methods used by Transparent Security are highlighted below. This list is not exhaustive.

Also examined in this section are the points on the network where the mitigation is performed.

4.2.1. Network Locations for Mitigation

4.2.1.1. Out of Band

Out-of-band DDoS mitigations come in two flavors, appliances and scrubbing services. When an attack is detected against a host in the network, the traffic is routed through the out-of-band device, where it can remove or re-route the malicious traffic. An appliance routes traffic to the target to itself then removes the malicious traffic and provides a clean flow to the target. A service functions similarly except it routes target traffic to a mitigation center that cleans the traffic and forwards it to the destination.

However, both methods of out of band DDoS mitigation have downsides. The appliance is generally costly, and deployment can be complex. The service has difficulty defending low-bandwidth slow attacks,

and every interface must be protected, or it can fail to stop some attacks. Both methods can add latency to network traffic during an attack, and they generally rely on other methods for attack detection.

4.2.1.2. In Band

Traditional in-band DDoS mitigations are appliance based and work in the network path, comparable to a firewall, allowing the method to see all network traffic and react accordingly. As a result, in-band mitigation can provide detection as well as mitigation. However, it is costly, adds to network complexity and introduces another point of failure in the network.

4.2.1.3. At the Source

Transparent Security enables network operators to provide DDoS detection and mitigation at the source. This method protects target organizations with little to no effort on their part. It can also provide insight into hacked devices on the customer premises. This information can go a long way into preventing future attacks. Mitigation at this location has the added benefit of limiting attack traffic on an operator's network by blocking malicious packets before they enter the core network. Transparent Security mitigates primarily at the source, but the method is compatible with additional in-band target-based mitigation methods.

4.2.1.4. At the Target

While DDoS mitigation can be deployed at the target, it is generally of little use unless the attack is small (less bandwidth than the access circuit and less packets per second than the CPE is capable of processing).

4.2.2. Methods for mitigating DDoS attacks

4.2.2.1. Scrubber

As the name suggests, a scrubber cleans traffic to a specific host that is under attack. When an attack is detected, the traffic to that host is diverted to the scrubber—malicious packets are removed, and clean packets are forwarded. Different types of scrubbers have issues with either high-volume attacks or low-volume attacks. See Section 2.2.2.1, "Out-of-Band," for explanations of the weaknesses of out-of-band packet scrubbing. Scrubbers also require additional hardware in the network, which adds capital and operational costs for the service provider.

4.2.2.2. Flowspec

BGP Flowspec is an IETF specification (https://datatracker.ietf.org/doc/html/rfc7674) for diversion and filtering of malicious traffic. With Flowspec, the DDoS mitigation solution sends a BGP message to the routers which are forwarding the attack traffic. That router then blocks, rate limits, or forwards the traffic matching the Flowspec pattern.

4.2.2.3. Proprietary Matching Engine

Switch vendors have started developing proprietary packet matching and manipulation engines. Many of these engines could be leveraged to support DDoS mitigation.

5. Transparent Security Architecture

Transparent Security uses programmable data plane capabilities to enable real-time packet processing, high-resolution packet inspection, and in-band network telemetry (INT). INT allows MSOs to identify the compromised devices in milliseconds rather than minutes. This technology is enabled by new, currently available programmable chips that can process packets at line speed and be deployed at any point in the network, from the core network to residential and business customer premises.

Transparent Security is focused on inspecting, finding, and blocking malicious packets as close to the source as possible by adding details about the packet's source device, exact route through the network, and travel duration. INT can then be used by upstream processes to identify traffic patterns and then act on that information.

5.1. Data Plane Architecture

Figure 1 shows a typical data plane architecture of customer premises connected to the access network into an operator's core network. Any combination of the customer gateways, switches, and routers can be P4 enabled or not. By enabling P4, the available use cases increase substantially, but this architecture can be implemented in stages. Since the INT metadata is encapsulated in a UDP shim, there is no impact to L3 routing. Packets with the INT header can traverse network devices which do not support INT header insertion. Firewalls and other devices which use L4 headers will need to be updated to support INT to inspect the original L4 header.

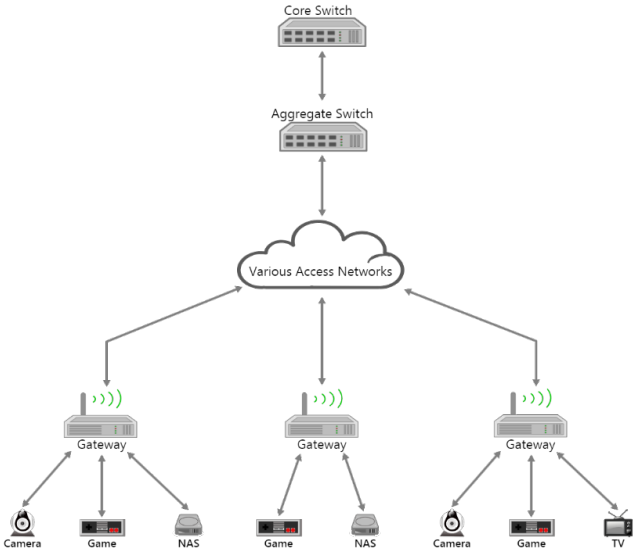


Figure 1 - General Architecture for the Data Plane in the Transparent Security Model

5.1.1. In-Band Network Telemetry

INT is a method for adding telemetry data to every packet at multiple points on the network. This approach can help determine things such as the path of a packet or the source device emitting packets. With INT, the packet only needs to be inspected at the edge of the network, which reduces the overhead and generates less traffic compared with sampling at multiple points in the network.

The P4.INT specification suggests a set of predefined fields:

- switch ID,
- control plane version,
- ingress/egress port,
- timestamp,
- RX packet count, and
- congestion status.

By leveraging P4 to implement INT, the data can be customized to meet the needs of a service provider. With this data, one can identify the source of a packet behind a firewall by including the originating MAC address in the INT header. With this data, one can obtain the specific source of a packet behind a firewall. Because customer premises and wireless networks are dynamic and multiple devices share the same ingress point on the gateway, knowing the ingress port is not sufficient when trying to identify the packet source. These extensions to the standard P4 INT data structure provide the packet source's MAC addresses, which are unique and are required to identify specific devices for customer premises and wireless networks.

One potential issue with INT is that it increases the size of the packet header. If this increase exceeds the frame size, it will cause packet fragmentation, which has a negative impact on network performance. This issue should not arise with Transparent Security, however, because the INT data is added at the gateway. With access networks, such as DOCSIS, the frame size is larger between the gateway and the core network than at the customer premises. Once the packet reaches the service provider core, INT metadata will not be added to the packet if doing so would cause the MTU size to be exceeded.

INT header uses UDP encapsulation and domain specific source only metadata. Transparent security is the reference example for source only metadata in the 2.1 version of the P4 INT specification (https://github.com/p4lang/p4-applications/blob/master/docs/INT_v2_1.pdf).

The INT header and metadata data will be removed after the telemetry report has been generated and before the packet leaves the network being monitored by transparent security.

5.1.2. Switch or Gateway with P4

The features available with networking hardware supporting the P4 language allow for the development of flexible, open, and consistent DDoS mitigation solutions. Using the information gleaned from the INT data, it is possible not only to detect an attack very quickly but also to identify the compromised device. Once identified, the switch or gateway can be notified to reroute or drop the problematic packets with a simple match rule performed at line speed when deployed with hardware acceleration in software on a gateway device containing a P4 chiplet. Additional information on P4 can be found at https://p4.org/.

5.2. Control Plane Architecture

Figure 2 depicts an example control plane architecture in which the analytics engine receives INT data from the core in-band as packets flow through the network. When malicious patterns are detected, the SDN controller is notified and updates the P4-enabled devices to handle the packets based on the pattern signature. The management interface between the controller and the P4-enabled devices can leverage a variety of protocols, including GRPC, Thrift, HTTP, or RPC. The protocol between the SDN controller and switches can vary, depending on the protocols supported by the switches and gateways. Telemetry data and alert notifications can optionally be sent to a dashboard or NOC server for integration with other analytics.

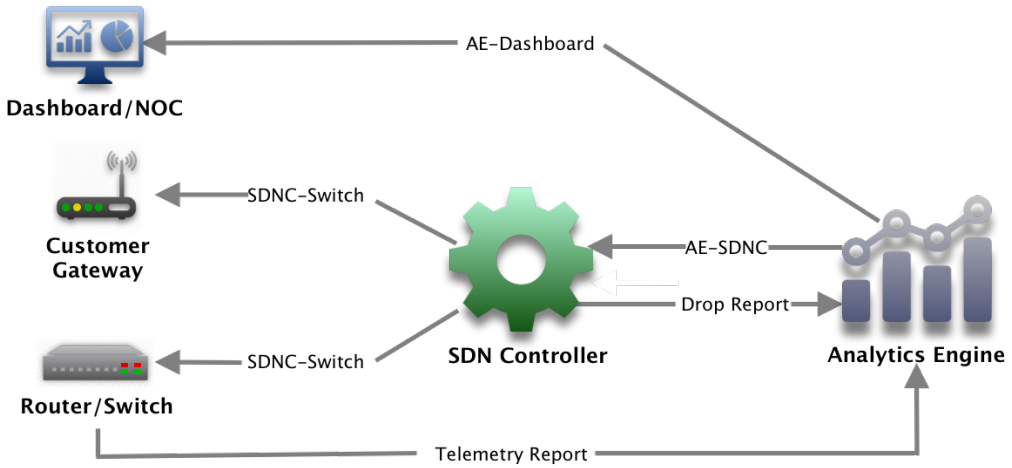


Figure 2 - General Architecture for the Control Plane in the Transparent Security Model

6. Components

6.1. Transparent Security Analytics Engine

The analytics engine (AE) serves as the intelligent core of the programmable data plane. Its purpose is to analyze telemetry reports and make inferences relative to generally defined patterns. For this use case, these patterns are limited to DDoS attack identification. The AE could be extended to manage Quality of Service (QoS), proactive network maintenance (PNM) or other use cases.

When an attack it identified, the AE informs the SDN controller, which is responsible for implementing network changes through the control plane.

The AE's include the following functions:

- identify DDoS attacks,
- mitigate the attacks, and
- remove inactive mitigations.

Multiple iterations of the AE have been tested as the project matures. The initial implementation was a Python service, which would parse the telemetry reports and identify the attacks. This solution was rigid and hard to scale. It was however able to detect and mitigate an attack in about 1 second and as such did meet our needs for detecting UDP flood-based attacks. This implementation can be found here: https://github.com/cablelabs/transparent-security/tree/master/tests/trans_sec/analytics

The second implementation leveraged a Painless pipeline to parse the telemetry reports and Elasticsearch was used to ingest the incoming data, visualize it and identity the attacks. This solution allows for multiple attack identification pipelines to use a shared parser. This architecture should scale without issues. The primary limitation of using Elasticsearch trend analysis is it took around 1 minute to identify an attack, which drove the minimum time to detect an attack outside of the target for transparent security. This implementation can be found here: https://github.com/cablelabs/transparent-security/tree/master/snaps-hcp

The current version of the AE is based on Siddhi. This solution provides modularity and scalability while still identifying attacks in about 1 second. Information on this implementation can be found here: https://github.com/cablelabs/transparent-security/blob/master/docs/SIDDHI_AE_SETUP.md

6.1.1. DDoS attack identification

Telemetry reports which include INT data are used to provide samples of the traffic running across the network. The telemetry reports use the telemetry report interface as noted in Figure 2 - General Architecture for the Control Plane in the Transparent Security Model. Telemetry reports are used instead of IPFIX since they contain the full packet header and a fragment of the data. With this richer data the AE will be able to employ attack identification techniques which are just not possible in DDoS identification applications that are limited to sampled flow data.

6.1.2. DDoS attack mitigation

The AE notifies the SDN controller of an active DDoS attack. This is sent over the AE-SDN interface as shown in Figure 2 - General Architecture for the Control Plane in the Transparent Security Model. This notification includes a signature of the attack, so that the SDN controller can mitigate it. The current implementation uses a rest call directly to the SDN controller. It is possible in production to have multiple SDN controllers. This interface may be updated to use a pub/sub model where SDN controllers can subscribe to message bus, such as Kafka (https://kafka.apache.org/protocol.html) or MQTT (https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html)

6.1.3. Remove inactive mitigations

The AE receives drop telemetry reports from the SDN controllers. These reports provide counters for all active mitigation rules on the network. These are sent over the drop report interface as shown in in Figure 2 - General Architecture for the Control Plane in the Transparent Security Model. The AE determines when the mitigation rule has not been used for a specific period which is an indication the attack has ended. Once an attack has ended, the AE will notify the SDN controller to remove the inactive mitigation rule.

6.2. SDN Controllers

The information used by the AE is captured by the P4 devices as part of the data plane and forwarded to the AE in line with the network traffic. When the SDN controller is informed by the AE that a DDoS attack has been detected, it updates action tables in the P4 devices through the control plane management interface (using protocols such as GRPC or Thrift). Thus, the P4 devices gain additional abilities in the data plane with a minimally intrusive controller activating them before consequences are experienced.

The controller's functions include the following:

- manage the network configuration on switches and gateways,
- push DDoS mitigation to managed devices,
- track which devices are participating in an attack by querying counters of dropped packets based on DDoS mitigation from the device preforming the mitigation,
- send periodic drop reports to AE with the dropped packet counters for each mitigation, and
- remove DDoS mitigation from managed devices.

6.3. Programable Data Plane

The programable data plane consist of routers, switches and customer gateways used in Transparent Security to add/remove the INT header, generate telemetry reports, and mitigate DDoS attacks. To date, the testing and development has focused on performance on devices which support P4. These devices provide a great deal of flexibility to quickly develop the data plane for transparent security. It is possible to deploy transparent security using other types of programable devices. See 4.2.2.3 for additional information on these types of devices.

The programable data plane functions include the following:

- manage the traffic across the network,
- add P4.INT data (source port, time, queue),
- add Transparent Security data (source MAC),
- remove P4.INT headers before they leave the service provider network,
- send telemetry reports to the AE and,
- mitigate DDoS attacks from the core network and the aggregate network.

6.4. User Equipment

The end user devices are the typical end units used by end customers, including IoT devices, phones, laptops, and an ever-increasing variety of devices. Many are connected to the gateway over Wi-Fi or LTE. Implementing transparent security does not require any changes to user equipment. The user equipment should not be able to detect if transparent security is deployed or not.

7. Message Flows

The following sections outline high-level message flows between the possible components.

7.1. Packet Flow and Alerts

7.1.1. Standard Packet Flow

Standard packet flow:

1. Customer device sends packets to the server provider network.
2. The packet will traverse a series of gateways, switches, and routers where some are aware of INT and others are not.
3. This first INT enabled switch will insert the INT UDP header and the metadata for that hop.
4. Incremental networking devices with INT support will add their node ID to the INT metadata.
5. Packet is forwarded to the final INT networking devices, such as a PE router.
6. Packet is cloned and sent to the analytics engine.
7. INT data are removed from the packet.
8. Packet is forwarded to the Internet.

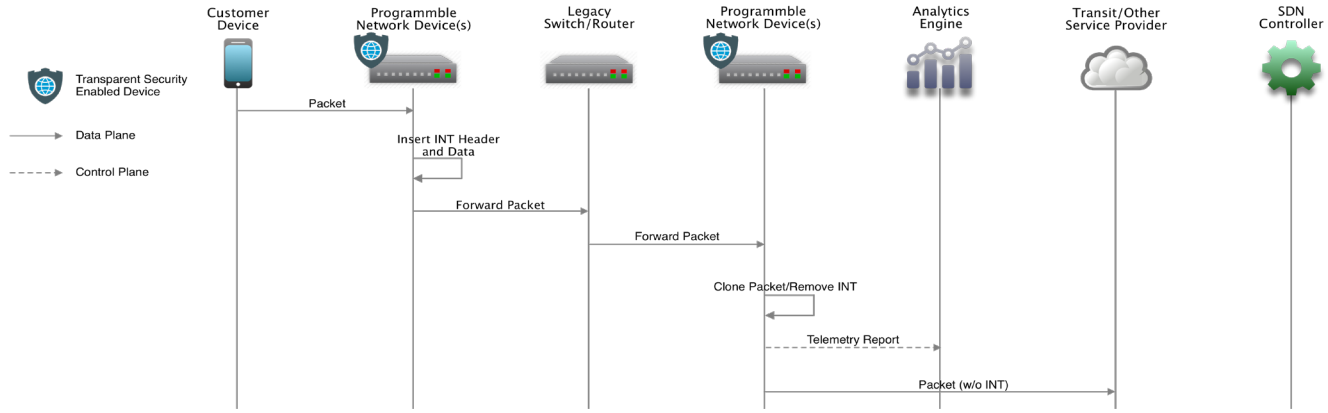


Figure 3 - High-Level Message Flow in the Control Plane and Data Plane when Packets Are Allowed Through.

7.1.2. DDoS Attack Identification and Mitigation

The following actions occur when an attack is detected:

1. Analytics engine recognizes an attack.
2. AE notifies the SDN controllers of the attack with its signature.
3. The SDN controllers notify the proper networking device to add an entry in the transparent security drop table for this specific attack signature.
4. The networking device add the entry into its transparent security drop table.

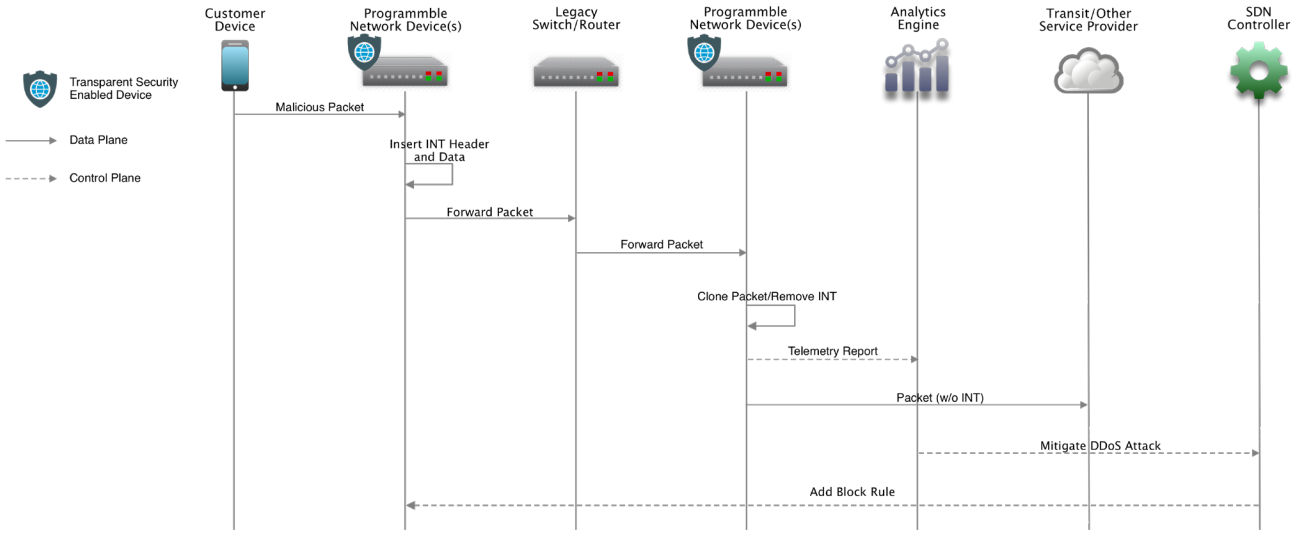


Figure 4 - High-Level Message Flow in the Control Plane and Data Plane during DDoS Attack Identification and Mitigation

7.1.3. Mitigated Attack Packet Flow

1. A packet as a part of the DDoS attack is sent to the service provide network.
2. The first transparent security enabled networking device with a drop rule drops the packet.
3. After the packet is dropped, the device increments the drop counter for that rule.
4. The SDN controller periodically collects the drop counters and sends them to the AE in drop telemetry report.

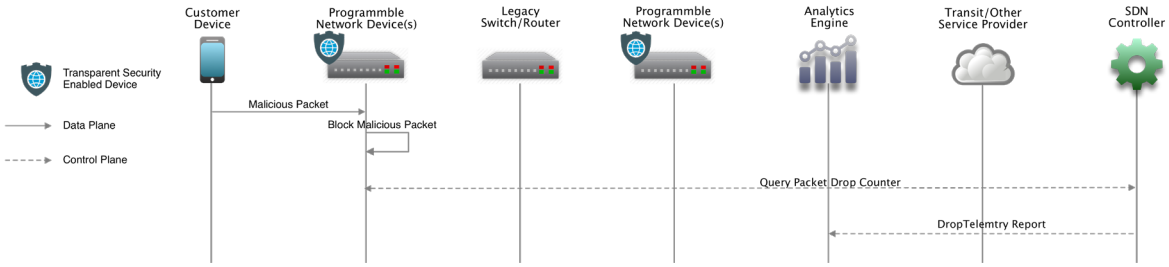


Figure 5 - High-Level Message Flow in the Control Plane and Data Plane while Mitigating an Attack.

8. Deployment options

Transparent Security can be deployed across the service provider network in several phases. A phased deployment makes it easier to deploy and realize many of the benefits while postponing any changes to cable modems (CMs). The process of a phased deployment begins at the hub or head end and moves to the customer premises at later phases.

8.1. Core/Head End/Hub Deployment

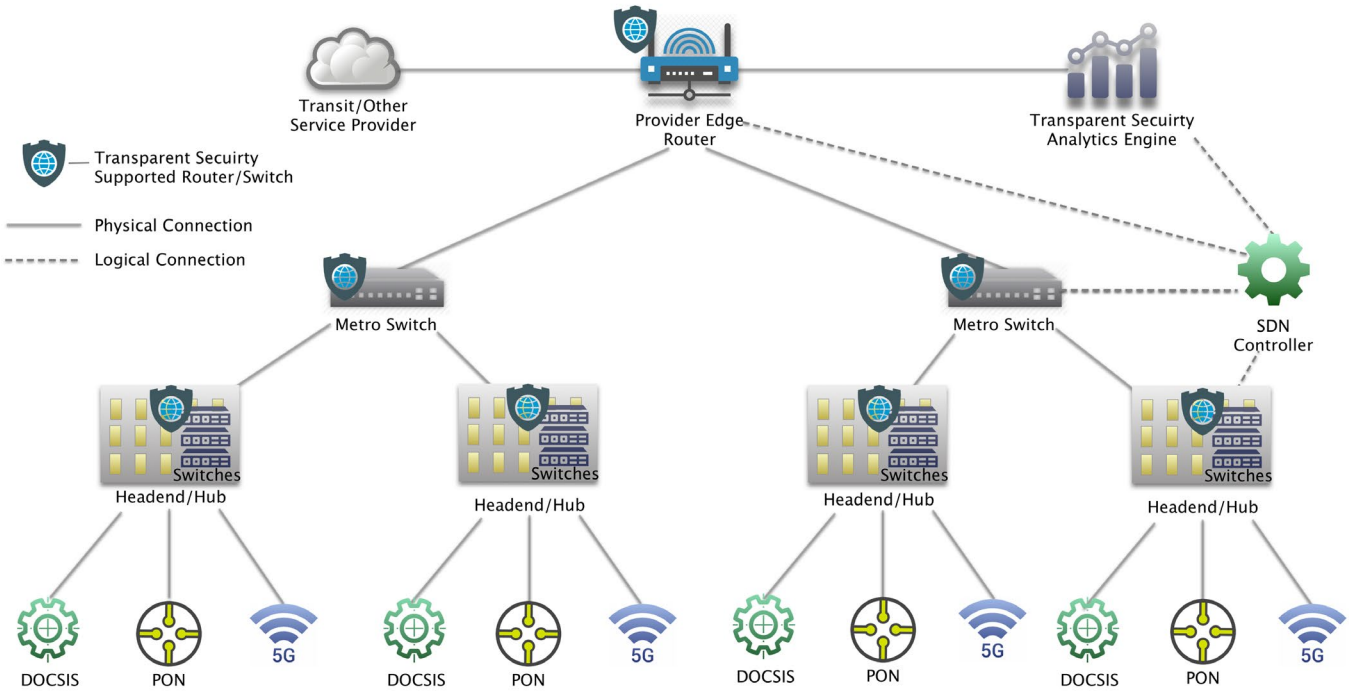


Figure 6 - Phased Deployment: Head End/Hub Updates

This deployment option focuses on the service provider's core network. Two classes of attacks can be mitigated. Attacks from this service provider's network against targets outside the network and attacks across the service providers network can be detected and mitigated with this model.

The first phase of deployment begins with upgrading the head end/hub by adding programmable switches with an analytics engine and an SDN controller. These switches can be included as part of distributed access architecture (DAA) upgrades. They provide the ability to mitigate a DDoS attack from a customer at the head end and identify which customer is the source of the attack.

This solution can also address ingress attacks from outside of the hub which is an advantage over a typical edge-based DDoS mitigation model.

There are no changes to the CM for this phase. The primary limitation to this solution is that it cannot identify the device originating the attack, only the customer. The compromised device remains active and continues to participate in ongoing attacks.

8.2. Customer Premisis

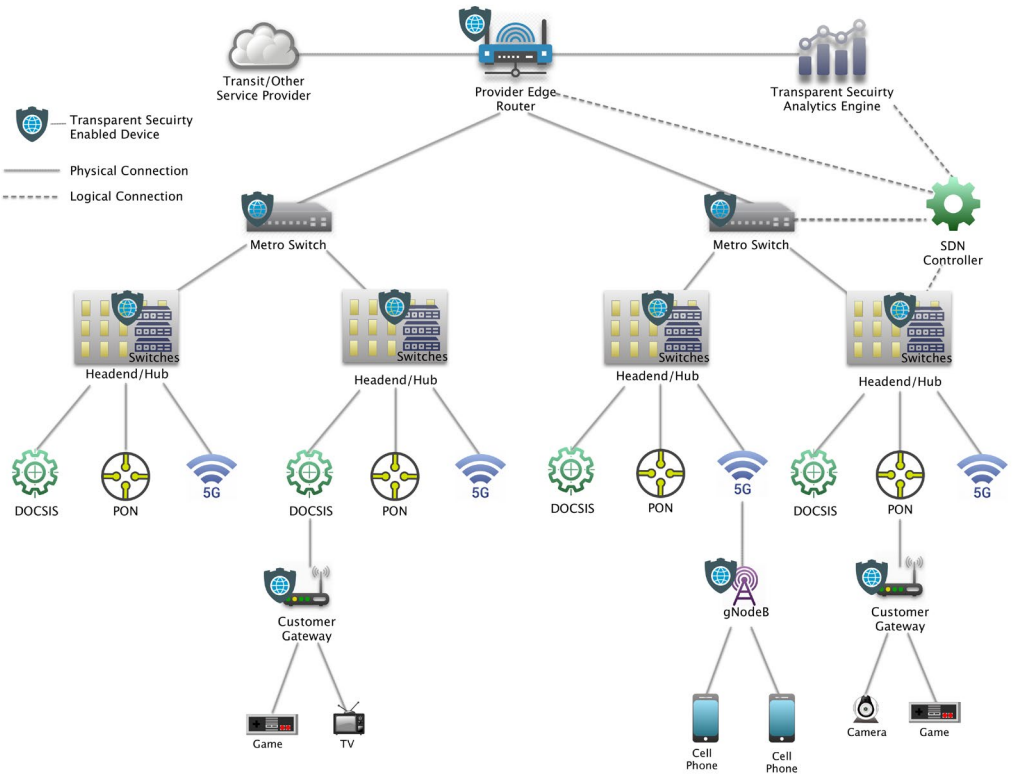


Figure 7 - Phased Deployment: Customer Premisis

In this deployment model, transparent security is deployed on a gateway device. For 5G, this can be deployed on a gNodeB. The initial INT insertion and DDoS mitigation is performed on the gateway device.

Once this model is deployed, the traffic on the access network will be scrubbed before it can impact any other customers, and customers will be able to address issues on compromised devices. The malicious traffic is dropped, and the count of blocked packets for each device is tracked. This model will not block benign traffic from the compromised device.

9. Lab Trial

To validate the viability and network impact for transparent security, Cox Communications conducted a lab trial with transparent security. This trial focused on a derivation of the Core/Head End/Hub deployment option. The goal of this trial was to validate that there is no adverse impact on network performance when adding transparent security. This includes latency, throughput and basic network routing.

9.1. Lab Trial Setup

The test environment was designed to simulate traffic originating from the access network, carried over the service provider's core backbone network, and targeting another endpoint on the service provider's access network in a different market (e.g., an "east-to-west" or "west-to-east" attack).

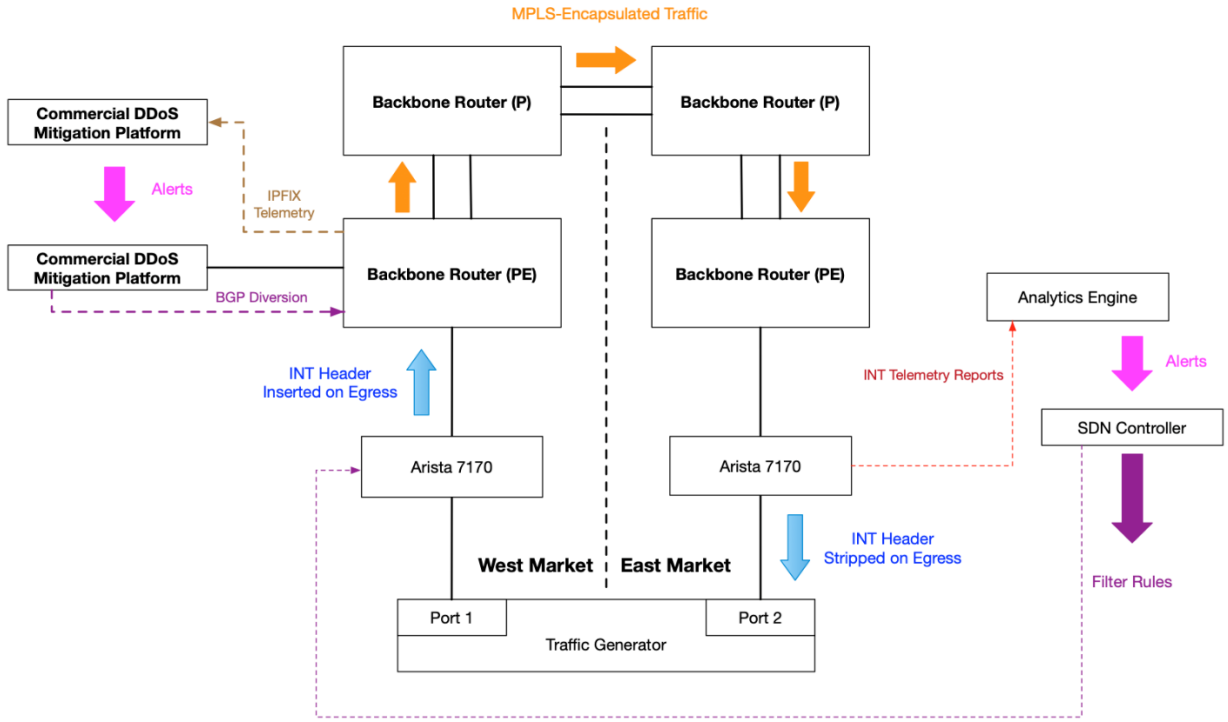


Figure 8 - High-level overview of the lab test environment

In the lab trial, various types of DDoS traffic (UDP/TCP over IPv4/IPV6) were generated by the traffic generator and sent to the West Market Arista switch, which used a custom P4 profile to insert an INT header and metadata before sending the traffic to the West Market PE router. The traffic then traversed an MPLS label-switched path (LSP) to the East Market PE router, before being sent to the East Market Arista, which used a custom P4 profile to generate INT telemetry reports and to strip the INT headers before sending the original IPv4/IPv6 packet back to the traffic generator.

9.2. Results

When comparing the performance of the Transparent Security solution against that of a leading commercially available DDoS mitigation solution, the lab test results were very promising. Detection and mitigation of outbound attacks was rapid, taking around one second. The commercial solution took 80 seconds to detect and mitigate the attack. Randomized UDP floods, UDP reflection and TCP state exhaustion attacks were identified and mitigated by both solutions. In this trial, only packets related to the attack were dropped. Packets not related to the attack were not dropped.

The Transparent Security solution was implemented on commercially available programmable switches provided by Arista. These switches are being deployed in networks today. No changes to the Networking Operations System (NOS) were required to implement Transparent Security.

The tests validated that INT-encapsulated packets can be transported across an IPv4/IPv6/MPLS network without any adverse impact. There was no observable impact to throughput when adding INT headers, generating telemetry reports, or mitigating the DDoS attacks. We validated that the traffic ran at line speed, with the INT headers increasing the packet size by an average 2.4 percent.

Application response time showed no variance with or without enabling Transparent Security. This suggests that there will be no measurable impact to customer traffic when the solution is deployed in a production network.

10. Conclusion and Next Steps

Transparent Security uses in-band telemetry to help identify the source of the DDoS attack.

This trial focused on using Transparent Security on switches inside the service provider's network. For the full impact of Transparent Security to be realized, its reach needs to be extended to gateways on the customer premises. Such a configuration can mitigate an attack before it uses any network bandwidth outside of the home and will help identify the exact device that is participating in the attack.

This testing took place on a custom P4 profile based on our open-source reference implementation. We would encourage vendors to add INT support to their devices and operators to deploy programmable switches and INT-enabled CPEs.

Take the opportunity today to explore the opportunities for using INT and Transparent Security to solve problems and improve traffic visibility across your network.

Distributed denial of service (DDoS) attacks and other cyberattacks can cost operators billions of dollars. With more and more devices coming into customers' homes and businesses, many of which with less-than-optimal security, this problem will only get worse. By quickly identifying attacks and blocking them before they reach the access network, Transparent Security can:

- reduce the operations impact of large-scale attacks by mitigating closer to the source within seconds of an attack starting,
- eliminate the risk of revenue impact resulting from a failure to meet service-level agreements,
- protect and enhance customer sentiments by avoiding large-scale DDoS attacks within a network, and
- provide a data stream from which new analytics and innovations can be built.

10.1. Alternate Applications for Programmable Data Plane

The Transparent Security source-based DDoS use case is just one of many that can benefit from the programmable data plane. Once deployed, this architecture can improve many of the operations performed today. It will also open networks to new waves of innovation and allow operators to provide such things as network optimization and additional customer services.

With the programmable data plane, it is possible to change the behavior of the network after hardware has been deployed. Use of an analytics engine and controller, as is done with Transparent Security, can

provide a closed-loop automation. Some of the network management capabilities available with the programmable data plane are listed below:

- packet flow tracking and optimization,
- prioritization on low-latency flows,
- active network monitoring/management, and
- traffic shaping.

Providing new services frequently requires installing new purpose-built hardware or deploying a virtual machine. With the programmable data plane, some of these services can be deployed on existing switches with very good performance. Some of the services that can be deployed with the programmable data plane include the following:

- firewalls,
- managed router as a service,
- Layer 4 load balancing, and
- SD-WAN (software-defined networking in a wide-area network).

Micronets is another CableLabs project that can leverage the programable data plane. Micronets creates additional layers of security within the customer premises and helps protect devices by using OpenFlow for the L3 traffic management. Transparent Security is focused on identifying devices participating in a DDoS attack and mitigating the attack at the source. The programable data plane, analytics engine, and controller used in Transparent Security can be leveraged by Micronets to improve packet processing performance and provide access to additional fields in the packet header. For more information on the Micronets project, visit www.cablelabs.com/micronets.

Abbreviations

INT	In-band telemetry
CPE	Customer premises equipment
DDoS	distributed denial of service
IoT	Internet of Things

Developing the DOCSIS 4.0 Playbook for the Season of 10G

A Technical Paper prepared for SCTE by

Dr. Robert Howald
Fellow
Comcast
1800 Arch Street, Philadelphia, PA 19103
robert_howald@comcast.com

John Williams
Vice President, Engineering & Architecture
Charter Communications
14810 Grasslands Drive, Englewood, CO 80112
John.Williams2@charter.com

Jon Cave Comcast

Olakunle Ekundare, Comcast

Matt Petersen, Charter Communications

1. Introduction

Nearly three years after the introduction of 10G, industry activity tied to this foundational next generation cable technology, known in technical circles as DOCSIS 4.0, is accelerating rapidly. The timing is well-suited to the moment. DOCSIS 3.1 has been in production for over 5 years. Newer features are now being exercised, such as 8k FFTs (25kHz) for OFDM and ODFDMA, Mid-Split and High Split networks to expand upstream, and the revolutionary Profile Management Application (PMA), the automated form of DOCSIS3.1 Multiple Modulation Profiles (MMP).

As DOCSIS 3.1 deployments continue, so do capacity growth and speed expectations for the network. DOCSIS 4.0 was developed to support relentless bandwidth consumption and speed trends, leveraging the DOCSIS 3.1 PHY basis for its bandwidth efficiency, resiliency, and backwards compatibility.

As with most cable technology evolutions, there is no "one size fits all" solution among, or often even within, a single operator. Each operator's current as-built networks have different starting points and range of other variables – region of the country, geography and topology, municipal and state make-ready differences and a range of aerial/underground construction techniques, to name a few.

In this paper, Comcast and Charter will provide a unified point of view on the rapidly approaching cycle of DOCSIS 4.0 upgrades. We will address:

- Implementation options in DOCSIS 4.0, including Full Duplex/FDX and Extended Spectrum/FDD
- Common use cases, with guidance on how DOCSIS 4.0 options may be applied
- The range of network variables and anticipated impacts
- How key dependent variables may drive evolution path decision criteria
- The complementary nature of the technologies

This is a useful and timely session for all operators, suppliers, and industry partners interested in how DOCSIS 4.0 will shape tomorrow, as envisioned by engineering leads of two largest cable operators in North America.

2. 10G Overview

2.1. Objectives

Cable operators have steadily increased bandwidth and speeds to subscribers since the launch of the DOCSIS high speed data services (HSD). Compound annual growth rates (CAGRs) and the appetite for more bandwidth-intensive applications, has made continuous investment in the network the cost of doing business. These trends have been reasonably predictable, allowing operators to be very efficient with their network investments and develop standard practices to manage growth. HFC continues to show the ability to adapt and increase capabilities to meet the needs of today's subscribers and businesses.

MSOs have historically been conservative in talking about their network capabilities. For example, technical terms such as "DOCSIS 3.1" and "DAA" represent some of the language of today's industry technology advances and foundational initiatives leading to 10G. However, the terms themselves are not particularly effective for describing what the network is capable of and what it can deliver for customers.

Building upon DOCSIS 3.1 and DAA, among others, the 10G aims to represent a technical benchmark but also be associated with what these new advances mean to services and to the customer experience. So what is 10G? The "pillars" of the 10G networks are defined as:

- Speed – Enabling of multigigabit symmetric speeds, raising the bar for consumer broadband.
- Low latency – Low Latency DOCSIS (LLD) is now incorporated into the DOCSIS 3.1 specification and carried forward into DOCSIS 4.0. Delivers a better customer experience, in particular for applications such as gaming and AR/VR.
- Reliability – Methods to proactively identify and address network issues before consumers are aware of them
- Security – Improve the confidentiality, integrity, and availability of safe communications.

Summarizing, 10G will offer up to 10 Gbps of intelligent, reliable, secure bi-directional capability that, coupled with decreased latency, will enhance the customer experience of today. The 10G network will unleash a new generation of capability that will drive innovative applications and create new digital opportunities, novel services, and impact lifestyles for the better.

2.2. HFC Challenges to 10G

With such a grand vision, what will it take for operators to move from PowerPoint (and white papers(!)) to implementation and field deployment?

Well, if 10G was simple to obtain and pedestrian in its objectives, it might already be available and there would not be multi-years of play-by-play media coverage and industry-wide excitement about its promise. Furthermore, painting the end state of a next generation network evolution is relatively straightforward. It is the transition from the current state to that end state that creates the challenges, consuming most of the energy that access engineers and business planners spend when debating network path and technology direction. The starting point for 10G comes with several interrelated obstacles:

1) *Coaxial Bandwidth* – First and foremost, the cable network spectrum is very broad, but there are still limits to the easily accessible bandwidth. For most HFC systems today, that bandwidth is either 750 MHz, 860 MHz, 1 GHz, and more recently to 1.2 GHz. The limit that applies for a particular plant is typically determined by the generation of RF amplifiers installed. With a downstream limit such as 750 MHz, for example, there is simply not sufficient spectrum to achieve 10 Gbps even in a complete DOCSIS 3.1 migration.

2) *Upstream Allocation* – Despite the scenario described above, the path to 10Gbps on the downstream is not too difficult to envision, even for 750 MHz. Simply add more very efficient DOCSIS 3.1 spectrum extending to 1 GHz or 1.2 GHz with readily available amplifier replacements, for example, and it is within reach, depending on other EOL fidelity variables. However, 10G is really setting its sights on massively expanding upstream bandwidth, which in North America is typically 42 MHz. Some MSOs have trialed upgrading to a Mid-Split Upstream, extending the return band to 85 MHz, to increase the capacity runway of the network and defer node splits. As can be observed above, the downstream-to-upstream (DS/US) ratio is highly asymmetrical. It is the primary goal of 10G to develop a more symmetrical DS/US ratio.

3) *Current DOCSIS Production Spectrum* – An assumption made above on the route to 10 Gbps is a "complete DOCSIS 3.1 migration." This is an unlikely scenario for many years. There will be relatively inefficient video QAM signals on the network for many years to support broadcast video services over many millions of existing set-top boxes (STBs) in the field today. In addition, while DOCSIS 3.1-capable devices are increasing in the field, DOCSIS 3.0 Cable Modems (CMs) still represent the majority in Comcast's footprint, and of the millions of DOCSIS 3.1 CMs in the field operate today, most operate in DOCSIS 3.0 mode, carrying the majority of traffic and accounting for most of the HSD spectrum. Other operators, including Charter, have activated more DOCSIS 3.1 upstream and are taking advantage of these additional efficiencies to add capacity and defer node splits. Existing services based on 256-QAM signals (as well as QAM VOD signals) put limits on how much spectrum can be allocated for DOCSIS 3.1, burdening some of the spectrum with less bandwidth efficiency.

4) *DOCSIS Capability* – The DOCSIS 3.1 standard limits the upstream spectrum to a maximum of 204 MHz – the "High Split" configuration. This configuration is designed to enable up to 1Gbps of upstream when carrying all DOCSIS 3.1 OFDMA signals. In this configuration, it is capable of up to about 1.5 Gbps. The definition of the DOCSIS 4.0 standard sets about to expand the amount of spectrum made available for Upstream, up to 684 MHz for either DOCSIS 4.0 option. This expansion is the major change that the standard is targeting towards the 10G network.

5) *Distributed Access Network (DAA)* – DAA is a definite "good" guy, so why is it listed here as a challenge? Well, the DAA journey has not begun with many operators, and is also fragmented in approach taken. For all of the discussion of DOCSIS 4.0, it is a consensus agreement that it will only be implemented via DAA. DAA is itself a complex, multi-faceted journey to production scale and as a prerequisite to DOCSIS 4.0, presents a large obstacle of a DAA plan is not first established and executed on.

6) *Business Case* – Plant upgrades have been part of cable network evolution for decades. Each upgrade cycle has undergone typical analysis of pro/con and meeting criteria for return on investment. Well-understood guidelines have been established over the years for node splits, spectrum addition, and other network augmentations, such as FTTH. Similar modeling will take place for DOCSIS 4.0 and be honed over time using empirical data.

2.3. Still Early in the Life of DOCSIS 3.1

Before expounding in the exciting technology and opportunities ahead with 10G, it is important to recognize that the DOCIS 3.1 journey is still early. Planning ahead for DOCSIS 4.0 does not mean things are standing still in the meantime – on the contrary, operationalizing and optimizing the advanced features enabled with DOCSIS 3.1 has really just begun.

2.3.1. 2020: Everything Changes

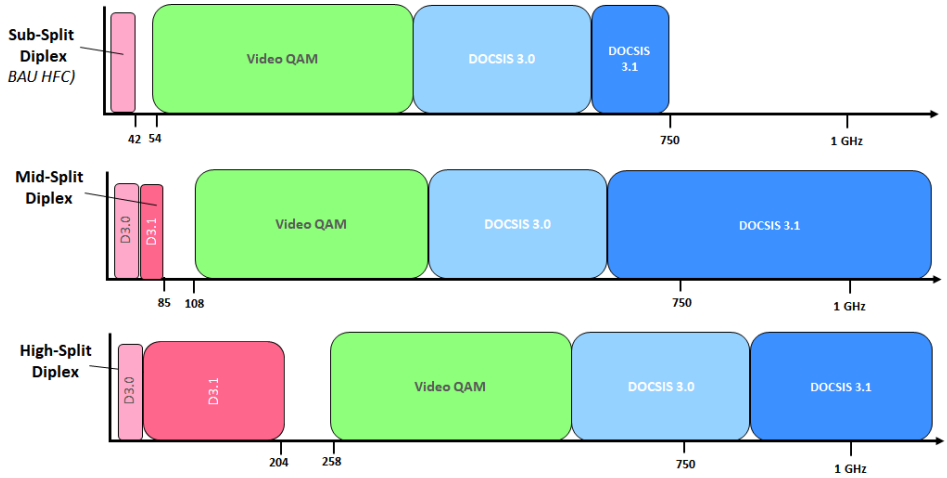
When it comes to the evolution of DOCSIS, each version seems to have a longer runway than the prior. That is certainly the case when we talk about DOCSIS 3.1, which can easily be considered to be in the infancy of its life span. Some operators began to roll out DOCSIS 3.1 OFDM downstream a few years ago, primarily to provide a Gigabit service offering. Like all versions prior, it is backwards compatible; however not until DOCSIS 3.1 CPE reaches meaningful penetration rates will operators be able to use it to its fullest capacity.

Recently with the COVID-19 pandemic we saw an increased number of people working from home and learning from home using video conferencing and real time communications tools. The industry shined in its ability to handle the increase in traffic due, but the pandemic did highlight the significance to optimizing how we leverage DOCSIS 3.1 increased bandwidth.

2.3.2. Pulling the Levers

A significant example of what is possible and still ahead for DOCSIS 3.1 is the expanded downstream and upstream spectrum it provides over DOCSIS 3.0. Current DOCSIS 3.1 has the capability to support a downstream frequency range up to 1.2 GHz and an upstream frequency range up to 204 MHz, better known as High Split. Today, a 42 MHz upstream bandwidth is the most common for North American cable operators, although some migrations to Mid-Split (85 MHz) have taken place. The expansion of frequency capability will allow operators more flexibility in how to leverage their networks for offering enhanced service tiers.

Figure 1 illustrates the differences between a standard HFC spectrum allocation today and the Mid-Split and High-Split architectures.



The majority of North American operators have deployed DOCSIS 3.1 downstream OFDM blocks. Some operators have also begun to leverage OFDMA in the upstream, which provides higher modulation schemes and greater bits per second per hertz (bps/Hz) when compared to the DOCSIS 3.0 A-TDMA Single-Carrier QAM (SC-QAM). As operators continue to increase the penetration of DOCSIS 3.1 CPE, some are opting to gradually remove ATDMA one service group at a time, replacing this spectrum with OFDMA because of the higher modulation profiles and increased throughput, and therefore longer capacity runway. Most node splits are driven by the upstream capacity utilization associated with the limited available 42 MHz of spectrum.

In the downstream path, operators are using the increased CPE penetration to further expand OFDM blocks and increase the total throughput made possible with 4096-QAM (12 bps/Hz) OFDM subcarriers as compared to 256-QAM (8 bps/Hz) of the legacy single carrier QAMs, a 50% increase in inefficiency. As downstream growth trends continue, there will be a reclamation of the legacy single carrier QAMs in the spectrum that are currently supporting DOCSIS 2.0/3.0 devices that, through attrition, are currently being

replaced with DOCSIS 3.1 CPE. The decreasing penetration of these legacy devices can be supported with fewer SC-QAMs occupying the downstream.

2.3.3. Increasing the Spectrum and Moving the Split

As shown in **Figure 1**, most HFC networks today are built with a 750 MHz or 860 MHz upper frequency range utilizing a 5 – 42 MHz "Sub-Split" return. This bandwidth cap of 42 MHz means a limited amount of DOCSIS carriers, and the amount of spectrum used for the upstream falls far short of the capabilities available in DOCSIS 3.1 devices that have now been available and deployed for years. There is ample opportunity to replace active components in the field with new nodes and amplifiers that support a downstream to 1.2 GHz, with either a Mid or High Split configuration. This extra spectrum will allow for the expansion of both OFDM and OFDMA carriers and greatly increase the total throughput capacity of the network in both directions.

Most operators, feeling the upstream growth constraints, now accelerated by COVID-19 and its implications [1][4] are already executing on these frequency expansion upgrades. Some are making these bandwidth changes in coordination with their Distributed Access Architecture (DAA) rollouts, minimizing outside plant impacts and making most efficient use of construction opportunities.

DAA itself can be considered another lever operators can use to increase signal fidelity and, with the available optimizations of bandwidth efficiency with DOCSIS 3.1, deliver greater capacity to the network, benefitting customers.

3. DOCSIS 4.0 Technology Options

3.1. Full Duplex DOCSIS (FDX)

Figure 2 illustrates the essential spectrum goal of FDX– enabling significantly more upstream. More interestingly, these new FDX upstream bands (Red "FDX" in **Figure 2**) are available for downstream. Huh? This is quite different than typical Frequency Domain Duplex (FDD) operation, the approach taken in DOCSIS 4.0 FDD. It is also begs the question – how can both downstream and upstream data exist in the same spectrum?

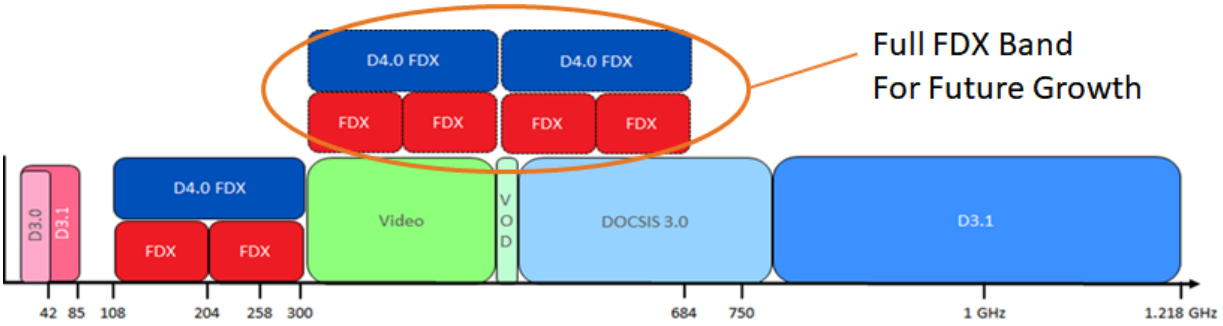


Figure 2 - Upstream Spectrum Added for DOCSIS 4.0 Full Duplex (FDX)

3.1.1. Key FDX Innovations

Although more upstream bandwidth is defined, it is the same 96 MHz OFDMA physical layer blocks as defined in DOCSIS 3.1. Thus, DOCSIS 4.0 leverages the power of the DOCSIS 3.1 PHY completely. Six additional 96MHz blocks are added across the 108-684 MHz band, complementing an 85 MHz "mid-split" system.

Downstream and upstream can occupy the same band with a technology known as Echo Cancellation (EC). Echo Cancellation in general is a mature technology used in other telecom networks, such as xDSL and wireless. It has not yet been implemented in cable networks. The EC concept is very similar, although the cable does introduce some new challenges. EC is the first of the two critical innovations that power FDX.

The second key innovation is based on an architectural difference in cable systems when compared to telco xDSL systems. Twisted pair telco networks are point-to-point from the DSL Access Multiplexer, or DSLAM, whereas HFC is a point-to-multipoint. This logical architecture difference creates the need for another layer of innovation for FDX. This is the creation of Interference Groups (IGs) and Transmission Groups (TGs) for the scheduler to manage.

Figure 3 illustrates these innovations from the CMTS perspective, using a passive coaxial network (i.e., N+0) for simplicity. N+0 is NOT a requirement for FDX, but the specifications were developed with N+0 as a baseline. FDX-capable amplifiers are being developed to support FDX signals over N+x networks, allowing FDX over a broader range of architectures. As in standard HFC networks, an RF amplifier cascade impacts quantifiably the network performance. With amplifiers for FDX, one of the impacts is the effect on IGs and TGs. As a result, additional considerations that account for the relationships among of maximum speed tier, penetration, and cascade depth have been developed. We will discuss FDX amplifiers in a subsequent section.

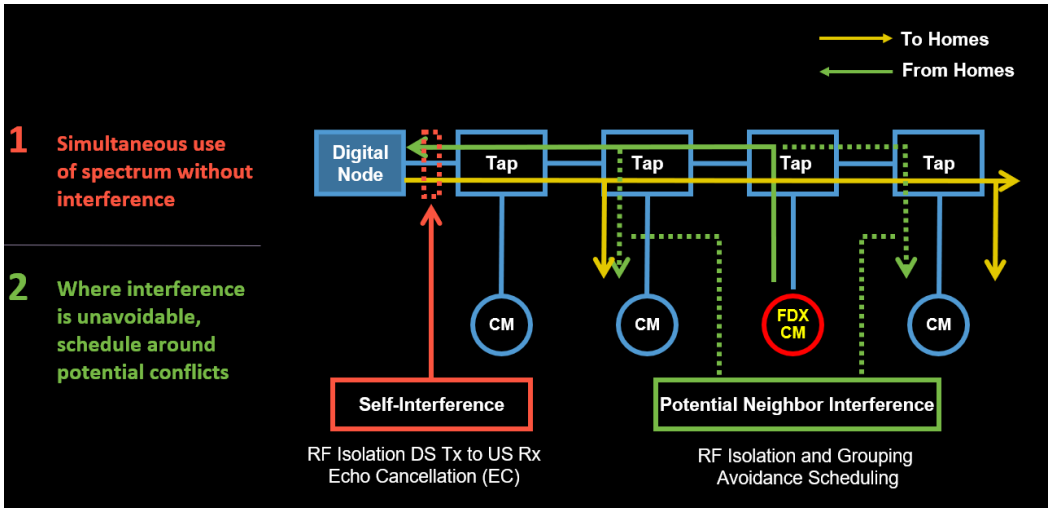


Figure 3 - Two Key New Innovations in DOCSIS 4.0 Full Duplex

3.1.2. A Closer Look Part 1 – Echo Cancellation

Referring to **Figure 3**, adding upstream where downstream exists requires that the downstream signal be "subtracted" before the US OFDMA receiver. This requires high RF isolation and strong EC of the much

higher downstream signal. While the implementation details may be complex, the EC concept is a quite simple, and the digital signal processing (DSP) principles to build it very mature. A simplified diagram illustrating the EC concept is shown in **Figure 4**.

The node downstream transmit signal will have some of its energy reflected back by the imperfect RF interfaces, such as described by the return loss of a tap, for example. These are the so-called "Echoes" that give the EC function its name. What is distinctive to EC for cable is the high cancellation required across a broad bandwidth. These cable specifics push the envelope.

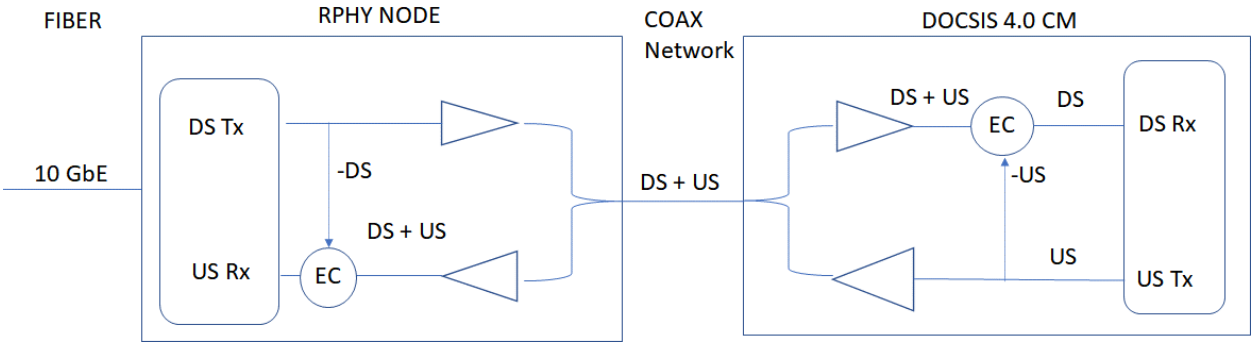


Figure 4 - Conceptual Basics of Echo Cancellation in a non-DAA Configuration

3.1.3. A Closer Look, Part 2: Interference Groups and Transmission Groups (IGs/TGs)

As noted, for HFC, the network is a point-to-multipoint architecture. While an FDX modem knows its own upstream transmission, it cannot know that of his neighbor. It requires sufficient RF isolation among neighbors to prevent FDX-band upstream users from interfering with a neighbor using that band for downstream. Unfortunately, RF isolation among homes cannot always be guaranteed to be high enough. The isolation relationships among homes a shared RF leg are determined as part of FDX "sounding" process.

Without sufficient RF isolation we can have the vCMTS scheduler designed to avoid this scenario. In an FDD system, the scheduler does not need to pay particularly close attention to the relationship of downstream and upstream access to the wire. This changes in FDX. During FDX "sounding," the FDX system determines these isolation relationships. Potentially interfering users are lumped into "Interference Groups," or IGs. A logical set of IGs is called a Transmission Group (TG), because not every IG needs to be treated independently – it depends on traffic. This scheduler assures that potentially interfering pairs are not subscribing the same spectrum in the same time slot.

Because there are six OFDMA blocks, the vCMTS can service multiple IGs with uniform capacity and speeds by assigning different Resource Block Assignments (RBAs) to each IG. **Figure 5** shows an example of how the 108-684 MHz FDX band might be allocated to simultaneously support a case with three TGs. These RBAs can adapt with time based on traffic demand.

Note that the FDX band is not all the DOCSIS spectrum available. Non-FDX DOCSIS 3.1 spectrum and DOCSIS 3.0 spectrum will also exist.

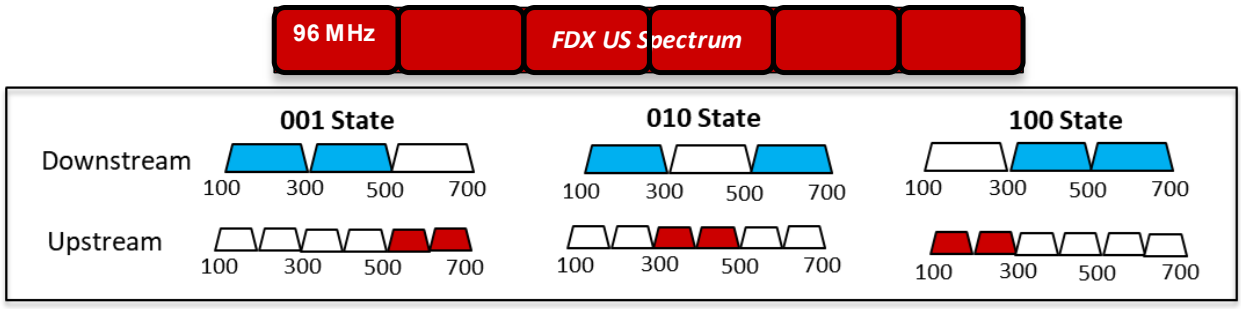


Figure 5 - OFDMA "Resource Blocks" Enabled in the FDX Band

3.1.4. DOCSIS 3.1 Compatibility / Coexistence

Because FDX is fundamentally based on DOCSIS 3.1, the FDX band can only be activated where the DOCSIS3.1 downstream is allocated. Existing DOCSIS 3.1 devices should be aware that they are part of an FDX system in order that they partake in IG/TG sounding and thereby be well-behaved co-existing devices in an FDX-enabled plant. This DOCSIS 3.1 device mode is known as "FDX-Light" or FDX-L. Obviously, DOCSIS 3.1 devices do not support FDX. They cannot transmit in all of the FDX band higher than 204 MHz. Nonetheless, it is important that the device be aware that it is operating on an FDX system. In this way, it can participate in network sounding, allowing the vCMTS to determine when it can schedule packets to be received by the device and avoid being interfered upon with by an FDX device that may be transmitting in the same band and at the same time.

FDX-L is a software-only upgrade to existing DOCSIS 3.1 devices. Without this mode, a separate DOCSIS4.0 band would need to be set aside, similar to what is done today with DOCSIS 3.0 and DOCSIS3.1 spectrum. Launching FDX and the much higher upstream speeds therefore implies a commitment to an amount of DOCSIS 3.1 downstream spectrum, within which the FDX upstream can operate. As with any spectrum allocation, this must be managed within the constraints of the HFC network's overall spectrum plan. Networks supporting FDX will all be a 1 GHz systems or 1.2 GHz systems, which are bandwidth upgrades from today's 750MHz systems and 860MHz systems. So, at first, the challenge to "find" spectrum is not as daunting considering that new spectrum that will be added as FDX becomes deployed when an amplifier is installed that adds new bandwidth above 750 MHz or 860 MHz.

3.1.5. DOCSIS 4.0 FDX is Coming to Life

The first DOCSIS 4.0 FDX RPD bring-up and characterization began in March of 2021, and the results were described in various media publications the next month [2].

Some hardware images of the first FDX RPD are shown in **Figure 6**. The RPD module shown on the right plugs into an FDX node, which is not shown.

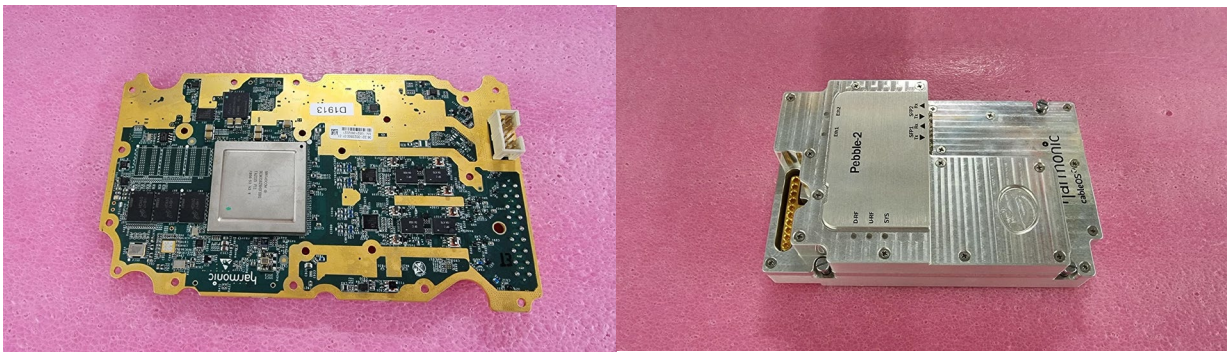


Figure 6 - DOCSIS 4.0 FDX RPD Hardware

Figure 7 shows the sample network's echo characteristics used for the test, as well as a simulation of how the FDX Band operates with both downstream and upstream present in such an environment. Also shown is the measured downstream and upstream throughput from the test, which includes block of DOCSIS 3.0 spectrum downstream and upstream (legacy), plus video carriers akin to today's channel line-up. The network achieved 7.6 Gbps downstream and 4.3 Gbps upstream, both in terms of net throughput. 1024-QAM was able to be successfully activated in the upstream, as shown, with performance similar to a typical DOCSIS 3.1 upstream receiver. The latter is of defined only to 204 MHz upstream.

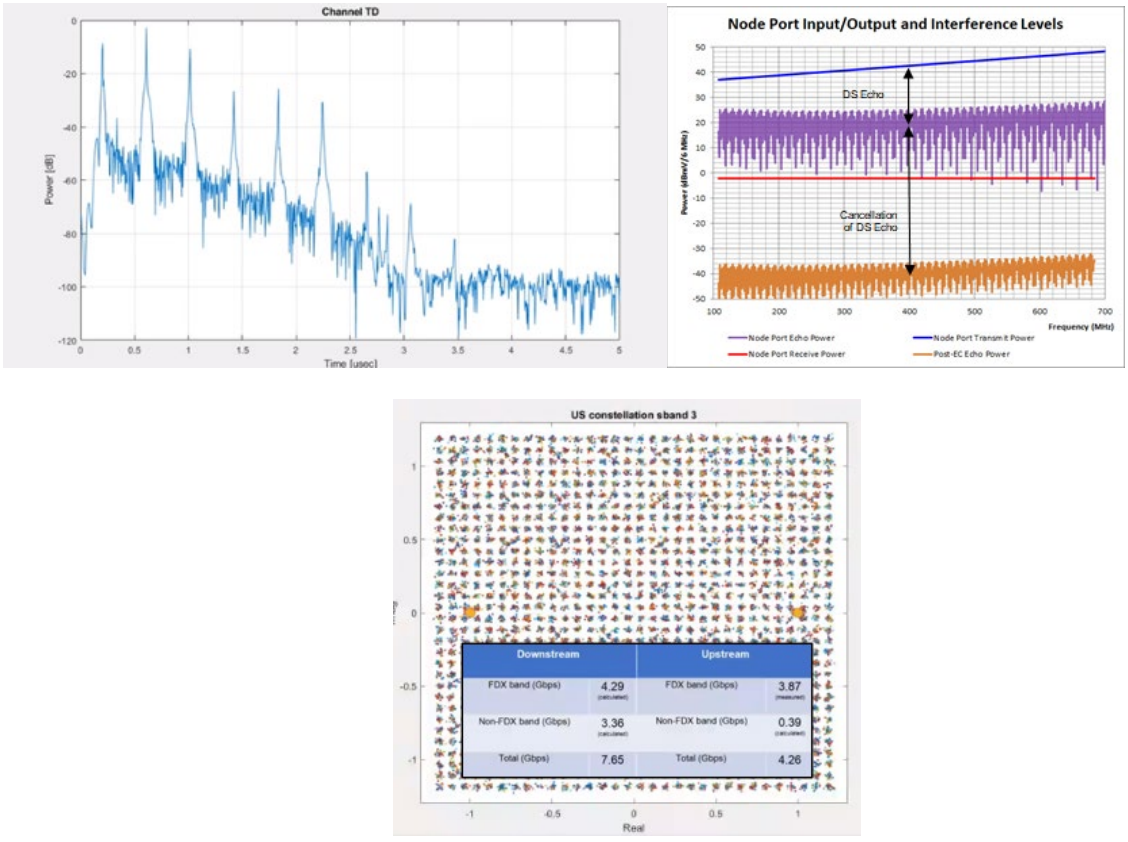


Figure 7 - DOCSIS 4.0 FDX RPD in Operation (clockwise from upper left): a) Plant Echo Response from Node Port b) Simulated Node Tx, Tx Echo, Node Rx, Cancelled Echo c) 1024-QAM US

The DOCSIS 4.0 specified minimum upstream MER requirements are higher than DOCSIS 3.1 in anticipation of noise contributions in the form of residual echo from the FDX band. Nonetheless, the test results observed and shown here using a "Model 1" plant and echo environment is delivering performance similar to a straight DOCSIS 3.1 US receiver, without needing any additional headroom for MER loss.

There will surely be more to come from the labs and shortly in the field as FDX is birthed in 2021!

3.1.6. FDX-Capable Amplifiers

The FDX system specifications were written using the assumption of an N+0 network. However, FDX is not technically limited to an amplifier-free plant and the specification does not prevent it. In fact, shortly after the FDX specifications were compiled, a CableLabs Study Group was formed to evaluate methods for implementing amplifiers that support FDX. The foundational EC technology developed for the FDX RPD can in principle apply at any point in the network to manage overlapping spectrum, and this can include amplifiers. Of course, these are therefore not traditional amplifiers, but a new class of device that includes this new digital signal processing (DSP).

An EC-based amplifier concept is shown in **Figure 8**. The nature of overlapping spectrum and gain in both directions creates a full-circle loop gain path. The EC must be capable of suppressing the FDX loop gain, such that the net gain around the path is < 0 dB to maintain a stable device. The EC must further be designed to act on the echo it is suppressing sufficiently that the aggregate residual echo noise, which becomes part of the amplifier's own noise floor, supports the US MER requirements effectively for DOCSIS4.0, without introducing unacceptable MER degradation and subsequent loss of bandwidth efficiency.

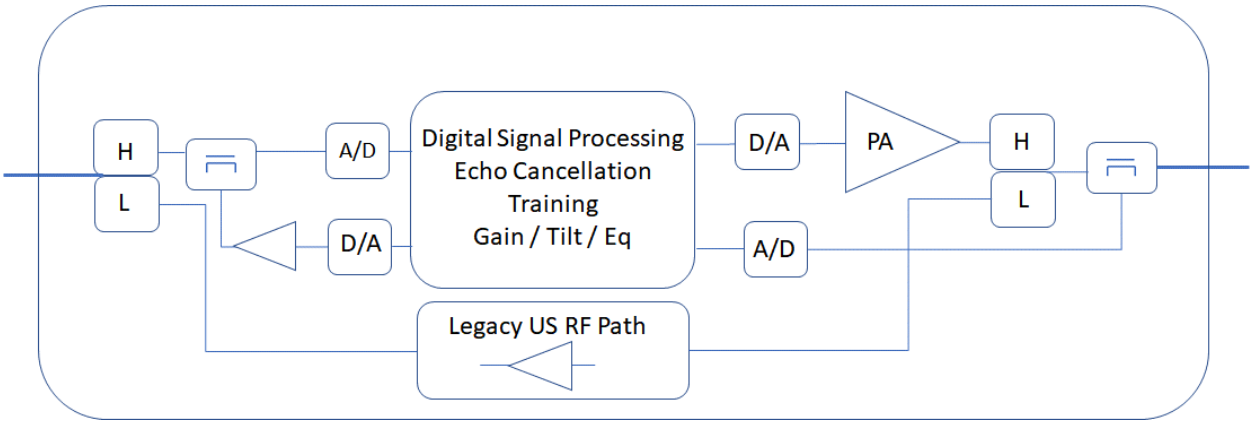


Figure 8 - FDX Over N+X Using an Echo Cancellation-Based Amplifier

The Echo Cancellation function required is similar to that of the node, except that for an amplifier it can exist on both sides. A sample of the upstream signal is taken so that an opposite phase, equal magnitude version can be added in front of the downstream amplifier. Similarly, the signal from the downstream amplifier is sampled and an anti-downstream version added at the input to the upstream amplifier.

As in HFC, as the cascade increases, noise contributions aggregate and the MER decreases. For an FDX amplifier, in the FDX band, there is an additional noise contributor to account for in the form of residual echo. The amount of acceptable degradation due to the amplifier is a system engineering parameter that flows from performance specifications ultimately to the performance of the EC itself. A significant advantage for amplifier EC when compared to N+0 is that the levels on the DS port are lower, and on the

US port higher. Thus, the DS to US level ratio is smaller, which is favorable for the EC that must subtract the downstream.

While noise analysis is relatively straightforward for N+x with FDX, the traffic engineering requires additional scrutiny due to the shared DS and US and the effects amplifiers can have in expanding the size of an IG.

Consider the N+1 system shown in **Figure 9**. When an amplifier is included, there is an expansion of Interference Group 4 (IG4) to the "south," or home-facing, side of the amplifier. These users become part of the last IG of the tap string before the amplifier. This is because of the limited drop-to-output isolation characteristics of today's taps. This parameter can be optimized for high drop-to-output isolation, but until there was FDX to consider, there was no reason to do so. Work in this area shows promise if the logistical and operational challenges can be managed.

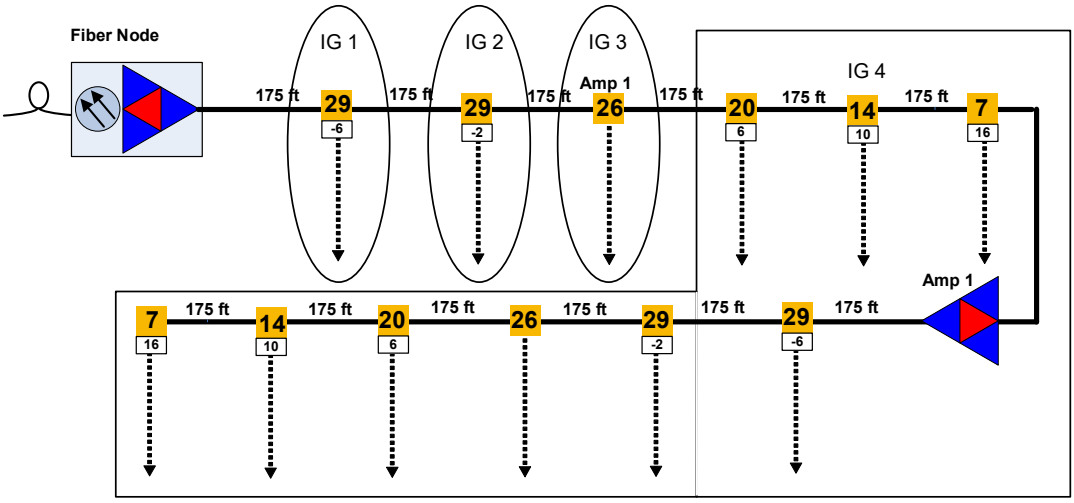


Figure 9 - Potential Interference Group Expansion due to Amplifier on an FDX Network

If we think about network segmentation triggers in current HFC networks, rules have been developed that link service group size, speeds, and percent capacity utilization and define total capacity required. Then, with an awareness of device penetrations, this can be translated to DOCSIS 3.0 and DOCSIS 3.1 spectrum requirements and the numerical thresholds that trigger a network augmentation.

Similar capacity-based analysis and empirical rule making will now apply to FDX, with one additional nuance. The FDX band is allocated for both downstream *and* upstream. The infrequent bursts of peak speeds can therefore be called upon to service the downstream *or* the upstream, but they have a new shared bandwidth dependency. Although from the node perspective there is full duplex spectrum operation, from the CM perspective the introduction of IGs places constraints on who can simultaneously access the spectrum. **Figure 10** depicts this observation.

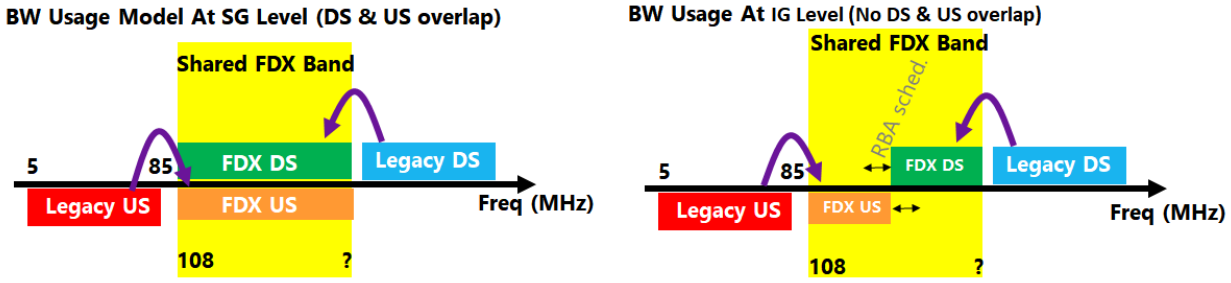


Figure 10 - Managing FDX Bandwidth to Guarantee Peak Speed Bursts [Courtesy CommScope]

The traffic engineering questions become:

1) How large can an IG be before there is an impact to the customer experience?
2) What service speed / IG size / spectrum rules exist when downstream and upstream traffic engineering become co-mingled in FDX?

Figure 11 shows a model assessing the subscriber count, spectrum, and QAM efficiency trade space for a set of input parameters projected to 2028 that includes average utilization with CAGRs of 35% downstream and 30% upstream, and speed tiers up to 4 Gbps/4 Gbps subscribed to by *every* user. The subscriber group (or IG size) sharing FDX spectrum can be as large as 64 and stay within 1200MHz of total spectrum, or even within 1 GHz for a relatively inefficient (for DAA) 9 bps/Hz net downstream throughput. The upstream, even for the 4 Gbps/4 Gbps (green) case, stays within the FDX band allocation for an IG size above 128.

What these studies reveal is that FDX bandwidth can be used extremely efficiently. Peak bursts are extremely infrequent and the collision of bursts from concurrent peaking users is rarer still. With an 85MHz legacy upstream, the upstream bandwidth is sufficient for peak-busy-hour (pbh) average utilization for up to 200 subscribers – where the red arrow points to the FDX bandwidth breach in **Figure 11**. The FDX upstream band turns out to be primarily a spillover reservoir for the occasional burst peaking user. Meanwhile, most of the daily grind for the FDX band is in delivering downstream capacity. This represents an extremely efficient use of precious HFC spectrum, which is precisely the principle that FDX was created on.

Reviewing, today's capacity management is based on assessing service group size vs available capacity vs utilization. At a certain empirically derived utilization threshold, a network augmentation is triggered. Typically, these are triggered by upstream utilization, and is a node split. However, it can also be new spectrum, more DOCSIS 3.1, or QAM reclamation.

The effect of FDX and the IG phenomenon is that now, in addition to the parameters above, network augmentation may be triggered by introduction of a new *speed tier*, and as a secondary factor the penetration of that tier over time. This will also ultimately be empirically derived, but initially be based on guidelines such as those given above. Network segmentation rules have been developed for services levels of 2/2 Gbps, 3/3 Gbps, and 4/4 Gbps.

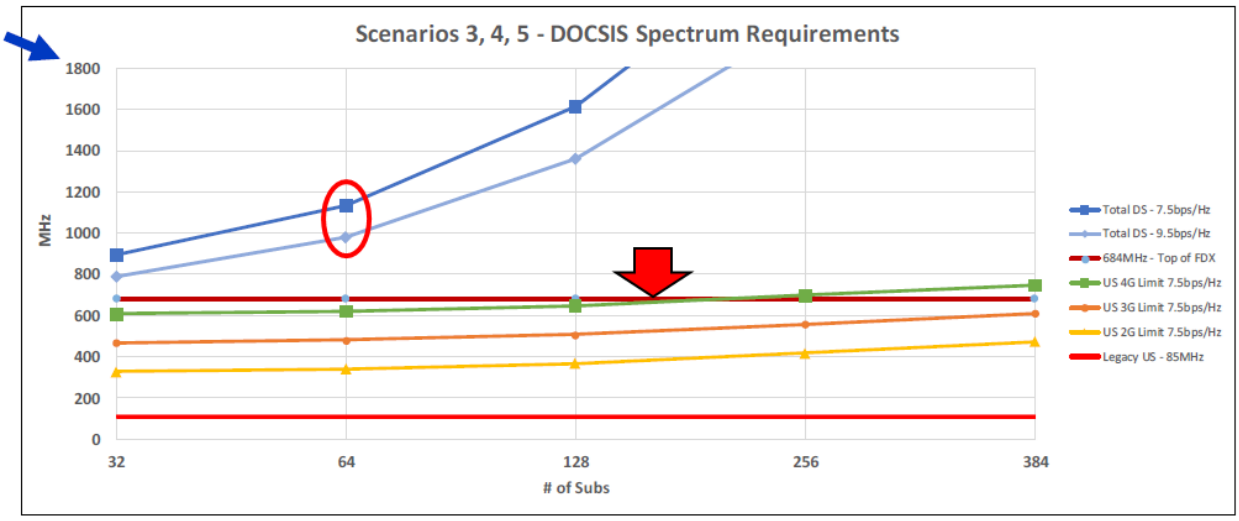


Figure 11 - Total RF Spectrum Required vs Subs Sharing a TG [3 Gbps/3 Gbps, 4 Gbps/2 Gbps, and 4Gbps /4 Gbps]

3.2. Extended Spectrum DOCSIS

3.2.1. What Is Extended Spectrum in the Plant?

The answer is simple! The cable industry has been executing HFC bandwidth expansion path for decades, so moving the spectrum to 1.8 GHz can be considered simply the next step in a well-established network evolution model. The next step in this path to higher bandwidth HFC has been coined *"Extended Spectrum"*. Looking back at the history of HFC plants, it was not so long ago that the upgrade was to 550MHz, then to 750 MHz, closely followed by 860 MHz, then 1GHz, and most recently to 1.2 GHz. Many "old-timers" can probably remember some of the incremental steps prior to 550 MHz, with long amplifier cascades – prior to adding the "F" in HFC. A step to 1.8 GHz from this perspective is just the next step in a spectrum and capacity expansion on the way to potentially a 3.0 GHz HFC network. In this manner we continue to leverage our valuable assets as has been done in HFC successfully for decades.

Recognizing the fact that HFC is going to be around for a very long time, cable operators are looking at what to do to keep moving on the path to 10G (10 Gbps). Prior spectrum expansion in the last two decades were primarily driven on the need for more downstream channel capacity; now, it is primarily upstream capacity, creating a more complex set of questions for what to do next.

There is increasing pressure on the physical layer of the network to create incremental capacity. Many networks may need additional interventions to keep up with new product and service offerings, but mostly to keep pace with the data traffic growth. The increase in Internet bandwidth demand is one of the primary reasons we touch the outside plant network today. Internet connectivity has become a very competitive landscape over the last decade, with the majority of the operators increasing data speeds on a consistent cadence. Had the role of data growth on network augmentation been understood 20 years ago at the dawn of HSD services, a more methodical and capacity engineering-based approach to place nodes could have produced more balanced service groups, evenly distributing data capacity. Nevertheless, over the last decade operators have learned the nature of capacity engineering and have been performing node segmentations, node splitting, and in some cases more aggressive steps, such as N+0 upgrades.

As expected, the results of these activities have benefited operators by systematically driving fiber deeper into their networks and closer to customer premises, thus reducing amplifier cascades and ultimately improving network performance and reliability. The segmentations have shrunk service group sizes sharing the bandwidth. Combined with aforementioned spectrum allocations, operators have a set of tools at their disposable to manage capacity, and with a range of impact suited to the incremental capacity upgrades of a simple node split, to tools suited to high growth, competitive footprint preparedness that include new spectrum, deeper fiber, and DAA.

3.2.2. Why Extend the Spectrum?

Leveraging the current assets of the HFC network in a manner that does not greatly change how operators invest capital is the key fundamental premise of *"Extended Spectrum."* Keeping network operations simple and familiar for the field operations teams is a key advantage of this approach. Extended Spectrum will allow cable operators to defer capital investments over a much longer period of time by using the "Business As Usual" (BAU) approach, while pragmatically driving fiber deeper into their networks. The improvements to amplifier technology will allow for continued use of most existing amplifier cascades, which fits into the current tree and branch topology of HFC networks.

Figure 12 shows how and Extended Spectrum allocation compares to today and to the High Split scenario that was shown in **Figure 1**.

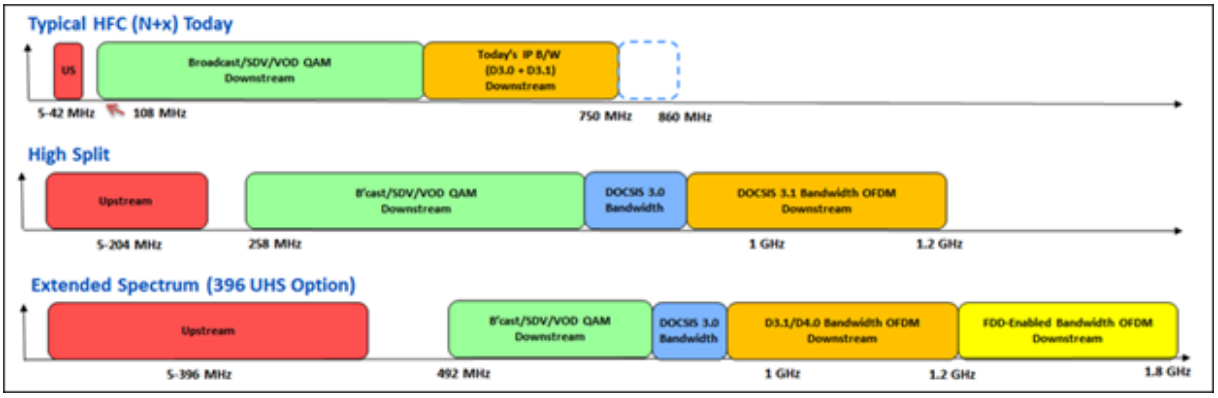


Figure 12 - Typical HFC vs High Split vs DOCSIS 4.0 Extended Spectrum

One of the biggest changes in future HFC plant upgrades will be pushing the upper limit of the carriers higher in the spectrum. Many details are still being worked out, but the goal is to perform drop-in upgrade to systems that were properly designed and built to a true 750MHz or higher. A properly designed and built system is highly dependent on the accuracy of the plant maps, the designer doing the work, the line equipment being used, and the distances in the spans of the coaxial cable. All of these potential uncertainties, plus the variables that went into the last system upgrade, can have an impact on this next step. With this in mind, operators will have flexibility to offer increased bandwidth in select locations to increase their speed to deployment.

So, how are operators looking at implementation of Extended Spectrum in practice? Below are some examples from an operator point of view in addressing some of the key changes introduced:

- *When is the right time to replace taps and passives*? The objective is for the frequency response of the housing to support 3.0 GHz now, with a 1.8GHz capability built into the device from day one. There is not an urgent need to put 1.8 GHz in place today, as most network upgrades are

beginning with a 750 MHz plant. Thus, even after an upstream split change, there is typically open and useable spectrum up to 1.2 GHz to consume, buying time to phase in taps and passives upgrades as the downstream bandwidth gets utilized.

- *How best to manage Total Composite Power over the extended bandwidth?* Maintaining a realistic Total Composite Power (TCP) while extending the spectrum to 1.8 GHz requires thinking about the tilt and output levels, because of the need to maintain legacy RF levels into legacy CPE. Working closely with the silicon vendors has been an important step to understand trade-offs between level and performance.

  Note that the RF amplification is being designed to drop into the current amplifier spacing and location currently utilize for 750 MHz plant. This creates the need for new RF hybrids that are capable of amplifying at much higher frequencies while maintain a Total Composite Power (TCP) that is acceptable to minimize all distortions while not exceeding a power budget. Given the magnitude of this challenge, alternatives are being considered while the amplifier efficiencies catch up, looking at ways to not exceed the TCP capabilities of the amplifier. A simple way this can be accomplished is to use a step-down approach of the signal at either 1 GHz or 1.2 GHz. This step down will maintain a TCP, while not impacting the amplifier linearity that would otherwise degrade amplifier MER we want to achieve.

- For the upstream split, what is the "right" ratio? 204, 396, 492, 684 MHz or beyond are all flexible and effective solutions for whichever direction various operators choose for their markets. Each must be evaluated comprehensively within the market of interest to ensure they do not break powering boundaries of the network, reduce overall performance, or dramatically increase complexity. **Figure 13** and **Table 1** [6] show the spectrum allocations for downstream and upstream based on the DOCSIS 4.0 standard "Ultra High Split "(UHS) options.

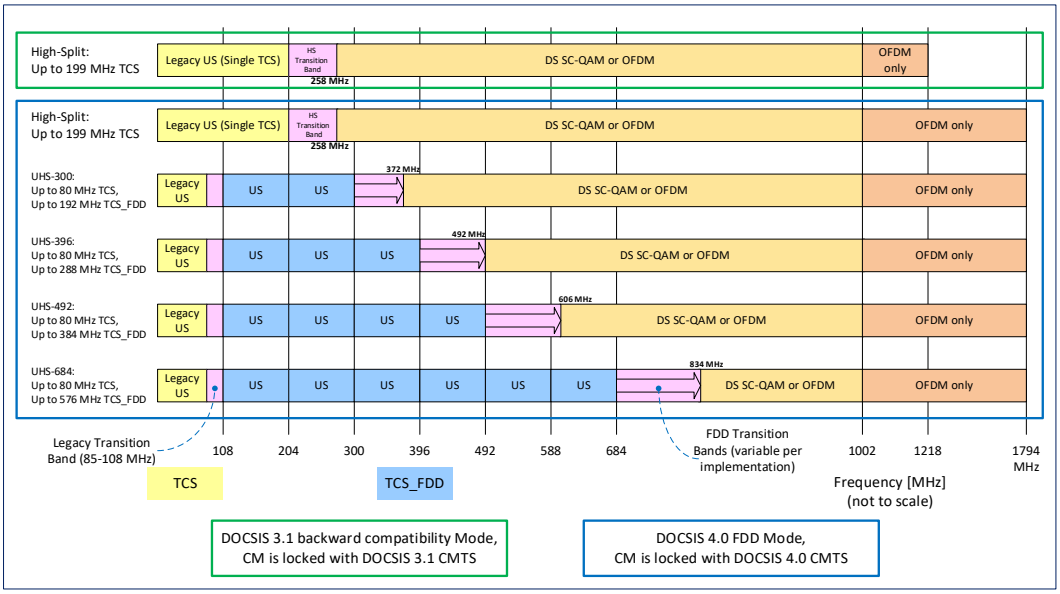


Figure 13 - DOCSIS 4.0 Extended Spectrum Allocation Options for Downstream and Upstream

Table 1 - DOCSIS 4.0 Extended Spectrum Ultra High-Split Options

Split Name	Diplex Filter Start Frequency (Upstream upper band edge) (MHz)	Diplex Filter Stop Frequency (Downstream lower band edge) (MHz)
High-Split	204*	258
UHS-300	300	300 MHz–372 MHz
UHS-396	396	396 MHz–492 MHz
UHS-492	492	492 MHz–606 MHz
UHS-684	684	684 MHz–834 MHz

Changes are coming to HFC networks over the next few years in order to deliver more bandwidth and prepare for 10G. Extended spectrum ensures that operators have a long-term cost-effective option. Continuing to leverage these assets in a strategic manner is one way in which we will remain competitive while meeting the growing demands and insatiable subscriber's appetite for more bandwidth. The path to 10G initiative announced by CableLabs has accelerated efforts to deliver 10 Gbps service over coaxial cable networks. This makes strategic research efforts a big part of how operators are planning to reach the 10G vision and which technologies will be best to use. **Figure 14** shows the first DOCSIS 4.0 Extended Spectrum DAA node SoC, now being brought up in labs.

As an industry, it is imperative to have a large selection of tools that allows cable operators flexibility and scalability for their specific operating models. When looking at all the different architectures and scenarios that exist across the globe, operators must consider how to best scale the available assets and when to do so. As has been the case for decades, this will include a mixture of tools that support implementing roadmaps, such as the path to 10G, with anticipation for key technologies on the horizon.

Figure 14 - First DOCSIS 4.0 FDD SoC

Recently CableLabs completed the specification work on DOCSIS 4.0, which includes Full Duplex DOCSIS (FDX) and Extended Spectrum DOCSIS (ESD). These DOCSIS 4.0 variants will support different options to increase bandwidth and capacity, and furthermore support convergence of the technologies over time.

3.3. FDD-FDX Synergies and Key Differences

The most important commonality between DOCSIS 4.0 FDX and DOCSIS 4.0 FDD is the use of the DOCSIS3.1 technology as the basis for their Physical Layers. DOCSIS 3.1, of course, was a major shift away from what had been exclusively Single-Carrier QAM (SC-QAM) modulation formats of limited maximum efficiency for both the downstream (256-QAM) and the upstream (64-QAM), and using a combination of Convolutional and Reed-Solomon coding for Forward Error Correction (FEC) to increase the robustness. DOCSIS 3.1 broke the precedent of compatible SC-QAM based versions of DOCSIS, embraced the fast-growing adoption of multicarrier modulation (OFDM/OFDMA), and updated the FEC to Low Density Parity Check (LDPC) that was, despite its computational complexity, now able be implemented in real time, delivering significant new coding gain.

Both DOCSIS 4.0 FDX and DOCSIS 4.0 FDD rely on these powerful DOCSIS 3.1 modulation formats and FEC. Furthermore, the same OFDMA and OFDMA "numerology" details are unchanged in DOCSIS4.0 – subcarrier spacings options, range of cyclic prefixes, FFT size, windowing, exclusion bands, pilots, PLC channel, etc.

Of course, the focus of DOCSIS 4.0 is the expanded spectrum range that these "DOCSIS 3.1" signals can be extended over, into frequency allocations where they had not been allowed before.

3.3.1. The Same Except Where they are Different

The discussion to follow will mostly take place from a network and DAA node point of view. However, a parallel discussion with many of the same characteristics apply on the CPE side, although the economic equation for CPE devices is very different. Technology parallels notwithstanding, the commercial differences and the premium value on the customer-facing LAN interfaces may lead to different business decision points on implementation and feature priorities.

For both FDX and FDD, upstream signals can extend beyond 204 MHz, up to 684 MHz. The required amount of OFDM and OFDMA resources is similar, as shown in **Table 2**. This is by design, enabling system-on-a-chip (SoC) suppliers to design and manufacture their chips which highly common blocks and functions for the most complex elements of the SoC. The common technology basis makes for an efficient and cost-effective ecosystem.

Where DOCSIS 4.0 differs from DOCSIS 3.1 more significantly are the differences between FDX and FDD that force changes at the "edge" of the SoC and outside of the SoC. However, the massive DOCSIS 3.1 processing engines required for either FDX or FDD are dominant features of these SoCs. Coupled with the common processing resources required, manufacturers of these devices may choose to implement the more moderate differences at the edge of the chip within a common SoC as configurable elements. This could be the case in the DAA node, or the CPE, or both.

Table 2 - Common DOCSIS Resources are Defined for FDX and FDD [6]

Item	Device	FDX OFDM/OFDMA	ESD OFDM/OFDMA	SC-QAM
Downstream Channel Support	CM	5 total OFDM channels; 3 channels capable of FDX operation; All channels capable of non-FDX operation up to 1218 MHz	5 total OFDM channels; All Channels capable of operation up to 1794MHz	32
	CMTS	6 total OFDM channels; 3 channels capable of FDX operation; All channels capable of non-FDX operation up to 1218 MHz	8 total OFDM All channels capable of operations up to 1794MHz	32
Upstream Channel Support	CM	At least 7 total OFDMA channels; 6 channels capable of FDX operation; 2 channels capable of non-FDX operation within the legacy diplexer configuration. (Some channels can be configurable to support either FDX or non-FDX operation. When supporting 6 FDX OFDMA channels, only 1 non-FDX OFDMA channel is required.)	7 total OFDMA channels	4 (or 8) SC-QAM channels, operating within the legacy diplexer configuration
	CMTS	8 total OFDMA channels; 6 channels with FDX operation; 2 channels capable of non-FDX operation based on operator deployment requirements.	8 total OFDMA channels	4 (or 8) SC-QAM channels, operation dependent on operator deployment requirements

The nature of the differences on the Original Equipment Manufacturers (OEMs) who build the FDX or FDD components around the SoC technology is more challenging than the SoC vendors. In FDD systems, the DAA node and CPE extends to 1.8 GHz, and includes diplex options for upstream settings. In FDX system, the typical 1.2 GHz bandwidth suffices, and there are no additional diplexers in the node beyond the standard Mid-Split one isolating the legacy upstream from the FDX band. FDX upstream is added by remotely configuring the vCMTS for the FDX band to activate the desired number of OFDMA blocks and extending the EC technology across these new blocks. The "Big Idea" is that this new massive upstream is added but where there is also downstream in use, thereby creating an extremely efficient use of coaxial spectrum without a large guardband penalty.

However, since OFDMA activation in the FDX band is not filter-based, an RF path at a node port, as shown in **Figure 15**, uses a *directional coupler* before the legacy diplexer in order to siphon off the upstream signal to send to an FDX band receiver. Thus, RF layout for an RF board in an FDX node is very different than in an FDD system.

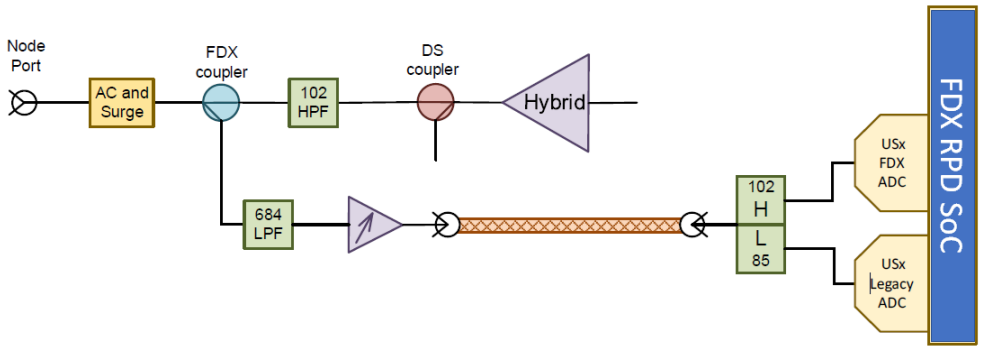


Figure 15 - RF Processing in FDX is Different than in FDD systems

Another change on the RF layout for FDX (not shown in **Figure 15)** are RF feedback traces from the downstream RF path, such as from the "DS Coupler" to sample and enable the signal so the EC can learn the interference and subtract it.

As described, in an FDD system, the RF processing chain is similar to today's nodes except for the much wider frequency response, and that it is expected that FDD nodes will incorporate a subset of the DOCSIS4.0 specification-defined diplexers to address upstream speeds by adjusting the diplex filter value remotely. The FDD "Big Idea" is then to enable this extended upstream-only spectrum, and downstream signals are extended up to 1794 MHz to make room for the shifted downstream to sit above the higher upstream plus ensuing guardband. FDD guardband is approximately 20% of the upstream diplex edge, for a maximum of 150 MHz in the 684 MHz case at the top end of the range. While 684 MHz is an available option, current usage and speed trajectories indicate this maximum upstream band edge may not be required, or at least not for many years.

Technically, 1794 MHz is the upper band limit defined initially for the DOCSIS 3.1 downstream. However, it had been largely deferred as DOCSIS 3.1 was developed (circa 2012) and the "I01" first release published. During that time, it was determined that quantifying the specification to this forward band limit could be deferred. This quantification is now taking place, and because of this, while the band edge of the frequency was identified by the DOCSIS 3.1 specification, it is the DOCSIS 4.0 work that is completing the requirements. Thus, 1794 MHz is typically identified with DOCSIS 4.0 Extended Spectrum and similarly, DOCSIS 4.0 FDX is associated with the 1218 MHz limit.

Figure 16 illustrates the commonality of DOCSIS processing resources from a spectrum utilization point-of-view between FDX and FDD.

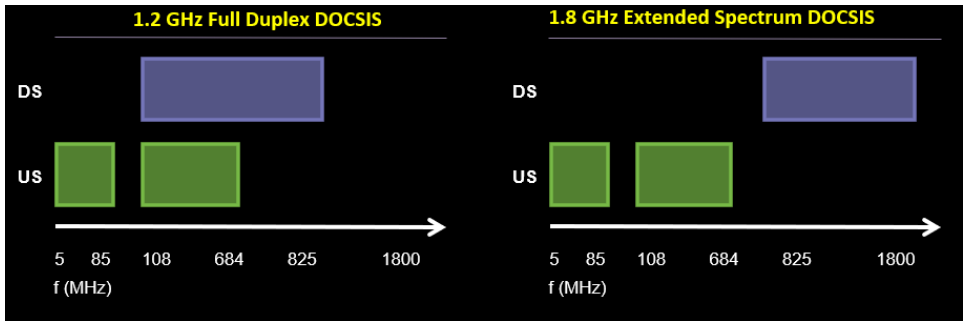


Figure 16 - DOCSIS Resources Downstream and Upstream are Common in FDX and FDD

3.3.2. Is this What DOCSIS 4.1 Might Look Like?

Observing the two DOCSIS 4.0 options, we see that both add a common set of upstream OFDMA blocks over a common frequency range. One, FDD, pushes a significant physical bandwidth extension into the network relying on common RF practices executed for several generations of HFC migration. The other, FDX, relies on the introduction of new DSP technology to complement the existing plant technology in order that the spectrum asset be operated more efficiently over a bandwidth that the plant is built to support.

Importantly, these are not mutually exclusive approaches! In fact, they are complementary technologies that, in principle, could be merged in some future DOCSIS extension, perhaps as shown in **Figure 17.**

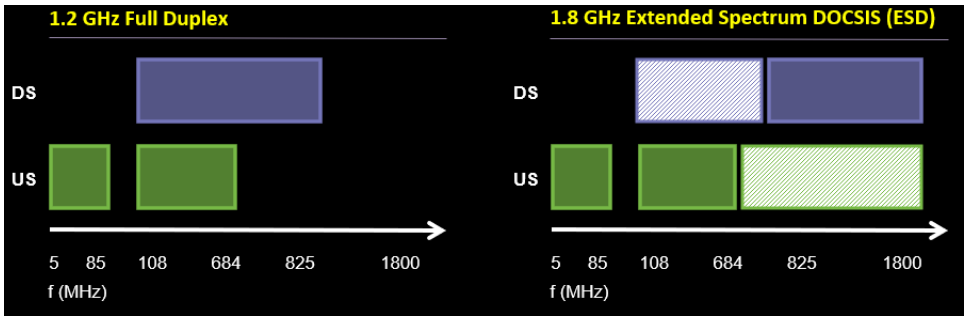


Figure 17 - The Potential Complementary Nature of FDX and FDD

With the DOCSIS 4.0 upstream capacity as defined for FDX or FDD in the 5-6 Gbps range, merging the two technologies may be next step toward push the upstream also towards the 10 Gig goalpost.

3.4. Characteristics and Attributes Summary

Table 3 summarizes DOCSIS 4.0 FDD and FDX comparative attributes discussed herein (most of them), side-by-side. This is a useful cheat sheet to have on a whiteboard (perhaps a relic of a pre-Covid era) to generate dialogue, inspire feedback, identify where information gaps exist, zero in on those that carry the most weight, and generally spark debate and pro/con scorecards . We will use it later to re-form from a side-by-side comparison into DOCSIS 4.0 NFEAQs – Not Frequently Enough Asked Questions.

Table 3 - FDD-FDX Attributes Comparison

Attribute	FDD/ESD	FDX
Strategy/Philosophy to 10G	• Based on access network BW extension upgrade, up to 1.8GHz, for existing actives and passives • Introduction of DAA to migrate to 10G, similar to previous HFC plant upgrades with a choice of diplex split configurations	• Based on access network technology upgrades to introduce new DSP (EC) into RPHY nodes and Amplifier platforms • Build on DAA production and scaling of vCMTS as the foundation for 10G
Migration Factors	• 1.8 GHz DOCSIS 4.0 DAA Nodes, Amps, Taps and Passives • Allows for cascade of amplifiers • New CPE	• RPD Nodes with DOCSIS 4.0 EC function • FDX-capable amps with DSP • New CPE
Complexity	• New tech challenges – BW and TCP extension • Use of "low power" amp extender for edge cases	• New tech challenges – EC function, CMTS scheduler, DSP-based amps • New capacity mgmt rule for IG/TG size for peak speed
Spectrum/Capacity	• 1536 MHz DS/656 MHz US (see **Figure 13** and **Table 1**) • Up to 15G/5G (all-DOCSIS 3.1) • DS/US: BW and Capacity per diplex selection	• 1110 MHz DS / 656 MHz US (see **Figure 2**) • Up to 11G/5G (simultaneous, all-DOCSIS 3.1) • FDX BW/speed by SW config
Operations	• Utilize existing common operational practices – FDD system with different possible split choices • New field tools	• New operational practices for handling of spectrum overlap and amplifier installations • New field tools
Network	• N+X • Cascade reduction/trade-off based market capabilities	• N+0 (optimal) • N+X – Cascade reduction/trade-off based on market speeds
As-Built Migration	• Continue node split and introduce DAA node splits, leverage for HFC migration activity, introducing components of FDD over time • Migration path and timing considerations for Underground vs Aerial and MDU vs SDU cost implications	• Introduce DAA for node splits with vCMTS, leverage for HFC migration activity and platforms that enable FDX activation • Migration path and timing considerations for Underground vs Aerial and MDU vs SDU cost implications

While capacity and data speeds often garner the most of the attention when discussing access network, other important attributes from **Table 3** that consume more attention when comparing the options above are alignment to long-term strategy for the access network, network upgrade costs, and operational implications. The latter two have corollaries the attributes "as-built" and "complexity," which we will dip into a bit deeper in the next section.

4. DOCSIS 4.0 Migration – Key Variables

Across a single MSO are a range of HFC architectures. Larger MSOs, such as Comcast and Charter, tend to have a very wider range of network variants, owing to the consolidation of many smaller operators over time and the exchanging of properties among MSOs to gain operating efficiencies. There are opportunities to reel in the range of variations with the introduction of new technology and defining new architectures as part of a Next Generation migration plan. It is an opportunity to build a more common end state. However, as noted, the difficult part is always in the transition *to* a desired end state. Several important network characteristics play a role in the cost, complexity, reliability, and performance of the end state achieved from a given HFC baseline architecture and physical network.

4.1. Network Bandwidth

Today's HFC networks come mostly in the 3 varieties of maximum bandwidth described earlier – 750 MHz, 860 MHz, and 1 GHz – with a fourth emerging at 1.2 GHz, which nearly all new actives and passives support today. Many of these networks, in particular 750 MHz networks, likely began their lives designed for much lower total bandwidth, such as 450 MHz or 550 MHz (even 330 MHz). RF signal loss over coaxial is frequency-dependent and has a predictable inverse root-frequency relationship. As such, the construction of the network led to a physical distance between actives and passives as well as feeder cable

requirements that could deliver a desired target end-of-line (EOL) performance – modest by today's standards – although at that time built around analog video requirements. Analog video is sensitive to noise and distortion, but these are pre-High Definition (HD) days with what would be considered low video quality expectations today.

This "spacing" generally was able to be held intact as RF amplification technology over time overcame the limitations of long spacing that was associated to an assumption of RF loss that may have fit for 450 MHz, but not for 750MHz. Bandwidth increments were relatively small steps, and broadband power amplifiers got better through the evolution from Silicon-based to Gallium Arsenide-based (GaAs) to Gallium Nitride (GaN)-based. By "better," we mean able to extend in bandwidth, but also able to extend in Total Composite Power (TCP), since more bandwidth to cover means more power to transmit. In addition, because RF transmission is launched on a tilt, the TCP is impacted disproportionately higher to the bandwidth added.

Lastly, most taps and passives on the network have upper frequency specifications such as 1GHz or 1.2GHz over which their RF parameter must be met. Passive devices can carry signals without guarantee of performance beyond these ranges with generally parasitic additional loss, to a point. With specification guaranteed only to 1.0 GHz or 1.2 GHz, there is a point above that at which the technology used in the design becomes incapable of supporting reasonable transmission with predictable loss.

The above variables have been part of the considerations for every HFC plant bandwidth upgrade over the years, and will be again as the network extension to 1.8 GHz is developed. How capable is the existing plant for supporting 1.8GHz and what are the cost / performance trade-offs and implications to prepare the network for this extension. There are no current HFC plants built with 1.8 GHz spacing in mind, and the disproportionate TCP due to tilted RF loading to 1.8 GHz places some constraints on the power spectral density profile.

On the FDX side, the RF design at the port of a node adds new passive components for feedback EC and a directional coupler to support the FDX band operation in both directions. These additional passive losses occur right at the output of the PA on the way to the port, contributing to launch power decreases that can reduce reach to homes on a passive network or decrease levels at the next active. The coupler designed loss also impacts the upstream signal at the node port in the FDX band that must be accounted for in the design and system engineering. These effects must be compensated for in the overall node design and accounted for in the system engineering. The trade-off space is typical in many ways - gains, levels, distortion, noise – towards an optimized FDX node design. What is different is that in the FDX band the trade-offs must have both DS and US in mind, since they share that part of the spectrum.

4.2. Cascade Depth

VERY long amplifier cascades were eliminated with the invention of HFC, allowing fiber to cover long distances and limit the coaxial cable portion of the plant. "Long" cascades today are much different. **Figure 18** shows a distribution of cascade depths across a sample within a region of the Comcast network used for a business modeling exercise. These distributions vary across regions with an operator as well as across operators. Comcast has builts million of homes passed to date in an N+0 configuration [3]. The sample selected in **Figure 18** is specifically outside of these N+0 build areas, but of course include some natural N+o as-built cases, such as MDUs. For this subset of the network, over 5 amplifiers in cascade would qualify as "a "large" cascade. Note that the "5" in N+5 merely refers to the maximum series cascade in the RF design. There could be 15 amplifiers off of a node port (near the median), but the number in cascade would be a maximum of 5.

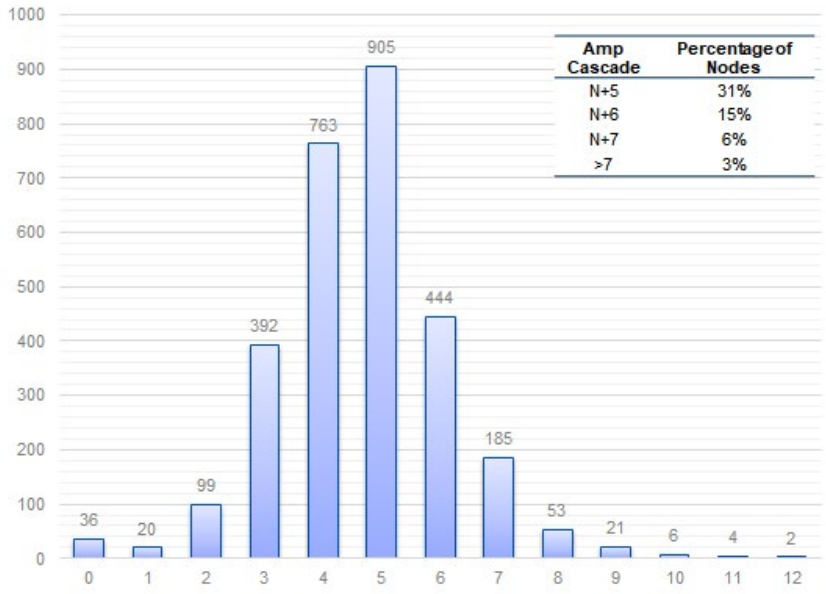


Figure 18 - Cascade Distribution – Comcast Sample

Cascade depth effects the same things it always has – MER degradation, frequency response, serving groups size (when combined with homes passed density).

For DOCSIS 4.0 cascade depth impacts are the same, only different. For FDX, as discussed, cascade depth includes the contribution of any residual noise of echo cancellation in addition to natural thermal noise. And, by the nature of the service group size relationship, cascade depth has a relationship to IGs and speed offering.

For FDD, the cascade depth drives the frequency response roll-off associated with the increased RF loss with frequency of existing feeder and drop cable, and taps and passives. Projected performance and the frequency extreme drive the minimum receive levels and ultimately bandwidth efficiency expectations.

For both cases, setting a maximum cascade design rule to offset the negative consequences of deeper cascades will drive cost into the migration plan in order to accomplish the segmentation needed. These costs can be modeled based on target objectives for the DOCSIS 4.0 network and the implications of those objectives on the cost of the network augmentation. In addition, as is always the case with network architecture, it can be phased over time to balance the deferment of upgrade cost with the need to achieve certain capacity and speed targets on Day 1 vs Day 1001. The flexibility to easily deploy new downstream and upstream capacity then becomes a consideration, always looking to eliminate or minimize physical plant touches.

4.3. Aerial or Underground Construction

For any network augmentation that involves pulling fiber or coax cable, there is a significant difference in the labor cost of this work between aerial and underground construction that is very favorable to the aerial plant style. The availability of conduit with unused carrying capacity underground helps to minimize this difference, where that is the case. Of course, not all underground is created equal either, such as augmentations that require tearing up concrete and crossing busy streets in metropolitan cores.

Operators know their own labor costs for these physical and regional variables, as well as any overhead costs associated with different municipalities, which can also vary widely. Sample network migration models can be created based on these known construction types. And, as pointed out in the discussion on "cascade depth," network augmentation can be phased as a function of time tied to capacity targets versus time, governed by CAGRs, and speed capability versus time, governed by the HSD business. The phasing has to be balanced by the opposing force of not creating too many network disruptions that are customer-impacting, and not creating excess inefficiency with too many ventures into the plant to take steps that were not sufficiently consequential the last time around in buying time to support traffic demand.

4.4. Homes Passed (HHP) Density

Not surprisingly, all else the same, it is more cost effective to upgrade a high-density area than a low density area, simply because the denominator is larger. Secondarily, there is a relative uniformity of higher density footprint that has a stabilizing effect on metrics, the customer experience and simplifying operations.

The boundary case of high density is the Multi-Dwelling Unit (MDU) environment. Unfortunately, and somewhat counterintuitively, such environments are also prone to be among the more challenging from an RF standpoint for a variety of reasons. MDUs will be discussed in more detail in the next section.

4.5. Network Powering

We will likely be asking more from the network power supplies that power the HFC plant when moving to DOCSIS 4.0. The DAA foundation of DOCSIS 4.0, the DSP that is introduced as part of FDX, or the extended bandwidth that is part of FDD all point towards taking a close look at the state of the existing network powering – voltage and amperage – as well as back-up power requirements.

Available power supply monitoring information has allowed Comcast to determine the percentage of upgraded footprint will require more current from existing supplies in a modest plant upgrade, and how many drive all-new power supplies to be added. Sensitivity analysis has been done around how many new amps of current drive power supply upgrades ($) or new power supplies altogether ($$$) given the known supply capacity margin going in.

4.6. Multi-Dwelling Units

When it comes to Multi Dwelling Units (MDUs), the "Good and Bad" are closely coupled. On one hand, an MDU can be considered the low hanging fruit for a DOCSIS 4.0 deployment given the high density and minimized cable lengths. It is that high density that makes these environments targets for competition. To add to the complexity, the FCC has rules (usually referred to as the "Inside Wiring Rules" 47 CFR 76.802) designed to enhance competition in MDU buildings. The FCC rules allow the MDU owner to gain control over Inside Wiring in order to make it available for use by a competitive service provider. It is these rules that make operators more reluctant to rewire the MDU or upgrade with fiber to each unit. A caveat to the FCC rules is that they are technology-agnostic and do not distinguish between the types of wiring that comprises the inside wiring. The applicability of the rules does not depend on whether the Inside Wiring is CAT-5, RG-6 or fiber optic cable.

The demographics of these properties vary greatly from location to location. Some utilize a campus layout, also known as a "garden style." Of course, the high-rise single building is usually what people think of when they hear "MDU." In any case, the density is typically much higher than serving single family units (SFUs), so an operator typically has opportunity for lower investment per living unit.

Servicing the MDU space is also unique in that the owner has the ability to grant exclusivity to the use of the inside wiring. This allows operators to sign "Bulk Agreements" to serve the entire building. Until recently buildings were often wired with coaxial cable. It is this wiring that DOCSIS 4.0 looks to leverage with symmetrical Gigabit speeds similar to fiber, without the costly expense of rewiring the building.

As with any HFC outside plant upgrade to increase bandwidth there will be some network and design challenges. With respect to MDUs, when compared to aerial or underground plant in easements or rights of ways, some of the challenges are similar. There are typically two types of environments, classified as either "Greenfield" new build or "Brownfield" existing network. The latter is the area that gives the greatest benefit to utilizing DOCSIS 4.0, given the re-use of the coaxial infrastructure that exists in a majority of buildings. The higher density of the MDU environment will allow operators to easily deploy in a cost effective and strategic manner. Operators have taken note of the power of fiber to the building, and thereby many sites are fed from a "Dedicated Node" that serves the complex only. These dedicated nodes can be upgraded incrementally by only making changes to that location. Even in the garden style layout there are very few actives and much less cable than single family units. The latter is of high importance given the higher attenuation of coaxial cable at upper frequencies. These shorter coaxial runs would also benefit from a future 3.0 GHz Extended Spectrum for the same reason.

In Greenfield, more builders and owners are opting to install fiber or Ethernet cable as they build the units. These will typically be fed using a Passive Optical Network (PON) technology which is not the focus for this paper. However, cable operators have equalized the conversation around fiber vs coax with respect to HSD speed offerings with DOCSIS 3.1 and DOCSIS 4.0.

Experienced field personnel will attest to the fact that MDUs have their own set of challenges related to maintaining the integrity of the RF performance. These units typically have high churn with people moving in and out more often. This creates more opportunities for loose connectors, open terminations, damaged inside wiring, etc. These types of issues can result in a trouble call or truck roll to resolve. Note for DOCSIS 4.0 Extended Spectrum, the new bands will be occupied by OFDM carriers only, which are much more resilient than SC-QAM signals, with better error correction and with the ability to change modulation profiles when needed.

The above MDU variables can be considered in the MDU design during the deployment. DOCSIS 4.0 has some requirements that will change how these buildings are served. Two of the biggest DOCSIS 4.0 changes are:

1) DOCSIS 4.0 must be deployed as part of a *Distributed Access Architecture* (DAA). With this requirement, there will no longer be a reliance on analog optics that serve nodes today. This will result in improved signal fidelity to the property and into the unit, and thus more DOCSIS 4.0 capacity.

2) The DOCSIS 4.0 modem will be a *Point of Entry device* when used in DOCSIS 4.0 mode, meaning that it will be the sole HFC-terminating device in the unit. There will be no need to be concerned with the splitter network to feed other boxes, such as STBs, and in many cases also the cabling inside the unit. These fundamental changes reduce the concern that we typically have when dealing with insertion and attenuation loses.

As stated in the beginning, MDU environments could be considered the low hanging fruit for initial DOCSIS 4.0 deployments to provide multi-gigabyte services without the need to re-wire the inside of the units. We even see a possibility of leveraging 3.0 GHz in the future in the MDU space.

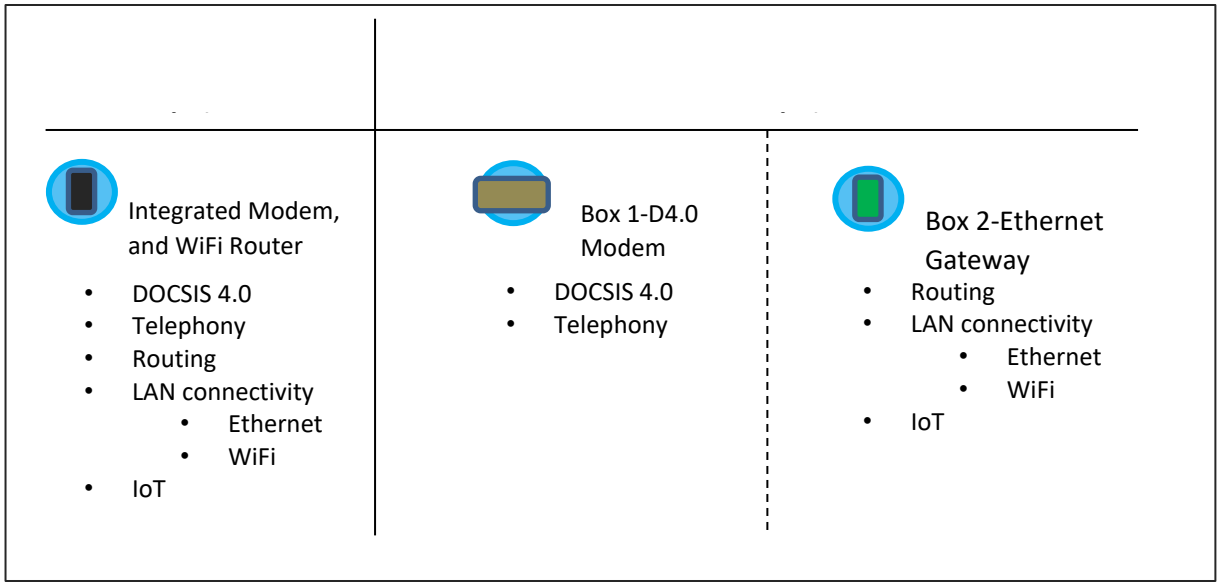
4.7. The Home Network

With increasing numbers of Wi-Fi devices within home networks (on average, there are currently 15.8 Wi-Fi connected devices per home, and current projections have these numbers doubling to above 36 devices per home by 2025), the strong preference for the convenience of wireless over wired by customers, the increase in IP traffic to and within the home, and the move to higher speed WAN solutions such as DOCSIS4.0, it is imperative that the industry collectively start assessing the various options for in-home devices to meet these evolving customer needs. Operators and technology partner experts who are focused on the customer experience, premise equipment, technical operations, and fulfillment operations are well aware of one of the fundamental questions around in-home device architectures; whether to deploy a single integrated Gateway box (device) or a dual box solution that separates the WAN modem from the LAN gateway. The advent of DOCSIS 4.0 has reignited that discussion, largely due to the value of locating the DOCSIS 4.0 modem near the home's demarcation for improved DOCSIS performance.

In considering the CPE options for a DOCSIS 4.0 solution, there are a number of tradeoffs that need to be contemplated. The primary consideration is the customer experience from ordering the service, to unboxing the devices(s), to installation and activation, to performance and reliability, and if there is an issue, how the customer can identify and quickly resolve the issue. Other key considerations include the cost of the device(s), as well as operational costs of managing SKUs including supply chain, Technical Operations, and care.

When talking about a 1-Box or 2-Box solution, the first consideration is to decide what functionality goes in each device. For the 1-Box solution, everything is integrated into a single device, so this is more straightforward. However, when this functionality is split across two boxes, the split of the integrated functionality across two devices as well as the connectivity between those devices must be considered. **Table 4** below shows one option for the separation of functionality; however, other options are possible depending on specific services that need to be supported as well as operational considerations.

Table 4 - DOCSIS 4.0 CPE 1-Box vs 2-Box Considerations



As noted previously, there are pros and cons to either the 1-box or 2-box approach. A single box configuration enables the Service Provider to offer a single device (gateway) to deliver the customer's connectivity and all of their IP-based services. This option makes an SIK installation process relatively simple, assuming no RF-related issues at the premises (note – a 2-box solution can also be based on an SIK model). However, it also can introduce some additional complexity around in-home/Wi-Fi connectivity. The connectivity issues are generally related to the location of the RF outlet within the premise, which is often not centrally located within the dwelling. The advancement in Wi-Fi and mesh technologies has helped alleviate the wireless connectivity challenges of the past.

Consider **Figure 20**. When considering a 2-box solution, one of the advantages, if properly executed, is that it can allow the router to be placed in a more central location in the home. This will likely provide a better in-home connectivity experience, as the Wi-Fi functionality in the gateway can be more centrally located within the home. Another advantage of the 2-Box configuration is the ability of the DOCSIS 4.0 device to be located at the demarcation, which will improve DOCSIS 4.0 performance by avoiding any additional passive losses or impairments from actives withing the home, as the in-home wiring is uncontrolled and can create installation and troubleshooting challenges. Avoiding this potential unnecessary degradation to the DOCSIS 4.0 solution can be beneficial to both operators (consistent performance, straightforward operations) and customers (higher speeds, more consistent performance, fewer technician visits). This is true from DOCSIS 3.1 as well, but the value of a demarcation installation with DOCSIS 4.0 is amplified since we are pushing the limits of the RF capabilities of the coax network.

The industry has been fortunate for some time in that the core WAN and Wi-Fi technologies have not been changing (individually or jointly) historically at the pace being observed today. The recent launch of WiFi6 (2019), the approval for Wi-Fi 6E (2020), and pending Wi-Fi 7 specification ratification suggest Wi-Fi is iterating at an accelerated pace, which is a good reason to consider a 2-Box solution. A 2-box solution allows the WAN and LAN functionality to be updated independently since this functionality is physically separated between two devices. As always, however, it comes with a cost that must be considered. These costs could manifest negatively on the operator side and the customer experience side. Some of these implications are discussed below.

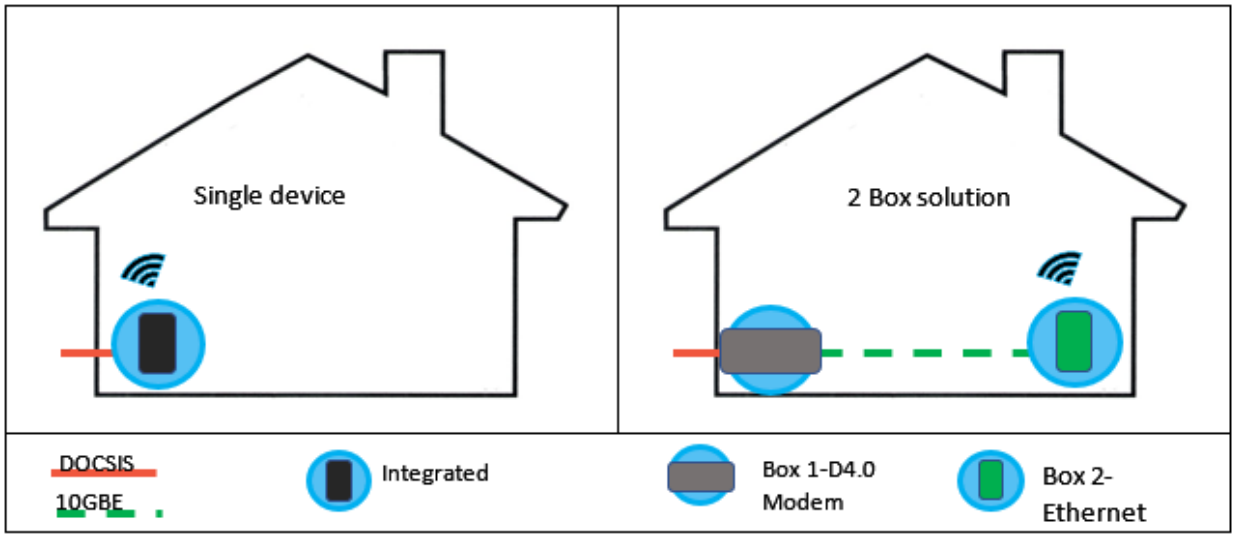


Figure 19 - DOCSIS 4.0 CPE 1-Box vs 2-Box Installation

As an example, an interesting challenge with a 2-box configuration is related to the connectivity between box 1 and box 2. As we look to moving to multi-gigabit speeds within the home we are faced with a couple of choices.

1. Leverage a 10Gbit Ethernet connection between the devices, which provides a robust connection but introduces complexity around running Ethernet cables and could compromise the ability to locate box 2 (router) with Wi-Fi in a central location within the home. This potentially undermines one of the key advantages of a 2-box solution, which is flexibility in the location of the Wi-Fi AP.

2. Consider leveraging a wireless medium (mmWave or Wi-Fi) for backhaul. Wireless backhaul introduces other challenges from cost on each device to managing Wi-Fi congestion within customer premises and guaranteeing reliable multi-Gigabit performance.

Another interesting item for consideration in the 2-box architecture, with its own "pros" and "cons," is whether the WAN (box 1) could be hardened and sit outside the customer premise or if it should be located within the dwelling. Locating the DOCSIS 4.0 eMTA outside the home has operational benefits as the Technician is now able to troubleshoot DOCSIS issues without having to enter the customer's home. However, in addition to the cost of hardening a modem to work in an outdoor environment, powering this device and how to get the LAN connected from outside the home to the Wi-Fi router in the home can be a challenge. Having the DOCSIS 4.0 eMTA outside the home also has some potential security risks because the LAN will be exposed outside the home. With all of these constraints in mind, it has been difficult to date to justify mounting the DOCSIS device outside the home.

With the increasing demands on our networks, the need to move to DOCSIS 4.0 longer-term is clear, and while this will take some time, it is important to plan how this technology will be deployed. While there are significant considerations for the DOCSIS networks as part of this transition, we must also be thoughtful about the CPE solution, as it is not likely to be cost-effective initially to deploy only DOCSIS 4.0 devices as this new technology is being launched. With that in mind, the customer impact and operational impact of the CPE decisions must be considered.

5. Planner's Guide to DOCSIS 4.0 Migration

There are reasons that North American cable operators, although generally aiming for similar service objectives, operating in similar competitive environments, and with common ecosystem technology options to choose from, have deviated in their architectural solutions and directions over time. Looking back to the mid-2000's, the popular debates of the era were many. Perhaps most prominent was around how to future-proof the downstream – the concern of the time – with increasing demand for HD channels, which consumed significantly more bandwidth per HD program (4-5x) than standard definition (SD) video. Back then, the discussions were around upgrading plant spectrum, deploying Digital Terminal Adaptors (DTAs), deploying Switched Digital Video (SDV), and transitioning to IP Video. Note these are also mostly complementary initiatives (SDV or IP Video being the exception).

More broadly, invest in FTTH now or double down on coax technologies was a hot topic. This perhaps sounds odd, given the 15+ years of continued successful coaxial strategy. But this period of time was also the launch of major Telco-based FTTH initiatives that signaled the "end" of cable services, again.

For MSOs, and again in particular large MSOs with broad and diverse geographical footprint, there are also many different starting points of HFC networks. Other initial conditions include existing portfolios of CPE devices with ranges of capabilities, critical OEM partners with varying roadmaps and unique expertise, different internal viewpoints on where the investment focus should be, and different perspective on near

term and long-term architecture. Capacity-building investment does not have simple, direct revenue tied to it the way new service opportunities do. Yet supporting capacity growth and demand is the cost of doing business as a network provider.

Furthermore, across operators and within a single operator, there are

- Different network architectures, often related to the range of per-homes passed (hhp) densities
- Different construction practices
- Zip-code, neighborhood, and property-specific demographics
- Different municipal operating environments
- Regionally varying competitive environments

These variables make it challenging for simple-to-state guidelines to easily apply and be executed. For example, a Comcast axiom previously mentioned was that the upgrade approach selected for an area must ensure at least a 5-yr lifespan before it would project to be augmented again. These types of principles are based on a combination of both network traffic and business modeling of executing various upgrade options.

The list above could surely be expanded upon, but it suffices to say, as the clichés go:

1) Cable solutions are rarely one-size-fits-all
2) Operators need a variety of "tools in the toolbox"
3) Evolution over Revolution

DOCSIS 4.0 is aligned to these well-worn principles.

5.1. DOCSIS 4.0 Begins with DAA

A major technology upgrade consensus "mandate" among operators is that the DOCSIS 4.0 roadmap will be based on DAA. Of course, this itself has tentacles that have been covered many times over in previous technical conferences, panels, the media, etc. There are different options within DAA. Comcast chose the Remote PHY (RPHY) path over 5 years ago, and has been successfully deploying DAA via RPHY in production scale for over 3 years. Some of this success can be attributed to moving the RPHY platform onto a virtualized CMTS core (vCMTS), decreasing interoperability permutations, and simplifying SW upgrades and changes through this centralized platform. With a node platform significantly more SW-based than its predecessors, this direction and migrating to agile development within the vCMTS was viewed as a critical step.

Since the standardization and deployments of RPHY, the Flexible MAC Architecture (FMA aka R-MACPHY) initiative at CableLabs has continued and matured. Some MSO's envision FMA as their DAA vehicle. There is a pro-con set of attributes to weigh between RPHY and FMA that is well-worn. For the topic of this paper, it is considered mostly orthogonal – DAA of one form or another must be in place for DOCSIS 4.0. RPHY vs R-MACPHY is not a significant consideration with respect to the DOCSIS 4.0 options. FDX or FDD can be implemented without any major dependencies.

5.2. Building a DOCSIS 4,0 Plan? Ask these Questions….

The debate over whether to move ahead with FDX or FDD is not likely to result in a crisp answer soon (if ever.) The reason is simple – there is no statistically meaningful data from which to make comparisons of the two, or to compare the projected results versus actual. There has been excellent DOCSIS 4.0 progress,

as we have discussed herein, but sample size and trial variety progress, not results of scale or statistical significance. FDX proof-of-concept field trials were executed in 2018 and 2019 using modified DOCSIS 3.1 devices with FDX EC designs from key technology partners. In addition, FDD field characterizations across multiple MSO networks to quantify the extended bandwidth behavior of various architectures and passive components of today's networks have also been completed.

A major milestone in 2021 shed some light on how the slideware is comparing to the reality using the first true DOCSIS 4.0 FDX RPD production SoC vehicle. Thus, the era of minimal information is coming to an end. The DOCSIS 4.0 RMD for FDD is now also in labs today. By the end of 2021, we will see DOCSIS4.0 FDX end-to-end modem registration, and in 2022 have a first look at FDX performance in the field from vCMTS to RPD to CPE. So.…we are on the verge of learning A LOT about the reality of both of these DOCSIS 4.0 options technically over the next 12-18 months.

Genuine technical capability and more confident, empirically-based extrapolations, based on real measurements, against the range of environments and conditions described will add significant insight in this time frame. However, it will be longer than 12-18 months to compare projected versus actual with respect to upgrade costs and operational challenges, since production at scale takes a few iterations to get right and production field teams take time to get ramped up, while incubation teams manage early technology introductions. Scale is a slowly ramping process where new technology and operational processes are concerned, as efficiency is a lower priority early on. And many fundamental components of construction, and upgrade costs, *are* well-understood. This includes items such as BAU node splits, DAA node splits, RF amplifier swaps, tap swaps, pulling fiber, upgrading coaxial cable, power supply augments, CMTS ports, spectrum addition, CPE installs, etc. Most MSOs have a good handle on these activities and mature budgeting around them. It is because of this that MSOs are able to form reasonable models, adding assumption figures or ranges to capture unknowns and sensitivity to unknowns of a DOCSIS 4.0 upgrade. And, like any new technology, DOCSIS 4.0 FDX and DOCSIS 4.0 FDD will bring their share of unknowns.

With this in mind, lets return to the attributes comparison of **Table 4** and try to reduce the broad implications of a single table of high-level attributes into a few succinct questions that can be used as a guide, as shown in **Table 5**. From there, we'll attempt to narrow these to what's really at the junction of Analysis Paralysis Avenue and Religious Belief Boulevard.

Table 5 - DOCSIS 4.0 Attributes Comparison - FAQs

Attribute	FDD/ESD	FDX
Strategy/Philosophy to 10G	• Based on access network BW extension upgrade, up to 1.8GHz, for existing actives and passives • Introduction of DAA to migrate to 10G, similar to previous HFC plant upgrades with a choice of diplex split configurations	• Based on access network technology upgrades to introduce new DSP (EC) into RPHY nodes and Amplifier platforms • Build on DAA production and scaling of vCMTS as the foundation for 10G
Key Questions	• What is the confidence level for efficient, quality spectrum to 1.8 GHz? • Is changing every active and passive in the plant non-regrettable capital?	• What is the confidence in broadband Echo Cancellation? • Is adding new EC to RF amplifiers too complex a device?

Attribute	FDD/ESD	FDX
Migration Factors	• 1.8 GHz DOCSIS 4.0 DAA Nodes, Amps, Taps and Passives • Allows for cascade of amplifiers • New CPE	• RPD Nodes with DOCSIS 4.0 EC function • FDX-capable amps with DSP • New CPE
Key Questions	• What are the projected new costs, cost premiums, and construction cost for FDD migration?	• What are the projected new costs, cost premiums, and construction cost for FDX migration?

Attribute	FDD/ESD	FDX
Complexity	• New tech challenges – BW and TCP extension • Use of "low power" amp extender for edge cases	• New tech challenges – EC function, CMTS scheduler, DSP-based amps • New capacity mgmt rule for IG/TG size for peak speed
Key Questions	• What are the practical limits of freq response and TCP for most of the targeted upgrade area? • Do all Taps and passives get swapped?	• What is the risk associated with the new technology of FDX? • What are the new algorithms for Capacity mgmt?

Attribute	FDD/ESD	FDX
Spectrum/Capacity	• 1536 MHz DS/656 MHz US (see **Figure 13** and **Table 1**) • Up to 15G/5G (all-DOCSIS 3.1) • DS/US: BW and Capacity per diplex selection	• 1110 MHz DS / 656 MHz US (see **Figure 2**) • Up to 11G/5G (simultaneous, all-DOCSIS 3.1) • FDX BW/speed by SW config
Key Questions	• What are the real BW efficiencies in the new US bands and DS bands for FDD?	• What are the real BW efficiency in the new overlapping US and DS bands for FDX?

Attribute	FDD/ESD	FDX
Operations	• Utilize existing common operational practices – FDD system with different possible split choices • New field tools	• New operational practices for handling of spectrum overlap and amplifier installations • New field tools
Key Questions	• What are the alignment and maintenance implications for support downstream bandwidth to 1.8 GHz?	• What are the setup and maintenance implications to supporting FDX actives and overlapping spectrum on the plant?

Attribute	FDD/ESD	FDX
Network	• N+X • Cascade reduction/trade-off based market capabilities	• N+0 (optimal) • N+X – Cascade reduction/trade-off based on market speeds
Key Questions	• What are the realistic N+x limitations for 1.8 GHz? • For different generations of plant design? • How does that translate to Capacity Q above?	• What are the realistic N+x limitation for FDX? • For what service speeds and penetration? • How does that translate to Capacity Q above?

Attribute	FDD/ESD	FDX
As-Built Migration	• Continue node split and introduce DAA node splits, leverage for HFC migration activity, introducing components of FDD over time • Migration path and timing considerations for Underground vs Aerial and MDU vs SDU cost implications	• Introduce DAA for node splits with vCMTS, leverage for HFC migration activity and platforms that enable FDX activation • Migration path and timing considerations for Underground vs Aerial and MDU vs SDU cost implications
Key Questions	• How are actives and passives upgraded over time? • How much additional cost is associated with N+x caps, new cabling, or remediating poor frequency response	• How are amplifiers upgraded to FDX? • How much additional cost is associated with Amplifier DSP or N+X caps?

Table 4 and **Table 5** cover many variables, and yet, as anyone who has engaged with field teams and network construction personnel can attest, beneath each of these are additional layers of detail owing to the aforementioned variability of architecture and plant in an MSO network. DOCSIS 4.0 implementation will at least establish an HFC demarcation at the DOCSIS 4.0 CPE device and eliminate all or most of the home coaxial network variability, which today is technically, unfortunately, part of the HFC network.

If we were to consolidate the 20,000 ft. list of **Table 5** list into what "really really" are DOCSIS 4.0 FDX and FDD decisions hinging on – what consumes the majority of the dialogue when drawing up the internal pro-con table for the mighty offsite whiteboard sessions (in a nod to the impact of Covid – this actually sounds attractive!)– it might look like **Table 6**.

Table 6 - Debate Kindling Top 3

Challenges	FDD/ESD	FDX
1	Total Composite Power limitations for 1.8 GHz	N+0 foundation implications for cost
2	As-built freq response over N+x	New technology risk
3	Upgrade and replacement of all taps and passives	FDX amp and N+x operation

6. Summary

The 10G network is perhaps the most recognizable industry-wide initiative today. Its vision has been organized around four key pillars of service – capacity/speed, latency, reliability, and security. For the access network, it is the capacity and speed objective, in particular symmetrical multi-gigabit capability that represents the most directly addressable. It is a shift from existing BAU network migration strategies, because of its dependence on physical network changes over and above node splits, and because of the massive service payoff in the form of significantly more capable HSD services.

From the first established at the Consumer Electronics Show (CES) in 2019, the details of the visions gave way to the development of the technical requirements to achieve it, which is DOCSIS 4.0. The FDX specification actually began its life as an Appendix to the DOCSIS 3.1 specification aimed at optimization capability for N+0 systems. The DOCSIS 4.0 specifications are now completed and released for both FDX and FDD. For "FDX, the Appendix," was a "lift and shift" operation into DOCSIS 4.0. The FDD specification was then completed, with the latest release that includes both (I03) publishing in December 2020.

In this paper, we tried to objectively, but (of course) through individual company lenses, articulate common and differentiating characteristics of FDD and FDX, compare the implications of the most important characteristics beyond the slideware and into real upgrade consequences, and bring to the external whiteboard some of the discussion points that have been occurring in internal and cross-MSO network brainstorming sessions. Hopefully this paper has provided a peek into these dialogues.

Lastly, while some of these activities are clearly multi-year endeavors and apt to adapt with learnings over time, there is always a need for a "North Star" target. Among most (but not all) cable operators, DOCSIS 4.0 is this North Star. There is an early fork of implementation paths, but many MSOs have assessed which trail makes sense for them at the outset of their DOCSIS 4.0 journey. Breadcrumbs recommended!

Abbreviations

A-TDMA	Advanced Time-Division Multiple Access
BAU	bits per second
CAGR	forward error correction
CES	high definition
CM	hertz
CMTS	International Society of Broadband Experts
DAA	kelvin
DS	Society of Cable Telecommunications Engineers
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DSP	Digital Signal Processing
EC	Echo Cancellation
ESD	Extended Spectrum DOCSIS
EOL	End-of-Life
FDD	Frequency Domain Duplex (aka DOCSIS 4.0 ESD)
FDX	DOCSIS 4.0 Full Duplex
FDX RPD	FDX Remothe PHY Device
FDX-L	FDX-Light
FEC	Forward Error Correction
FTTH	Fiber-to-the-Home
HSD	High Speed Data
IG	Interference Group
LDPC	Low Density Parity Check Code
LLD	Low Latency DOCSIS
MDU	Multi-Dwelling Unit
MER	Modulation Error Rate
MMP	Multiple Modulation Profiles
OFDMA	Orthogonal Frequency Division Multiple Access
OEM	Original Equipment Manufacturer
OFDM	Orthogonal Frequency Division Multiplexing
pbh	Peak Busy Hour
PMA	Profile Management Application
PON	Passive Optical Network
RBA	Resource Block Assignment
SDV	Switched Digital Video
SoC	System-on-a-Chip
STB	Settop Box
TCP	Total Composite Power
TG	Transmission Group
UHS	Ultra High Split
US	Upstream
xDSL	[any variant of] Digital Subscriber Line

Bibliography & References

[1] Barker, Bruce E, and Claude Bou Abboud, Erik Neeld, "Access Capacity Planning: Staying Well Ahead of Customer Demand Helped Ensure Stability During COVID-19," SCTE Cable-Tec Expo, Oct 13-16, 2020.

[2] Baumgartner, Jeff, "Comcast Full Duplex DOCSIS trial pumps out 4-Gig symmetrical speeds," LightReading, 4/2/2021.

[3] Howald, Dr. Robert, The Fiber Frontier, 2016 INTX Spring Technical Forum, Boston, MA, May 16-18.

[4] Howald, Dr. Robert, Repair the Ides of March: COVID-19 Induced Adaption of Access Network Strategies, 2020 SCTE Expo, Oct 13-16.

[5] Howald, Dr. Robert, Roaring into the 20's with 10G, 2020 SCTE Expo, Oct 13-16.

[6] CableLabs DOCSIS 4.0 PHY Specification: CM-SP-PHYv4.0-I03-201202

Distributed Access Architecture is Now Widely Distributed – And Delivering on its Promise

A Technical Paper prepared for SCTE by

Dr. Robert Howald
Fellow
Comcast
1800 Arch Street, Philadelphia PA
robert_howald@comcast.com

Frank Eichenlaub
Executive Director, Access Network Technology
Comcast
frank_eichenlaub@comcast.com

Tobias Peck
Director of Product Management
EnerSys
tobias.peck@enersys.com

Adi Bonen
Executive Director
Comcast
frank_eichenlaub@comcast.com

1. Introduction

A large-scale production Distributed Access Architecture (DAA) footprint in terms of both homes connected and the number of digital nodes is being built by Comcast right now. Comcast began deploying DAA in 2018, focusing on the mature, multi-vendor Remote-PHY (RPHY) standard. The company continues to aggressively deploy DAA across all regions. A core premise of RPHY is that because it is a CableLabs standard, devices from different manufacturers of CMTS and Digital node are interoperable. Comcast has succeeded in delivering on this promise with Digital Nodes from three technology partners connected to its virtual CMTS (vCMTS) platform.

Many of today's DAA deployments are in areas with an N+0 (node plus zero amplifiers) architecture, which is sometimes referred to as "Fiber Deep." There is also an existing footprint of N+0 based on traditional analog fiber nodes. In both cases, there are no actives between the fiber node and customer homes. This creates the perfect opportunity to make A/B comparisons between the best-in-class analog HFC and DAA technology. Without RF amplifiers "watering down" any differences between the two networks, such as might exist in N+x deployments, we can attain a clear and unvarnished, side-by-side view of yesterday's and today's cable access technologies in production scale.

In this paper we will compare observations and numerical results obtained from operating these two N+0 variants. We will:

1) Describe lessons learned building both architectures, and adjustments made along the journey
2) Describe the transitional operational challenges and best practices
3) Quantify key performance metrics
    a. EOL MER (End of Line Modulation Error Ratio)
    b. Capacity
    c. Network power consumption – a critical area of focus for 1.2GHz DAA over N+0 in particular

This paper will address the journey, results, and future expectations of the industry's largest known initiative to take DAA from concept to reality..

2. Why DAA?

A Distributed Access Architecture (DAA) breaks the CMTS function into two components and distributes one part of it into the HFC network, typically the fiber node. The functionality distributed has narrowed to two DAA options. These are Remote PHY (or RPHY) and Remote MAC-PHY (known interchangeably as "RMD" for Remote MAC Device, or FMA = Flexible MAC Architecture). Figure 1 depicts the variations. Also shown is the fully Remote CCAP (R-CCAP) approach, which has faded from consideration among most cable operators due to the complexity.

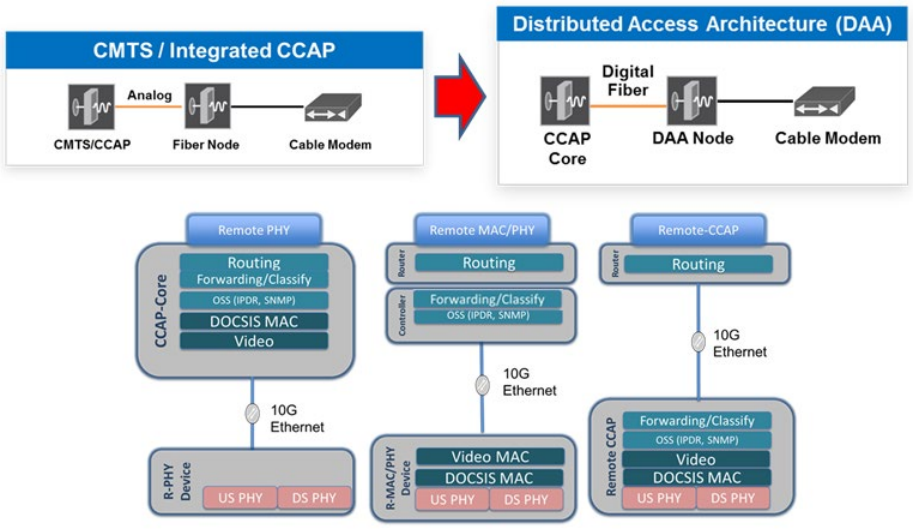


Figure 1 - Distributed Access Architecture Variants: RPHY, RMD, and R-CCAP

Remote PHY is currently the dominant deployed architecture, with tens of thousands of RPHY DAA nodes deployed in production networks in the US market alone, while only a relative handful of RMD devices have been installed. The FMA specification governing RMD interoperability got off to a later start, and this architecture takes on additional HW and SW complexity at the node. These two factors and forward-looking capacity advantages are contributing to the slower uptake of this technology. Whether this technology is brought to market in volumes will be determined over the next few years as more MSOs embrace DAA in their next generation architecture plans.

2.1. Key Benefits

Regardless of the amount of DOCSIS processing functionality distributed, DAA delivers powerful advantages associated with the standard digital Ethernet optics that are fundamental to the architecture. Until DAA came along, every node was implemented with Amplitude Modulated (AM) optics, created by, and unique to, the cable industry. Until that time, fiber optics had been based on binary digital transport only. A DAA node, fed by digital optics, offers significant network, infrastructure, and performance benefits over traditional AM-based optical links.

These chief advantages are:

1) *DWDM Wavelength Efficiency*

The use of digital optics means that there is a path to 80 DWDM wavelengths/nodes on a single fiber. Typical HFC AM optics are restricted to 16 wavelengths, depending on reach, to manage nonlinear effects that degrade MER. More than 16 are possible but they come with compromises in performance and distance. These nonlinear effects are much less significant in digital optics.

2) *Link Reach*

Using AM optics, as the link from Headend to node increases, fewer wavelengths can be used to meet a fixed end-of-line (EOL) MER requirement. HFC networks are constructed around these limitations. Digital optics eliminate this dependency for typical HFC reach requirements and enable link distances much longer than AM optics. The above aspects of reach and wavelength multiplication taken together simplify

and add significant flexibility to network architecture, promote cost savings due to the reduced fiber construction entailed, and enable the potential for consolidation of physical sites.

3) *EOL MER performance improvement*

DOCSIS 3.1 enables higher order modulation profiles, increasing the bandwidth efficiency up to 50% over DOCSIS 3.0 downstream and 67% in the upstream. Typically, EOL MER in HFC is dominated by the performance of the AM optics and degrades a little with each ensuing amplifier. In an Integrated CMTS, the fidelity requirement at the RF port feeding the optical transmitter is very high – 43 dB minimum for DOCSIS 3.0 and up to 48 dB for DOCSIS 3.1. The AM optical link then degrades the MER delivered to the node. It is a common node requirement to achieve a 38 dB minimum MER at the RF output port. This represents a significant fidelity loss of 5-10 dB. The MER degrades more as it is passed through HFC amplifiers.

Implementing DAA eliminates this AM Optical MER degradation situation completely. Instead, the CMTS fidelity requirement is met at the RPHY device (RPD) RF interface inside of the node. This increases the margin available to deliver DOCSIS 3.0 capacity, while increasing the capacity available for DOCSIS 3.1.

A similar situation exists for the upstream. Elimination of HFC upstream technology, whether an analog or digital return, occurs by placing the DOCSIS US Receiver in the node within the RPD, which recovers significant (67%) upstream MER.

4) *Space, power, cooling efficiencies in the Hubs and Headends*

Moving CMTS functionality to the node leaves less Headend components to power, cool, and consume space in a Hub site. Some of this power is distributed into the plant, which, as we shall see, quantifiably can provide a net positive return to the outside plant (OSP) due to the overall technology upgrade over the equipment it is replacing.

With respect to space, with the RF ports of the CMTS distributed, the connector density of the DOCSIS processing core can now be bounded by a higher density of optical connectors.

As traffic continues to grow, and new ports and computational power are necessary to support new demand, the incremental increase in footprint, power, and subsequent cooling needed to deliver the necessary processing power can be more granular, in particular as DAA leads to a virtualized processing core.

5) *Alignment with virtualization and convergence across last-mile access technologies*

As the HFC architecture evolves to DAA, nodes are more readily adaptable to other last mile access technologies that leverage Ethernet connectivity. Wireless, PON, and direct Ethernet services are all applicable, enabling convergence opportunity in the processing core.

Figure 2 shows the simple example of a Passive Optical Network (PON) Remote Optical Line Terminal (R-OLT) module. The concept is similar in principle to a plug-in RPD module, distributed to and backhauled from the same Ethernet-connected DAA architecture to deliver last mile residential and business services, but over an FTTH network.

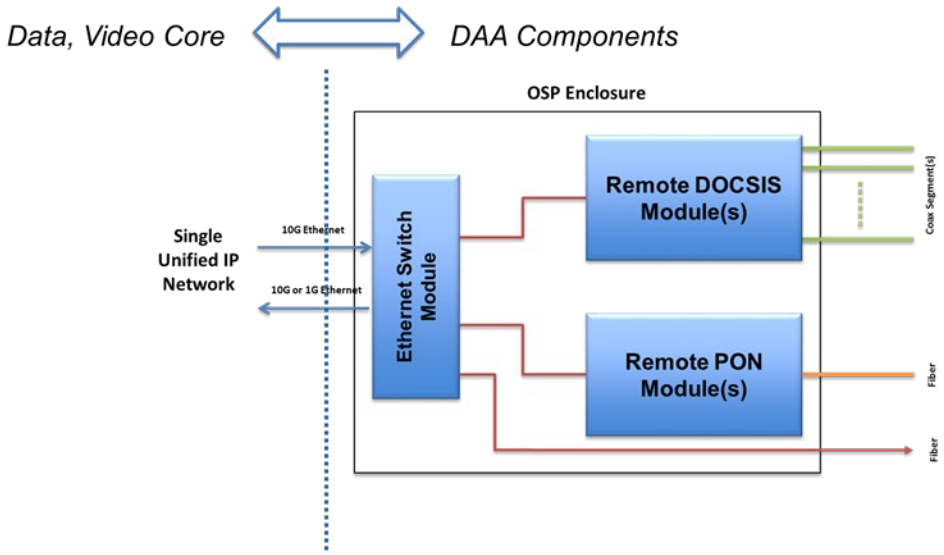


Figure 2 - DAA Enables Access-Agnostic IP Network Architecture Convergence

2.2. DAA as a Path to Virtualization

DAA can be considered the starting point of a larger shift in the HFC technology ecosystem away from purpose-built cable hardware where possible, and embracing the broader trends towards digital processing, software, and cloud-based services and applications. With the CMTS packet processing function in the headend separated from major DOCSIS-specific functions now implemented in the RPD, the CMTS can be revisited in the context of packet processing, switching, storage, and its DOCSIS scheduling role, all of which can be translated into virtualized real-time computational resources. With today's compute power, and the capability of software to deliver real-time services, a purpose-built DOCSIS machine is no longer required.

Instead, commercial off-the-shelf (COTS) servers and switches, combined with software running the required DOCSIS MULPI layer, can be used to implement the CMTS function. This leads to a simplified and more flexible virtualized CMTS platform, or vCMTS. Significant benefits can be leveraged over time through Moore's Law, delivering ever more compute power in increasingly dense physically dense footprints. Comcast introduced its first production DAA deployment hosted by a vCMTS in 2018, and since that time every DAA node has been connected to the Internet via a vCMTS, as the architecture has rapidly scaled.

Figure 3 conceptualizes DAA and its alignment to virtualization and vCMTS. As noted in the figure and pertaining to the idea of an access-agnostic last mile enabled by DAA, virtualization can apply to DOCSIS or PON / FTTH last mile, or other natural alternatives such as wireless. The virtual platform itself is not access-specific when it comes to the switching and routing of packets, simplifying the path to convergence of access technology at the packet processing layer when the access technology specific-SW is abstracted.

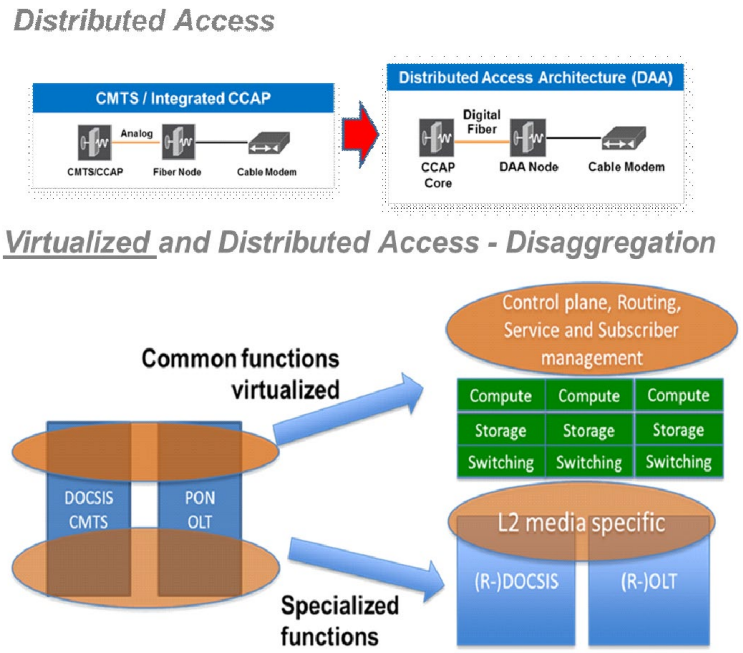


Figure 3 - DAA and Alignment to a Path Towards Virtualization and vCMTS

3. DAA-Powered Next Generation Access

3.1. Architecture Migration

Like most MSOs, Comcast is constantly assessing its access network upgrade strategy, adapting accordingly to changes in services, customer demands, traffic, and business priorities. The environment has never been more dynamic and perhaps – as the pandemic-related impacts blend into a new normal – never less certain. And, like all MSOs, while many network tasks are well-exercised muscles that are developed over years of experience, upgrades effecting the access network are generally not quick to implement nor simple tasks. Adding to the degree of difficulty is the need to make these upgrades and transitions as seamless as possible to customers.

Given that access network upgrades are multi-year initiatives, there is a need to prioritize accordingly. There is also a need to recognize that the pace of upgrades plays a role in applicable "tools in the toolbox." Considering the relatively long-term nature of an upgrade cycle, it is important there are not "starving" areas of the network left unattended until the upgrade comes along. No part of the network can afford to stand still. While it varies across the footprint, the pace of traffic growth and service demand outpaces upgrade pace, forcing repeated touches. This limitation of physical hardware and construction is also one of the drivers for introducing more software into the network through DAA and virtualization, which we will discuss shortly. Lastly, it is important that there be simplicity and repeatability in the operation to make these upgrades seamless, efficient, and delivering the benefits expected.

With the above dynamics in mind, a multi-pronged strategy was developed to make sure network upgrades are in-tune with growth and service requirements, as they vary by region, and the projected timing and scale at which the network must be upgraded to do so. **Figure 4** outlines a strategy based on these objectives.

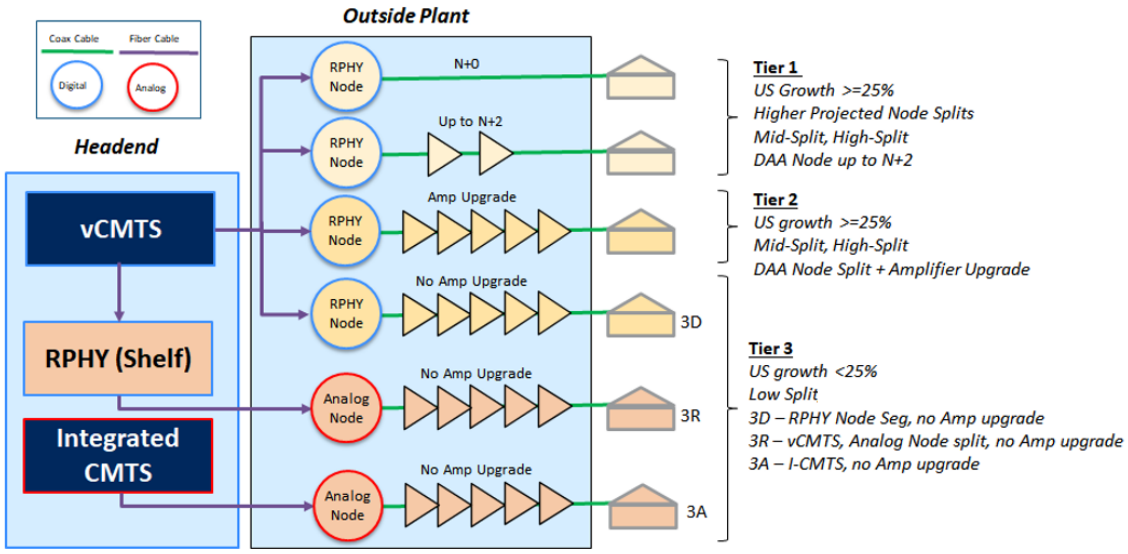


Figure 4 - Breakdown of Upgrade Approaches by Tier, Criteria, and Technology

In **Figure 4** we can observe that, across the roughly 60M homes and businesses passed by the Comcast network, the variations have been categorized into three Tiers of defined activity. Within the three Tiers, five migration options depend on various criteria and existing conditions. These are summarized as follows:

Tier 1 – The very highest HSD utilization and highest growth areas require the most aggressive change to network to stay ahead of demand – pulling new fiber deeper into the network, maximizing the downstream available bandwidth out to 1.2GHz, and decreasing the node size to a maximum allowable homes past while building more headroom into the capacity runway, deferring subsequent upgrades.

Tier 2 – High utilization in % growth, although not as high overall as the Tier 1 category. This approach defers subsequent upgrades for at least five years. For this category of network, this deferment can be accomplished with less aggressive means and can be done more quickly and more broadly by continuing to use node splits, but in doing so convert the network to vCMTS and DAA, while adding spectrum downstream and upstream via RF amplifier upgrades.

Tier 3 – Lower upstream growth and utilization than Tier 1 and Tier 2. Splitting of nodes "BAU" is sufficient to stay ahead of demand. However, while doing this work, begin introduction of key new enabling technologies of Tier 1 and Tier 2 (DAA and vCMTS) to place the foundation for a subsequent technology upgrade wave via Tier 1 or Tier 2 when necessary.

3.1. Spectrum Migration

As mentioned above, while migrating the access network to DAA, taking the opportunity to adjust the spectrum allocation is a practical and efficient way to increase long-term capacity. In addition, in order to provide higher upstream speeds and service tiers, a wider upstream path is required to deliver the instantaneous capacity necessary. With the commitment to install new equipment into the access network, it is the ideal opportunity to ensure that this equipment has a long runway of traffic capacity.

Figure 5 depicts stages or options available within the DOCSIS standards by which the upstream and downstream can be defined. It also indicates the HSD speeds available with these implementations for residential services under typical serving group sizes and capacity utilizations. Cable operators have been living in the "Low Split" (5 MHz – 42 MHz)_world for decades and have now begun moving to these higher upstream spectrum options.

While the Mid-Split architecture (upstream extended to 85 MHz) will cover an extended capacity runway for typical node sizes, its maximum speed, when carrying DOCSIS 3.0 traffic in the Low Split band, will be limited to about 300 Mbps. This increases to about 500 Mbps with an all-OFDMA upstream in the years ahead as DOCSIS 3.1 penetration continues to grow and eventually dominates the available bandwidth.

Both High Split (with an upper boundary of 204 MHz) and DOCSIS 4.0 Full Duplex (10G) provide the opportunity to offer, respectively, Gigabit and Multi-Gigabit symmetric services. Not shown below is a DOCSIS 4.0 Extended Spectrum DOCSIS (FDD) spectrum allocation, which, like FDX, can take the upstream as high as 684 MHz. However, as a Frequency Domain Duplex (FDD) architecture, it requires replacing all actives and passives in the plant, to extend the network to 1.8 GHz. That's necessary to provide sufficient downstream spectrum to offset the loss due to the added upstream allocation, plus the increased diplex crossover guardband.

By contrast, FDX uses common spectrum for both downstream and upstream to maximize efficiency in this known, mature, well-behaved RF spectrum.

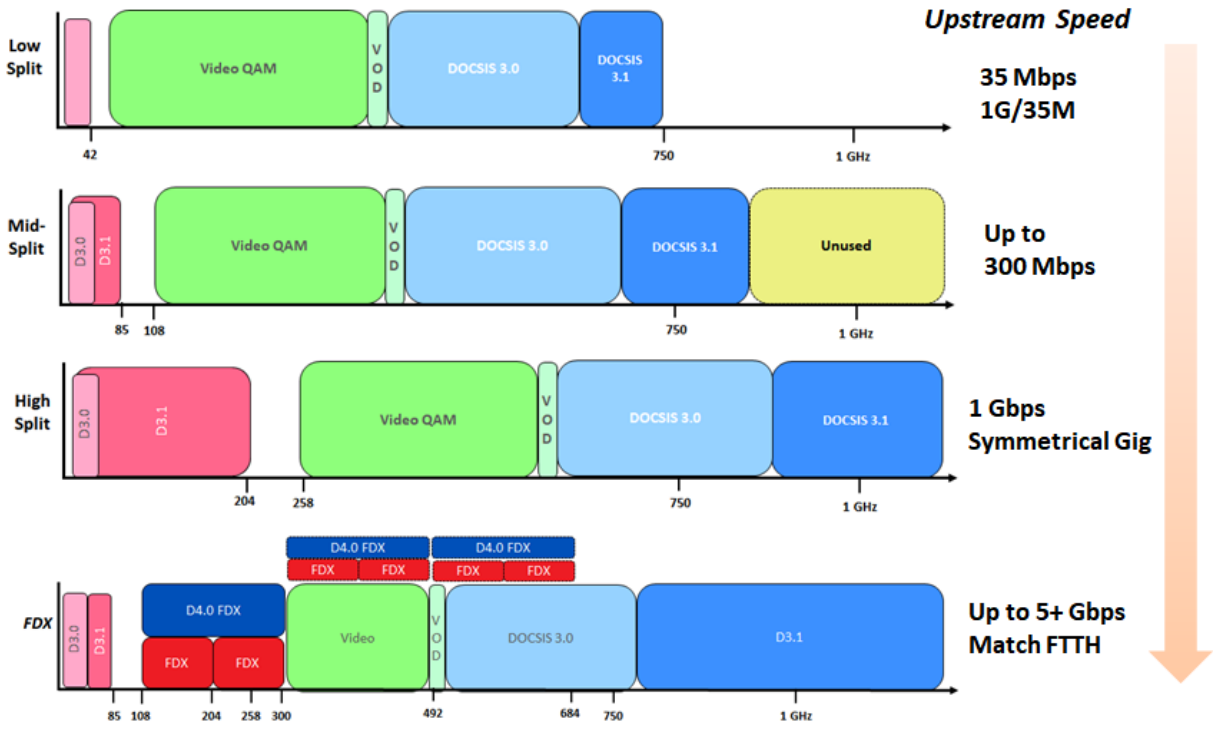


Figure 5 - Migration Options for HFC Spectrum Allocation

4. DAA Node Design Principles

4.1. Leveraging a Connected Platform

A DAA node has many of the expected essential features of an HFC node on the coaxial side of the device. It must place high fidelity RF signals onto the coax to reach the customers connected to its port, and it must receive RF upstream signals from those same customers. However, a "BAU" HFC node is a "dumb" transducer device, converting an analog optical waveform on fiber to an RF waveform on coax, and the opposite in the return path. Digital return upstream systems are a nuanced exception – they translate the RF waveform, but do so by digitizing it into a series of numbers and transmitting those numbers digitally before re-creating the waveform at the end of the optical link.

By contrast, a DAA node's RPD is a smart and IP-connected device. The DAA node design can take advantage of this connectivity. To start with, common node functions can be accommodated without the RF plug-ins that are often used in traditional HFC nodes. These plug-ins can tempt technicians, resulting in excess tweaking flexibility that can be misused without careful guidance. This is particularly the case for attenuation pads and tilt equalizers. Furthermore, local control using plug-ins results in loss of traceability, and unknown equipment settings. Instead, a DAA node can implement remote configuration of gain and tilt, enabling centralized power profile management, including outage-free spectrum enhancement and lineup changes. This provides state visibility and tracking, and can be enabled (or not) to technicians at the expertise depth of the control desired.

In the RF chain leading to and from the node ports, a DAA node that provides per node-port US attenuation can be controlled to identify a port experiencing an abnormally large ingress, minimizing triage and recovery. Moreover, if the RF launch amplifier includes an US RF switch matrix, mitigation schemes can be used to minimize the group of subscribers experiencing service reduction. For example, in a 4-port node segmented to 2 US service groups, the node port suffering ingress can be segregated to be on its own US service group while the other 3 node ports are combined to share the other US service group without interference from the ingress. This preserves the highest possible level of service for the most customers until the issue can be resolved.

In terms of inventory management, a "smart" DAA node can be designed to enable remote module inventory by reporting the model and serial numbers of every pluggable module in the node as well as other factory parameters. In addition, status monitoring of various module parameters enables tracking the component and node health.

Lastly, and perhaps the most powerful benefit of a DAA architecture and as previously described, is introducing the ability to easily alternative last-mile access technologies, such as PON, Ethernet services, or wireless. DAA can go beyond just DOCSIS-based RPHY and RMD nodes. Devices implementing an outdoor switch (aka "Switch On A Pole" or SOAP) functionality can provide OSP IP-link aggregation for other OSP devices, as well as access points for Metro Ethernet subscriber lines. A SOAP device can share the same DAA node with an RPD or be independently housed in an RF-less node housing. In both cases, the enclosure for the SOAP device should have similar traits to a "regular" DAA node.

Moreover, as MSOs deploy more FTTH PON to their customers, a Remote-OLT device housed in an OSP DAA node is becoming many MSOs' preferred approach. It provides both success-based deployment granularity, and serves as a more localized intermediate point from which to launch fiber-budget limited PON signals that may not be able to be passively delivered from the site where a typical chassis-based OLT would reside for a cable network. Similar to the Ethernet case, whether the R-OLT device shares the same DAA node with an RPD or is independently housed in an RF-less node housing, the DAA node should utilize the same key principles of resilient powering, remote monitoring, and control.

In short, the smart and connected nature of the DAA node, in particular the use of standard Ethernet optics for transport and virtualization, opens up a wealth of new architecture flexibility and operational benefits to DAA nodes that cannot be achieved in a typical HFC node.

4.2. New Powering Challenges

DAA nodes can drive an already existing N+x HFC system, be part of the more aggressive "Fiber Deep" migration that removes RF amplifiers (N+0) and extends downstream and upstream spectrum, and of course can also support anything between those two boundaries of plant state. The most challenging of these cases is the DAA node that powers the N+0 system, because these systems rely on higher RF output power to maximize reach, have the highest tilt to cover the longer length of coaxial cable that accompanies this reach, extends the downstream spectrum to 1.2 GHz, driving the Total Composite Power (TCP) higher than any previous node. It subsequently draws the most AC power to achieve these goals. Therefore, if a DAA node can support fiber deep, it will support all possible N+x variants it is likely to be deployed in with different RF configurations aligned to those architectures.

4.2.1. Node Power Design Considerations

The upgrade of a traditional N+X HFC system to a N+0 DAA requires paying even closer attention to plant powering. The DAA nodes that effectively replace (not 1:1) existing amplifiers are likely to consume significantly more power, and care should be taken not to surpass the line power supply's rating or the hardline current carrying limit. DAA nodes can implement several key features to help mitigate such issues as part of an upgrade.

In a typical older N+X system, power is passed from a line power supply to a node, and through it to multiple amplifiers. A node is rarely fed power which first passes through another node. If, as a result of bringing down a node for maintenance that power to the amplifiers is also lost, no additional subscribers are affected. However, in N+0 systems with smaller node domains, a single line power supply often feeds multiple nodes, often passing through one node to reach another. In such cases, it is very desirable to maintain uninterrupted power passing through a node even when it is undergoing maintenance (such as a module replacement), such that a service interruption due to maintenance is not extended to other node domains that are not undergoing maintenance.

Note that while the customer-facing blast radius may be similar from a numerical perspective because of the smaller number of customers per N+0 node, DAA nodes are less tolerant to short outages, due to the possibility of a reboot being triggered, than their traditional HFC node or amplifier counterparts. In addition, node outages create new software challenges in the back office for state management and recovery that could be otherwise avoided with thoughtful OSP design practices. Accordingly, it is desirable for fiber deep nodes to support power passing even while their internal modules are swapped.

4.2.2. Node Power Efficiencies

In a node's RF sections, most of the power is consumed by the output hybrid amplifiers. However, due to the insertion loss of a 4-way split as well as gain and tilt temperature compensation circuitry normally built into such an amplifier, it typically requires significant power to be allocated to drive these output hybrids. A DAA node with a smart RPD can significantly reduce this power requirement by implementing gain and tilt control in the RPD itself. Similarly, in the US, including configurable step attenuators prior to the return amplifiers enables the RPD to keeps tight control on the signal level at these amplifiers, thus reducing the dynamic range requirement for them, allowing them consume significantly less power. In an N+0 node, about 15~20 watt can be saved with such schemes, which is significant savings for a device that in the 100-140W range, depending on configuration.

Power delivery of both line power supplies and hardline cable is limited by delivered current, not delivered power. Implementing power factor correction (PFC) in the DAA node, which aims to align the phase of the AC voltage and current waveforms, reduces the quasi-square-wave current at the given power it consumes. Therefore, a DAA node implementing power factor correction enables more power to be delivered by the existing system. Moreover, by implementing power factor correction to reduce the quasi-square-wave current, additional efficiency is achieved by reducing the power loss on the hardline itself due to its loop resistance.

4.2.3. Practical Considerations and Solutions

One of the lessons learned in early DAA deployments is the importance of immunity to short power interruptions, as mentioned previously. The HFC plant is prone to split-second power interruptions, with most happening as a result of tap faceplate removals. In a traditional node such an interruption likely results in a split-second signal loss to customers, an event which is hardly noticeable. However, in a DAA node this can cause the RPD to reboot, extending the signal loss to several minutes. A solution we Comcast utilizes a capacitor-based scheme which can maintain power to the RPD for a few seconds during power interruption. Unlike battery backup, capacitor-based backup is maintenance-free and its lifespan surpasses the node's.

Another lesson learned is the need for a temporary node powering option *during* installation, burn-in and provisioning. Upgrading a traditional HFC plant to a DAA system typically involves installing a new DAA node in parallel to an active old node, burning in the new DAA node, provisioning it, and only then cutting it in instead of the old node. Often, there is not enough available power in the active system to power the DAA node on top of the existing system, and thus typically a generator is used to power the DAA node during these stages. However, it is desirable that the node be independently ready for plant power in advance, without causing interruption when moving from temporary generator power to permanent plant power. A solution we used involves a special temporary-powering power pack (PP) that is mounted in the second PP position in the node. Examples of the temporary power packs are shown in **Figure 6**.

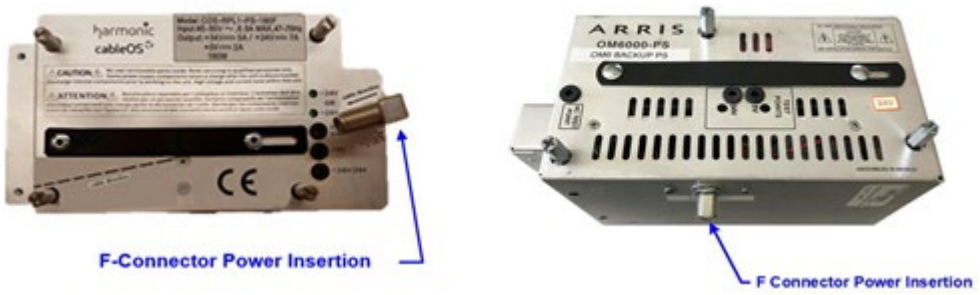


Figure 6 - Temporary Power Packs for DAA Nodes Simplify Cutover

The temporary PP drives the node's internal DC power rails as any PP, but gets its QSW power through a directly connected coax from an external generator, and not through the plant's and node's permanent power. During cutover, when the primary PP first gets its power from the plant, a node is operated briefly with redundant power (permanent and generator), and thereafter the temporary PP is removed (to be reused in the next node to be installed). **Figure 7** illustrates this temporary installation arrangement.

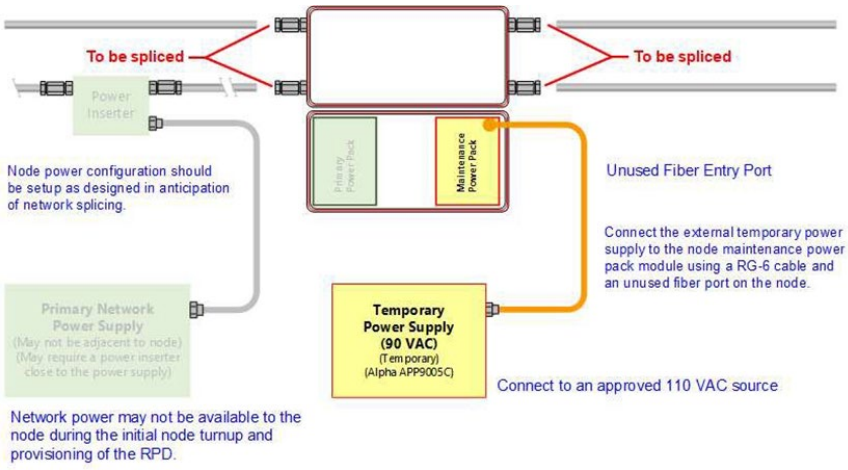


Figure 7 - Deployment Model of Maintenance Power Supply for DAA Node

5. Network Powering Implications for DAA

In a typical Comcast N+X deployment, it is common to have three power supplies providing AC to the network actives. Management of the distribution of power is provided by blocking power at strategic points to best utilize the available power. In an N+0 architecture there are three primary methods of distributing power to the nodes and other active devices (non-amplifiers) attached to the plant:

1. Centralized power using parallel feeds of 0.875" or 23-ohm power hardline cable to distribute power independently of the RF distribution.
2. Utilize power straps to connect node service areas to existing power supply locations.
3. Deploy independent, low wattage power supplies for each node service area.

There are advantages and disadvantages to each power architecture. To minimize the amount of parallel cable placement, permitting and construction activities, power straps were used to distribute power in the Comcast N+0 network. Simply put, a power strap is a power-only connection using 0.875" cable or 23-ohm power cable to route power from node area to node area. **Figure 8** is a simplified drawing of a power strap.

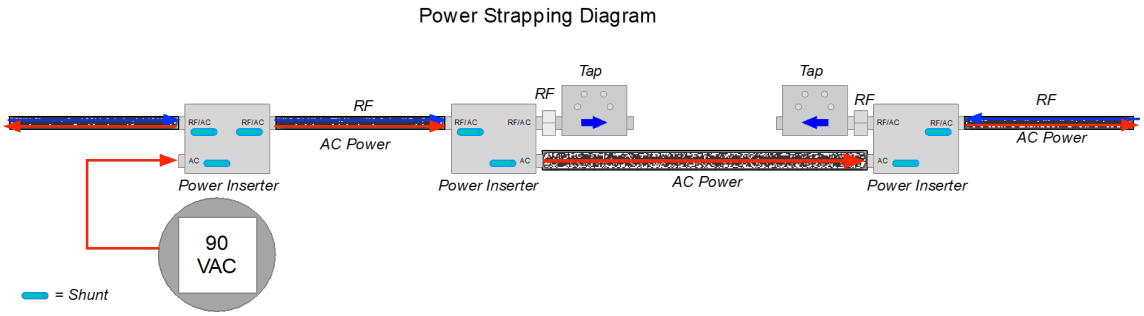


Figure 8 - Use of Power Straps in HFC Networks

Power straps provide a low-cost solution to manage power in a N+0 architecture but introduce a more complex network for maintenance teams to manage, troubleshoot and maintain. In addition to the potential for power glitches previously described, the common practices of pulling fuses or shunts to isolate plant issues now can shut down many nodes at once. Because of this, the power network design for DAA requires some additional levels of resilience to ensure a more robust experience.

5.1. Example Power Interruptions Example Power Interruptions

Typical holdup time for the AC power packs in DAA nodes extends from about 75ms to 250ms. The holdup time is defined as how long the power supply can maintain the DC voltage in the node after the input AC is interrupted. Loss of DC to the RPD initiates a reboot that can take many minutes to re-establish service to devices in the serving area. **Figure 9** is a capture of an oscilloscope displaying the impact of loss of input AC waveform and the ability of the node power supply to hold up the DC voltage.

Figure 9 - AC Interruption to Node Power Supply

Several common maintenance practices can also interrupt the AC waveform. Replacement of a tap faceplate or shorting a nut driver while tightening a seizure screw can occur during maintenance. **Figure10** illustrates examples where a tap faceplate removal or installation can impact the AC waveform enough to cause an RPD to reboot.

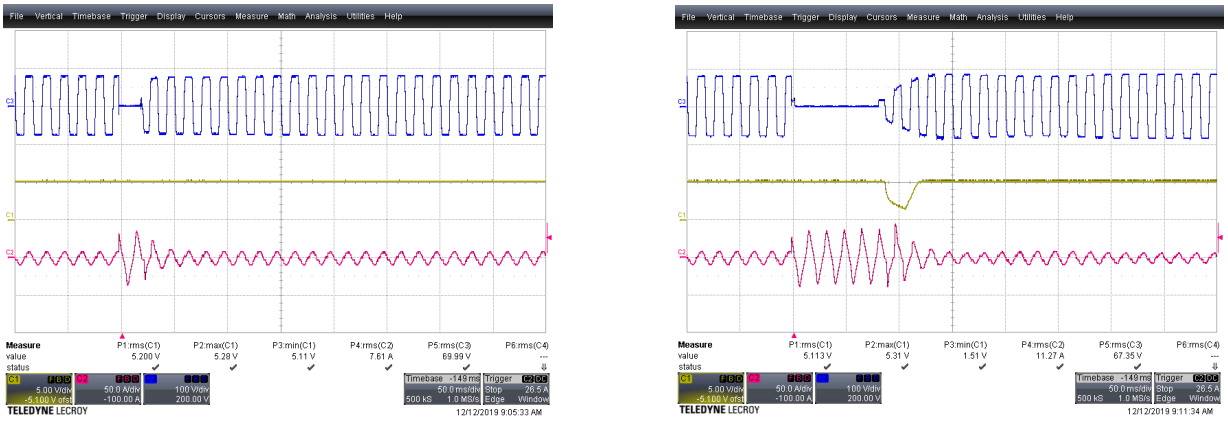


Figure 10 - AC Interruption to Node DC Power Supply Caused by Tap Plate Removal

In the example on the left, the faceplate was removed with little interruption to the AC waveform and no RPD reboot occurred. The example on the right demonstrates where a tap faceplate change took longer,

and the DC voltage was not able to be maintained to the RPD. In this case, an RPD reboot would be initiated.

AC interruptions impacted the nodes downstream of where the maintenance was being performed. Shorts, however, impact all nodes being fed from a common network power supply. **Figure 11** is an example of a technician accidentally shorting a nutdriver while tightening a seizure screw. As shown, there are several interruptions to the AC waveform eventually leading to an RPD reboot.

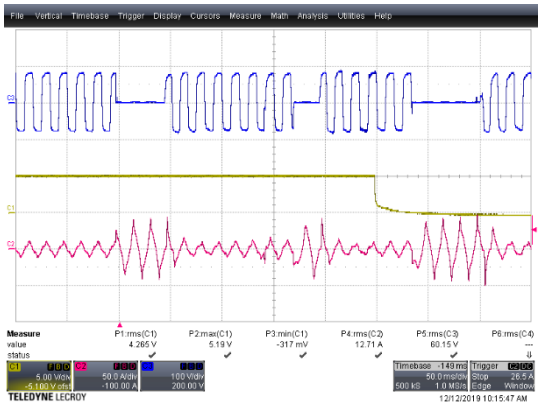


Figure 11 - AC Interruption to Node DC Power Supply Caused by Short at Seizure Screw

5.2. Design and Processes for DAA Power Resilience

To reduce the impact of normal maintenance procedures on the DAA nodes and improve the customer experience, several elements of the RPD deployments were focused on to improve the resiliency of the power network.

The impact on power from network passives and taps was reviewed. Currently available network passives are designed such that when the faceplate is removed, RF and Power continuity is interrupted. This is an undesirable characteristic in DAA deployments. Comcast is working with vendors to redesign the network passives to allow for cover removal and troubleshooting without impacting RF or Power to all ports.

In addition, taps were evaluated to measure their make-before-break characteristics. These characteristics vary by manufacturer and specific tap design. Processes were established to ensure that taps have no more than an 8 ms impact on the AC waveform when removing or installing a faceplate.

From a network powering point of view, some power supply models can be deployed with a fault-isolating dual output controller option that can limit the impact of a short. This controller allows the nodes to be split into two fault-isolated groups. When a short is detected on one leg, the device isolates that leg from the others in order to protect one group from the effects of the short. To fully utilize the feature some additional coax should be placed to maximize the coverage of the power supply.

6. DAA Production Findings

As of June 30th, Comcast has thousands of DAA RPHY nodes deployed, with the majority of these in an N+0 architecture, servicing over 1.5M homes passed across about 10k miles of plant. Beginning in 2021, DAA nodes are being deployed in traditional HFC architectures to accelerate capacity expansion and vCMTS footprint. These will number in the thousands this year and will consist of nodes with a single or two RPDs installed.

6.1. Fidelity Metrics

As noted in the overview describing the advantages of DAA, one of the key benefits is End-of-Line (EOL) MER. By eliminating the analog optical link, the MER loss that accompanies that fiber connection is eliminated. The MER generated by the RPD is essentially the same as the MER that would be delivered from a legacy I-CMTS port in a Hub site prior to traversing the AM Optical link into the field. It is therefore a high-fidelity signal placed onto the coax by the DAA node.

With the large scale of production of DAA nodes in the field there is now sufficient statistical sample size to do A:B MER comparisons among BAU HFC, N+0 with analog nodes, and DAA nodes.

Figure 12 shows distributions of Rx MERs of DOCSIS 3.1 OFDM for "BAU" HFC networks (top), analog N+0 networks, and DAA nodes (bottom).

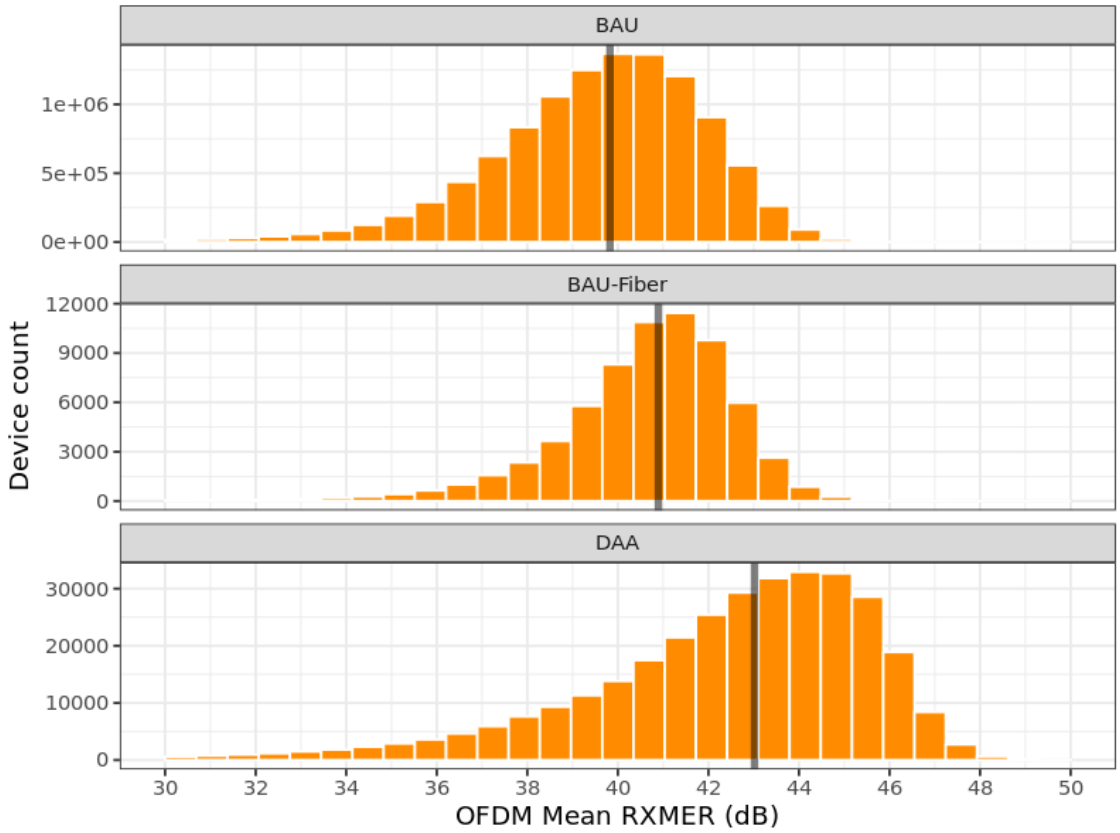


Figure 12 - Downstream Rx MER Comparison: N+x HFC, N+0 Analog, and N+0 DAA

Figure 12 trends exactly as would be expected. Analog N+0 has a better EOL MER than BAU HFC N+X MER, and DAA N+0 has a better EOL MER than Analog N+0 MER. There is a noticeable and quantifiable MER improvement in the bottom plot of **Figure 12**. The average MER increases with Analog N+0 over HFC N+X is a little over 1 dB. The MER gain from HFC N+X to DAA N+0 is over 3 dB. A 3 dB increase is equivalent to one modulation order of efficiency, such as the difference between the fidelity needed to support 512-QAM, for example, and 1024-QAM, which is a more efficient modulation in bps/Hz.

The numbers shown here are measured at the CM downstream receiver, and as such include all channel impairments up to that point, including the Noise Figure of the CM itself. The minimum DOCSIS 3.1 MER requirement for 4096-QAM is 41 dB at -6 dBmV/6 MHz. The minimum requirement for 2048-QAM is 37 dB @ -9 dBmv. Thus, these distributions would suggest that the DAA N+0 case will enable the 4096-QAM modulation profile *on average*, while for non-DAA there is a bandwidth efficiency penalty of about 9%. Whereas MER differences did not translate to anything except for performance margin in DOCSIS 3.0 systems, the use of DAA can add nearly 200 Mbps per OFDM block allocated in the downstream. With the use of multiple modulation profiles supported by the DOCSIS 3.1 Profile Management Application (PMA), this incremental new capacity is added to the network, optimizing coaxial bandwidth efficiency.

Lastly, note that the DOCSIS 3.1 downstream has MUST QAM formats up to 4096-QAM, but also offers optional profiles of 8192-QAM (8K-QAM) and 16,384-QAM (16k-QAM). The incremental capacity of increasing QAM formats is relatively modest (about 8%) for the increased MER needed to achieve them. However, with DAA, once impossible MERs are now within reach. While the full system specifications were not completed at the time of the DOCSIS 3.1 standards development for these QAM formats, based on the 4k-QAM requirements for the DOCSIS 3.1 downstream, the 8k-QAM MER requirement is likely to be in the range of 45 dB. From the distribution in **Figure 13**, a substantial number of CM MERs could reach this profile if it became available.

Now, pivoting to **Figure 13**, we see the side-by-side for the upstream, in this case only comparing BAU HFC vs DAA N+0.

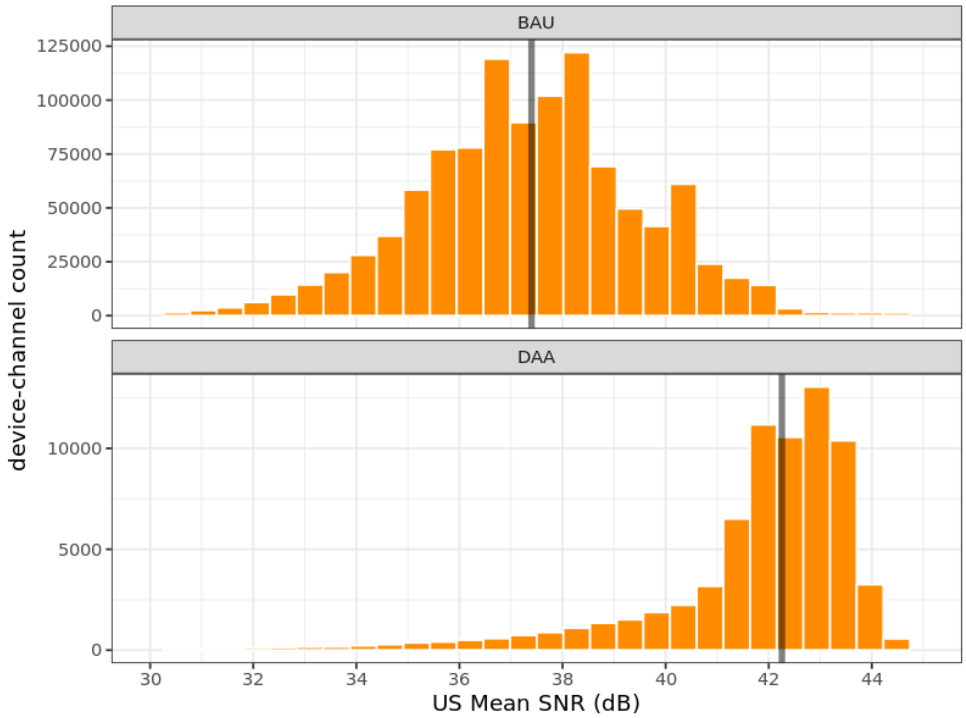


Figure 13 - Upstream Rx MER Comparison: N+x HFC (top) and N+0 DAA (bottom)

The large majority of node splits are driven by upstream capacity limitations, so we will consider the impacts of higher MER on capacity of the upstream.

The increase in bandwidth enabled by DAA in a production upstream is not determined completely by the capacity multiplier noted above for DAA and DOCSIS 3.1. There is an existing DOCSIS 3.0 traffic load in the Low Split band that will work to limit the percent capacity gain, as these carriers deliver a fixed DOCSIS 3.0 64-QAM maximum capacity regardless of network architecture. However, since we are considering Mid-Split, which delivers a long-term capacity runway, we can project that these DOCSIS 3.0 QAMs will be gradually reclaimed for DOCSIS 3.1 as those devices begin to dominate the CPE footprint, and base the analysis on the all-DOCSIS 3.1 case.

Thus, considering the use of Mid-Split supporting OFDMA, the DAA benefit of 3 dB can be used to estimate the impact on timing of a node split due to Compounded Annual Growth Rate (CAGR). For exemplary purposes, assume that the DOCSIS 3.1 upstream QAM format can be 256-QAM, where once it was 64-QAM for DOCSIS 3.0. The more powerful FEC of DOCSIS 3.1 delivers most of the additional QAM efficiency for "free" due to the extra coding gain. Improvements over time in the network RF quality that occur via node segmentation and noise funneling effects since the launch of DOCSIS 3.0 create additional dB benefits. Introducing the DAA gain to this situation, and the MER assessment above, suggest that we can improve to 512-QAM, or a 12% capacity gain.

An US CAGR of 25% is the current working assumption for business planners and capacity managers. This is a monthly growth rate of about 1.02%, and 12% capacity translates to 6-7 months. When considering the expense of tens of thousands of node splits per year, along with factoring in the time value of money, while it may be undramatic, this represents a quantifiable savings of capital by deferment to the node split budget.

In addition to capacity implications, this unavailable bandwidth must be accounted for when considering the penetration of speed tiers in a service group. As higher speeds are offered in the Upstream via Mid-Split, for example, increasing the peak burst, the remaining capacity during peak bursts must be shared by the other users in the service group. When the total available capacity is lower, there is less headroom above the peak, and this can impact the guidance for the number of users that can be added at the peak available tier. In addition, empirical data shows that when a subscriber upgrades to a higher speed tier, there is a predictable increase in their average utilization. As an example, the net result of these phenomenon, based on modeling for 300 Mbps US speeds, suggests that up to 6 users can have this speed tier in a Mid-Split system if utilization is below 60%. However, if the available payload is 9% lower, this number may be reduced to 4 or 5 users in a SG.

Furthermore, increasing US speed tiers has historically also increased utilization by freeing up latent demand. This again puts a premium on maximizing available capacity.

6.2. Customer Experience Metrics

Now consider Figure 14, which shows a common maturation trend observed when introducing the DAA platform followed by increasing production volume over the course (in this example) of a year. Statistically, the majority of customer trouble calls (TCs) are associated with issues that occur within the home, not the network. However, changes of any type on the network often results in interim jumps in TC activity. It is typical that as the DAA ecosystem components – vCMTS and RPHY Nodes – are rolled into new areas, there is an activation and support learning curve before returning to a "business-as-usual" state of health with respect to the customer experience. In some cases, this manifests itself as a temporary uptick in TC rate, which over time, as shown, returns to the mean.

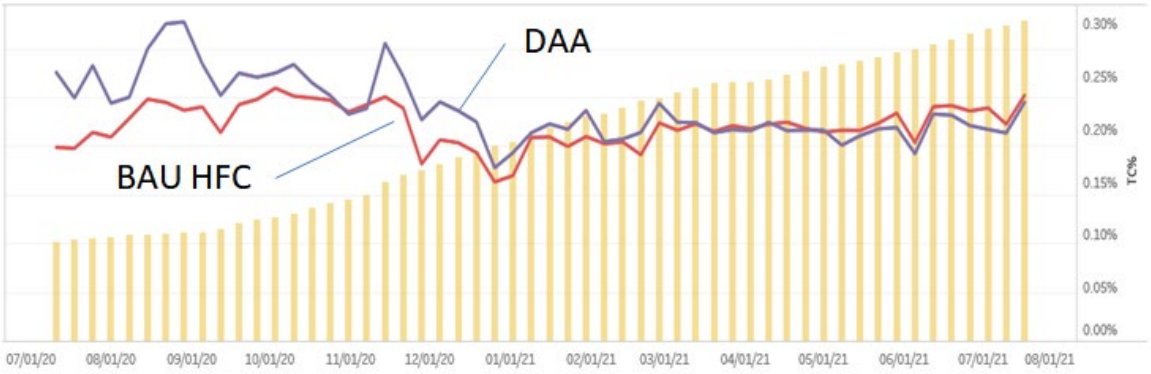


Figure 14 - Operational Metrics Proceed to Maturity as the DAA Platform Grows in Scale

As the DAA footprint continues to expand, availability and customer experience metrics are being carefully tracked and itemized so that any network-related causes are understood and addressed. Decreasing network maintenance activity and associated operational costs over time is cautiously expected, and saving targets established to quantify the costs removed from the operation. Observing network-related ticket root causes, improvements are anticipated for several reasons:

- Analog fiber links, prone to thermal variations and RF alignment sensitivity, become digital
- Large and more bug-prone monolithic SW releases of I-CMTS platforms become more incremental and agile
- The virtual implementation lends itself better to proactive measures and self-healing
- The virtual implementation lends itself better to smaller blast radii
- More robust DOCSIS 3.1 signals take over more of the spectrum allocation in both directions

We will continue to update the industry as DAA continues to scale across the enterprise, improving the network everywhere it is deployed.

6.3. Network Power Consumption Metrics

Returning to the topic of power consumption, one important original question about DAA networks, and in particular the very challenging N+0 DAA network – was what the impact would be on network power usage. RMD devices, with the added processing capability, distibuted into the node, would increase the power consumption challenge further.

A power study was devised to measure the true impact of the DAA architecture on a section of an N+0 DAA network built in Denver in 2020. Five nodes were selected which contained 18 network power supplies. The 18 power supplies contained a mix of Alpha XM2 and XM3, 15 and 18 amp, 120 and 240 VAC variants. At the input and output of each power supply, measurements taken pre- and post-construction:

- Power
- Voltage
- Current
- Power Factor

All measurements were taken with a FLUKE 345 PQ CLAMP METER.

The N+0 architecture took advantage of the high RF output levels previously discussed. Four specific drop models were developed based on a majority of residential builds. All taps include the capability to shape the RF to provide an optimized DS and US RF profile.

Table 1 depicts characteristics of the five nodes selected. The power boundaries are restricted to the node boundaries in the design effort, so power distribution was not changed from a pre and post construction point of view allowing for an apples to apples comparison.

Table 1 - Node Statistics

					HFC Legacy Design					xNET Design										
PS	Node ID	Model		VAC	node	LE	MB	WIFI	ACTIVES	NODES	Wi-Fi	AER HHP	UG HHP	HHP	Aer Miles	UG Miles	MILES	% ports	HHP/MILE
A	11	915	XM2	120	0	3	6	2	11	3	2	340	43	383	1.5	-	1.5	92%	255
B	11	918	XM3	120	1	9	6	4	20	3	7	343	241	584	1.0	0.1	1.1	100%	510
C	11	915	XM2	120	0	4	4	1	9	4	0	382	190	572	1.6	0.0	1.6	81%	349
D	11	915	XM2	120	0	7	3	0	10	2	0	187	0	187	0.8	-	0.8	50%	238
					1	23	19	7	50	12	9	1252	474	1726	4.9	0.2	5.1	83%	340
A	12	915	XM2	120	0	8	4	0	12	2	2	342	7	349	0.9	0.1	1.0	100%	335
B	12	915	XM2	120	1	6	8	2	17	2	0	317	16	333	0.6	0.3	1.0	100%	343
C	12	915	XM2	120	0	6	5	2	13	4	1	529	137	666	1.4	0.1	1.5	100%	441
					1	20	17	4	42	8	3	1188	160	1348	2.9	0.6	3.5	100%	383
A	17	915	XM2	240	0	14	5	0	19	4	0	490	30	520	2.4	0.8	3.2	100%	164
B	17	915	XM2	120	1	7	4	1	13	1	1	25	13	38	0.8	0.2	1.1	100%	36
C	17	918	XM3	120	0	12	5	1	18	5	0	464	87	551	1.9	0.9	2.8	100%	198
D	17	915	XM2	120	0	7	10	1	18	4	1	167	294	461	0.9	0.5	1.5	88%	316
					1	40	24	3	68	14	2	1146	424	1570	6.0	2.4	8.5	96%	186
A	8	915	XM2	120	0	4	3	1	8	3	0	438	11	449	1.2	0.2	1.4	92%	322
B	8	915	XM2	240	1	11	7	3	22	2	0	274	0	274	0.7	0.2	0.9	100%	292
C	8	918	XM3	120	0	5	3	3	11	4	5	353	34	387	1.5	0.4	1.9	88%	204
					1	20	13	7	41	9	5	1065	45	1110	3.4	0.8	4.2	92%	262
A	T0	915	XM2	120	0	12	7	0	19	4	0	0	440	440	-	2.0	2.0	94%	223
B	T0	915	XM3	120	0	10	3	0	13	3	0	0	268	268	-	1.6	1.6	75%	171
C	T0	915	XM2	240	0	8	8	0	16	6	0	0	802	802	-	2.8	2.8	92%	36
D	T0	915	XM2	120	1	20	6	0	27	5	0	0	459	459	-	1.9	1.9	70%	242
					1	50	24	0	75	18	0	0	1969	1969	-	8.2	8.2	83%	239

One of the most significant items from the data in **Table 1** is the reduction in the number of actives, which is facilitated by the high RF levels utilized in the fiber deep design. This reduction creates several operational benefits in the plant, including lower mainteance, less failure points and lower power consumption.

The results of this study are shown in **Figures 15-19.** Assessing the results of the trial, there was a significant reduction in overall utility power usage in the plant. We can see a net reduction of input and output power from most of the power supplies studied and a net reduction in input and output power in every original node boundary area. This is most welcoming news. While using less power is always an excellent result, another practical point of view exemplified by this result is that it has once again been shown that over time we can accomplish much more for a fixed amount of power, as network services continue to be added. Effectively, the bits-per-second-per-kilowatt-hour (bps/kwh) continues to improve. This has led to the deployemnt, for example, of outdoor WiFi Access Points (APs) or APs supporting CBRS (Citizens Broadband Radio Service) coverage.

Additionally, while cutting the total number of actives is the cause of the most significant decrease in power usage in this study, the effect of reduced transmission power loss in the cable plant is also to be noted. When current is drawn through the coax network, power losses defined by Joule's Law ($P=I^2R$) are incurred based on current (I) and loop resistance of the network (R). As the number of actives is reduced, and the current required to run actives in the plant is also lessened, so is the total power dissapated by transmission. Reducing this dissapated power is key to overall plant efficiency because, unlike power driving nodes or other actives, the power consumed due to transmission creates no value.

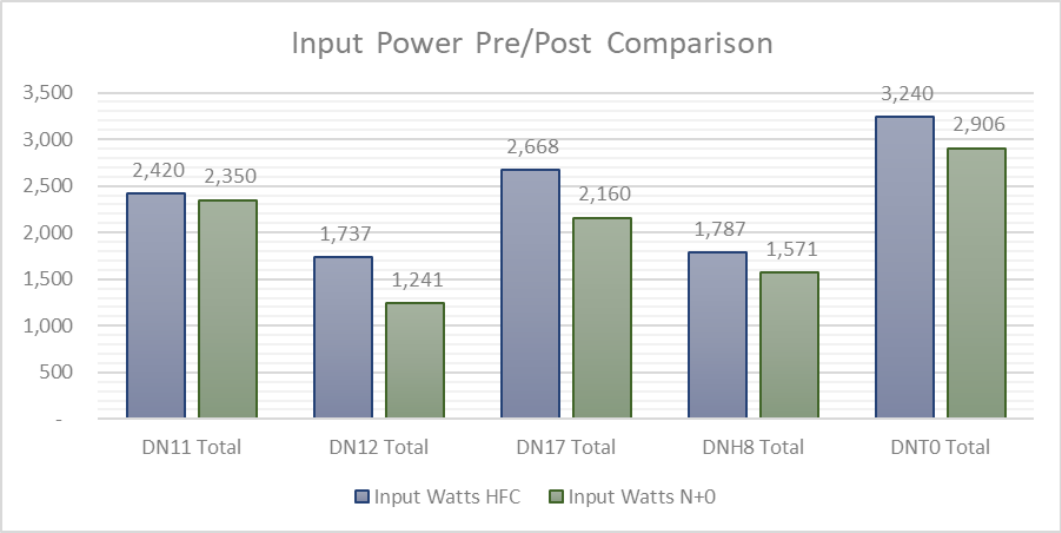


Figure 15 - Pre/Post Input Power Comparison by Original Node Boundary

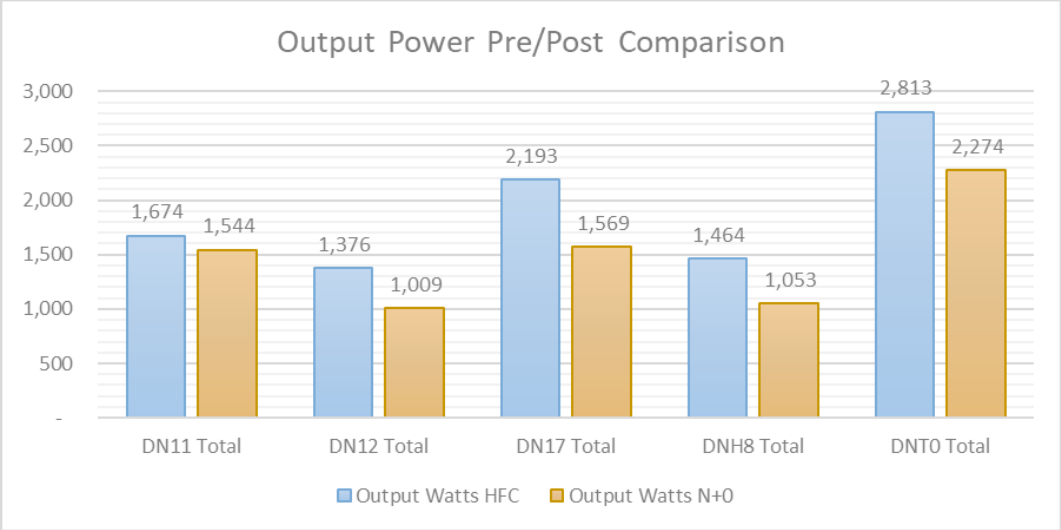


Figure 16 - Pre/Post Output Power Comparison by Original Node Boundary

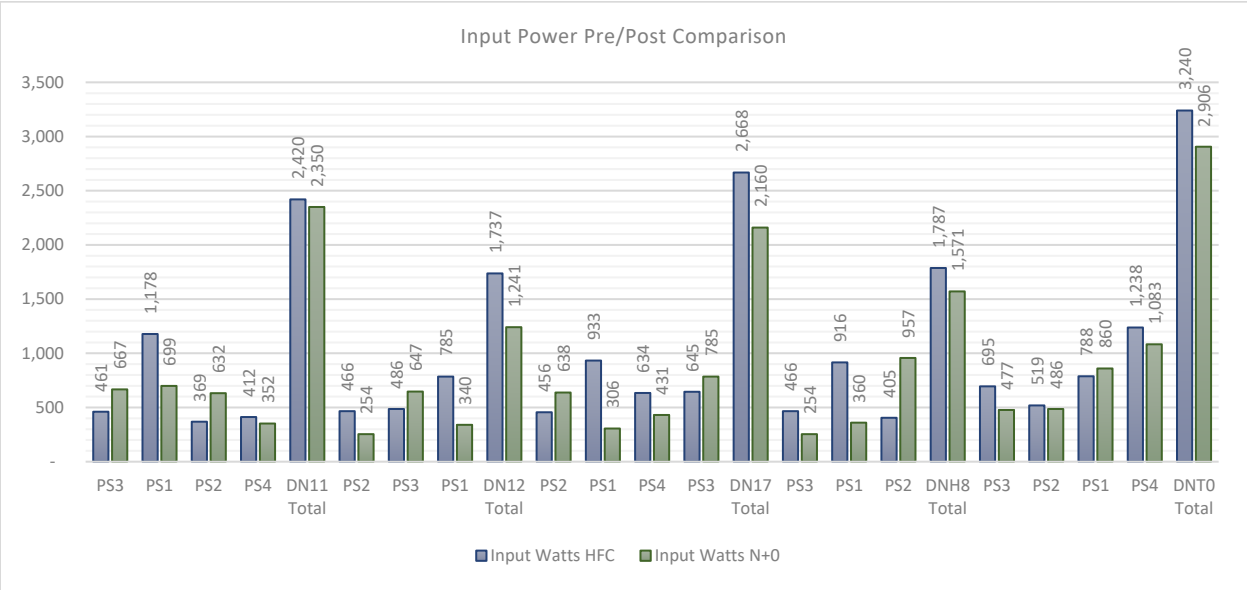


Figure 17 - Pre/Post Input Power Comparison by Power Supply

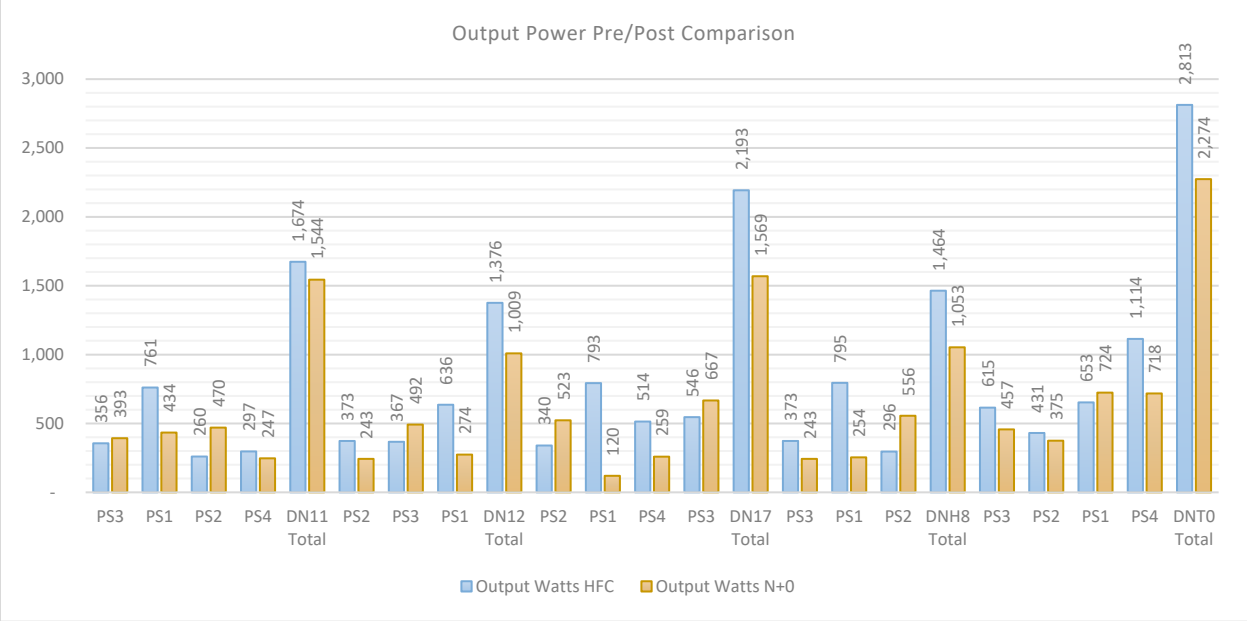


Figure 18 - Pre/Post Output Power Comparison by Power Supply

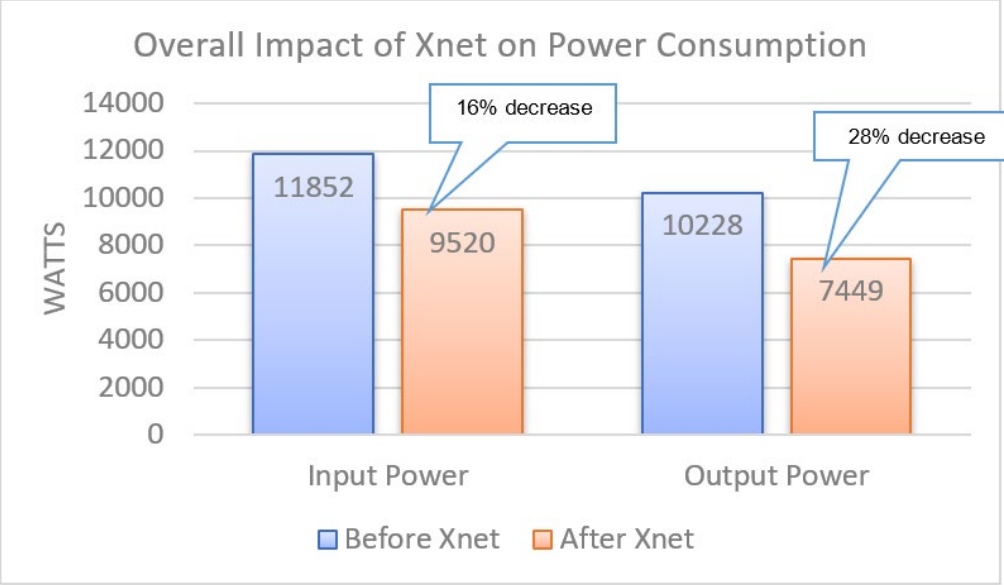


Figure 19 - Pre/Post Input Power Comparison Consolidated

An additional benefit derived from a reduction of plant actives and the resulting power savings is the positive impact on plant backup times. Using the powering data collected, and a 3-battery string of 114 Ah batteries, the sites in this study would experience an average of a 30% increase in battery backup time. This is shown in **Figure 20**. This additional runtime will reduce operational costs of additional backups to these sites during outages and decrease the risk of these outages becoming customer affecting.

Alternately, where current backup times are adequate, there may be opportunity to realize savings by either deploying a smaller, lighter, more cost-effective battery or by reducing battery string counts.

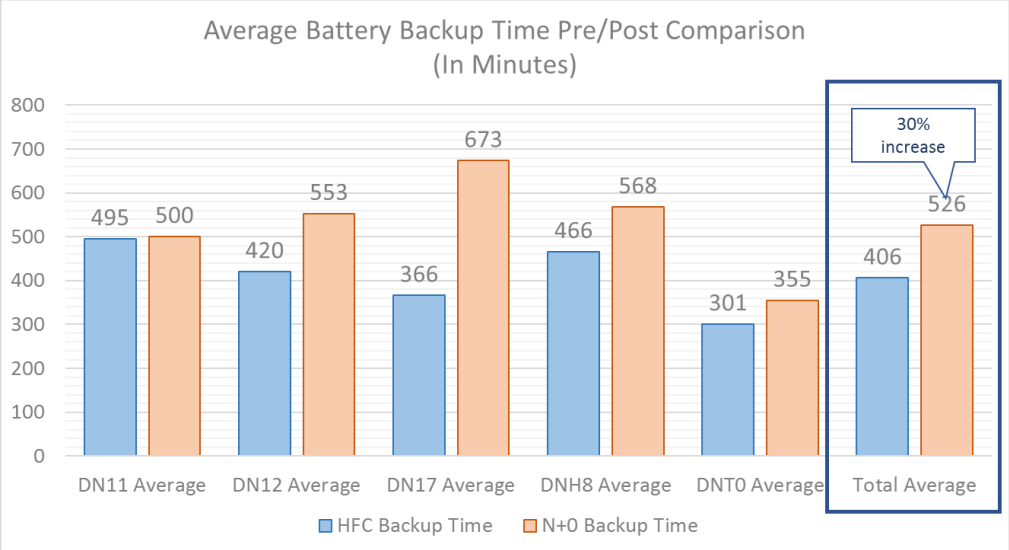


Figure 20 - Average Battery Backup Time Pre/Post Comparison

Finally, beyond the savings in network power, the reduction in network actives decreases the number of active failure points in the network. As with any system, reducing the number of potential failure points within the system, reduces the overall likelihood of system failure and, in this case, impact to customers.

7. Conclusion

Comcast began its DAA journey into the production network in mid-2018. By 2019, in every location that a Tier 1 or N+0 upgrade took place, it was done with vCMTS and DAA nodes. Remote PHY went from a CableLabs standard and PowerPoint to reality. There are now three production OEM partners producing DAA nodes in this interoperable ecosystem.

The scale of the deployment is now significant, yielding scale metrics of statistical significance, and validating projected performance improvements. We observed MER improvements that were anticipated with DAA, driving more capacity into the plant when it is combined with the more bandwidth efficient capabilities of DOCSIS 3.1. The capacity gains translate both to deferred costs of node splits, or other augmentations, that are driven by high utilization thresholds.

This added bandwidth can also translate to service speed scalability. While the added capacity technically could represent an opportunity for higher speed offerings, the difference is relatively modest for it to be a meaningful difference. For example, does it make very much difference to the customer or to the business if a service tier is 250 Mbps vs 280 Mbps (a 12% difference)? Probably not. However, the added capacity headroom will mean that more users can be added onto a higher speed tiers before a plant augmentation must occur to support the penetration – again, savings by deferring cost.

One of the initial uncertainties, and ultimately better than anticipated result, is how the introduction of DAA via RPHY nodes, as part of an N+0 upgrade, is, as it turns out, actually reducing power consumption of the serving area. This is a combination of the evolution of available technology replacing much older equipment in the field and incorporating efficient power network design as a priority from the outset. In the case of RF amplifiers that are eliminated in an N+0 build and effectively replaced by nodes, the amplifiers being replaced can be over 20 years old, and thereby relatively inefficient technology by today's standards.

Comcast is still relatively early in the DAA journey, but the engine is in 5th gear. The data from these upgrades is and will continue to be being collected and analyzed. This will drive optimizations and investment decisions going forward, such as the path to deployment of 10G and the evolution of the access network edge for alternative last mile technologies. Comcast will continue to keep the industry abreast of the observations and findings from the world's largest cable DAA network, fed by the world's most subscribed virtualized broadband platform.

Acknowledgments

A big thank you to the two Comcast technicians: Seth Williams, CommTech 5, Network Maintenance and Mike Torrez, CommTech 5, Network Maintenance. These talented gentleman collected all of the data for the power consumptions study.

Special thanks to Maher Harb and Ramya Narayanaswamy from Comcast's Next Gen Access Networks team for their support in providing dashboard readouts for the DAA production fidelity and customer experience metrics.

Abbreviations

AM	Amplitude Modulated
APs	Access Points
BAU	Business-As-Usual
CBRS	Citizen Band Radio Service
CMTS	Cable Modem Termination System
COTS	Commercial-off-the-Shelf
DAA	Distributed Access Architecture
EOL MER	End-of-Line Modulation Error Ratio
FMA	Flexible MAC Architecture
I-CMTS	Integrated CMTS
N+0	Node plus Zero Amplifiers
N+X	Node plus X Amplifiers
OFDMA	Orthogonal Frequency Division Multiple Access
OSP	Outside Plant
PFC	Power Factor Correction
PMA	Profile Management Application
PON	Passive Optical Network
PP	Power Pack
QAM	Quadrature Amplitude Modulation
QSW	Quasi-Square Wave
R-CCAP	Remote Converged Cable Access Platform
RMD	Remote MAC Device
RPHY	Remote PHY Device
vCMTS	Virtual CMTS

Bibliography & References

Howald, Dr. Robert, Aboard the Technology Wave: Surf Report, 2016 SCTE Expo, Philadelphia, PA, Sept 26-29.

Nandiraju, Dr. Nagesh, Distributed Access Architecture – Goals and Methods of Virtualizing Cable Access, 2016 SCTE Expo, Philadelphia, PA, Sept 26-29.

DOCSIS 4.0 - A Key Ingredient of the 2030s Broadband Pie

A Technical Paper prepared for SCTE by

Zoran Maricevic, Ph.D.
Engineering Fellow
CommScope
Wallingford, CT
203-303-6547
zoran.maricevic@commscope.com

James Andis
General Manager – HFC Technologies |
CTO Networks, Engineering & Security (NES)
nbn™ Australia
Melbourne, Victoria - Australia
+61 409 066 745
JamesAndis@nbnco.com.au

Tom Cloonan, Ph.D.
Chief Technology Officer CommScope
Lisle, IL
630-281-3050
tom.cloonan@commscope.com

John Ulm
Engineering Fellow CommScope
Moultonborough, NH
978-609-6028
john.ulm@commscope.com

1. Introduction

There is no doubt that Fiber to the Premise (FTTP) is the very long-term end goal of every operator. The questions of when and how to move to FTTP within a cable brownfield is the billion-dollar question. It is widely recognized that FTTP is expected to cost more than a comparable upgrade to DOCSIS 4.0 using *Extended Spectrum DOCSIS (ESD)* – yet both will yield similar speed tiers and revenue from a residential service perspective during the 2030's and into the early 2040's.

In Australia, nbn™ (National Broadband Network) is a government-business enterprise that has recently completed the rollout of a ubiquitous broadband network. This national broadband network is leveraging copper, coax, fiber, Fixed Wireless and Satellite assets, and has used a truly technology agnostic approach to servicing 100% of the population. Like other operators in the post-COVID world, nbn is driving its broadband infrastructure efficiently to meet traffic demands and expected performance, and nbn has one of the most extensive DOCSIS 3.1 deployments in the world. Coupled with fully segmented (4x4) nodes over a fully analogue forward and return path architecture, nbn is now evaluating what the next significant investment cycle, expected in 2025/26, will entail and is asking the hard question as to whether this is the right time to overbuild FTTP to every residential home or to do another round of DOCSIS investments using Distributed Access Architectures (DAA) and DOCSIS 4.0.

This is a multi-billion-dollar question that has implications over the next two decades. nbn's investigations and analysis has concluded that the overall cost of deploying FTTP to their residential footprint would be about 5 to 6 times the cost of rolling out DOCSIS 4.0 with DAA. Figure 1 shows the FTTP journey that nbn envisions for its residential subscriber base.

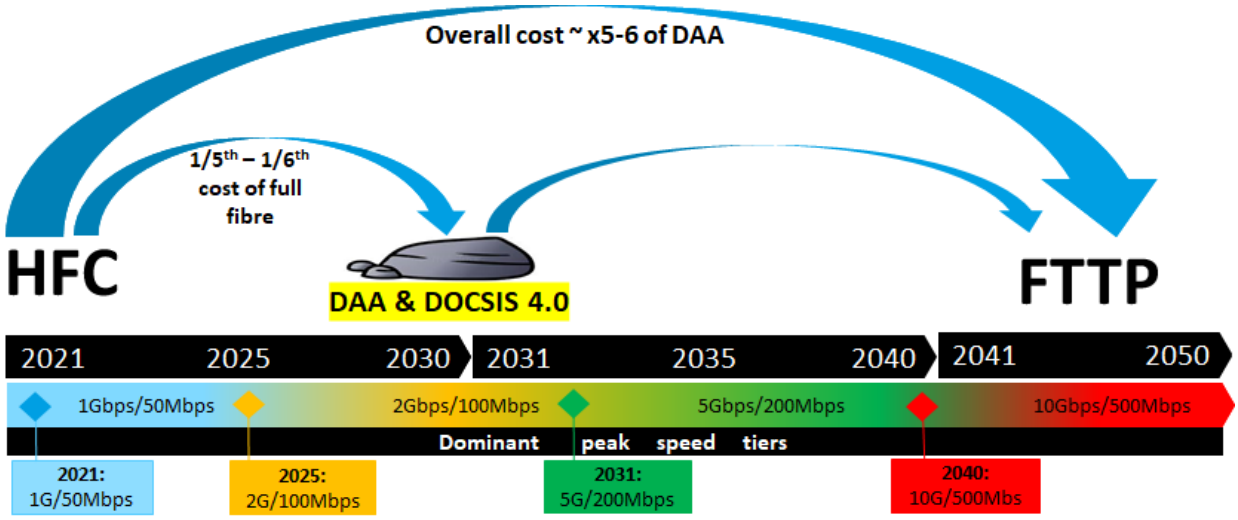


Figure 1 – The *nbn™ Residential* FTTP journey

In Australia, nbn still sees very asymmetric service tier offerings in the future serving the residential market. The main benefit that DOCSIS 4.0 brings to their table is the extended 1.8 GHz spectrum and additional downstream (DS) capacity to get through this decade and next. For other operators in highly competitive environments with fiber-based competitors, DOCSIS 4.0 enables multi-gigabit upstream (US) tiers.

For the next 5-10 years, nbn will be prioritizing major uplifts to FTTP in area's served by copper infrastructure (Fiber-to-the-Node/FTTN), namely, to address demand for higher speed services (> 50 Mbps or > 100 Mbps) which come with associated revenue uplifts and ceases further investment in copper remediation required to maintain the network. These copper areas will utilize existing Distribution Fiber Network (DFN) that nbn has already deployed to serve the DSLAM nodes, essentially extending the fiber deeper to pass all the premises served by the copper Distribution Area (DA). Fiber drops to homes will be built when customers order services of 100 Mbps or more; essentially using a demand driven approach.

Unlike copper-based services, nbn's HFC services do not suffer from the inability to reach Gigabit speeds or require intense and expensive remediation activities to maintain those speeds. For this reason, there is no major uplift in revenue expected from offloading HFC services to FTTP. When coupled with the recent large scale DOCSIS 3.1 investment and deployment, the nbn HFC network has adequate headroom in capacity to defer the next major investment cycle for several years.

There is no doubt that fully symmetric services will become an important offering for business in the near future. Fibre based 1 G and 10 G Ethernet are commonly deployed for businesses as well. However, DOCSIS-based alternatives with lower deployment costs might also have a great appeal - to business customers and operators alike.

For operators, DOCSIS 4.0 is a compelling proposition. It allows them to remain competitive, by delivering reliable high-speed residential services well into the 2030's across a wide footprint. As an interim technology, it allows operators to strategically deploy deep-fibre and FTTP where needed, while enabling the evolution of the HFC network towards N+0/N+1/N+2 DAA deployments. The reality is that most operators don't have the funds and resources to deploy FTTP to every customer in a short time frame – this is evident from the efforts of many companies who have attempted full fibre overbuilds.

This paper provides a case study of possible network evolution steps on an actual node; and then looks at the Total Cost of Ownership (TCO) for each option. This will show others the same steps that nbn went through in their decision-making process. The three approaches considered are:

- Full FTTP overbuild and offload.
- DOCSIS 4.0 Extended Spectrum DOCSIS (ESD) using DAA
- A hybrid approach - where deep-fibre DOCSIS 4.0 DAA deployments are coupled with selective/targeted FTTP offloads – to provide an optimum outcome that minimizes upfront CAPEX and maximizes the longevity of operator Investments in the DOCSIS part of network

The paper also considers network power consumption, material and truck roll costs for actives, passives, fiber and coax maintenance, as well as for the drop line maintenance and replacement, and give overview on how to make estimates for the costs over an extended period of performance.

The net present value (NPV) of the expected OPEX savings is the dollars-and-sense measuring stick against the investments required to achieve the savings - similar to calculating a payoff for a solar-powered home system without giving credit to all the environmental benefits. (Going green while collecting greens, that is.)

2. Network Evolution: Drivers and Timing

2.1. The 10G™ Initiative

The 10G Initiative within the Cable Industry is a key focal point for future vendor product developments and for future Multiple System Operator (MSO) architectural plans. At a high level, 10G defines the simple goal of providing 10 Gbps to subscribers in the future. In addition to speed increases, 10G also defines important goals of improving latency, security, and reliability within future Cable Industry service platforms. As a result, 10G will undoubtedly drive innovation on many fronts into the Cable Industry.

The upcoming arrival of DOCSIS 4.0 equipment (both Full Duplex, FDX, and Extended Spectrum DOCSIS, ESD) within the next few years will mark the next step in the quest for 10G - with a focus on providing the bandwidth via HFC plant augmentations. Since DOCSIS 4.0 equipment deployments will begin in the near future, it is imperative that we begin studying the timing and magnitude of the associated bandwidth capacity needs to ensure that the DOCSIS 4.0 equipment can support the required bandwidths.

These future capacity needs are being partially driven by subscriber demands (higher average bandwidth consumption resulting from higher resolution IP Video and larger subscriber numbers). But as we move forward towards 10G services, the biggest drivers for 10G operation will likely materialize from market challenges instead of subscriber demands. These challenges may result from competitor technologies that are capable of offering service level agreements (SLA) with bandwidths exceeding the typical maximum SLA of 1 Gbps offered by many Multiple System Operators (MSOs) today. Last-mile technologies that may be competitive to DOCSIS over the next 10-15 years include Passive Optical Networks (PON), 5G wireless/fixed wireless, and satellite services.

Satellite services are becoming more ubiquitous with more broadband-oriented, low-earth orbit satellites being launched on a regular basis. However, large latencies (due to the lengthy round-trip path) and relatively low bandwidth capacities (< 100 Mbps) would likely preclude these systems from competing with fiber and DOCSIS in a high-bandwidth marketing war of the future [SATELLITE].

5G wireless/fixed wireless can potentially offer competitive bandwidth capacities, and it has much more bandwidth capacity than satellite services. For sub-6 GHz 5G operations, capacities are expected to reach 500 Mbps. However, mmWave 5G technologies using 24-39 GHz ranges could offer 1.5 Gbps service level agreements or higher that could leap-frog the current DOCSIS 1 Gbps capacity [FORBES]. Limitations in current DOCSIS upstream capacity may be another area where fixed wireless service providers may attack.

PON is the most competitive to DOCSIS of these last-mile technologies. Several variants of PON already support 10 Gbps services. These include 10G EPON (10G x 10G), XG-PON (10G x 2.5G), XGS-PON (10G x 10G), NG-PON2 (multiple 10G lambdas), and NG-EPON (25G x 25G or 50G x 50G) [ITU]. Coherent PON services are also being studied that may eventually support even higher capacities.

Thus, 5G and PON are both technologies that could launch capacity-oriented marketing campaigns against DOCSIS in the coming years. It may prove beneficial to use forward-looking estimates of likely bandwidth capacity rollouts to predict when DOCSIS 4.0 augmentations may be required in the future.

2.2. Traffic Engineering For the Future

Previously, [CLO_2014] introduced traffic engineering and quality of experience (QoE) for broadband networks. From there, the paper went on to develop a relatively simple traffic engineering formula for

service groups that is easy to understand and useful for demonstrating basic network capacity components. It is still extremely relevant today.

The "Basic" formula shown below is a simple two-term equation. The first term (Nsub*Tavg) allocates bandwidth capacity to ensure that the aggregate average bandwidth generated by the Nsub subscribers can be adequately carried by the service group's bandwidth capacity. The first term is viewed as the "DC component" of traffic that tends to exist as a continuous flow of traffic during the peak busy period.

The "Basic" Traffic Engineering Formula (Based on Tmax_max):

$$C \geq (Nsub*Tavg) + (K*Tmax\_max) \quad (1)$$

where:

C is the required bandwidth capacity for the service group
Nsub is the total number of subscribers within the service group
Tavg is the average bandwidth consumed by a subscriber during the busy hour
K is the QoE constant (larger values of K yield higher QoE levels),
where $0 \leq K \leq$ infinity, but typically $1.0 \leq K \leq 1.2$
Tmax_max is the highest Service Tier (i.e. Tmax) offered by the MSO

There are obviously fluctuations that will occur (i.e., the "AC component" of traffic) which can force the instantaneous traffic levels to both fall below and rise above the DC traffic level. The second term (K*Tmax_max) is added to increase the probability that all subscribers, including those with the highest service tiers (i.e., Tmax values), will experience good QoE levels for most of the fluctuations that go above the DC traffic level.

The second term in the formula (K*Tmax_max) has an adjustable parameter defined by the K value. This parameter allows the MSO to increase the K value and add bandwidth capacity headroom that helps provide better QoE to their subscribers within a service group. In addition, the entire second term is scaled to be proportional to the Tmax_max value, which is the maximum Tmax value that is being offered to subscribers.

In previous papers [CLOONAN_2013, EMM_2014], found that a K value of ~1.0 would yield acceptable and adequate QoE results. [CLOONAN_2014] goes on to provide simulation results that showed a value between K=1.0 and 1.2 would provide good QoE results for a service group of 250 - 400 subscribers. Larger service groups (SGs) would need even larger values of K while very small SGs might use a K value near or less than 1.0.

2.3. Some Potential Future Service Tier Use Cases

Forward-looking predictions are always difficult to make, but we can bound these future challenges with reasonable assumptions for various cases. Figure 2 displays some downstream (DS) and upstream (US) average peak busy hour bandwidth consumption data that has been collected from the same four MSOs over a 12-year period. This data has then been corroborated with other MSOs globally. This gives us an unprecedented look into Tavg history. The DS compounded annual growth rate (CAGR) has been gradually slowing over the last half-decade. Based on this, the Tavg growth estimates are expected to be ~20% CAGR for both US + DS over the foreseeable future.

Figure 3 shows the extrapolated predictions for Tavg. DS Tavg is expected to grow from ~3 Mbps in '21 to almost 40 Mbps in 2035. US Tavg starts at 0.25 Mbps in 2021 and ends up around 3 Mbps in 15 years.

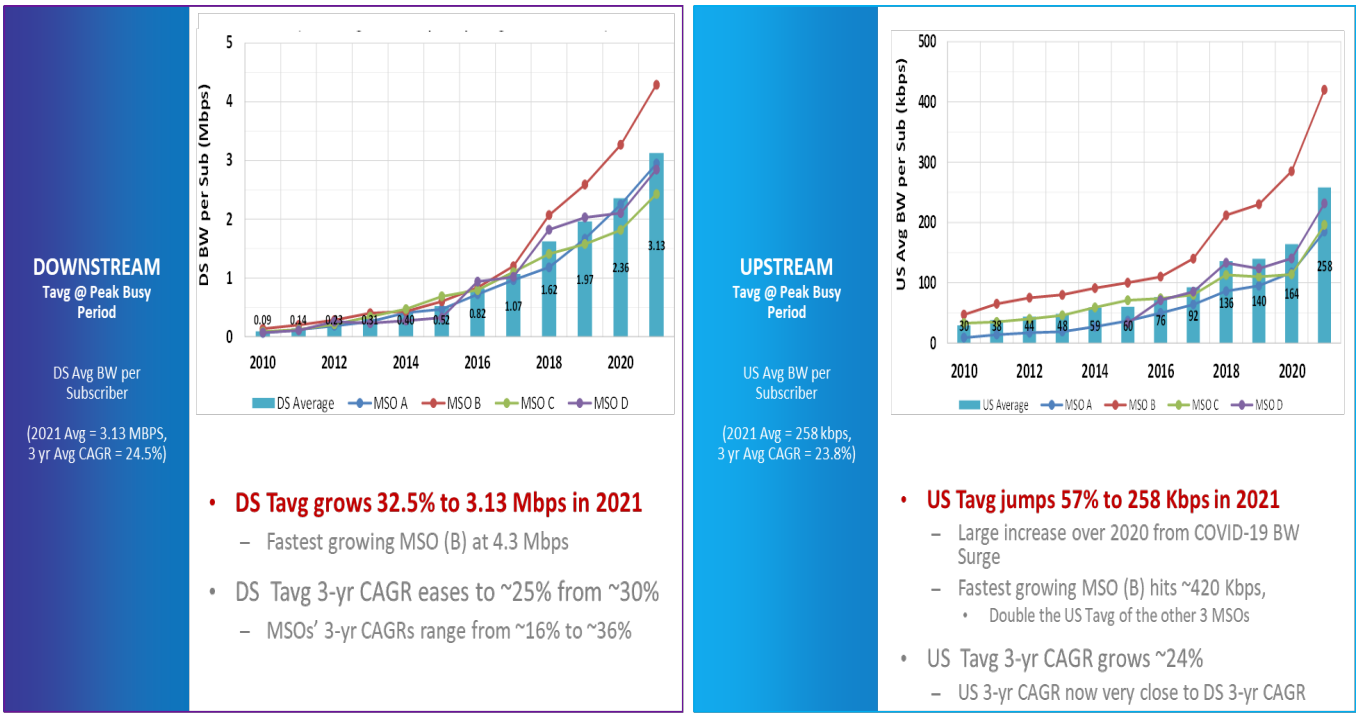


Figure 2 – Past Downstream and Upstream Average Bandwidth Usage Trends

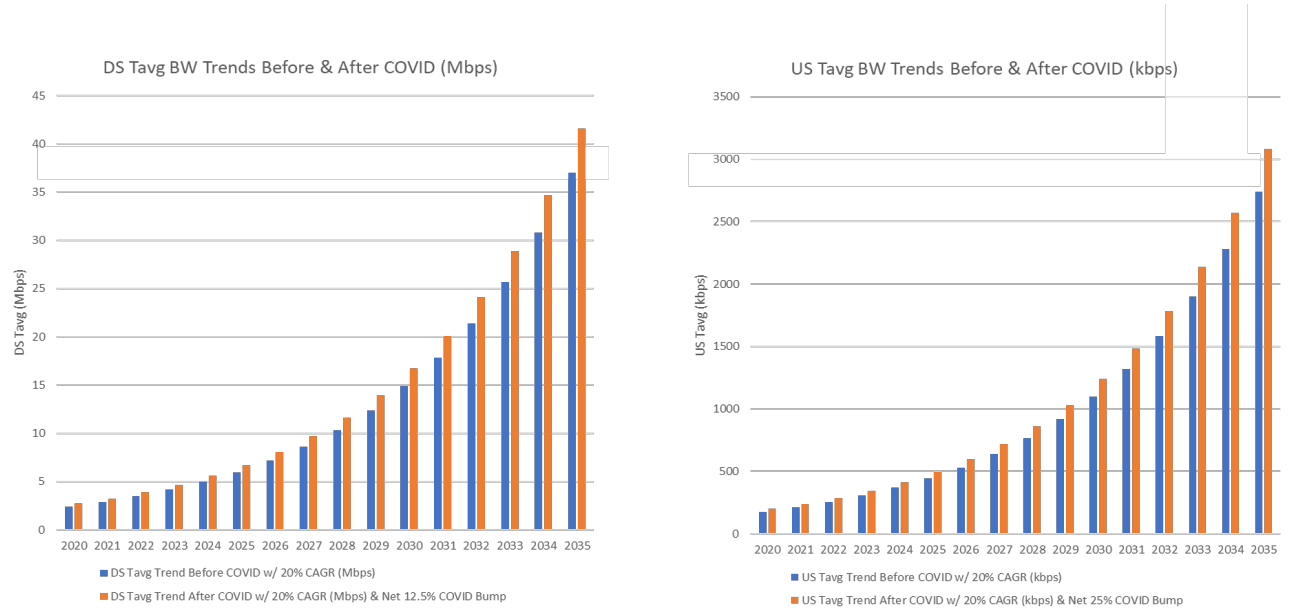


Figure 3 – Future Downstream and Upstream Average Bandwidth Usage *Predictions*

The maximum offered service tier, Tmax_max, had followed Nielsen's Law for many decades, growing at a 50% CAGR. However, this rate has significantly tailed off once Gigabit tiers were reached. The US tiers have been held artificially low due to the 42 MHz constraints. Typical DS:US tier ratios have been in the 10:1 to 25:1 range. As the US split is moved higher, the expected US tiers will jump as well. We do

not expect it to become fully 1:1 symmetric, but rather a much closer 2:1 to 4:1 ratio of DS to US. Going forward, it is estimated that Tmax_max will double roughly every five years for an effective 15% CAGR. This is shown in Figure 4.

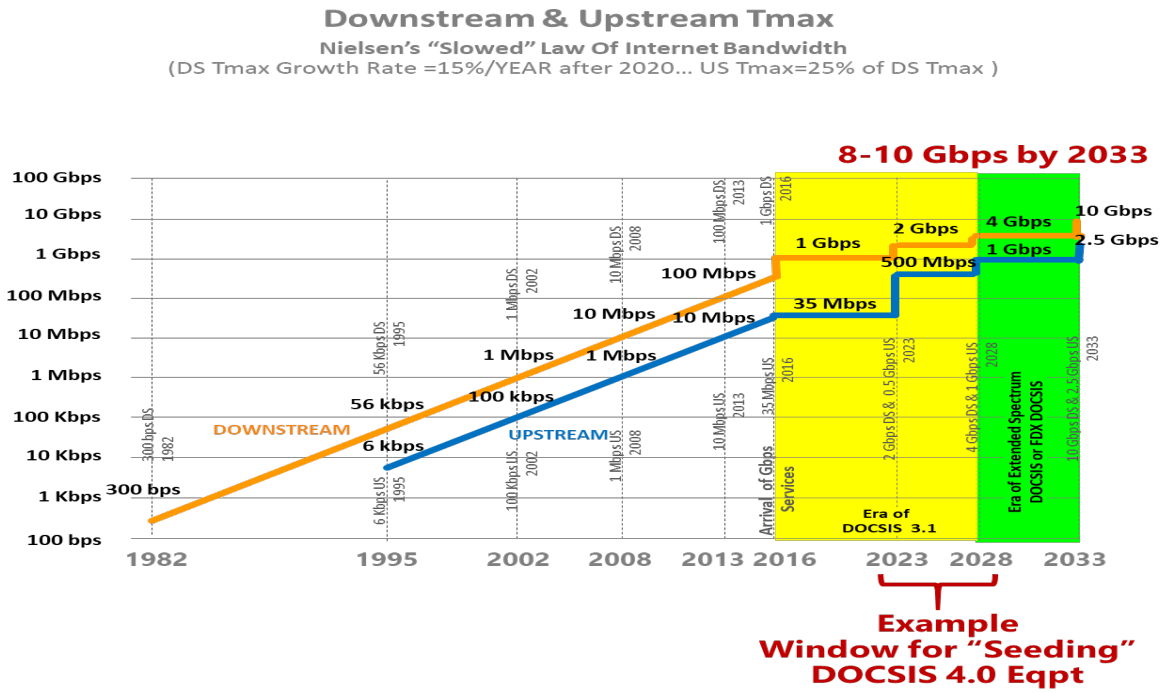


Figure 4 – Future Downstream and Upstream Maximum Service Level Agreement Trends

Previously, [ULM_2019] discussed how a 1218/204 MHz plant can support a 10G migration over the '20s decade. The 1218/204 MHz upgrade is still a strong viable option for networks needing immediate bandwidth capacity relief. However, with DOCSIS 4.0 products becoming closer to reality, some operators are wondering if they can get by with existing plants and then make a giant leap up to a 1794 MHz ESD plant. The CommScope Network Capacity Planning model that leverages the QoE-based Traffic Engineering formula helped us analyze several interesting use cases as the network migrates from 860 to 1794 MHz.

For our analysis, the timeframe is extended out 15 years, to 2035. The network starts with 300 MHz of legacy video spectrum and 30 bonded DOCSIS 3.0 SC-QAM channels. By 2033, the legacy video spectrum is removed in favor of IPTV over DOCSIS. Also, by this time, all 2.0/3.0 modems are assumed to be removed such that the network is now 100% OFDM/OFDMA channels. The model uses an effective DS OFDM bit rate of 8.8 bps/Hz and an US OFDMA bit rate of 7.7 bps/Hz.

The first two use cases shown in Figures 5 and 6 are for a CMTS Service Group (SG) with 120 subscribers (e.g., 240 HP, Homes Passed, @ 50% penetration). The next two use cases in Figures 7 and 8 assume a much smaller 60 subs per SG that an MSO might see from a fiber deep deployment.

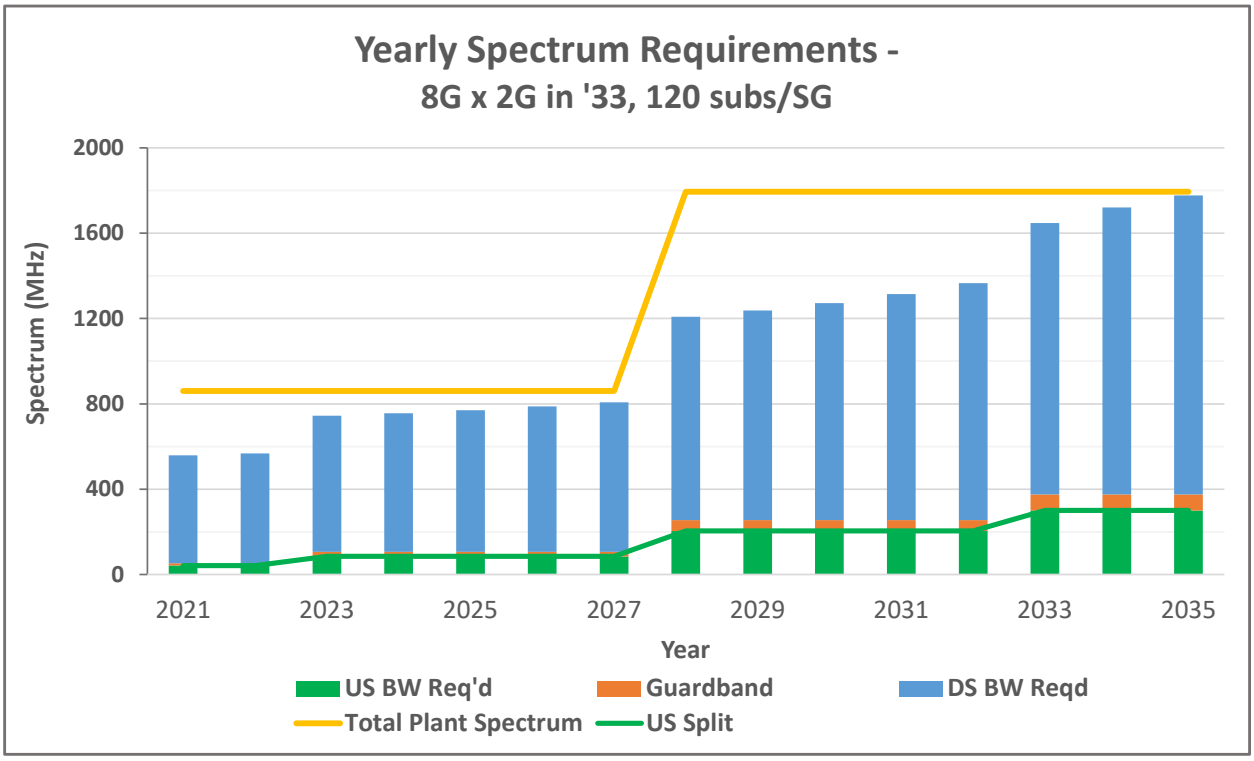


Figure 5 – 1794/300 MHz migration with 8G x 2G Service Tier, 120 subs/SG

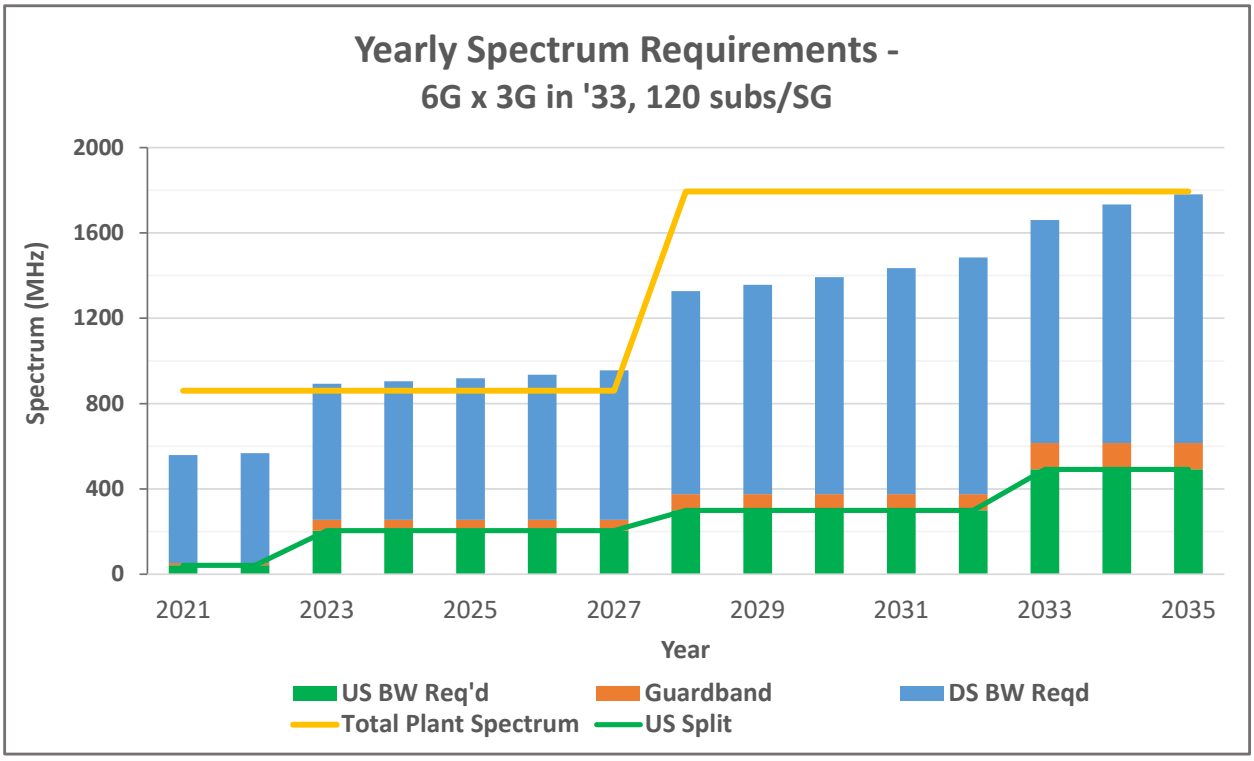


Figure 6 – 1794/492 MHz migration with 6G x 3G Service Tier, 120 subs/SG

The use case in Figure 5 assumes a 4:1 DS:US ratio with 120 subs/SG. The max service tier progresses to 2G x 500M, 4G x 1G and finally 8G x 2G. The 860/85 MHz plant can carry the operator through 2027. In 2028, the introduction of the 1G US tier forces an upgrade to 204 MHz at which time the downstream is also upgraded to 1.8 GHz. Note that a 1218 MHz plant would still be viable through the end of the decade. By 2033, the 2G US tier requires the US split to be changed to 300 MHz. Room is made for the 8G DS tier by the IPTV savings and removal of 2.0/3.0 SC-QAMs. By 2035, both the US + DS spectrum are filling up for this SG size.

Operators will need to install DOCSIS 4.0 equipment prior to enabling it, because the installation of the various components may take years. That installed equipment may be operated in a DOCSIS 3.1 mode for a while before being configured to enable DOCSIS 4.0 bandwidth capacities whenever the needs arise. Also note that some existing “1GHz” taps can actually operate at much higher frequencies. Depending on the tap type, they may not need to be replaced until the year 2033 jump in service tiers. This would give the operator a 12-year window to make the tap replacements.

The use case in Figure 6 assumes a 2:1 DS:US ratio with 120 subs/SG. The max service tier progresses to 2G x 1G, 4G x 2G and finally 6G x 3G. This use case requires a 204 MHz US split once the 1G US tier is introduced. An 860/204 MHz plant can be stretched for a couple more years by using the roll-off region above 860 MHz. Note that only the limited number of 2G subscribers might need to infrequently use that roll-off bandwidth. By 2028, the introduction of the 4G x 2G tier forces the network to upgrade to 1794/300 MHz. Later, the 3G US tier causes the US split to move to 492 MHz. This reduces available DS spectrum which is why the DS tier is limited to 6G.

The use case in Figure 7 assumes a 4:1 DS:US ratio with a fiber deep 60 subs/SG. The max service tier progresses to 2G x 500M, 4G x 1G and finally 10G x 2G. Even though the SG size is half of that in Figure 5, it does not change the date on when the 1.8 GHz upgrade is needed. That is because Tmax_max is dominating our traffic engineering formula rather than Nsub * Tavg. The SG size is hitting the point of diminishing returns for node splits as usual. The smaller SG size does buy some added capacity in the later years, enabling the operator to offer a true 10 Gbps DS service with a 1794/300 MHz plant!

The use case in Figure 8 assumes a 2:1 DS:US ratio with a fiber deep 60 subs/SG. The max service tier progresses to 2G x 1G, 4G x 2G and finally 7.5G x 3.5G. This use case is similar to Figure 6 except the smaller SG size provides enough capacity usage savings in later years to enable the higher 7.5G x 3.5G tier. Note that this is effectively what 10G PONs can offer too.

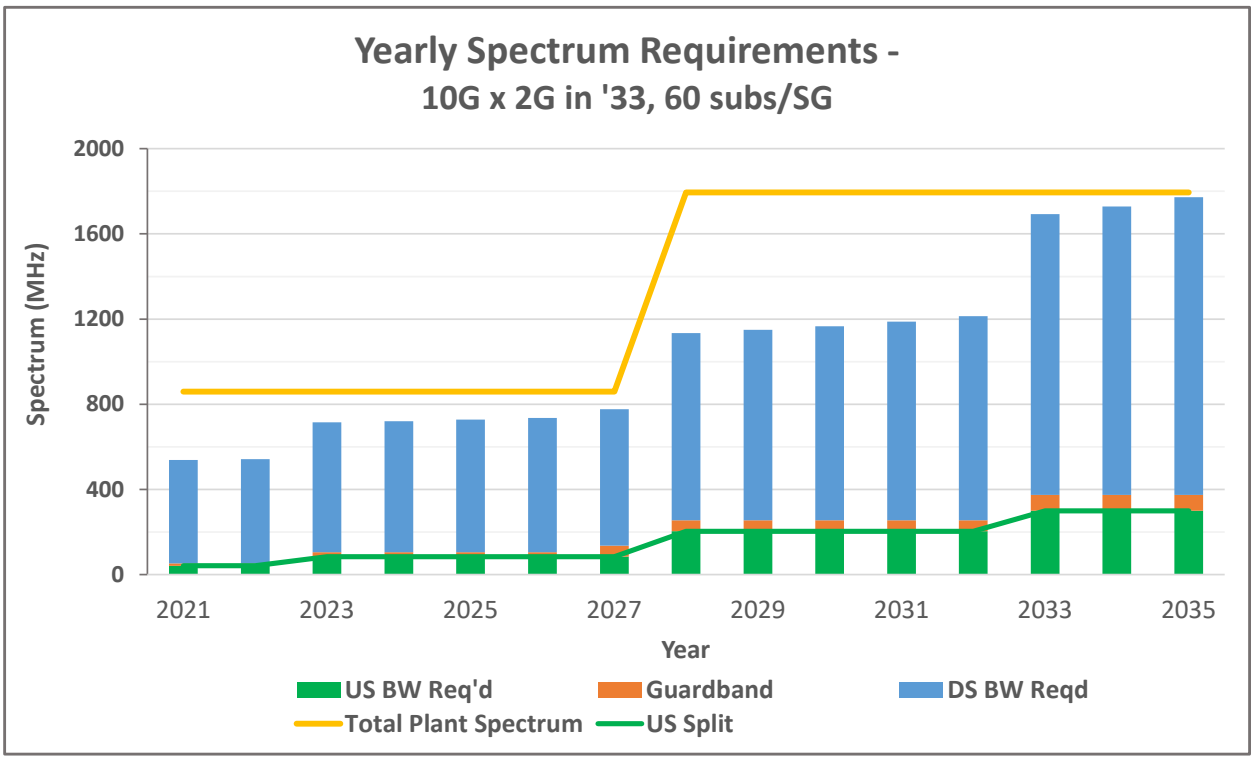


Figure 7 – 1794/300 MHz migration with 10G x 2G Service Tier, 60 subs/SG

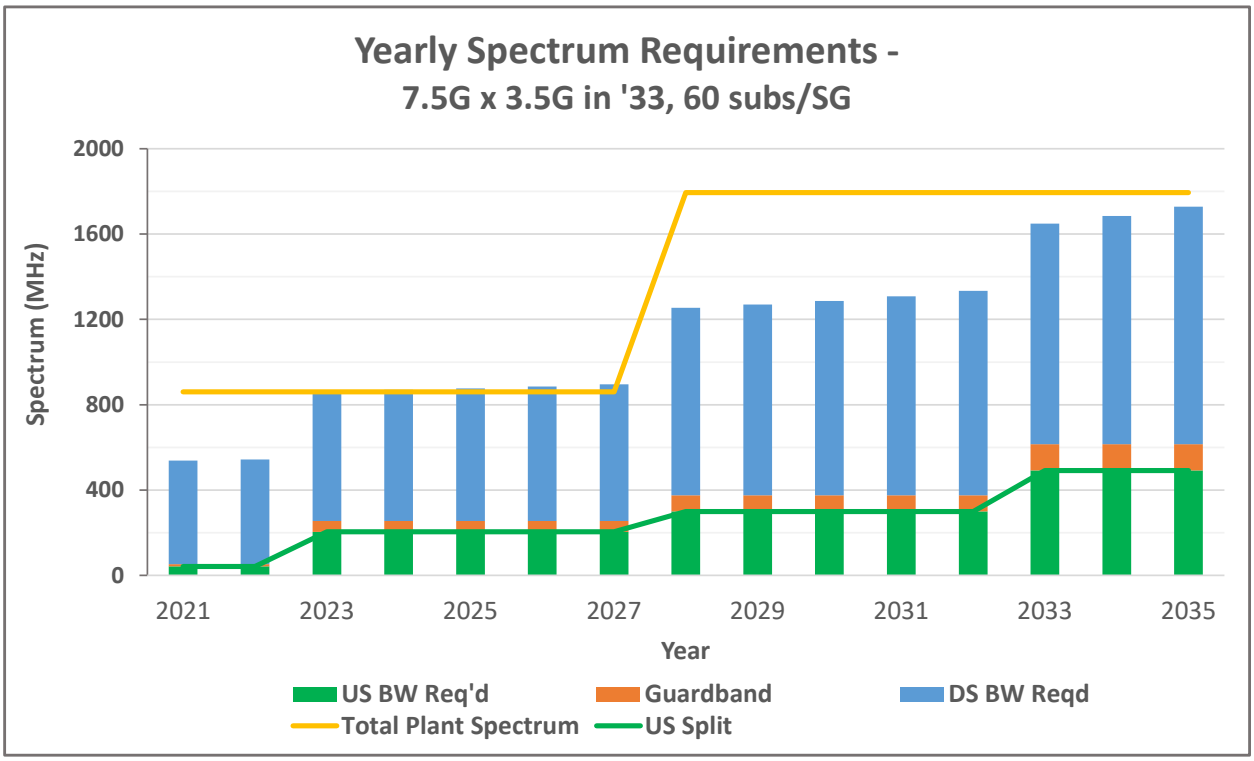


Figure 8 – 1794/492 MHz migration with 7.5G x 3.5G Service Tier, 60 subs/SG

3. Network Evolution: Various Paths Considered

In this section, we examine a sample HFC network area and look at the ways it can evolve, to keep up with subscriber capacity demand.

3.1. Baseline – an N+5 node area to start from

Figure 9 shows a single service group comprised of a single fiber-node serving 945 HP, in "N+5" topology, with 5 to 42 MHz upstream and 54 – 860 MHz downstream frequency bands. 300 MHz of DS is dedicated to legacy digital video, as in 50 SC-QAM channels of 6-MHz width. Assuming 50% subscription take rate, there would be 473 subscribers (Nsub) off this node, but in some scenarios, it could be even higher. These are very high subscriber counts for a single service group (SG), so our baseline scenario starts with 2x2 segmentation of the node into two roughly equal SG. This can be a challenge when the number of homes passed (HP) varies significantly across each RF leg of the fiber-node.

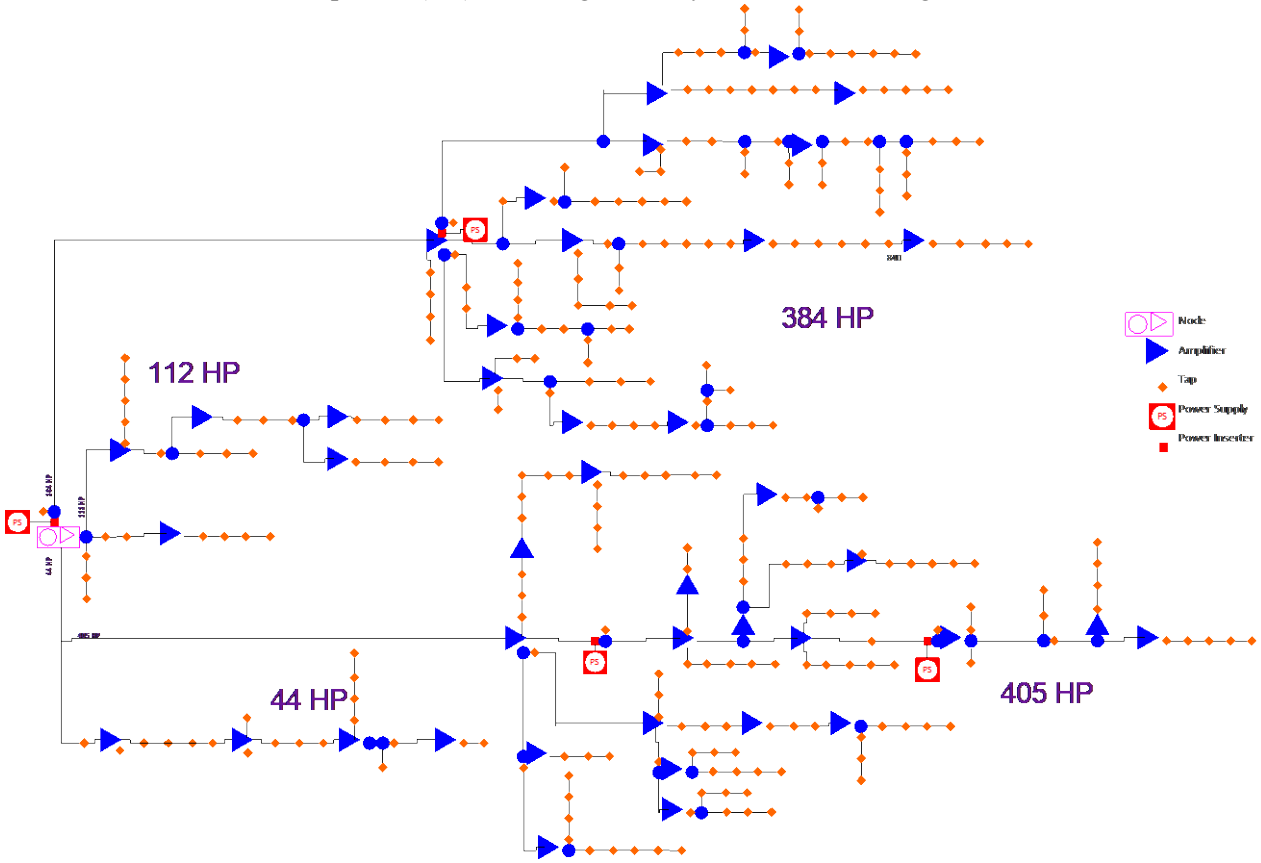


Figure 9 – HFC area under study, 945 HP, uneven distribution across node ports

Grouping node legs with 112 HP and 384 HP in one case, and 44 HP and 405 HP in another gives a balanced 496/449 HP split and is a reasonable step to make. This might mean two SG with roughly 250 subs per SG. As shown in Figure 6, a SG with 113 subs needs an 85 MHz upstream by 2023 and then jumps to a 1794/204 MHz plant by 2027. A larger SG size might push these dates even sooner.

If our example fiber-node is segmented again, going from 2 to 4 SGs, this step would not yield as much of a capacity improvement, given the unevenly distributed number of HPs on the four ports. The two larger RF legs might still have close to 200 subs/SG each, or more. Other options besides business-as-usual node splits need to be considered to meet our bandwidth needs.

3.2. Network evolution options

What upgrade options are available then? Figure 10 shows 3-dimensions to consider: US bandwidth augmentation, DS bandwidth augmentation, and how deep into the network the MAC/PHY functions are placed. There is, however, a network aspect to consider – to reduce "X" in the "N+X" topology, by stringing "fiber-deeper" into the network, and thus to reduce what was here an N+5 topology to N+2 and N+0, for example.

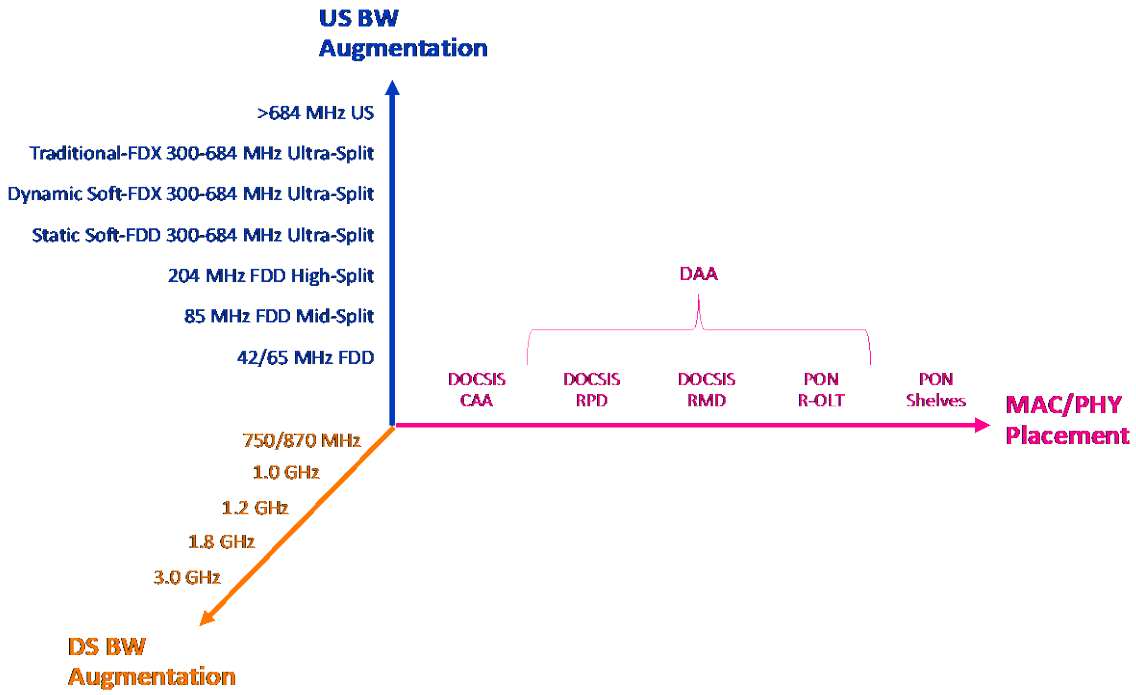


Figure 10 – Possible future evolution path directions for a typical HFC network

3.3. N+2 and N+0 upgrade options

Figures 11 and 12 show how the above baseline fiber-node area under study might evolve to get to an 8-node N+2 or a 15-node N+0 topology, respectively. Lots of changes happen with this topological progression. Table 1 lists the key ones and helps us comprehend the trade-offs involved. As fiber goes deeper, and the RF cascade gets shorter, from N+5, to N+2 and N+0, the number of nodes goes up, from 1 to 8 to 15, and the number or RF amps goes down from 42 to 34 to 0, respectively.

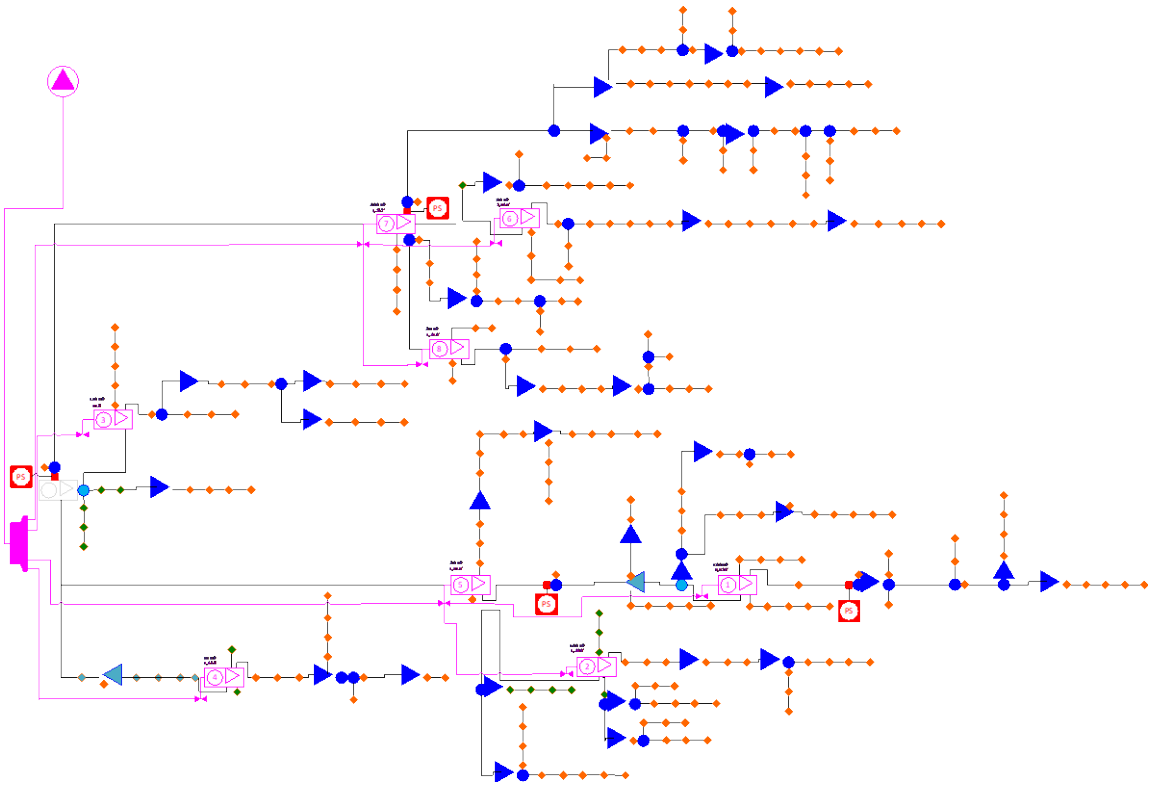


Figure 11 – Single node area under study converted from N+5 to N+2 topology

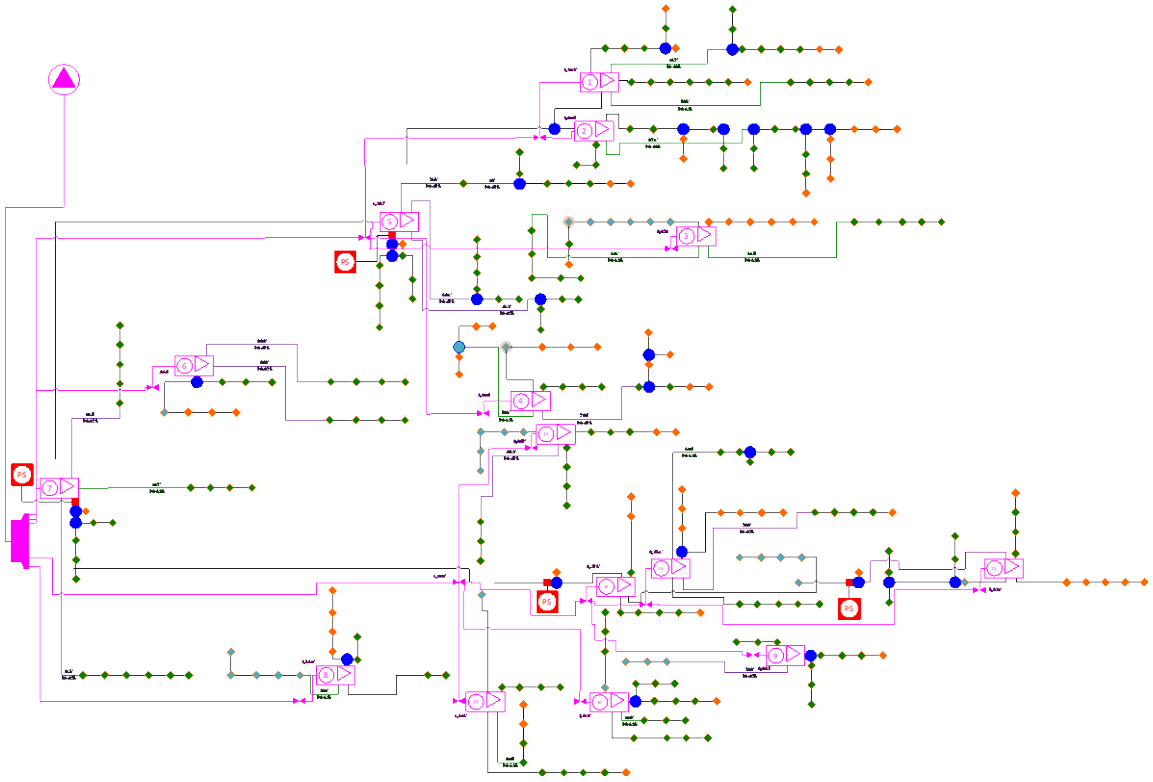


Figure 12 – Single node area converted from N+5 to N+0 “fiber-deep” topology

Repositioning of actives also causes changes to the taps: 34 out of 286 need a new faceplate for N+2 while as many as 208 out of 286 taps need the same for N+0. A 1.9 miles portion of the hardline coax needs fiber overlash for N+2, and as many of 6.5 miles of new coax/fiber for N+0. As a result, the distance from the furthest subscriber to the nearest fiber point/node reduces from 7,000 to 2,500 and 1,600 feet away, for N+5, N+2 and N+0 respectively.

Table 1 – How HFC network attribute change with upgrading N+5 area to N+2 and N+0

Topology:	N+5	N+2	N+0
Number of Standard **Nodes**	**1**	**8**	**15**
Number of RF amps	42	34	0
Number of tap faceplate changes out of # of taps	0/ 286	15/ 286	208/ 286
New plant; miles/ %	0	1.9 miles/ 19%	6.5 miles/ 67%
Fiber to the last subscriber	<7,000 ft	<2,500 ft	<1,600 ft

Note, this is an important point to consider for a future FTTP evolution. Installing new fiber is beneficial, in the sense that it gets the operator closer to the goal of getting fiber all the way to the customer premise. Figure 13 shows what percentage of the hardline plant gets over-lashed with fiber for various N+X scenarios: N+5, N+2, N+1, N+0, fiber to the last active (FTTLA) and fiber to the tap (FTTT), with some of these explained in detail in reference [Venk_SCTE_2016]

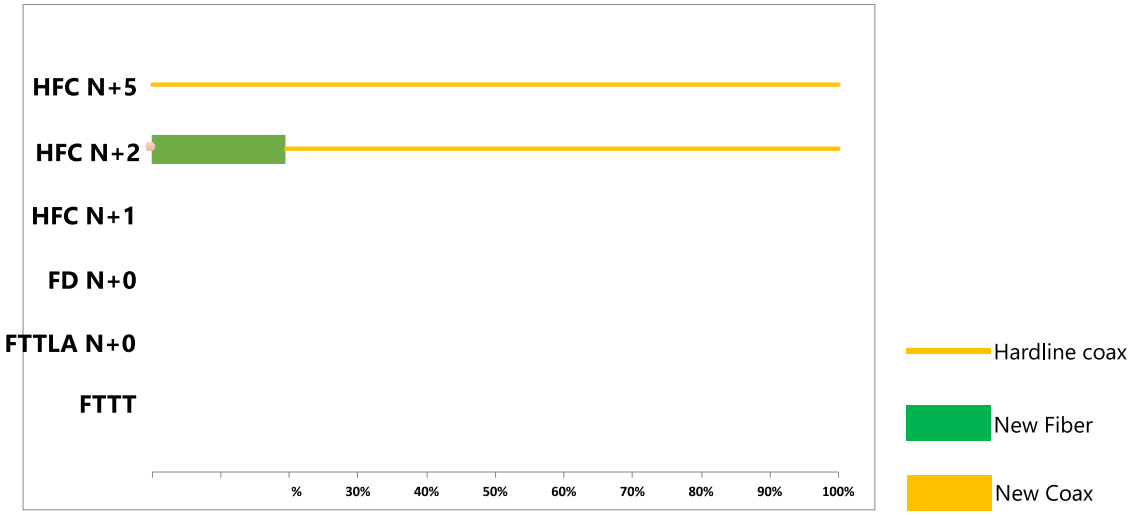


Figure 13 – Percentage of New Fiber + Coax required for various Upgrade Options

However, installing new fiber can be a costly proposition. This is especially true when dealing with underground plant that does not have any preinstalled conduits and thus where substantial digging and trenching is required. How does this cost compare to making improvements along the three dimensions shown in Figure 10? That is what the next section considers.

4. Network Evolution: Total Cost of Ownership Compared

To make sense of the myriad of tradeoffs involved, we present a simplified "total cost of ownership" (TCO) model for various upgrade options. Figure 14 shows an overview of the access network elements accounted for in this cost exercise, as well as a starting point from which to migrate this 5-42 MHz upstream, 54-860 MHz downstream, I-CCAP network. Furthermore, the taps in the network are assumed to need a faceplate upgrade for 1.2 GHz and a complete tap upgrade for 1.8 GHz.

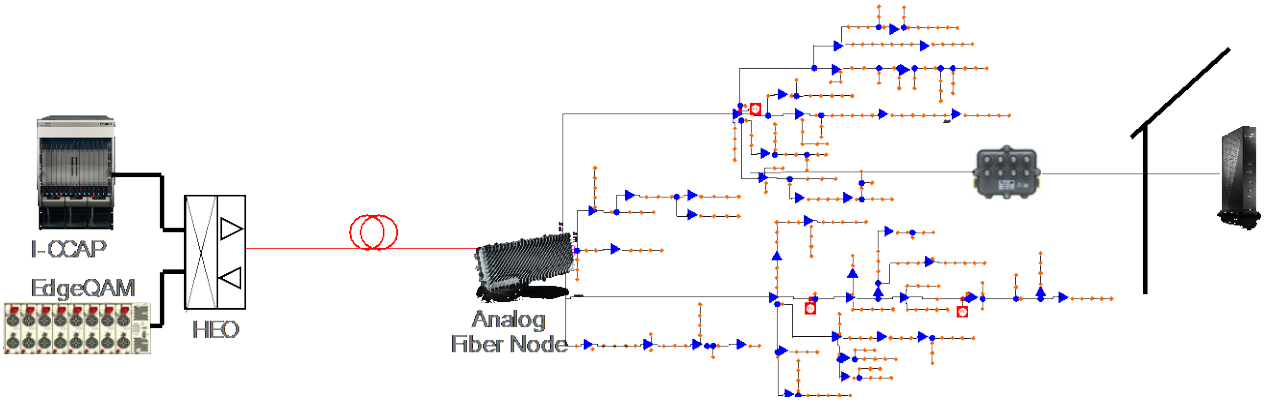


Figure 14 – Overview of Upgrade elements considered in CAPEX calculation

If the topology were to change from N+5 to N+2 or to N+0, the field portion of the network would get upgraded. The field portion would change from that in Figure 9 to that of Figures 11 and 12, respectively. If the I-CCAP were to get converted to DAA, the left-hand portion of Figure 14 would change to that of Figure 15.

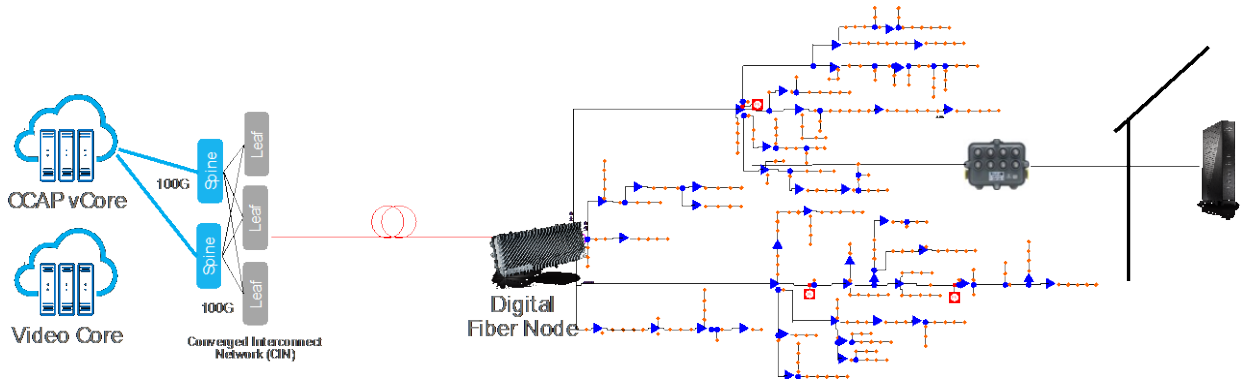


Figure 15 – Network upgraded from I-CCAP to DAA

4.1. Network upgrade paths considered

The nine upgrade scenarios analyzed are shown in Table 2. The names were selected for brevity, so they correspond to labels in the resulting plots. Other elements of the table include: architecture type (i.e. centralized I-CCAP, distributed DAA and PON), number of serving groups (SG), homes passed per SG (HP/SG) and number of nodes. The RF split and top DS frequencies, if applicable, are also shown.

Table 2 – Network upgrade scenarios considered

Name	Architecture	# of SG	HP/SG	# of nodes	RF split	DS BW
CCAP N+5	I-CCAP	2	~480	2	Mid or high	1,218 MHz
CCAP N+2	I-CCAP	4	~240	8	Mid or high	1,218 MHz
CCAP N+0	I-CCAP	4	~240	15	Mid or high	1,218 MHz
FDX N+0	DAA	4	~240	15	108-684	1,218 MHz
FDX-Lite N+5	DAA	2	~480	2	108-396	1,218 MHz
ESD N+5	DAA	2	~480	2	396/492 UHS	1,794 MHz
ESD N+2	DAA	4	~240	8	396/492 UHS	1,794 MHz
10G PON	OLT in hub	15	64	N.A.	N.A.	N.A.
10G R-PON	OLT in node	8	128	N.A.	N.A.	N.A.

4.2. Capital Expenditures (CAPEX)

Figure 16 shows the one-time capital expenditures (CAPEX) for each of the options. These are then normalized to the highest-cost case – that of the 10G Remote-OLT PON (R-PON). The legend shows a breakout for various cost components. On the headend side, there are CMTS (or OLT if PON) and head-end optics (HEO) – which includes SFPs for DAA and PON. On the field side, categories are: node hardware; field hardware consisting of RF amplifiers and fiber enclosures; labor expense to install nodes/amplifiers/fiber enclosures; then new fiber, material and labor; and taps, material and labor (or splitters, material and labor, if PON). On the customer premise side drops, material and labor and CPE – the ONUs for PON – are included.

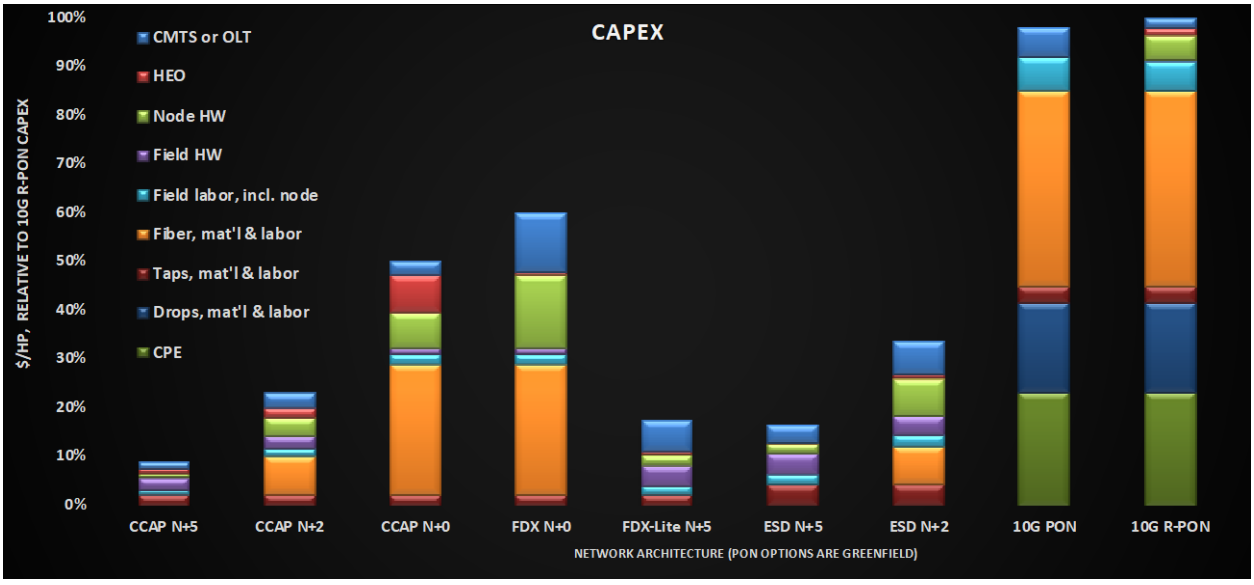


Figure 16 – Normalized CAPEX in $/HP for each of 9 upgrade paths

For I-CCAP, each service group gets 32x4 DSxUS D3.0 SC-QAM channels and 192x48 MHz DSxUS wide OFDM/A D3.1 channels. For DAA, it's 32x4 DSxUS D3.0 SC-QAM channels and 192x96 MHz DSxUS wide OFDM/A D3.1 or D4.0 channels. The fiber – material & construction labor estimates - shown in orange, are dominant in some upgrade cases and are based on $2 per foot for aerial and $12 per

foot for underground, with 80/20% mix of aerial/underground assumed, for a blended cost of $4 per foot. For CCAP N+5 and N+2, only the RF amplifier modules get replaced while the housings remain as they are. For all other cases which contain RF amplifiers, the complete old housing cutout and replacement with a whole new RF amplifier station is assumed. For FDX N+5, ESD N+5 and N+2 cases, given that no FDX nor ESD field RF hardware exists yet, a 40% HW cost premium is assumed over a comparable 1.2 GHz RF amplifier unit. For the PON cases, only the cost of subscriber drops and ONU hardware and install is included – which represents 50% of homes-passed - in order not to overburden PON CAPEX estimates.

4.3. Operating Expenditures (OPEX)

Operating costs are another significant part of the total costs of running a network and are calculated for a certain period of time. When combined with capital expenditures, CAPEX and OPEX add up to the "total cost of ownership" (TCO) over that same period. We select 15 years as a reasonable timeframe to consider, in part because the field hardware upgrades approximately coincide with this 15-year window. The operating costs considered include field network components power consumed, field actives hardware maintenance, hardline coax and/or fiber maintenance and drop lines repair and maintenance. Figure 17 shows these costs normalized to the same CAPEX 10G R-PON costs the y axes of Figure 16 were normalized to.

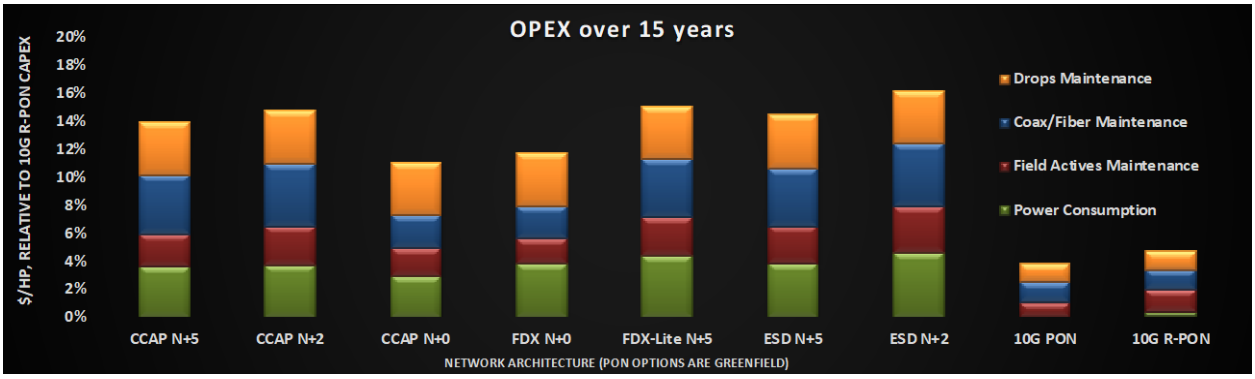


Figure 17 – 15-year Normalized OPEX, in $/HP, for each upgrade path

Power consumption is based on $0.12 per kWh; field actives (nodes and RF amps) are assumed to exhibit % failure over time as shown in figure 18; old coax replacement is based on 1% per year and the new coax/fiber at 0.35% per year; drops maintenance similarly is based on replacing 1% per year. The full costs from the year one is included, however, the costs from years 2 to 15 are time-value-of-money discounted with 5% per year discount rate. The values in Figure 17 is the OPEX cost normalized to the same value the CAPEX costs in Figure 16 are normalized – that of 10G R-PON CAPEX.

With these assumptions, Figure 17 shows that FTTP/PON upgrades have the lowest OPEX coming in at 3% - 4% normalized to the CAPEX costs while the cable options are in the 12% - 16% range. This is not surprising, as the PON options have a significant ~70% reduction in OPEX. So, how critically important is this?

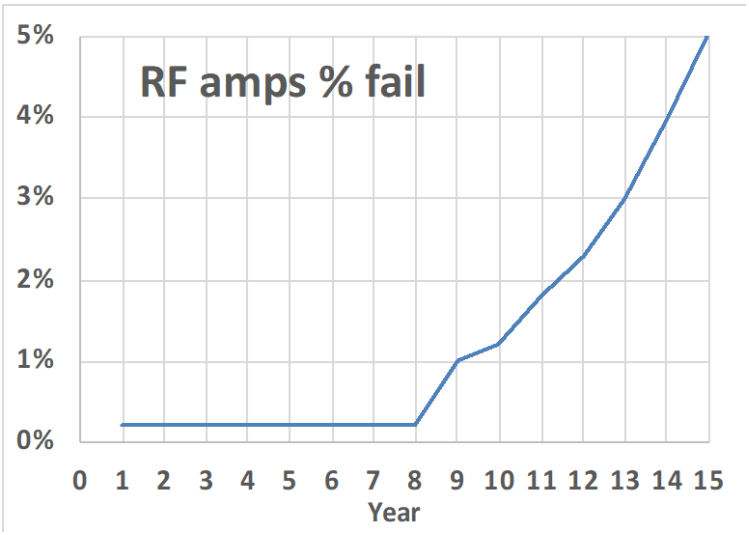


Figure 18 – Field actives failure over time, for estimating maintenance cost

4.4. Total Cost of Ownership (TCO)

The answer is provided by adding CAPEX and OPEX together, resulting in the TCO, as shown in Figure 19. At a first glance, no free lunch opportunity is detected, because those low operating expenses for the PON upgrades sit on the top of a very tall "candlestick" of the PON capital expenses. Nevertheless, TCO ratios did significantly reduce, as compared to those of CAPEX: take 10G R-PON vs. ESD N+5 for example: the CAPEX ratio was ~6, while the TCO ratio is down to ~3.

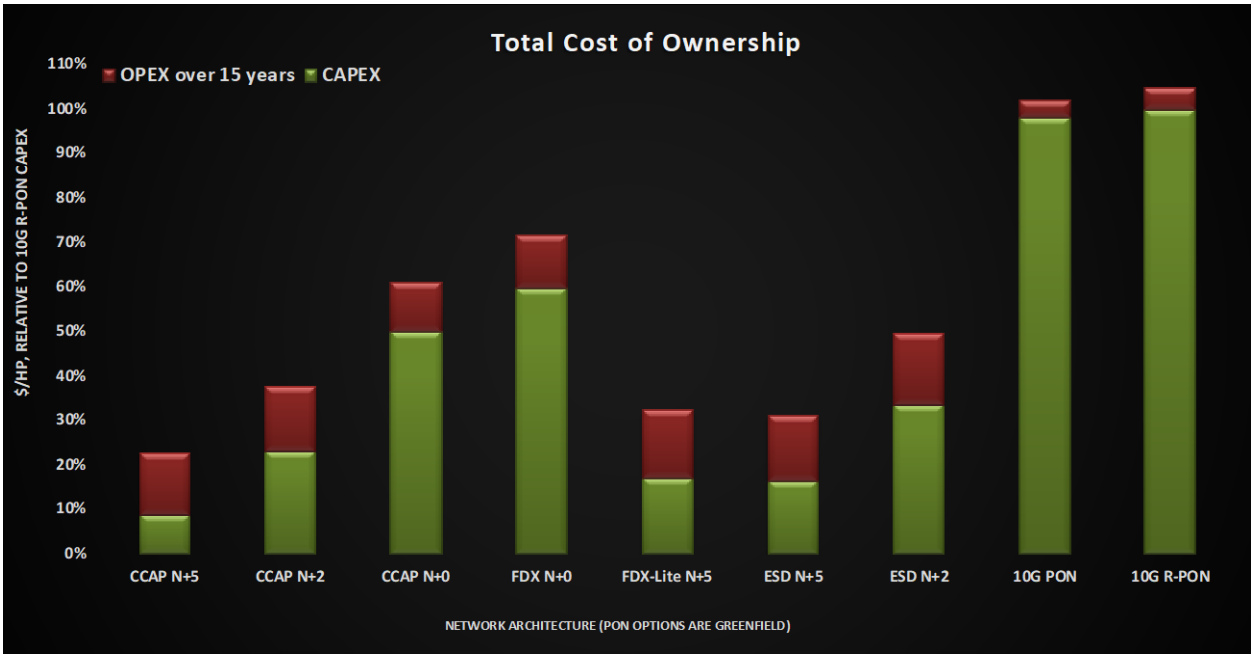


Figure 19 – 15-year Normalized TCO, in $/HP, for each upgrade path

Figure 19 gives the impression that staying with longer cascades (N+5) is the lower cost thing to do, as compared to progressively more expensive N+2, N+0 and FTTP. However, this does not consider the

potential network capacity gains from the other options. Would the network capacity gained justify these additional upgrade costs? The answer to this question is in a later section. However, let's first consider why should one trust these comparisons and numbers anyhow?

There are many factors that go into calculating TCO with many associated assumptions. Isn't there a way to give the reader some additional insights into the assumptions made, and what happens if those assumptions are changed? Monte-Carlo variability analysis comes to the rescue – and to getting these questions answered.

4.5. Sensitivity Analysis - CAPEX

The left sides of Figures 20, 21 and 22 show results of a 100,000-trial run Monte-Carlo analysis for CCAP N+5, ESD N+5 and 10G R-OLT CAPEX, comparatively. The blue areas give "95% confidence interval", based on input parameter assumptions made. Also shown are "base case" numbers – the numbers of Figure 16. What was 100% base case for the right-most 10G R-PON CAPEX, actually ranges from 86.8% to 126.8% for that 95%-confidence-interval, with an average of 105.8%. The ESD N+5 reports base case of 17.1% but ranges from 15.5% to 20.6%; the CCAP N+5 reports 6.7% base case but ranges from 5.9% to 8.5%. Thus, taking the numbers from Figures 16, 17 and 19 as fixed would indeed be a wrong thing to do.

Showing the variability helps but showing where the variability is coming from helps even more. That's why we include the "Tornado charts" on the right side of Figures 20, 21 and 22 – charts that show how much a change in the assumed input variables affect the results. The top variable for the two HFC cases is number of homes passed per study area. We used the 945 HP as the "base case" but allowed the range of 700 to 1,100 HP to also be considered, and that range changes HFC CAPEX significantly. The quantity of bridger and line extender amps follows; then, for ESD N+5 there is "ESD HW addition factor", followed by the labor cost to replace RF amps. For PON, the base case of 20% of underground plant, if changed from 0% to 40%, is the most CAPEX affecting, followed by fiber material and labor and how large of an area under study.

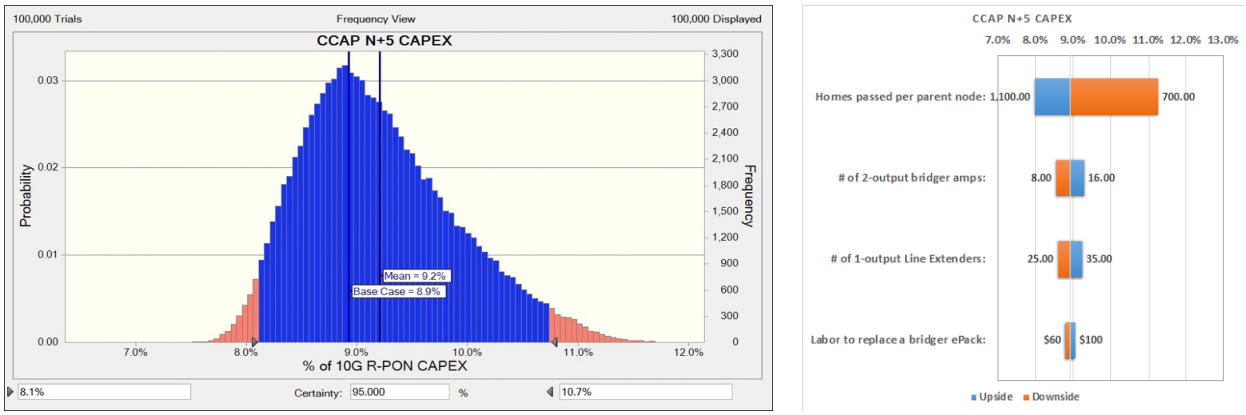


Figure 20 – Monte-Carlo variability and sensitivity analysis of CAPEX for CCAP N+5

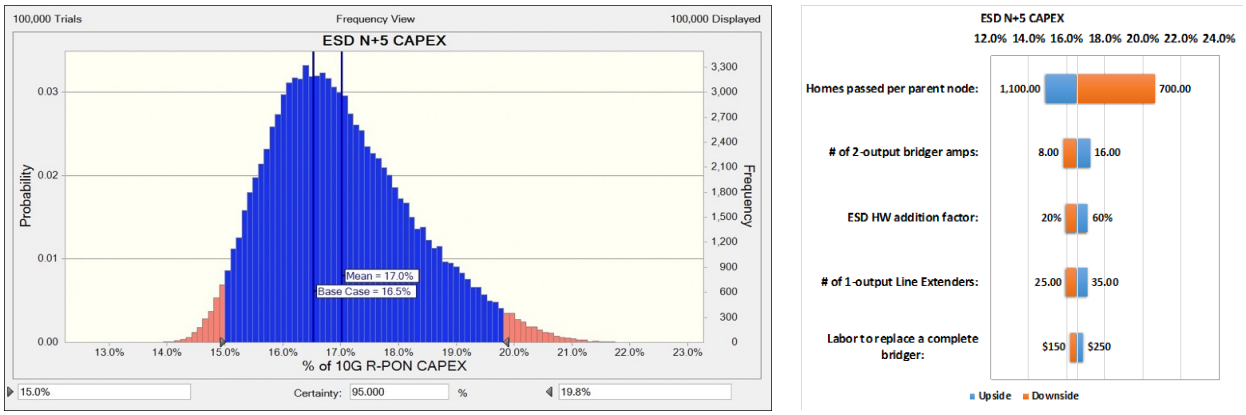


Figure 21 – Monte-Carlo variability and sensitivity analysis of CAPEX for ESD N+5

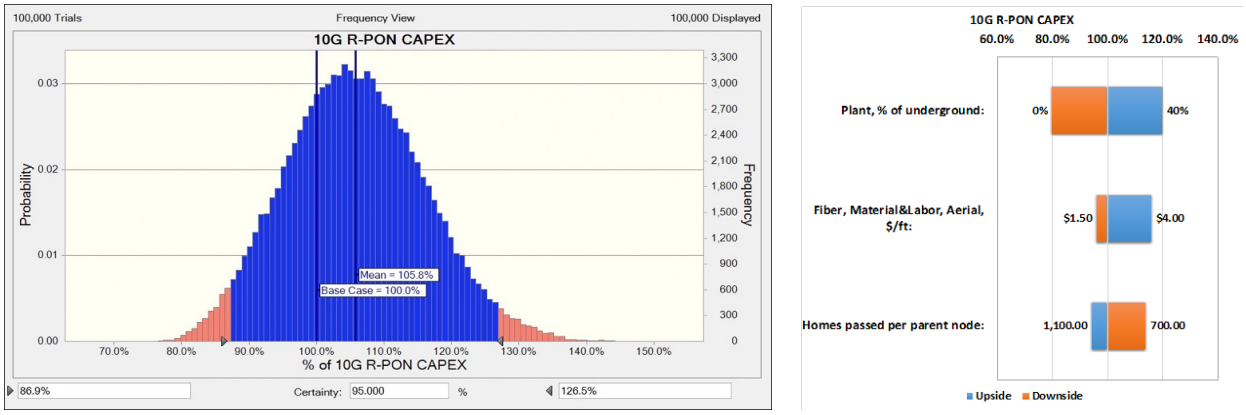


Figure 22 – Monte-Carlo variability and sensitivity analysis of CAPEX for 10G R-PON

4.6. Sensitivity Analysis - OPEX

Given the above insights about CAPEX variability and the causes of the same, it's only fair to look at variability of the OPEX too. Figure 23 shows variance of the average of all HFC upgrade cases and compares it to the variance of the average of the two PON cases. PON average base case numbers were 4.3% (of the 10G R-PON) for PON and 13.9% for the HFC upgrades overall. The 95%-confidence intervals, however, range from 3.1% to 6.4% for PON upgrades and from 11.2% to 19.1% for the HFC ones.

The sources of this variability are shown in Figure 24 – via tornado chart sensitivity analysis for PON and HFC upgrades on average, respectively. Some of the top affecting OPEX variables were educated guesses: how much less of repair and maintenance is the new fiber/coax going to have in comparison to previously installed one, what % of existing plant mileage gets maintained/repaired each year, and the same for the drop cables. In any case, even if some inputs were pure guesses, Figure 24 answers how those factors influence the overall OPEX outcome.

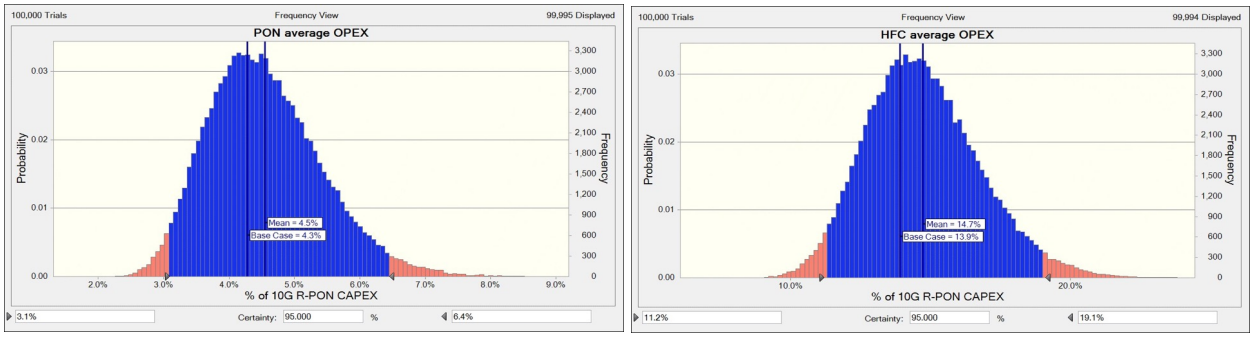


Figure 23 – OPEX variability of PON upgrades and HFC upgrades

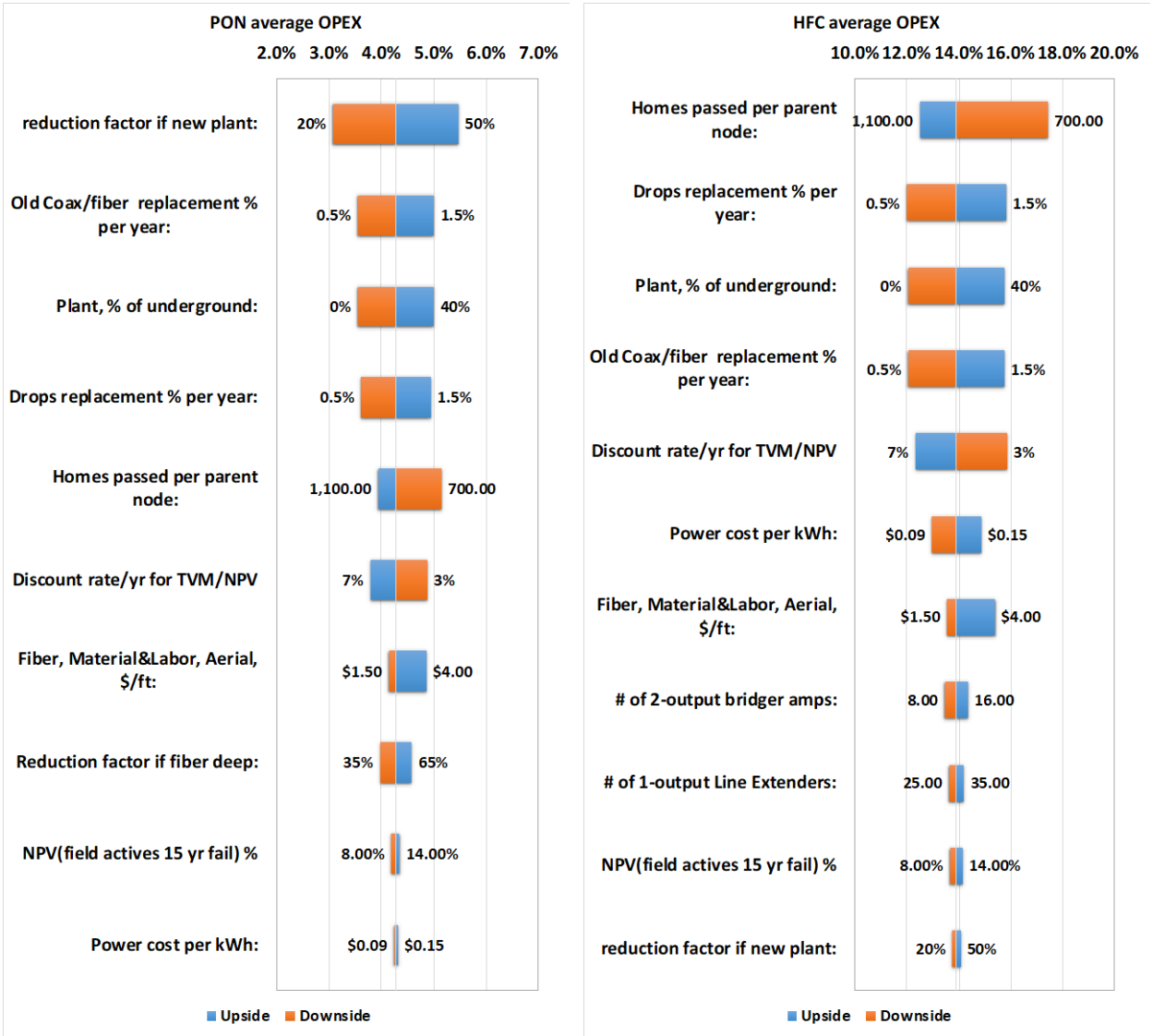


Figure 24 – OPEX sensitivity analysis for PON and HFC upgrades

4.7. Sensitivity Analysis – TCO

With separate variability analysis of OPEX and CAPEX, it may be interesting to see how the variations and sensitivities add up. Figure 25 shows variability of ESD N+5 for CAPEX next to the OPEX, and the resulting overall TCO. The sum of CAPEX & OPEX mean (17% & 15.4%) and base case (16.5% & 14.5%) do add up to the mean and base case of the TCO – 32.4% and 31%, respectively.

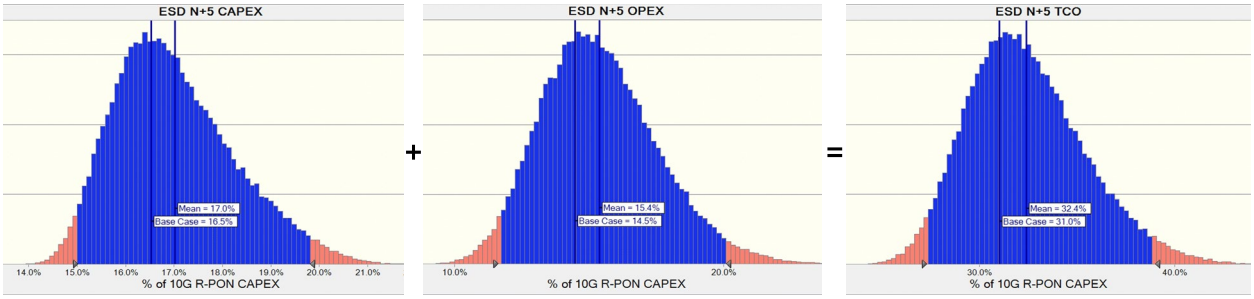


Figure 25 – CAPEX + OPEX variability adding to TCO variability

Figure 26 – Sensitivity of ESD N+5 TCO to various assumption ranges

The most contributing assumptions to the overall variability of ESD N+5 TCO are shown in the "tornado chart" in figure 26. The divisor of homes-passed per parent node features prominently, primarily because of the wide 700 - 1,100 HP range considered, as compared to the fixed 945 HP in the node area under study. Five other factors are OPEX; the rest are CAPEX driven. Hardline coax % replaced, 0.5% to 1.5% range per year, is the next most contributing, the % of underground plant enters via cost of old coax replacement – more costly if more underground; and so on.

The above sensitivity analysis, done by introducing variability into model's assumptions and then detecting what changes and by how much in the model's outputs, helped us gauge the reasonableness of the modeling process. That was but one of the ways of keeping "GIGO" (Garbage In, Garbage Out) from affecting our reasoning. The other way was to "keep it sophisticatedly simple (KISS), but not too simplistic". Thus, KISS and no-GIGO were our guiding principles in making these models.

4.8. Will network capacity gains justify various upgrade costs?

It depends… Figure 27 displays the 2021 max DS + US service tier rates (Tmax_max) that each of the upgrades can potentially support. Since no FDX nor ESD products are available yet, those estimates are left blank. Tmax_max rates are calculated by rearranging basic traffic engineering SG capacity formula of section 2.2 into:

$$\mathbf{Tmax\_max = (C - Nsub * Tavg) / K}$$

The 2021 Tavg values, as in Figure 2, are taken to be ~3 Mbps DS and 0.25 Mbps US, with K = 1; the CCAP 1.2 GHz mid-split options potentially give 6 to 7 Gbps Tmax_max in DS and 500 to 600 Mbps Tmax_max in US – and much more than that (8 x 8 Gbps DS/US) with the two FTTP options.

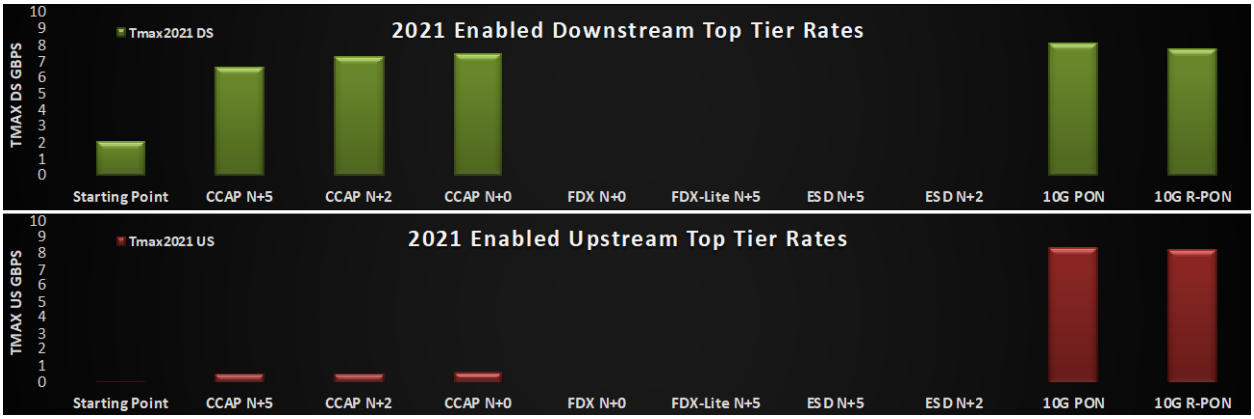


Figure 27 – Year 2021 DS/US Top Tier Rates enabled (FDX, ESD not yet available)

Figure 28 shows what the possible Tmax_max rates would look like in year 2036, for the 9 upgrade options, if next 15-year CAGR were 20% in DS and US, as was assumed in Section 2.3. The 2021 DS Tavg of ~3 Mbps thus grows to ~40 Mbps by 2036, the 2021 US Tavg of ~0.25 Mbps grows to ~1.3 Mbps by the year 2030 and to ~3 Mbps by the year 2036. In the DS, only the ESD version of N+5 has >1Gbps of capacity left for Tmax_max. The other two N+5 cases, CCAP and FDX-Lite, have their capacity drained by Tavg * Nsub (Nsub = 237 subscribers/SG for N+5, while N+2 and N+0 have Nsub = 118 subs). However, both N+2 and N+0 options in the DS seem fine and will still provide 5-6 Gbps Tmax_max. In the US, all three mid split CCAP upgrades tap out. Under the 20% US CAGR, mid-split will hold for US Tmax of 400 Mbps until 2030 but will run out of gas by 2036 – as the lower part of the

Figure 28 shows. This, as well as the ability to offer one gigabit US service, are the reasons to consider going high-split vs. mid-split for the 1.2 GHz upgrades.

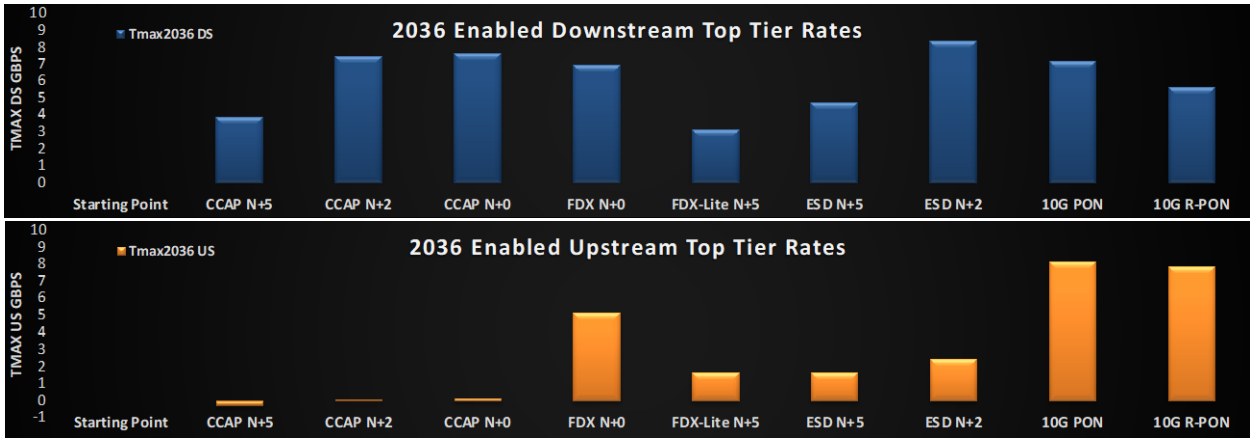


Figure 28 – Year 2036 DS/US Top Tier rates enabled; CCAP mid-split US out of gas

For the next scenario shown in Figure 29, the CCAP supports high-split US instead of mid-split. This step makes the gigabit upstream rates still a reality in 2036 with N+2 and N+0, and while N+5 CCAP, limited by its large SG size, is limited to a Tmax_max of only 600 Mbps.

Furthermore, the DS scenario in Figure 29 shows what happens if the DS CAGR slows and only reaches Tavg = ~25 Mbps. At this DS Tavg, all DS options still have multi-gigabit Tmax_max rates to offer, in 3 to 8 Gbps range. For FDX and ESD, both DS and US look promising – able to support max advertised rates up to 3-8 Gbps in DS and 2-4 Gbps in US. For FTTP/PON, 5-7 Gbps Tmax_max is available in the DS. In the US, Tmax_max of PON is 4x all the other options, other than FDX N+0, where the ratio is closer to 2x. It is interesting and not intuitively expected, however, that by 2036 the 10G PON Tmax in the DS is going to lag - ever so slightly - behind Tmax of ESD N+2 with 396/492 MHz ultra-high-split – granted; under the above assumptions.

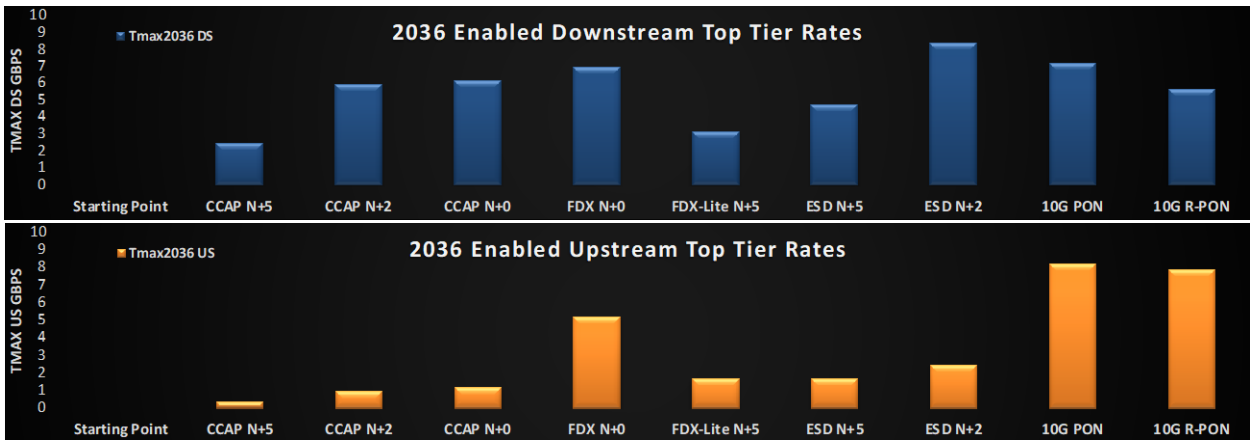


Figure 29 – Year 2036 DS/US Top Tier rates enabled; 15% DS CAGR + high-split CCAP

4.9. TCO analysis Takeaways

If DS capacity remains king, then ESD N+2 provides the highest DS capacities, even better than 10G PON, yet only has a TCO that is half that of PON. The N+2 architecture also leaves the network in an easy position to segment SG size if needed. Fiber is now much closer to every home, so a selective subscriber migration to FTTP now becomes feasible for any customers needing >10 Gbps connectivity. Finally, the deeper fiber from the N+2 HFC network aligns fiber nodes nicely with 5G Mid-band small cells. [ULM_2021] details several case studies showing the convergence of HFC and 5G Mid-band small cells.

The above analysis (with a DS Tavg = 25 Mbps and US Tavg = 3-4 Mbps over the next 12-15 years) shows that the CCAP N+2, FDX-Lite N+5 and ESD N+5 upgrade options are still able to deliver multi-gigabit service DS and at least 1 Gbps (CCAP high-split) or 2 Gbps (FDX-Lite, ESD) in the upstream. TCO for those three options are each roughly at 1/3 of the TCO of the PON - yet are expected to get the job of keeping up with the BW demand done over the next 12-15 years. This in short is the business case for “DOCSIS is here to stay”.

D3.1 or D4.0 or both? Under certain Tavg CAGR assumptions, D3.1’s “fuel tank” has enough gas to last into mid 2030’s, and especially so with the high-split 204/258 MHz option. Under those same assumptions, D4.0 has a potential to meet customer demand into 2040s.

As for comparing the two D4.0 options - The TCO for FDX-Lite N+5 is slightly higher than that of ESD N+5. The key difference, however, is in the DS Tmax_max – ESD N+5 will support 5 Gbps in 2036 while FDL-Lite N+5 can do ~3 Gbps. Upon closer inspection of the CAPEX cost contributors for these two (shown in Figure 16), the CCAP Core for FDX-Lite shows as more costly (and complex) than that for ESD, while upgrading taps as faceplate only for FDX-Lite shows as less costly (and complex) than the whole housing taps upgrade required for ESD. This tradeoff still made ESD CAPEX lower than that of FDX-Lite. The same follows for the TCO comparison, given relatively comparable OPEX for the two.

Why do we call FDX N+5 the FDX-Lite? Because, technically speaking, FDX has been envisioned as applicable to the N+0 topology only, and in order to use it with N+X, some type of “FDX amplifiers” would have to be required. Those could be implemented with echo cancelation, or, given the interference group (IG) elongation issue, via other options, as elaborated in [AYHAM_SCTE_2019]. As a result of IG elongation, however, with these amplifiers, some portions of FDX spectrum may end up getting used in the half-duplex mode only; therefore, the FDX-Lite designation. If the duty-cycle and simultaneity of Upstream and Downstream peak bandwidth bursts ever increase due to new applications (ex: AR or VR apps), then this half-duplex mode of operation required for FDX-Lite may become an impediment in the future that is not existent in the ESD solution. Nevertheless, time will tell when either ESD and/or FDX amplifiers will become available and what the burst requirements will look like for future applications. Additional conclusions can be drawn as this information becomes available.

4.10. Scenarios in which going straight to FTTP/PON makes more sense

There always will be some low-hanging-fruit opportunities in the networks, for example:

- where the total mileage of fiber build required is much lower than in the above study
- where the HP/mile density is much higher
- where the fiber construction cost may be lower than assumed above, either because of lower labor rates and/or because pre-installed conduit runs already exist and running fibers through those will cost significantly less and/or because of predominantly aerial plant

- Where existing fiber routes, perhaps installed with some other purpose - like those fiber routes to the two wireless towers in Figure 30 – can be repurposed for FTTP/PON use
- Innovation driven "disruptive" change in how fiber routing is done, as in this [THEGUARDIAN] article titled: "UK launches £4m fund to run fibre optic cables through water pipes"

Under those conditions, going FTTP would be the right thing to do, both for capacity gained and total cost of ownership reasons. Some of these conditions, dense urban customer base and pre-existing fiber conduits especially, may be behind the recent UK Virgin Media announcement "*Under an ambitious scheme announced today, Virgin plans to upgrade its entire network to full-fiber technology by 2028*", [VIRGIN].

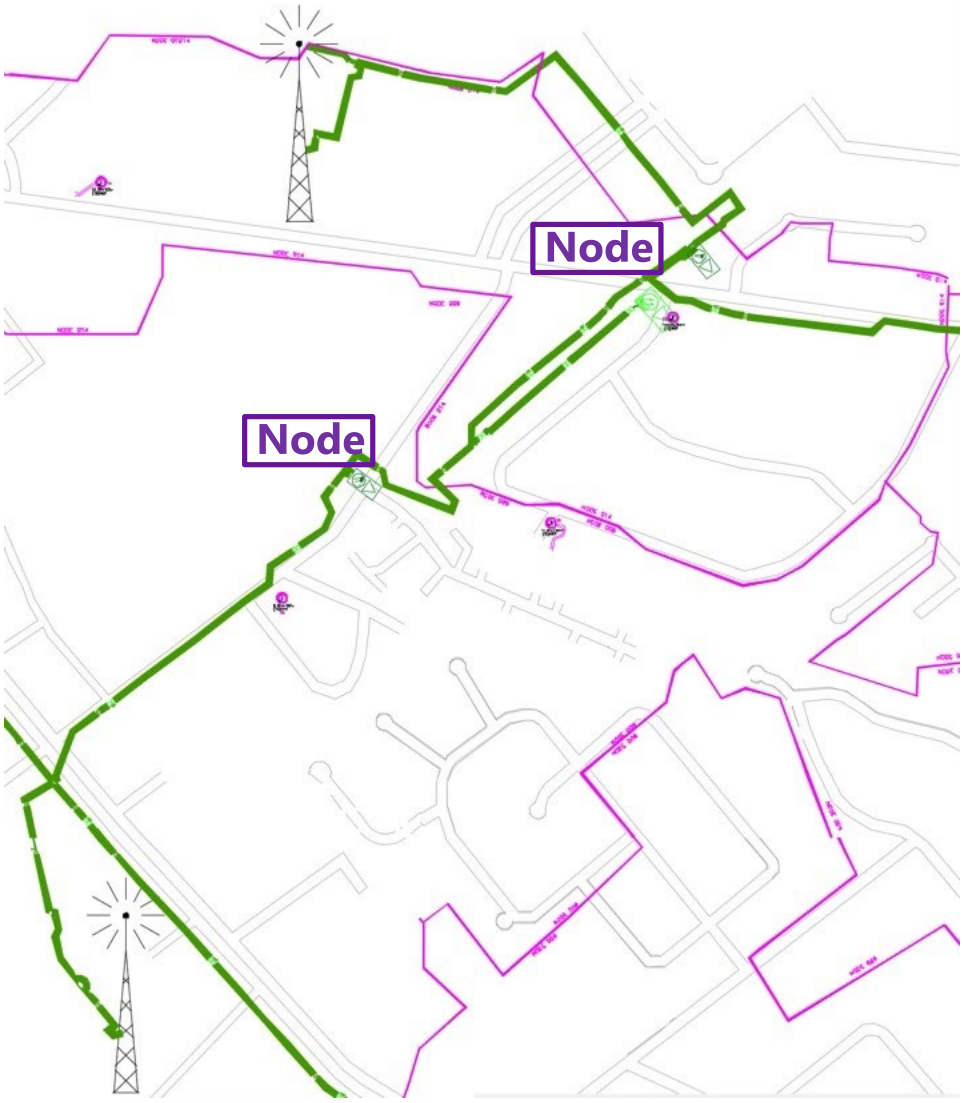


Figure 30 – Example of pre-existing cell-tower fiber routes, with "dark fiber" strands available to support FTTP/PON

4.11. Greenfield scenarios – business as usual or FTTP/PON?

Furthermore, what to build in the green field is the question to revisit often and to revisit hard. Based on our analysis, CAPEX for new builds HFC vs FTTP way - are on par – provided legacy video issues and back-office compatibility are taken out of the discussion. It comes down to the operations folks deciding to continue business as usual or to remove all the obstacles for going the "brave new world" way, and to start doing fiber all the way for all the new / greenfield builds

5. Discussion – The Steps to Get There

5.1. The name of the game is optionality

As many Operators contemplate the jump to FTTP, DOCSIS4.0 continues to offer an extremely viable and cost-effective upgrade path, comfortably enabling multi-gigabit services, including symmetric gigabit services for business and enterprises.

To maximize their optionality in the coming years, operators are beginning to seed new Extended Spectrum DOCSIS (ESD) passive components in their networks as part of their regular preventive network maintenance (PNM) programs. These new passives are emerging with 1.8GHz rated faceplates and housing supporting 2GHz+, and will enable operators to expand their plant's spectrum as required; and as they evolve from centralized CCAP platforms; which today are limited to 1.2GHz operation, to distributed access technologies like Remote MACPHY which are evolving to support ESD operation to 1.8GHz.

The deployment of distributed access architectures will drive fiber deeper into HFC footprints and strengthen the future economics behind fiber to the home and business. This non-regret investment in their footprints will also provide operators with powerful optionality that allows them to choose the right technology for the right place at the right time.

Operators have many options to choose from as they upgrade their networks to support the bandwidth demands of the future. Their choices will likely be different depending on their particular constraints, challenges and expected Return-on-Investment (ROI). In addition, they may opt to utilize one option for some period of time, and then switch to a different option at a later point in time. Some operators may even choose two or more options to utilize in different markets at the same time.

Faced with the potentially expensive decision of deploying FTTP, one approach to making a sound decision may include an economic cutoff analysis where the level of investment required for FTTP overbuild deployment in a HFC node serving area is compared against the uplift in revenue expected from that investment of fiber coupled with the OPEX savings from not having to maintain an HFC plant; which can be subject to variation by different area's depending on the geographic and environmental factors (e.g. is the plant overhead or underground? Is it in or adjacent to a coastal location? Is it a high traffic area prone to damage?). This investment analysis should be performed over the expected life of any uplift in HFC plant, which is typically 10-15 years.

Below is a list of potential options, divided into several different constraint types.

- Near-term and Long-term Greenfield Deployment Areas
  - In these areas, operators may opt to roll out FTTP. This is the logical choice, as it provides long-term bandwidth capacities for the deep future. Since connections must be pulled to each

home anyway, it makes the most sense to utilize the technology with the longest lifespan ahead of it (FTTP).

- Near-term Non-Greenfield Deployment Areas with D3.1-capable HFC plants already in place with needs for ~3 Gbps (or less) Downstream SLAs and ~1 Gbps (or less) Upstream SLAs
  - In these areas, operators may decide to utilize combinations of several near-term D3.1-enabled techniques to augment both SLA capacity and per-subscriber average capacity to extend plant life. These techniques can include:
    - Traditional Node-splits that move towards N+3 or N+2 cascade depths (which can especially help give more per-subscriber average Upstream Capacity which may be required in the future due to COVID bandwidth demand increases)
    - Moving the Upstream split from 42 (or 65) MHz up to 85 MHz or 204 MHz (high-split)
    - Moving the top-end Downstream frequency from 750 (or 860) MHz to 1.2 GHz (Note: Getting to 1.2 GHz Downstream/204 MHz Upstream is a valuable interim step, as it will yield:
      - ~1.1-1.3 Gbps of Upstream Capacity (supporting ~1 Gbps Upstream DOCSIS SLAs)
      - ~5.5-7.5 Gbps of Downstream Capacity (maybe limited to ~2-3 Gbps DS DOCSIS SLAs due to some of the capacity being used by Legacy QAM Video))
  - In these areas, operations can use combinations of several approaches to reduce the spectrum required to support video, including:
    - Using Switched Digital Video (SDV) to reduce spectrum that must be dedicated to Legacy QAM Video
    - Moving to IP Video (which offers some of the same benefits of SDV with the transmission of only requested video streams).
    - Use more efficient video coding schemes to reduce spectrum that must be dedicated to video
  - In these areas, operators may use "Selective Subscriber Migration" to move selected subscribers from the HFC network to a parallel FTTP network that only needs to be installed with adequate infrastructure to support the small subset of subscribers selected for this treatment. These selected subscribers will tend to be those who select the highest SLAs and/or those who consume inordinately large amounts of average Bandwidth Capacity. This approach will extend the lifespan of the HFC network for the majority of "normally-operating" subscribers.

- Long-term Non-Greenfield Deployment Areas with D3.1-capable HFC plants already in place with needs for ~4 Gbps (or greater) Downstream SLAs and/or ~2 Gbps (or greater) Upstream SLAs.
  - Several approaches are permissible, depending on the constraints and challenges facing the MSO.

  - Approach #1 (HFC Focus):
    - Operators can begin by deploying 1.8 GHz-capable taps (in 3 GHz housings) to seed the network. This may be a lengthy activity to ubiquitously cover a majority of HFC plants, so it may need to be started many years before the actual 1.8 GHz service is to be enabled. Prior to D4.0 enablement, these taps can be used to transport traditional D3.1 bandwidth capacities peaking at 1.2 GHz Downstream frequencies and 204 MHz Upstream frequencies.
    - Alternatively, MSOs can begin deploying 1.6 GHz-capable faceplates in existing tap housings to seed the network. This reduces the maximum Bandwidth Capacity by a

small amount, but it could reduce costs and reduce installment times. Prior to D4.0 enablement, these taps can be used to transport traditional D3.1 bandwidth capacities peaking at 1.2 GHz Downstream frequencies and 204 MHz Upstream frequencies.

- At the same time (or slightly later), MSOs can begin deploying D4.0 1.8 GHz-capable/Ultra-High-Split-capable amplifiers (in 3 GHz housings or in existing housings that can support 1.8 GHz operation) to seed the network. This may be a lengthy activity to ubiquitously cover a majority of HFC plants, so it may need to be started years before the actual 1.8 GHz service or Ultra-High-Split service is to be enabled. Prior to D4.0 enablement, these amplifiers can be used to transport traditional D3.1 bandwidth capacities peaking at 1.2 GHz Downstream frequencies and 204 MHz Upstream frequencies.
- At the same time (or slightly later), MSOs can begin deploying D4,0 1.8 GHz-capable/Ultra-High-Split-capable DAA Nodes (in 3 GHz housings or in existing housings that can support 1.8 GHz operation) to seed the network. This may be a lengthy activity to ubiquitously cover a majority of HFC plants, so it may need to be started years before the actual 1.8 GHz service or Ultra-High-Split service is to be enabled. Prior to D4.0 enablement, these DAA Nodes can be used to transport traditional D3.1 bandwidth capacities peaking at 1.2 GHz Downstream frequencies and 204 MHz Upstream frequencies.
- At the same time (or slightly later), MSOs can begin deploying D4.0 1.8 GHz-capable/Ultra-High-Split-capable CM/Gateways into homes to seed the network. This activity can be done in a ubiquitous fashion or can be done in a targeted fashion, targeting the high-end users who are likely to require D4.0 BW capacities in the future. Prior to D4.0 enablement, these CM/Gateways can be used to transport traditional D3.1 bandwidth capacities peaking at 1.2 GHz Downstream frequencies and 204 MHz Upstream frequencies. It is even possible that these D.40 modems can be used on the D3.1 networks to increase the number of bonded OFDM blocks that the CM can feed into a single home.

- Once the D4.0-capable taps and amplifiers and Nodes and requisite CM/Gateways are deployed within a particular HFC network, the higher capacity 1.8 GHz Downstream operation and Ultra-High-Split Upstream operation (up to 684 MHz) can be enabled whenever subscriber bandwidth demands and SLA requirements demand the higher capacities. The need for higher SLAs is likely to dominate this activity. Higher SLAs could be required to support subscriber demands, but it is more likely and expected that higher SLAs will typically be required to respond to marketing challenges from other competing Broadband operators offering high-bandwidth service (>3 Gbps) in the same geographical area. The time-frames for the development of these marketing challenges (and the time-frames for the enablement of D4.0 bandwidth capacities) will obviously vary from area to area depending on the nature of the competition in each area. Each MSO will need to predict when these challenges will arise to permit them to properly phase and schedule their D4.0 rollout activities with adequate lead time.
- As was the case for near-term HFC plants, MSOs can still use combinations of several approaches to reduce the spectrum required to support video in their long-term HFC plants, including:
    - Switched Digital Video or a move to IP Video (which offers some of the same benefits of SDV) to reduce spectrum that must be dedicated to Legacy QAM Video

        - Use more efficient video coding schemes to reduce spectrum that must be dedicated to video
    - As was the case for near-term HFC plants, MSOs can still use "Selective Subscriber Migration" to move selected subscribers from the HFC network to a parallel FTTP network that only needs to be installed with adequate infrastructure to support the small subset of subscribers selected for this treatment. These selected subscribers will tend to be those who select the highest SLAs and/or those who consume inordinately large amounts of average Bandwidth Capacity. This approach will extend the life-span of the HFC network for the majority of "normally-operating" subscribers.

- Approach #2 (FTTP Focus)
    - Operators can begin by targeting high-density subscriber areas (ex: MDUs, high-rises, some city dwellings) for 10G-capable FTTP deployments where the business case analysis indicates there is value. (Note: The business case analysis should be focused on subscriber density, fiber-pull complexity, cost of labor, and whether opportunities exist to share costs of fiber-pulls with other initiatives such as fiber deployments to support cell-tower installations). This FTTP network can initially be built as an overlay to the already-existing HFC network.
    - Operators can extend the above FTTP deployment activities to all other areas, extending the 10G-capable FTTP connectivity to all subscribers in high-density subscriber areas, medium-density subscriber areas, and low-density (rural) subscriber areas. This FTTP network can initially be built as an overlay to the already-existing HFC network.
    - Operators can begin deploying PON ONT Gateways into homes to seed the network. This activity can be done in a ubiquitous fashion or can be done in a targeted fashion, targeting the high-end users who are likely to require higher BW capacities in the future. Since having both PON and DOCSIS CPE equipment co-existing within a home is probably undesirable to subscribers, the FTTP system must be enabled and operational the moment that the first subscriber in the area receives their PON ONT Gateway equipment. It should be clear that during the transition from HFC to FTTP, the two networks will both be operational and running in parallel. A subset of subscribers will be connected to the HFC network, and a subset of subscribers will be connected to the FTTP network. Over time, more and more subscribers will be moved from the HFC network onto the FTTP network.

- Approach #3 (Blended HFC/FTTP Focus)
    - This approach permits the MSO to use both Approach #1 and Approach #2. Each of their subscriber areas can be upgraded using a different approach depending on the particular constraints and challenges in that particular area. Business case analyses will need to be done to determine which approach is most suitable for any particular area.

6. Conclusions

This paper has studied many different evolutionary paths that MSOs can choose to utilize as they evolve their networks to the higher bandwidth capacities needed for future 10G operations. Each operator will clearly choose the path that yields the best return on investment for their own situation, and each operator

will undoubtedly find themselves with different sets of constraints and challenges as they select their future paths.

In general, this study tended to show that for markets where HFC networks already exist, operators will likely want to augment their existing HFC infrastructure to provide the types of bandwidth capacities that are anticipated to compete in the late 2020's and 2030's. Several phases of upgrades are likely. Operators may find a very good return on their existing investment by utilizing DCOSIS 3.1 technologies and traditional node-splitting activities for as long as permissible. This should carry them a long way into the 2020 decade. But to support the bandwidths of the late 2020's and 2030's, operators will likely need to begin seeding DOCSIS 4.0 technologies coupled to DAA architectures into their existing HFC networks, perhaps operating the equipment in a DOCSIS 3.1 fashion for a while before enabling the DOCSIS 4.0 operations. This permits the MSOs to seed the DOCSIS 4.0 equipment over multiple years before having to enable it, which is required since it does require changes to many network elements (Nodes, Amps, Taps, and CPEs). Each DOCSIS 4.0 technology has its own strengths and weaknesses. FDX requires less work on taps (although taps will still likely need to be upgraded to 1.2 GHz operation) at the expense of an N+0 plant. ESD operation may be simpler to diagnose (since it uses familiar FDD technologies) and it may yield higher downstream throughput capacities when N+2 plants are permitted. ESD may also provide cost benefits over FDX.

For Greenfield markets, the study showed that it is beneficial to consider FTTX as a starting point, because the higher cost of initial deployment is a given, and the FTTX technology provides larger bandwidth capacities for the long-term.

In either case, it is clear that MSOs have many good options from which to choose as they augment their networks to support the 10G services of the future.

Abbreviations

10G	10 gigabits per second
5G	5th generation mobile network
AR	augmented reality
BW	bandwidth
CAGR	compound annual growth rate
CAPEX	capital expenditures
CM	cable modem
CMTS	cable modem termination system
COVID	corona virus disease
CPE	consumer premise equipment
D3.1	DOCSIS 3.1
D4.0	DOCSIS 4.0
DAA	distributed access architecture
DFN	distribution fiber network
DOCSIS	data over cable service interface specification
DS	downstream
ESD	extended spectrum DOCSIS
FDX	full duplex
FTTN	fiber to the node
FTTP	fiber to the premise
FTTT	fiber to the tap
FTTX	fiber to the "X"
GIGO	"garbage in, garbage out"
HEO	head-end optics
HFC	hybrid fiber coaxial
HP	homes-passed
I-CCAP	integrated converged cable access platform
IG	interference group
IP	internet protocol
IPTV	internet protocol television
KISS	keep it sophisticatedly simple; or, keep it simple and straightforward
MAC	media access control
Mbps	megabits per second
MDU	multi dwelling unit
MSO	multiple system operator
NPV	net present value
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
OLT	optical line terminal
ONT	optical network terminal
ONU	optical network unit
OPEX	operating expenditures
PHY	physical layer
PNM	proactive network maintenance
PON	passive optical network

QoE	quality of experience
RF	radio frequency
ROI	return on investment
SC-QAM	single carrier quadrature amplitude modulation
SDV	switched digital video
SFP	small form factor pluggable
SG	service group
SLA	service level agreement
TCO	total cost of ownership
US	upstream
VR	virtual reality

Bibliography & References

[CLO_2013] T. J. Cloonan et. al., "Advanced Quality of Experience Monitoring Techniques for a New Generation of Traffic Types Carried by DOCSIS," NCTA Spring Technical Forum 2013, NCTA

[CLO_2014] T. J. Cloonan et. al., "Simulating the Impact of QoE on Per-Service Group HSD Bandwidth Capacity Requirements," SCTE Cable-Tec 2014, SCTE

[EMM_2014] "Nielson's Law vs. Nielson TV Viewership for Network Capacity Planning," Mike Emmendorfer, Tom Cloonan; The NCTA Cable Show Spring Technical Forum, April, 2014

[VENK_SCTE_2016] "Cable's Success is in its DNA: Designing Next Generation Fiber Deep Networks with Distributed Node Architecture" Venk Mutalik, Zoran Maricevic; 2016 SCTE Cable-Tec Expo

[AYHAM_SCTE_2019] "Operational Considerations & Configurations for FDX & Soft-FDX - A Network Migration Guide To Converge The Cable Industry" Ayham Al-Banna, Frank O'Keefe and Tom Cloonan; 2019 SCTE Cable-Tec

[ULM_2019] J. Ulm, T. J. Cloonan, "The Broadband Network Evolution continues – How do we get to Cable 10G?", SCTE Cable-Tec Expo 2019, SCTE

[ULM_2020] J. Ulm, Z. Maricevic, F. O'Keefe, "Is "Unity Gain" Still the #1 Objective? – Maybe YES!", SCTE Cable-Tec Expo 2020, SCTE

[ULM_2021] J. Ulm, M. Zimmerman, S. Eastman, Z. Maricevic, "Overlaying Mid-Band Spectrum Backhaul/Fronthaul onto HFC – A Symbiotic Convergence of Cable & Wireless", SCTE Cable-Tec Expo 2021, SCTE

[SATELLITE] Satellite Internet Access, https://en.wikipedia.org/wiki/Satellite_Internet_access

[FORBES] How Fast Will 5G Really Be?, https://www.forbes.com/sites/bobodonnell/2019/11/19/how-fast-will-5g-really-be/?sh=199eb2cc5cf3

[ITU] ITU-T PON Standards: Progress and Recent Activities, https://www.itu.int/en/ITU-T/studygroups/2017-2020/15/Documents/OFC2018-2-Q2_v5.pdf

[VIRGIN] “Virgin dooms DOCSIS in UK with fiber-for-all plan”, 7/29/2021, https://www.lightreading.com/opticalip/fttx/virgin-dooms-docsis-in-uk-with-fiber-for-all-plan/d/d-id/771160?

[THEGUARDIAN] “UK launches £4m fund to run fibre optic cables through water pipes”, 08/08/2021, https://www.theguardian.com/technology/2021/aug/09/uk-launches-4m-fund-to-run-fibre-optic-cables-through-water-pipes

DOCSIS Time Protocol Proof of Concept

A Technical Paper prepared for SCTE by

Ruoyu (Roy) Sun
Ph.D., Lead Architect
CableLabs
858 Coal Creek Cir, Louisville CO, 80027
303-661-6789
r.sun@cablelabs.com

Jennifer Andreoli-Fang, Aaron Quinto, Mark Poletti, CableLabs, Inc.

Charles Cook, Ryan Tucker, Vikas Sarawat, Praveen Srivastava, Charter Communications, Inc.

John Chapman, Eric Houby, Cisco Systems, Inc.

Elias Chavarria Reyes, Wen Chun Wei, Vincent Cho, Hitron Technologies, Inc.

1. Introduction

Demand for broadband services continues to grow. While wireline access technologies have supported traditional broadband services, a significant amount of broadband traffic is supported by wireless access technologies. Multiple System Operators (MSOs) have deployed Data-Over-Cable Service Interface Specification (DOCSIS) networks for many years to provide wireline broadband access for their customers. Because of the increasing demand for wireless broadband access, MSO interest has grown in deploying and backhauling their own 5G wireless access networks using their existing DOCSIS infrastructure.

Radio access technologies such as the fifth generation (5G) New Radio (NR) require accurate alignment in frequency, phase, and time of day to minimize interference and improve efficiency. Several MSOs have recently acquired Citizen Broadband Radio Service (CBRS) spectrum licenses to deploy time division duplex (TDD) 5G base stations (BSs). 3rd Generation Partnership Project (3GPP) standards specify a time difference of no more than 3 µs between cells which requires timing accuracy to be within 1.5 µs of a Primary Reference Time Clock (PRTC) [1].

Outdoor base station antennas can be oriented to be line-of-sight with Global Positioning System (GPS) satellites for the reception of timing signals. If line-of-sight cannot be achieved, as in the case of indoor base stations (e.g., Femtocells), an alternative timing source with 1.5 µs accuracy is required. IEEE 1588 Precision Time Protocol (PTP) [2] can provide this level of accuracy.

PTP was developed and specified frequency and phase synchronization across Ethernet transmission links using timestamps to address latency and jitter issues. PTP enables Ethernet-based networks to be used as backhaul links for 3GPP Long-Term Evolution (LTE) and 5G systems. However, PTP was not designed for DOCSIS networks.

When DOCSIS is lightly loaded, it is possible to support PTP "over-the-top" of DOCSIS [3]. However, DOCSIS was not designed to support highly accurate PTP "over-the-top". Therefore, a highly loaded DOCSIS system will not be able to support PTP to the accuracy levels required by TDD 5G base stations without the use of DOCSIS Time Protocol (DTP). DTP was introduced in DOCSIS 3.1 to reliably support PTP on DOCSIS networks.

DTP, invented by Cisco [4] and included in CableLabs DOCSIS specifications [5] and [6], enables DOCSIS networks to deliver PTP to wireless base stations. DTP establishes PTP-to-DOCSIS interfaces at the Cable Modem Termination System (CMTS) and at the Cable Modem (CM). DTP allows the timing and frequency system of the CMTS, the Hybrid Fiber Coax (HFC) plant and the CM to be a timing bridge. DTP accurately takes the PTP timing source at the input of the CMTS, and replicates it at the output of the CM with the correct timing offsets to take into account all the delays through the DOCSIS system. Figure 1 shows PTP and DTP operating on a DOCSIS system. Components of a generalized DOCSIS system are shown in green.

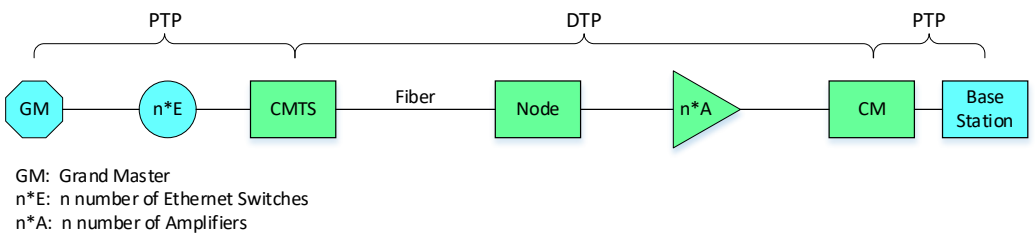


Figure 1 - PTP and DTP on an Integrated CMTS Architecture

The DTP protocol runs between the CMTS and the CM. The CMTS receives a PTP timestamp on a PTP slave port and synchronizes its internal clock to that timestamp. The CMTS synchronizes all DOCSIS timestamps to this internal clock, making the DOCSIS timestamp traceable to a PTP timestamp. The CMTS uses the ranging capabilities of DOCSIS, and delays through the CMTS and CM, to calculate timing offsets for an accurate PTP timestamp at the CM. The CM then regenerates PTP and sends it to the BS.

Many vendor companies have developed or are developing solutions that support DTP. Support for DTP is planned for the Cisco CMTS (cBR-8), and remote PHY device (RPD) products in a future software release (targeted for 2022). In January 2021, Hitron launched the ODIN-1112, the world's first DOCSIS 3.1 modem to support DTP. The ODIN-1112 supports operating as either a DTP master or a DTP slave. By pairing the ODIN-1112 with a small cell gateway, cable operators can leverage their existing DOCSIS networks to offer 4G/5G services. Hitron is dedicated to helping cable operators capture new opportunities in 5G and will continue expanding its product portfolio to enable not only outdoor but also indoor small cell deployments.

CableLabs, Charter Communications, Cisco and Hitron initiated proof-of-concept (PoC) testing for DTP in Q2 2020. This paper presents the DTP PoC test plan, methodology and up-to-date status.

2. DTP PoC Test Plan

DTP PoC testing started in September 2020 at both CableLabs and Charter. Cisco also conducted tests in their lab. The DTP PoC testing has three phases. Phase 1 evaluates the DTP time error in a lab environment with minimum fiber and coaxial cable length, without amplifiers, without traffic load, etc. The CM synced with DTP is plugged into an LTE base station to test the wireless signal time accuracy over the air (OTA). We also verified three manual calibration methods that allow changing the true ranging offset (TRO) or DTP time adjustment in the CMTS and the CM. The time error (TE) for all test scenarios and cases is compared with the DTP time error budget as defined in [5]-[9]. Phase 1 testing concluded in July 2021. The phase 1 methodology, setup, and results are reported in [9].

Phase 2 is designed to evaluate DTP performance in sophisticated configurations that are representative of anticipated field deployments. Different downstream (DS) and upstream (US) loads will be added to the

Table 1 - Phase 2 Test Plan

Parameter		Baseline test value	Comparative test values	Extreme value (optional test)
DS load		0	25%, 50%	75%
US load		0	25%, 50%	75%
Coax length (R-PHY to CM)		a few meters	1/4 and 1 mile	
Fiber length (Router to R-PHY)		tens of meters	25 km	
Number of amplifiers		0	1, 2	
CMTS config change (DS)	Interleaver	2	1	16
	Modulation	4096-QAM	1024-QAM, 256-QAM	
	Cyclic prefix	1 (1.25 µs, 256 samples)	2 (2.5 µs, 512 samples)	3 (3.75 µs, 768 samples)
CM config change (US)	Frame size	K = 6	K = 9, BW ≥ 72 MHz K = 18, BW < 48 MHz	
	OFDMA modulation	256-QAM	64-QAM	1024-QAM
	Cyclic prefix	6: 256 samples	4: 192 samples	

HFC plant to assess DTP and PTP performance. The coaxial cable and fiber length and number of amplifiers will be adjusted to determine DTP performance. The impact of HFC network configurations will be evaluated. These configurations include: DS interleaver, modulation and cyclic prefix, US frame size, modulation, and cyclic prefix. The phase 2 test plan is summarized in Table 1. For each parameter, a set of baseline, comparative, and extreme test values is defined. The extreme values are optional for phase 2 testing. Only one parameter will be changed for each test case to reduce the number of network configuration combinations. Phase 2 testing is planned for Q3 and Q4 2021.

The DOCSIS 3.1 timestamp transmitted from the CMTS to the CM is delayed while propagating downstream through the HFC network. DTP is designed to calibrate the DOCSIS 3.1 timestamp by using the TRO. DTP automatically compensates for the symmetrical (identical in DS and US) time error in the HFC network. However, the CMTS, RPD, and CM could introduce asymmetrical time errors that reduce the time accuracy of DTP. If pre-calibrated asymmetry values are known, DTP also compensates for asymmetrical time errors. Such asymmetrical time errors need to be measured in the lab before deployment for each combination of CMTS, RPD, and CM hardware and software versions. CableLabs/Kyrio established a Network Timing Lab to conduct these kinds of tests and collect data to calibrate the asymmetrical time error. This calibration data will be distributed by an Amazon Web Service (AWS) cloud server to enable the CMTS to calibrate the asymmetrical time error in the field automatically. The AWS cloud server design is presented in [10]. Once the AWS cloud server is developed and the corresponding automatic calibration feature is added to the CMTS, phase 3 tests will be started to validate the concept of this feature.

3. DTP Performance

3.1. Test Setup

The measurement setup for DTP performance testing is illustrated in Figure 2. Because there is no time measurement equipment available that supports DTP, DTP performance is measured between the input PTP timestamp to the CMTS and the output PTP timestamp of the CM. The Calnex Paragon-X is used to measure the DTP performance. Port 1 (PTP master) on the Paragon-X is connected to the Cisco integrated CMTS (I-CMTS) cBR-8 using PTP. The cBR-8 is connected to a Network Convergence System (NCS) router and a Cisco RPD via fiber. The RPD connects to the Hitron ODIN-1112 CM via a coaxial cable.

DTP is used between the cBR-8 and the CM. The CM is connected to port 2 (PTP slave) on the Paragon-X using PTP. The Paragon-X uses GPS and a Rubidium (Rb) clock to calibrate frequency. The Paragon-X compares the PTP timestamp received on port 2 against the timestamp generated by port 1.

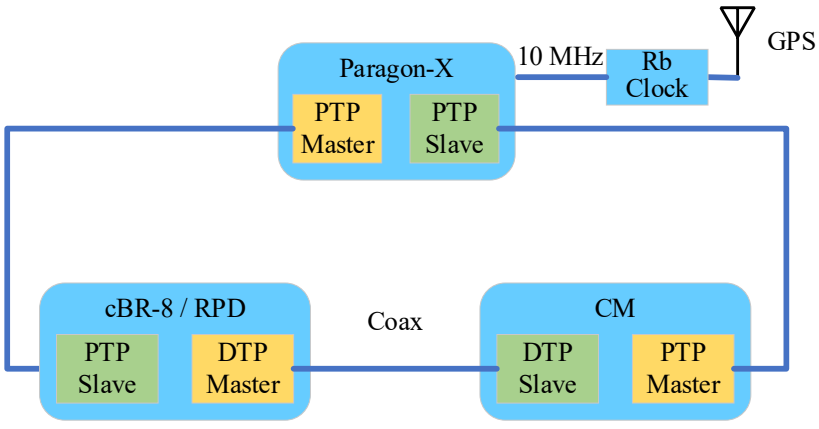


Figure 2 - Block Diagram of DTP Performance Test Setup [9]

The measurement accuracy of the Paragon-X is ±5 ns. CableLabs verified the performance of the Paragon-X before the DTP performance test. The method and results are reported in [9].

3.2. Time Error Budget

3GPP technical specifications 36.133 [1] and 38.133 require LTE and 5G NR base stations to have phase synchronization better than 3 µs between BSs. It indicates that each BS must have synchronization better than 1.5 µs. PTP (IEEE 1588) and ITU-T G. 8271 require a synchronization better than 1.5 µs.

The DTP TE budget for a typical Distributed Access Architecture (DAA) scenario is provided in Table 2, along with the TE budget calculated for the actual DTP test setup with an RPD used in the PoC testing. The clock used in the Paragon-X is compared with a delayed version of itself, so the PRTC does not apply. The test setup only uses one Class B boundary clock (BC) in the NCS. The total "Ethernet and Dynamic Aspects of Ethernet TE Budget" is 470 ns in total. The RPD is Class A. No node or amplifier is used. Hence, the "DOCSIS Network TE Budget" is 510 ns. A base station is not included in this test setup. The total TE budget is 980 ns which is much smaller than the required 1500 ns in PTP and 3GPP specifications. The DTP performance TE test results will be compared with this 980 ns TE budget as a pass/fail criteria.

Table 2 - DOCSIS and HFC TE Profile for the DTP Performance Test Setup with RPD [9]

Budget Component	DAA			DTP test setup		
	n	@	TE	n	@	TE
PRTC *(Class A is 100 ns, Class B is 40 ns, ePRTC is 30 ns)*	Class A		100	Class A		0
Network holdover and PTP rearrangements			200			200
Network dynamic TE and SyncE rearrangements			200			200
T-BC *(Class A is 50 ns, Class B is 20 ns)*	4	50	200	0	A@50	0
T-BC *(Class C is 10 ns, Class D is 5 ns)*				1	B@20	20
Link asymmetry			50			50
Ethernet and Dynamic Aspects of Ethernet TE Budget			**750**			**470**
I-CMTS/RPD/RMD *(Class A is 200 ns, Class B is 100 ns)*	Class A		200	Class A		200
DTP			50			50
HFC path			10	DAA		10
HFC node			10	DAA		0
HFC amp/LE	N+3	10	30	N+0	10	0
CM *(Class A is 250 ns, Class B is 100 ns)*	Class A		250	Class A		250
DOCSIS Network TE Budget			**550**			**510**
Rearrangements and short holdover in the end application			0			0
Base station slave or intra-site distribution	Class A		50	Class A		0
Base station RF interface			150			0
Base Station Network TE Budget			**200**			**0**
Total TE Budget			**1500**			**980**

3.3. Phase 1 Test Results

DTP performance tests were conducted in CableLabs, Charter, and Cisco. The testbeds used an upstream Orthogonal Frequency-Division Multiple Access (OFDMA) channel. DTP is manually calibrated by setting the TRO at the CM to compensate the asymmetrical constant TE. Five runs of data were collected at CableLabs, five were collected at Charter, and one was collected at Cisco. Each test is set to either 1076 s, one hour or three hours. The results are summarized in Table 3.

Many TE statistical results were analyzed by the Paragon-X including two-way time error, constant time error (cTE), which is the average two-way TE, maximum and minimum two-way time error, dynamic TE (dTE), maximum time interval error (MTIE) and time Allan deviation (TDEV). Descriptions of these concepts and results were presented in [9]. In this paper, we only focus on the most important parameters, as listed in Table 3.

The results show that the cTE is smaller than 31 ns. The max TE and min TE results are within ±200 ns, which meet the 980 ns TE budget requirement discussed in subsection 3.2. The only exception is run 4 in the Charter testbed. A PTP Delay_Response message in run 4 arrived at the Paragon-X approximately 9 s later than expected. This was not observed in the other four runs in the Charter and CableLabs data and is likely a test anomaly that can be discarded.

Table 3 - Time Error Results with RPD [9]

All TE results unit in ns			Test setup with RPD				
			Two-way Time Error				Peak-to-peak dynamic TE
	Run	Time duration	Mean (cTE)	Max	Min	Max-Min	
CableLabs	1	3600 s	30	46	-47	163	157
	2	3600 s	13	146	-94	240	220
	3	3600 s	31	118	-47	165	144
	4	3600 s	21	138	-67	205	183
	5	3600 s	29	125	-81	206	190
Charter	1	3 hours	-29	97	-151	248	226
	2	3 hours	-30	102	-146	248	225
	3	3 hours	-19	110	-183	293	231
	4	3 hours	13607	9,404,370,078	-147	9,404,370,224	9,449,878,963
	5	3 hours	-26	115	-141	256	222
Cisco	1	1076 s	-20	121	-122	242	225

4. LTE Timing Performance Using DTP in Backhaul

Section 3 verified that DTP provides synchronization much more accurately in a lab environment than the required TE budget of 980 ns. In this section, DTP is used in mobile backhaul to check the LTE OTA signal time accuracy.

4.1. Test Setup

The test setup is illustrated in Figure 3. The grand master (GM) clock connected to a GPS receiver provides the PTP time source for the DOCSIS system. Both I-CMTS and DAA R-PHY architectures use PTP as an input timing source. In the I-CMTS architecture, the CMTS exchanges DTP messages with the CM. In the R-PHY architecture, the CMTS Core exchanges DTP messages with the CM. The CM uses DOCSIS 3.1 timestamps and DTP as its timing source, and provides a synchronization signal for LTE BS1 by PTP. We set up another LTE base station, BS2, that is synchronized to a GPS clock to compare

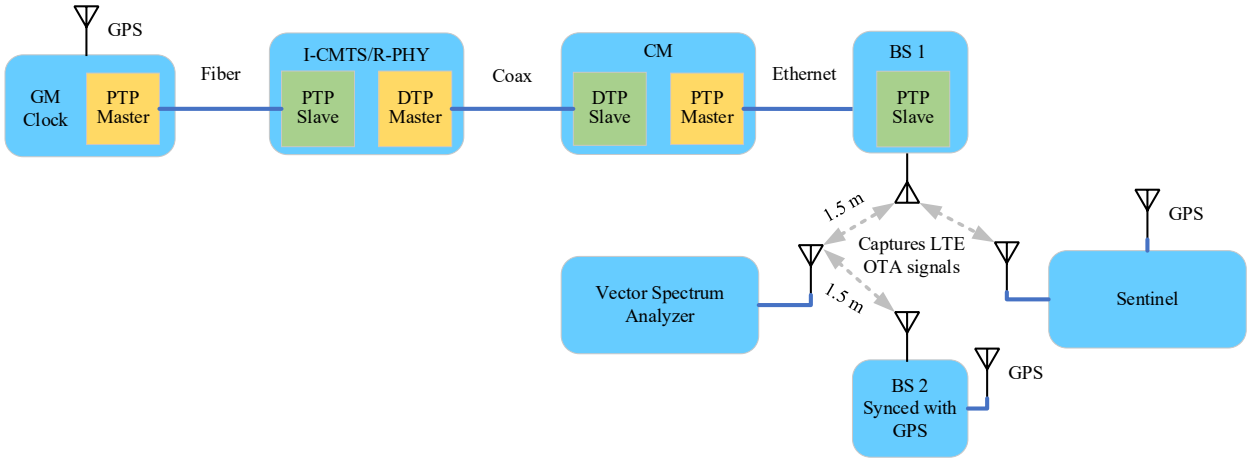


Figure 3 - Block Diagram of LTE OTA Setup

Table 4 - DOCSIS and HFC TE Profile for the LTE OTA Test Setup [9]

Budget Component	VSA			Sentinel		
	n	@	TE	n	@	TE
PRTC *(Class A is 100 ns, Class B is 40 ns, ePRTC is 30 ns)*	Class A		100	Class A		100
Network holdover and PTP rearrangements			200			200
Network dynamic TE and SyncE rearrangements			200			200
T-BC *(Class A is 50 ns, Class B is 20 ns)*	0	A@50	0	0	A@50	0
T-BC *(Class C is 10 ns, Class D is 5 ns)*	1	B@20	20	1	B@20	20
Link asymmetry			50			50
Ethernet and Dynamic Aspects of Ethernet TE Budget			570			570
I-CMTS/RPD/RMD *(Class A is 200 ns, Class B is 100 ns)*	Class A		200	Class A		200
DTP			50			50
HFC path	DAA		10	DAA		10
HFC node	DAA		0	DAA		0
HFC amp/LE	N+0	10	0	N+0	10	0
CM *(Class A is 250 ns, Class B is 100 ns)*	Class A		250	Class A		250
DOCSIS Network TE Budget			510			510
Rearrangements and short holdover in the end application			0			0
GPS receiver PRTC clock	1	A@100	100	0	A@100	0
Base station slave or intra-site distribution	2	A@50	100	1	A@50	50
Base station RF interface	2	150	300	1	150	150
Base Station Network TE Budget			500			200
Total TE Budget			1580			1280

with BS1. The two BSs radiate LTE signals over the air. The Calnex Sentinel collects the LTE primary synchronization signal (PSS), the secondary synchronization signal (SSS) and decodes the time of day from the BS1 LTE signal. The Sentinel uses GPS and an internal Rubidium clock as a reference to evaluate the accuracy of LTE timing. The measurement accuracy of the Sentinel is ±100 ns.

A vector spectrum analyzer (VSA) is used as another LTE OTA measurement method. It collects the spectrum of both BS1 and BS2, then converts them to the time domain by an inverse fast Fourier transform (IFFT). Thus, the downlink (DL) time-domain bursts in the TDD LTE signals can be compared. The VSA antenna is equal distance from both BS1 and BS2 antennas. By properly selecting the measurement bandwidth and fast Fourier transform (FFT) size, the VSA measurement accuracy achieved was on the order of tens of ns. The measurement accuracy of the VSA method is constrained by the LTE signal burst uncertainty due to the BS amplifier and local oscillator performance, which can be off by as much as 10 µs. BS1 and BS2 are the same model from the same manufacturer using the same hardware and software, so the relative uncertainty is much smaller than 10 µs. With the VSA method, it is straightforward to check the TDD-LTE signal bursts in the time domain, but it is not a high-accuracy method to judge if the LTE OTA signals meet the 3GPP synchronization requirement.

4.2. Time Error Budget

The TE budget for the LTE OTA setup is listed in Table 4. In comparison to the TE budget for the DTP performance test listed in Table 2, the LTE OTA setup includes extra TE budget for the PRTC and base station. The TE budget for BS1 is 200 ns, and for BS2 it is 300 ns. Since BS2 uses GPS, that introduces an additional 100 ns into the TE budget. Given that the VSA test compares the relative TE between BS1 and BS2, the total TE budget is 1580 ns which is smaller than the 3 µs (air to air) TE budget required by 3GPP specifications. The Sentinel only uses BS1. The total TE budget is 1280 ns which is smaller than the 1.5 µs (air to GPS) TE budget required by 3GPP specifications.

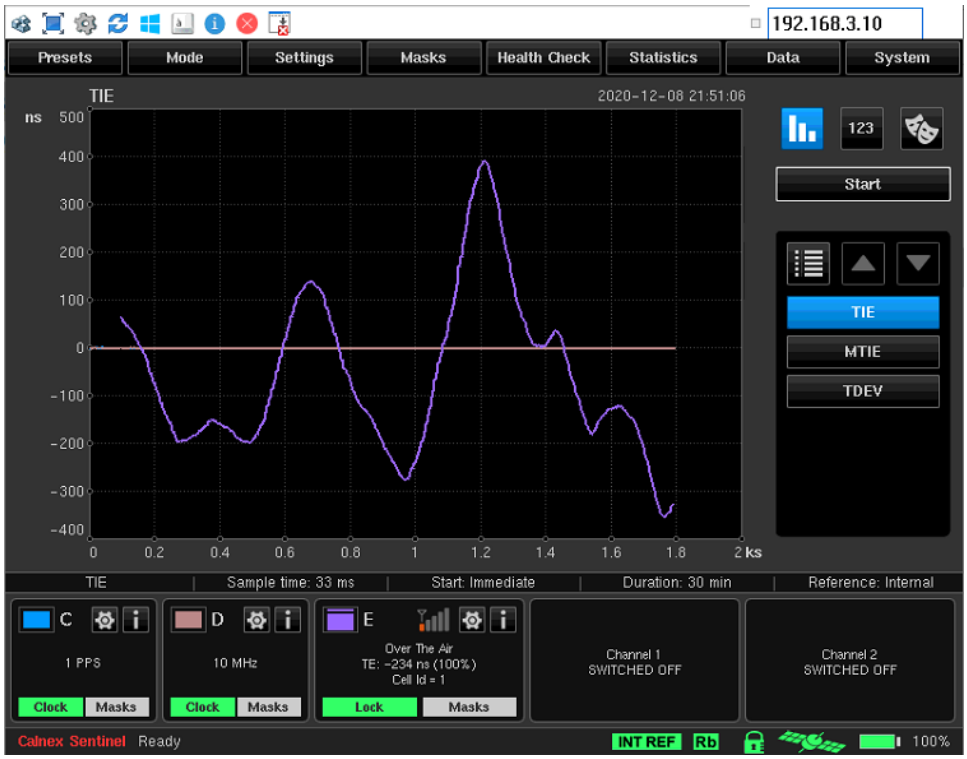


Figure 4 - RPD Sentinel OTA Measurements [9]

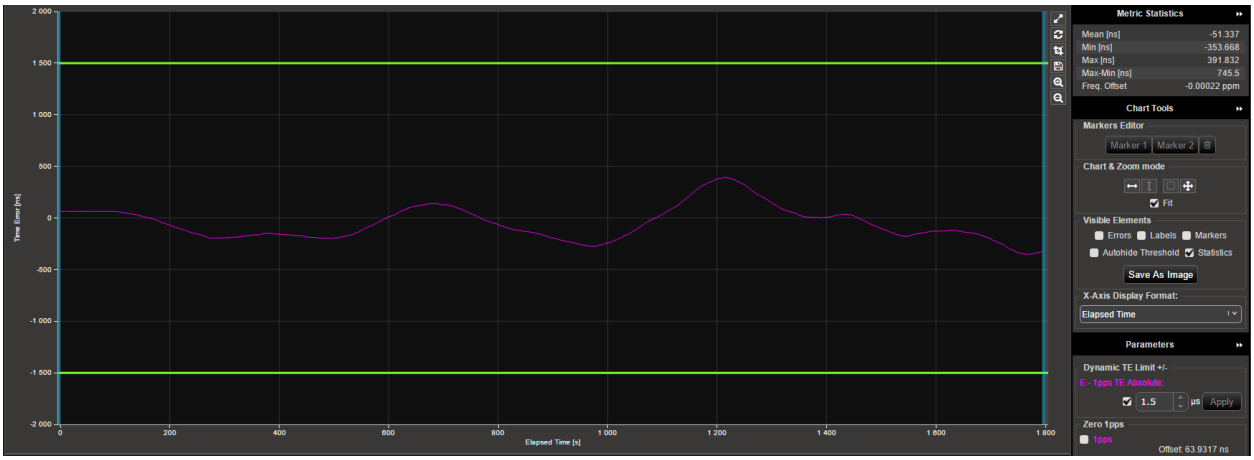


Figure 5 - Sentinel OTA Results [9]

4.1. Sentinel Measurements

The eNodeB (eNB), using DTP in the DOCSIS backhaul, radiates an LTE signal from 3620 to 3630 MHz with a cell ID of 1. Channel E on the Sentinel decodes the time of day on the LTE signal. This time of day is compared with the GPS time of day. Figure 4 shows an example of Sentinel measurement data. The OTA LTE signal TE varies from -354 to 392 ns, with a mean value of -51 ns. Figure 5 shows statistical results that are further processed by the Calnex Analysis Tool (CAT).

Five sets of Sentinel data were collected. Table 4 lists the results. The average LTE signal TE is between -71 and 9 ns. The largest peak-to-peak variation is 746 ns. All the TE results meet the ±1.5 µs 3GPP requirement.

Table 5 - Sentinel OTA Results Summary [9]

Run	Time duration (s)	Two-way Time Error (ns)			
		Mean (cTE)	Max	Min	Max-Min
1	1800	-51	392	-354	746
2	1800	-25	269	-255	524
3	1800	-41	186	-286	472
4	3600	9	159	-107	266
5	3600	-71	131	-332	463

4.2. VSA Measurements

The VSA measures LTE signals in the frequency domain and converts them into the time domain. To avoid mutual interference between the two LTE signals, the two BSs are configured on two separated channels: BS1 uses 3620-3630 MHz and BS2 uses 3690-3700 MHz. The VSA compares LTE signals from BS1 and BS2 in the time domain. As shown in Figure 6, the upper two subfigures (yellow and blue) are the frequency domain magnitude spectrum, and the lower two subfigures (green and red) are the time domain waveforms. The left side two subfigures (yellow and green) are BS1 signals using DTP in the backhaul, and the right side two subfigures (blue and red) are BS2 signals synced to GPS.

Both BS1 and BS2 employ LTE TDD configuration 2 and special subframe configuration 7. The subframe structure of TDD configuration 2 is provided in Figure 7, where D represents downlink, U represents uplink, and S means special subframe. The time duration of each subframe is 1 ms. The special

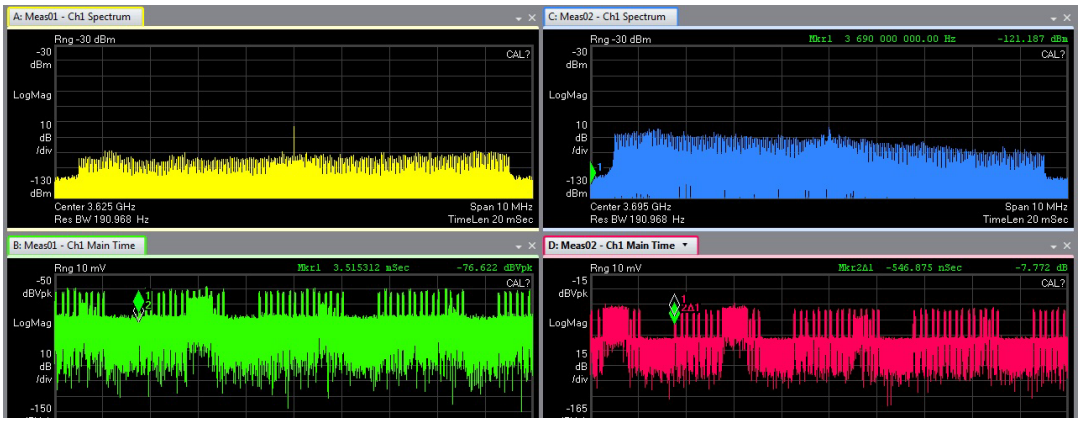


Figure 6 - VSA OTA Results [9]

Subframe number									
0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
D	S	U	D	D	D	S	U	D	D

Figure 7 - LTE TDD Configuration 2 Subframe Structure

subframe consists of 14 symbols, where 10 symbols are allocated for the downlink in special subframe configuration 7. So each of the downlink signals should last $3\frac{10}{14}$ ms. There were no UEs in the lab, and no traffic in either the downlink or the uplink during the VSA OTA measurement. BSs only transmit reference signals and control channel information in downlink subframes.

The lower two subfigures in Figure 6 present bursts with a period of 5 ms, each group of bursts lasts for less than 4 ms, which agrees with theoretical TDD LTE downlink signals. Marker 1 is placed on the rising edge of the burst for the BS1 signal (green), and marker 2 is place on the rising edge of the burst for the BS2 signal (red). VSA syncs markers are placed in both channels so that they are comparable. "2Δ1" represents the time difference between markers 1 and 2. 2Δ1 is 529 ns. The relative time error between BS1 and BS2 LTE signals is 529 ns which is much smaller than the required TE budget of 1580 ns as listed in Table 3 and the 3 µs requirement in the 3GPP specifications.

5. Conclusion

DTP is designed to provide accurate synchronization for the backhaul of TDD mobile networks. DTP PoC testing was conducted by CableLabs, Charter, Cisco and Hitron. PoC testing was divided into three phases:

- Phase 1 validates that DTP works in a basic lab environment.
- Phase 2 evaluates DTP performance in sophisticated environments that mimic field deployments.
- Phase 3 verifies automatic DTP calibration in field deployments by using an AWS cloud server to distribute calibration data.

This paper reported up-to-date progress of DTP PoC testing, and key findings in phase 1 testing. The results successfully demonstrated that DTP works in a lab environment. The measured DTP time error results meet the time error budget. Using DTP and PTP in the backhaul, LTE over-the-air signals meet the 3GPP synchronization requirement. DTP is being evaluated in various HFC network configurations. An AWS cloud server is being developed to enable automated DTP calibration.

Abbreviations

3GPP	3rd Generation Partnership Project
5G	fifth generation
AWS	Amazon Web Service
BC	boundary clock
BS	base station
CAT	Calnex Analysis Tool
CBRS	Citizen Broadband Radio Service
CM	cable modem
CMTS	cable modem termination system
cTE	constant time error
DAA	distributed access architecture
DL	downlink
DOCSIS	Data-Over-Cable Service Interface Specification
DS	downstream
dTE	dynamic time error
DTP	DOCSIS Time Protocol
eNB	eNodeB (LTE base station)
FFT	fast Fourier transform
GM	grand master
GPS	Global Positioning System
HFC	hybrid fiber-coaxial
LTE	long-term evolution
I-CMTS	integrated cable modem termination system
IFFT	inverse fast Fourier transform
MSO	multiple-system operator
MTIE	maximum time internal error
NCS	Cisco Network Convergence System
NR	new radio
OFDMA	orthogonal frequency-division multiple access
OTA	over the air
PoC	Proof of concept
PRTC	primary reference time clock
PSS	primary synchronization signal
PTP	precision time protocol
QAM	quadrature amplitude modulation
Rb	Rubidium
RPD	remote physical layer device
R-PHY	remote physical RF layer
SSS	secondary synchronization signal
TDD	time division duplex
TDEV	time Allan deviation
TE	time error
TRO	true ranging offset
UE	user equipment
US	upstream
VSA	vector spectrum analyzer

Bibliography & References

[1] 3GPP Technical Specification 36.133 v16.9.0, Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for Support of Radio Resource Management, June 2021. [link]

[2] IEEE 1588-2008, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," July 2008. [link]

[3] Dave Morley, "5G Small Cells and Cable – Realizing the Opportunity," *Cable-Tec Expo 2018*, Atlanta, GA, October 22-25, 2018. [link]

[4] John T. Chapman, Rakesh Chopra, Laurent Montini., "The DOCSIS® Timing Protocol (DTP), Generating Precision Timing Services from a DOCSIS System," *INTX/SCTE Spring Technical Forum*, 2011. [link]

[5] Cable Television Laboratories, Inc., "Synchronization Techniques for DOCSIS® Technology Specification," CM-SP-SYNC, April 2021. [link]

[6] Cable Television Laboratories, Inc., "DOCSIS® MAC and Upper Layer Protocols Interface Specification", CM-SP-MULPI, December 2020. [link]

[7] Elias Chavarria Reyes, John T. Chapman, "How the DOCSIS® Time Protocol Makes the SYNC Specification Tick," *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, Oct, 2020. [link]

[8] Jennifer Andreoli-Fang, John T. Chapman, "Mobile Backhaul Synchronization Architecture," *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, October, 2017. [link]

[9] Cable Television Laboratories, Inc., "DOCSIS® Time Protocol Proof of Concept Phase I Technical Report CM-TR-DTP-V01-210915," September, 2021. [link]

[10] Ruoyu Sun, Rahil Gandotra, Jennifer Andreoli-Fang, Elias Chavarria Reyes, John T. Chapman, Mark Poletti, "Designing a Cloud-Based DOCSIS Time Protocol Calibration Database," in *SCTE-Expo 2021*, Atlanta, GA, October 11-14, 2021.

Don’t Throw Away Your Shot: Rise Up to Change the Narrative for Construction Management

An Operational Practice prepared for SCTE by

Mindy Kang
Vice President, NGAN Product Development
Comcast
1800 Arch Street, Philadelphia, PA 19103
215-286-7570
mindy_kang@cable.comcast.com

Jennifer Smardo, Vice President, NGAN Implementation /Comcast

Yael Futer, Sr. Director, NGAN Product Delivery /Comcast

1. Introduction

It was the summer of 2017, Hamilton was the hottest show on Broadway. Data hungry consumers were straining the network and there was one looming question the Broadband provider couldn't answer: How can I build my existing network to be bigger, better, faster?

In a company built by acquisitions, network construction knowhow was managed by local knowledge, teams and antiquated systems running DOS. The organization was anxious for change. The company had to prepare and scale the next generation of construction to accelerate growth.

Enter a young(ish), scrappy and hungry crew that never spliced cable but needed a job. Management gave them 5 months and bowls of cashews to fuel their mission. Look three years into the future, assess how the birth of a workflow management platform became the center of the construction universe and, in the words of the C level execs resulted in "not too many complaints."

This paper will test a basic hypothesis of organizational change management: do people, process and technology have to move together in the same direction to drive change? It will argue that technology can drive organizational change and it will outline the inputs necessary to do so. It will further demonstrate that a group does not need to "own" the work to transform the work. It will articulate a bold approach to the routines and rituals required for agile technology development to translate into incremental organizational changes. Most importantly, it will challenge its readers to re-think their methods of driving change with construction and design resources in an ever-evolving race to construct the fastest data delivery network.

2. Building a National Tool Started With a Pivot

The first attempt to build a tool was called Polaris, referring to the North Star sailors would use to set their course. Polaris was conceived to set the course for how capital expenditures would be managed to build out network infrastructure. This tool was intended to help standardize the way the enterprise works with 3rd party Business Partners, and to become the source of truth for all construction activity, capital dollars spent, duration and quality of work completed. But after some time and substantial software investment, Polaris was still just an idea – or rather, a conglomeration of ideas collected over time, without any unifying vision or purpose.

While the Organization developed a nationwide infrastructure strategy, its 15 regions had uniquely different ways of managing the day-to-day of it, from walkout surveys to permitting and plant construction. Every region agreed that a national tool could be helpful, but none wanted to change the way they were operating. Regions would commit to getting their teams to use Polaris only if their specific, ever-growing list of features were delivered. In an effort to gain user adoption, the Polaris team had implemented somewhat arbitrary capabilities, based on disparate requests, from the most vocal regional users. The irony is that while trying to build a tool that would work for everyone, they built something that didn't quite work for anyone. The pressure to on-board Regions into Polaris was mounting, but the tool had only a handful of test users and the executive team couldn't get a clear answer on what was needed to launch.

In the summer of 2017, Polaris was handed off to a new software development team, with instructions to "fix it, fast." A technical deep dive yielded the brutal truth: it was not usable. Not even the technical foundation was salvageable. Neither the development team nor potential users seemed to know the answer to the question: what problem are we trying to solve? Without any reusable code, clearly documented requirements in place, or shared vision, the new leadership team recognized that "fixing" Polaris was an impossible task.

Lesson 1: Know When to Pivot

In August, Polaris was shut down. In September, P2 was born. By January 2018, P2 would need to support a Pilot market with an actual workflow. The bold decision to shut down a 3-year effort and task a new team to build a working tool in 4 months' time was the catalyst to *fan this spark into a flame*.

P2 (short for Polaris 2) began with a clear mission in mind: build a national tool that would help manage the flow of construction, with a common set of goals and language.

> Mission Statement: P2 is an enterprise-wide integrated workflow orchestration tool that tracks the progress of all construction job types – providing visibility into status of milestones, accurate cost of each project, the data to build forecasts, and the ability to roll up those metrics to a national view.

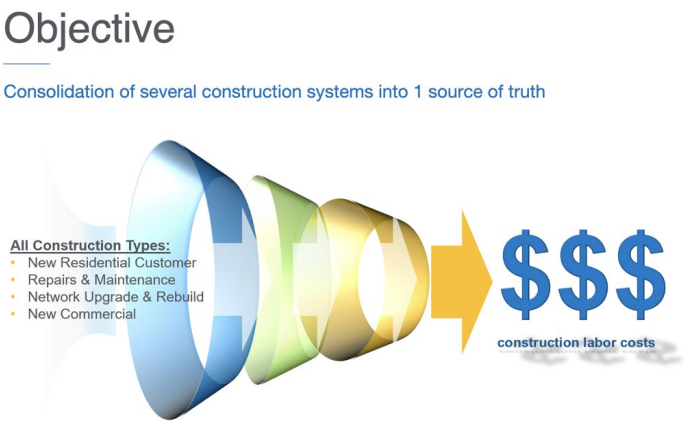


Figure 1: P2 Platform Objective

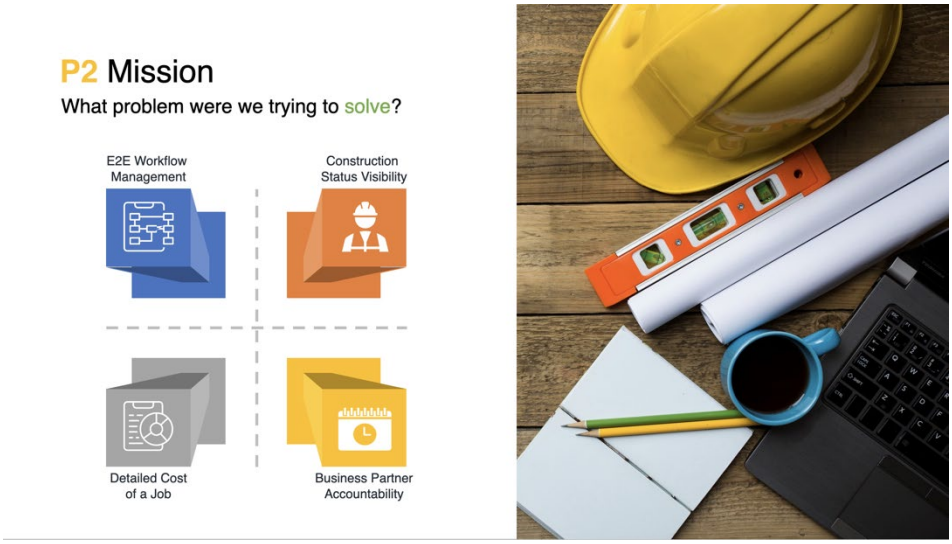


Figure 2: P2 Problem Statement

The plan was to build on that belief shared by all 15 regions that a single tool, that provides a national view of all construction builds, would be useful. But, moving 5,000+ users onto a new standard platform meant that many of the teams across those regions would need to change certain processes and ways of doing things. And change is hard.

Comcast's construction teams had been building out the network for years, which meant they had well established ways of managing phone calls, marked-up paper maps, post-it notes, spreadsheets, and local databases to keep the flow of production moving. In their minds, this insider know-how, sometimes collected over decades, had been optimized for their circumstances. Aerial coax construction to replace a span? Who else could more efficiently complete this build than the guys who pull cable in their towns every day? The original Polaris team struggled to find a meaningful benefit for end users to transition to a new platform; what incentive existed to drive change? With a pivot to P2, there was an opportunity to clearly identify and communicate "the why" and benefits of change.

The "why" was a parallel initiative called Fiber Deep. Comcast was about to deepen its investment in constructing a proactive network upgrade architecture that would increase capacity in the short term and pave a way for growth in the long term. The challenge was that this was a new kind of cable construction that was unfamiliar to many. It was massive in scale, impacted entire geographic areas, and the volume of this type of work was expected to grow with time. As work was increasing, it was clear that phone calls and post-it notes wouldn't be enough to keep these large-scale projects moving on time and budget, and the existing databases couldn't be refactored quickly enough to support this use case. Visibility of related work and automation to speed up data entry or help with calculations were identified as critical needs. The development team realized that building an intuitive way of managing these types of projects in a set of workflows would be compelling enough to persuade users to adopt the tool.

Lesson 2: Talk Less, Smile More… Ask Many Questions

The P2 team set out to build this new tool, that would support a new process, managed by new teams. Given the specific scope (Fiber Deep), but with many unknowns and open questions, it was a perfect situation to apply a *Design Thinking* framework. In Design Thinking, the goal is to use the following process to design a solution:

1. Empathize – to think about the community of people needed to be served, the roles they play, the challenges they'll have and what they'll care about
2. Define – based on their challenges, identify what will that community need, what is a problem that needs to be solved?
3. Ideate – brainstorm to come up with a wide range of ideas to tackle the problem identified
4. Prototype – build a small proof of concept that can be demoed, and that allows user interaction
5. Test – run experiments to test the hypothesis, allow users to engage with the prototypes to validate if the idea really solves the problem

In the months that followed, the P2 Team spent hours, days, weeks with the Construction experts. This was a small team of software developers who didn't know the first thing about construction, but knew how to really listen to people in order to *Empathize, Define,* and *Ideate.* These rich discussions gave the team enough direction to start iterating on a prototype of a single Fiber Deep workflow. During this time together, the P2 Team continuously strove to do two things – 1) show incremental progress, no matter how small, and 2) build trust by always delivering on a commitment and being transparent about the process and any mistakes or missteps. These frequent discussions, small feature demos and iterative development based on their feedback, gave the Regional Construction teams a sense of ownership in the workflow that was being built. By December, the P2 Team had built enough of a workflow framework, and enough advocates, to bring key Regional team members together in Philadelphia to demo the *Prototype* to prep for a January trial.

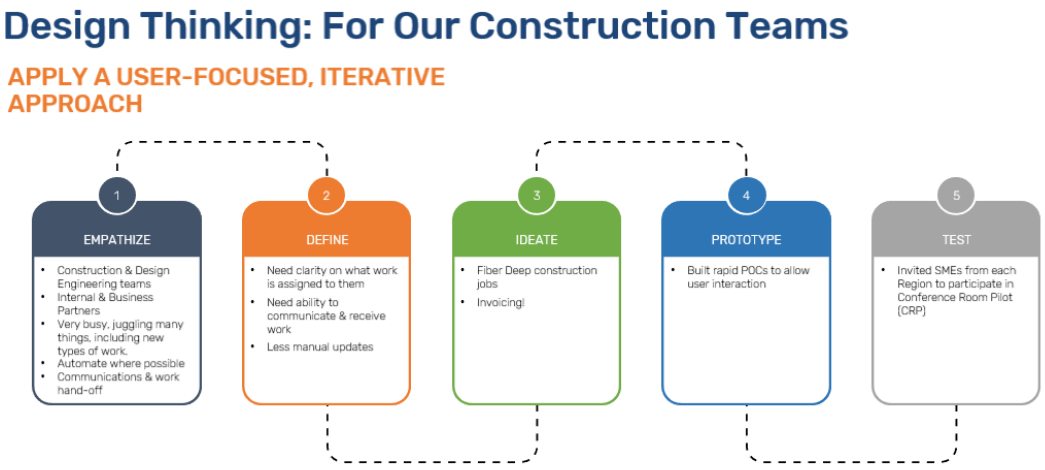


Figure 3: Design Thinking Applied to Construction Workflow

Lesson 3: Build Advocates… *in the Room Where It Happens*

Although constructing Fiber Deep builds involved a new process, a few Regional teams had already begun their projects and had strong opinions about how to implement these outside plant changes. The P2 Team knew that in order for a National tool to succeed, these Regional teams would have to develop a common language and come to an agreement on what really matters when managing these projects. The purpose of this large meeting was not only to demo the prototype, but to have the local experts and decision makers from each Region sit in the same room, debate the controversial topics, but leave with a shared commitment to live with whatever compromise they made. These sessions, which were called *Conference Room Pilots (CRPs)* allowed stakeholders to have a seat at the table, to voice their point of view on construction nomenclature, or the kind of specific attachments a vendor should load when submitting an invoice, and everything else. But, no matter what was debated, a decision was required at the end.

When decisions were needed on topics that reached an impasse, the only way to move forward was to vote. The vote was a sacred ritual, each of the three divisions got two votes, representing their Finance and Construction organization, as well as the voice of their regions. Once a vote was cast, and a decision was made, it was prioritized. Some change requests, like field labels, were quick to change and if a decision was reached in the room, an engineer made the change on the spot. Some requests, like standardizing construction quality audits, were agreed to be important but could be addressed outside of P2 in the short term and added later. Finally, a handful of decisions would require Executive Leadership input prior to implementation.

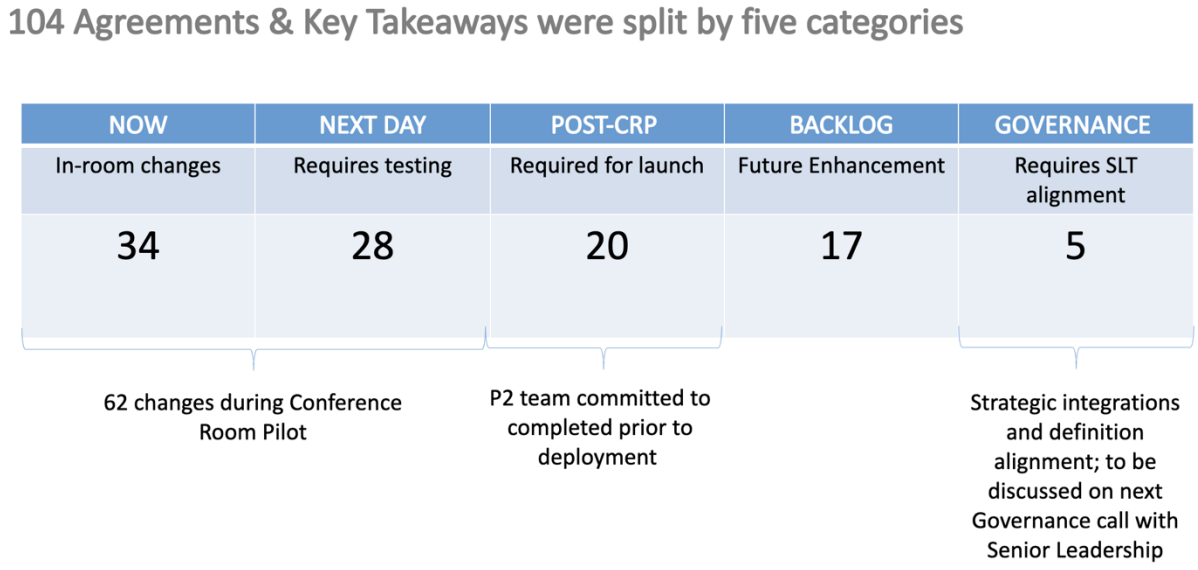


Figure 4: Conference Room Pilot Agreements Output

The Conference Room Pilots gave an opportunity for the people *in the room where it happened* to own the process that was being built into the tool. Debating the topics made their ideas and concerns feel heard, and voting gave them agency, even if the outcome wasn't exactly what they had proposed. As the team closed out on its first of many CRPs to come, they got the commitment from one of the Fiber Deep Project Leads to run a Pilot of the tool in their Region. A Pilot would allow the team to take their prototype and *Test* whether they got the solution right. What many folks did not know is that that in those early days, the P2 Tool only supported one Workflow, in one Region, for one Design Business Partner, and one Construction Business Partner. Yet, it was enough to run a Pilot. In software product development, there is a concept called "MVP," which stands for Minimum Viable Product. Building an MVP workflow, with continuous demos of small incremental progress, and by building trust in local advocates, was enough to prove out the value and benefit of managing Fiber Deep in a single tool. At the launch of the Pilot, the tool still had many gaps in functionality and many unanswered questions about process, but it didn't matter. The development team continued to iterate, develop and deploy new functionality each week, closing these feature gaps in both small and meaningful ways.

The Pilot began with a handful of Users in the production tool and the P2 team talked with them every single day. Users would join each day to let them know which buttons they clicked, which ones they couldn't find, which headers were confusing, and about a plethora of missed requirements. Each Pilot User had the personal cell phone numbers of the product development team. The development team ate feedback from this pilot community for breakfast, lunch and dinner. The P2 team continued to build on the trust and transparency established during the initial engagements. They always fully embraced and encouraged criticism and feedback about the tool, the process, the interactions with the Users. Both internal Construction team members and Business partners trusted the team enough to immediately inform them if a feature wasn't working. They all knew that sharing that feedback, however rough, was the only way to get better.

Lesson 4: ***When You Got Skin in the Game, You Stay in the Game***

As the Pilot continued, additional Regions agreed to come on-board. New requirements were identified with each new project, but decisions were always brought back into weekly calls where all Regional experts had a chance to weigh in and vote. The development team continued to work closely with each new Regional team to understand their challenges and needs, and fix things that weren't working for them. By November of 2018, the P2 Tool had become the national workflow tool for all teams that had Fiber Deep projects to manage. It had grown from 10 Users to about 150 Users across the country, all executing the same construction workflow in in P2.

With the Pilot, and incremental, iterative changes, P2 had managed to drive standardization across 15 Regions using technology to pave the way. Change was starting to happen. But the majority of new build construction workflows still needed to be accounted for. The number of Regional teams, decision makers, feature gaps, process changes needed to support all construction, increased exponentially. The development team continued to leverage the CRP format to quickly *Empathize, Define, Ideate, Prototype,* and *Test* workflow changes needed to support the variety of workflows needed. They spent the final few months of 2018 in a full-blown marathon to develop and deploy all the capabilities needed to manage all construction in P2. Local advocates built over the course of the Pilot evangelized the benefits of the Tool and helped strengthen the call for everyone to transition over to the new way of working.

By 2019, all three Divisions agreed to launch P2 across all their Regions in the first quarter of the year. The transition to National deployment, where all 5000+ Users across every Region and Business Partner were on-boarded to this new platform was not easy. For one Division in particular, adopting P2 meant migrating away from an existing legacy tool that had been in use for over 10 years. But the Construction

teams that had helped to build this tool could see the value of a single platform, and made the commitment to push through the challenges and adapt to the process changes that came along with it.

Lesson 5: "Aim" for "Not Too Many Complaints"

Immediately following the full adoption of the P2 Tool, the development team met with the Cable Division Finance Executives for a progress briefing. As the team gave a demo of the tool and an update about the launches, an Executive remarked – "based on where you are in this process, I haven't received too many complaints." In this meeting, it became clear that one indicator of success was not how many praises P2 received, but that it didn't draw, "too many complaints...." The motto became an anthem.

3. Piloting is Easy, Scaling is Harder

Lesson 6: When You need to Operate at Scale, Process Must Also Be a Product

The successful launch and adoption of the P2 Tool had much to do with the trust that was built between the Users and the development team. Thanks to the frequent touchpoints through the CRPs, daily chats and regular meetings, the Users felt like their needs and requests were always being heard. This feedback loop gave the Users and the development team comfort in knowing there was clarity on the list of problems that needed to be solved, and the path to ideate and validate different approaches. However, as P2 was adopted as the National Construction Tool, the User community grew from 150 to 5,000 users in a matter of weeks. At that scale, personal phone numbers and chat messages to the development team would not work; the product development process needed to maintain that trust and User engagement also needed to evolve.

To that end, the P2 Team added a Product Operations group, whose objective was to build the processes and communications channels needed to support a User community at scale. Rather than have people pick up a phone or send an email, the Product Operations team built out a self-service ecosystem that gave every User a path to look up release notes, FAQs, read documentation, watch training videos, and submit enhancement ideas. CRPs evolved into Product Trials, where the *Test* part of the product development lifecycle now included more structure around how to collect feedback, measure success, and how to troubleshoot issues when something didn't work as expected.

Product operations is critical to ensuring the platform can build with empathy at scale – supporting all Users where they are and giving them an equal voice, even if they don't have the phone number for a member of the product development team.

Lesson 7: *Every Action's an Act of Creation* – The Criticality of Process, Post-Launch

By using software to identify the work being performed, P2 became the center of the construction ecosystem. Technology had fanned the flame of change by bringing all regions together to loosely follow a high-level set of processes. Furthermore, having everyone work in a single tool provided an opportunity to use consistent language to talk about construction activity. The P2 platform provided the foundation to align around one national strategy to deploy a proactive network upgrade architecture. While P2 could provide data on how construction was progressing, it became clear that a governing body was needed to determine how to measure if the enterprise was successful in its implementation and create a forum to adjust policies around process and compliance collectively.

In early 2019, just as P2 was getting deployed nationally, Construction 2.0 (C2.0) was born. Just like P2, Construction 2.0 had a clear mission statement:

> "Construction 2.0 is a construction business operations ecosystem that compliments the strategy of fast, efficient growth through standardization and alignment of roles and responsibilities, processes and procedures, tools and reporting, and performance management."

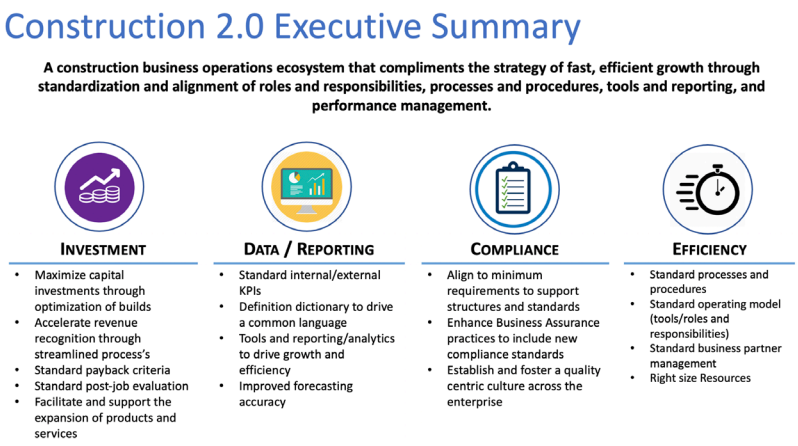


Figure 5: Construction 2.0 Charter

The inaugural C2.0 summit was the first time in the enterprise's history when Regional and Divisional representation came together to prioritize a construction process improvement roadmap and establish and agree on an approach to execution. The summit was the first deliberate step towards organizational sameness.

After the initial summit, the C2.0 team quickly discovered that gaining alignment to process, policies and standards would take a different level of effort to achieve sameness across the enterprise. Rather than starting with a focus on Roles and Title alignment, C 2.0 rationalized that by prioritizing the standardization of processes most critical to the business, those politically stickier issues would naturally follow.

Working teams were established with key Divisional and Regional representation, and facilitated by a member of C2.0. The working sessions were bucketed into four main categories to work through the prioritized list of national process changes and included:

1. Business Partner Process Standards
2. Sales Interaction Process Standards
3. Construction Standards
4. Strategic Software Integrations Standards with other National tools that impact Construction

The P2 Team was brought in if an organizational process change necessitated changes in the tool. This integrated approach allowed C2.0 to marry process and policy decisions back into the existing P2

workflows. For example, in the early development stages of P2, there was general consensus that quality audits were critical to construction workflow but it would require a process change. That process change was managed in the Construction Standards workstream discussions and the P2 team was present to collect requirements for development.

In addition to the working meetings where all the "sausage making" was occurring, governance and update forums were stood up. The Governance meeting facilitated fast alignment, and brought the most important decisions to leadership coming out of the working teams. The update meetings consisted of monthly program calls to keep the broader stakeholders informed and quarterly updates to keep senior leadership in the loop.

The governance structure was critical. Each Division had a voter and a proxy voter for both Construction Ops and Finance to ensure the organization was making well-rounded decisions and created more meaningful meetings; it was consensus driven decision making.

Lesson 8: Make Progress to *Get a Lot Farther by Working a lot Smarter*

At first, the governance decision making process was clunky, and the conversations were spent trying to understand the problem and the solution being solved for, versus making a decision to implement a new national process or standard. To address agility, C2.0 had to change the way it presented recommendations during governance.

To address quicker decisions making monthly Governance forums were limited to one hour and only included two topics: Decisions and Ideation items. For quicker decision making, a "pre-voting" process was implemented. The process involved all governance decisions and supporting documentation to be sent to leadership one week in advance of the meeting. Each division workstream lead was instructed to meet with their leadership on upcoming topics to inform the vote. Each voter would respond to the template via e-mail with their verdict. Anything aligned to pre-governance no longer required discussion, which saved time for topics where alignment was more challenging to garner. The ideation section consisted of strawman proposals to ensure the workstreams were solving the highest priority items with a general identification of the problem that required solving. Once prioritized and ideated, most all of the heavy lifting and debate on process change happened within the workstream teams.

Once decisions were made, several ways to communicate the change were implemented. To support the change management process more locally, a core group of division leads received launch documents. Each document provided pertinent information to support the communication of the change such as background of what is being launched, intended audience and FAQs. Internal stakeholders had access to a standards portal housing national standards and The Construction Hub portal that housed national policies. A Business Partner Portal providing standards and policies for design and construction business partners was also created. Where process changes impacted P2, the Product Operations team would build Job Aids or Awareness documents to summarize impact to P2 Users.

These structures enabled C2.0 to solution quickly, garner the appropriate approval from leadership to implement, disperse information to stakeholders and partner with P2 team as the solution unfolded. This cross-pollination of process and technology allowed C2.0 to care for the most important items in a manner that would yield quick wins in terms of user functionality, tool compliance and overall standardization sameness. It created a formalized way to share best practices and either adopt or augment those practices to fit the enterprise. It pushed the organization towards the sameness in a way that was agile and manageable for the ecosystem to consume.

4. You fought in this war, what was it all for?

Once the entire construction community is in one system, it created fractal changes within the organization. The C2.0 governance forum created a place to assess priority of new system integrations to trigger an action or provide status to other orgs that had interaction with the outside plant design and construction community. Within the 18 months of launch, P2 integrated with other national applications to process:

- Sales orders that required construction plant extension to new customers
- Capacity augmentation work that required outside plant design/construction activities
- Survey requests to determine total cost of construction prior to a customer sales commitment
- Procurement warehouse data to cost out materials required for a construction build
- Purchase order balances ensuring fund availability to complete a build
- Visibility of construction status to all construction activity to anyone in the enterprise
- Maintenance requests requiring outside plant construction resources to address

Three years post the initial pilot, the C2.0 governance process is continuously prioritizing additional integrations to ease communications across the organization.

5. Conclusion

Change is often difficult to accept. When a large organization is facing the need for meaningful change across the enterprise, it can be a challenge to get existing teams on-board to implement them effectively. In more traditional approaches to change management, the people, process and technology need to be fully defined in order to execute the asks. With P2, the team proved that Technology can be used to drive organizational change, even if the organization wasn't ready to answer every process and people related question. By allowing Technology solutions to remain agile, to flex and change to prioritize the highest impact needs of the organization first, you can get people to join the movement. With time, the incremental changes in behavior and process, driven by small, iterative changes in the Technology, will build up to the meaningful change you were striving for.

Using an agile approach, local knowledge can be leveled up and turned into enterprise best practices. This can, in turn, be used to propel the organization towards sameness, and drive compliance of tool usage and standards application. Facilitation of a collaborative approach, fostering a growth mindset, making it safe to fail fast and course correct while in flight can drive consensus and alignment needed to manage change effectively.

As your organization and solutions scale, process becomes more critical to the success of its growth. At a certain inflection point, a more formal process will begin to drive the change, with Technology supporting both. The results of weekly incremental change over a period of three years for the enterprise resulted in an ability to scale the network nationally, in a standard way for years to come.

6. Acknowledgements

The P2 and C2.0 efforts wouldn't have gotten off the ground without the vision, courage and executive decision prowess of Elad Nafshi, Tim Nester, Meena Soleiman, Larry Beauchamp and Doug Czekaj. The P2 platform could never have become a reality if it wasn't for the sheer brilliance, continuous cheerleading, coaching and sponsorship of Bruce Bradley. The P2 platform wouldn't have accelerated its growth and scaled properly without the incredible product operations team, envisioned and led by Chau

Phan. This paper wouldn't have been written without the encouragement of Bob Gaydos. This journey would have been lonely if it wasn't for the friendship the co-authors developed because of it.

Abbreviations

NGAN	Next Generation Access Network
SCTE	Society of Cable Telecommunications Engineers

Bibliography

Inspirational quotes courtesy of Miranda, Lin Manuel. Hamilton: An American Musical. Atlantic Records, 2015, MP3

Edge Computing Architecture

A Technical Paper prepared for SCTE by

Umamaheswar (Achari) Kakinada
Director, Wireless R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
847-544-6560
Achari.Kakinada@charter.com

Deh-Min Richard Wu
Director, Wireless R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
256-763-1202
deh-minrichard.wu@charter.com

Curt Wong
Senior Director, Wireless R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
425-395-4379
Curt.Wong@charter.com

Yildirim Sahin
Director, Wireless R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
720-536-9394
Yildirim.Sahin@charter.com

1. Introduction

Trillions of gigabytes of data is being generated/captured by devices and network systems, which need to be analyzed and processed. Forbes estimates [1] that in the year of 2025, 150 zettabytes of data will need to be analyzed and processed. Over half of this data is expected to come from the edge of the network (79.4 Zetta bytes by 2025 [2]). The world's most valuable resource is no longer oil, but data [3, 4]. While oil is a limited natural resource which needs to be preserved and conserved. On the contrary data is experiencing explosive growth and showing no signs of slowing down anytime soon. Data need to be managed, processed and analyzed to derive value from it. Conventionally, this data is sent to the centralized systems usually in the cloud for processing, analyzing and deriving insights. It is an enormous task at hand. This method of processing the data incurs significant latency and huge amount transport cost. Which often renders the derived stale insights not useful for most latency sensitive applications.

Edge computing (EC) addresses and mitigates some of these issues. In this paper we look at different aspects of EC, comprising of addressing the latency in data processing, analyzing and deriving insights; architecture, standards and deployment considerations.

2. Motivation For Edge Computing

EC supports placement of compute, storage and other processing resources needed for performing analytics and deriving insights in a given scenario, along a wide spectrum of deployment scenarios ranging from centralized data center to individual devices or somewhere in between either extreme, to address the latency, storage and any special processing needs of a given application or service. The key motivators of EC are latency reduction, enhanced privacy and security, backhaul bandwidth optimization and enabling autonomous decision making at the edge of the networks [5]. This is depicted in Figure 1.

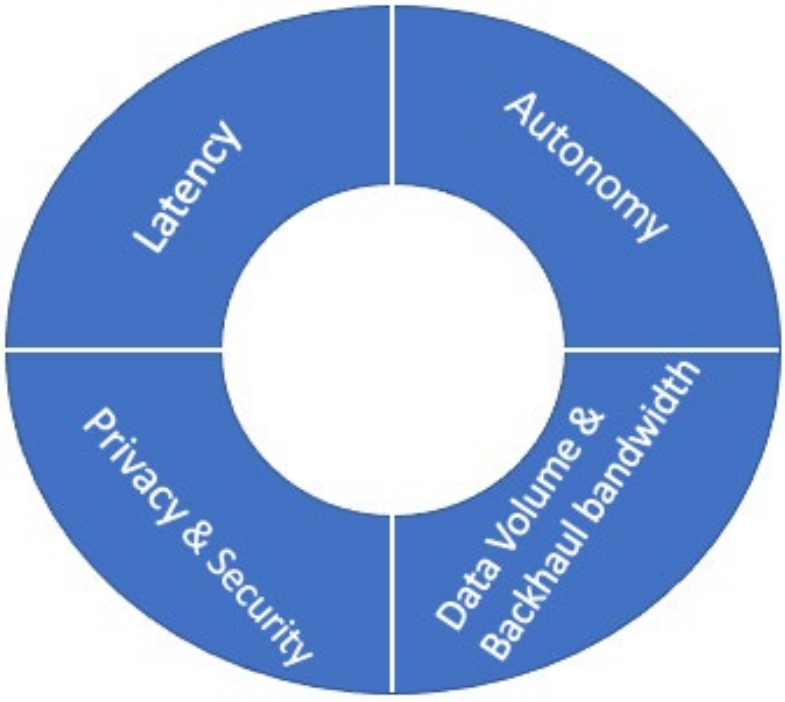


Figure 1 - Key motivators for Edge Computing

2.1. Latency

Zettabytes of data is being generated [2] at the network edge. Often there is a time-critical need to process and analyze these data, derive insights and take actions based on the learnings from these insights. It may not be viable to transfer this data to a central core and wait for the decisions to be made as there may be opportunity costs and safety issues because of this latency. Reducing latency is imperative[9] for many applications in Industry 4.0, healthcare, smart cities, aviation, autonomous driving, enterprises, entertainment, and augmented reality/virtual reality (AR/VR). Data need to be processed with a deterministic latency in a timely manner and EC addresses this critical need. We will look into various EC architectures and deployment scenarios currently being explored in various standards bodies and in the industry at large to meet different latency requirements. The EC enables agile service response and facilitates logic execution closer to the end users and devices both temporally and spatially.

2.2. Privacy & Security

In many centralized data processing scenarios such as cloud-based services, the user and/or device-generated data is transported to the central data center for processing, deriving analytics and taking actions. There are many regulatory requirements to ensure the privacy of the data as to origin, identity of the users/devices, etc. For instance, the energy/water usage reports from individual meters, if associated with an individually identifiable user may impact the privacy and security of this person. Another instance can be getting a count of number of vehicles in a road transport network without sending individually identifiable information about any vehicle or person, aggregating their number per segment/area, deriving analytics and insights to formulate a strategy for optimal traffic management. Further, these regulations vary widely across different sates/counties/municipalities, each requiring compliance with its own set of rules and regulations. This phenomena in not unique to smart cities scenario; it extends to other verticals such as healthcare and enterprises.

EC can act as an intermediary between the user data and centralized servers in the cloud or on premises. The EC can provide the desired anonymity for the sensitive data of the individual devices and users; and aggregate such data to minimize the processing at the central servers and reduce the costs of raw data transportation. It also can facilitate implementation of local rules and policies, and ensure compliance with regulations of the individual administrative domain (cities, healthcare systems, enterprises etc.). Additionally, EC provides autonomy and control often needed by these entities which are responsible for the data generated by users and devices in their respective domains.

2.3. Data Volume And Backhaul Bandwidth

Large, continuous data streams from huge number of devices/end-points can be burdensome on backhaul networks [5]. The overhead associated with the transfer of data from the devices to the central application servers, increases manyfold. Each data point being transported incurs overhead at each of the protocol layers, often these are small amounts of data compared to the transport overhead associated with it. In many instances, applications can benefit from aggregation of these data in edge nodes before being sent to the central servers. For example, a water meter can continuously collect the consumption data, while it probably needs to send them to the central server few times a day to meet the desired application processing needs. In some industrial 4.0 applications, the control may need to be gathered and logged, but may not be need to be reported to central servers for each individual data point. However, any exceptions, such as the temperature of a control unit becoming unusually high and needing immediate attention, may be sent immediately as a high-priority message by the edge node. Also, during normal course of operation, many systems collect large number of data points which are of interest for a short duration, after short life this data may lose it relevance, EC can address processing this type of data effectively. The

EC can act in this manner and conserve bandwidth usage on the backhaul, reduce the transportation costs and decongest the transport network while effectively meeting the application needs.

2.4. Need For Autonomy

Often the regulatory requirements for data autonomy, where the data is generated vary widely. This is the case for smart cities in which each city and local governments are responsible to define the regulation for maintaining the data and ensuring that it is secured. This calls for autonomy to do certain processing locally where the data is generated. Also, many applications in areas such as industry 4.0, need local context aware data processing, which are not suitable for central cloud computing. The EC enables meeting the local processing needs, also facilitates data processing per local policies and compliance with local regulations. The same needs exist in many other verticals, such as hospitality, health care, industry 4.0, enterprises and smart cities.

With the above-mentioned characteristics, EC is an enabler and plays a pivotal role in transitioning from centralized to distributed computing. It also makes it a key component of transition from centralized Industry 3.0 to distributed, data-driven and latency-sensitive Industry 4.0 paradigm. The same can be applicable to many other verticals. It minimizes time-to-insight, time-to-action by an order of magnitude, also minimizes cost-of-insight, in the process enabling timely and effective decision making and opening many new avenues of opportunities.

3. Edge Computing Architecture & Standards

3.1. 3GPP

Figure 2 and Figure 3 [14] provide a reference architecture for supporting EC per 3GPP specifications. The Figure 2 reference architecture depicts 5GS architecture for roaming scenario supporting EC with uplink classifier/branch point (UL CL/BP) support.

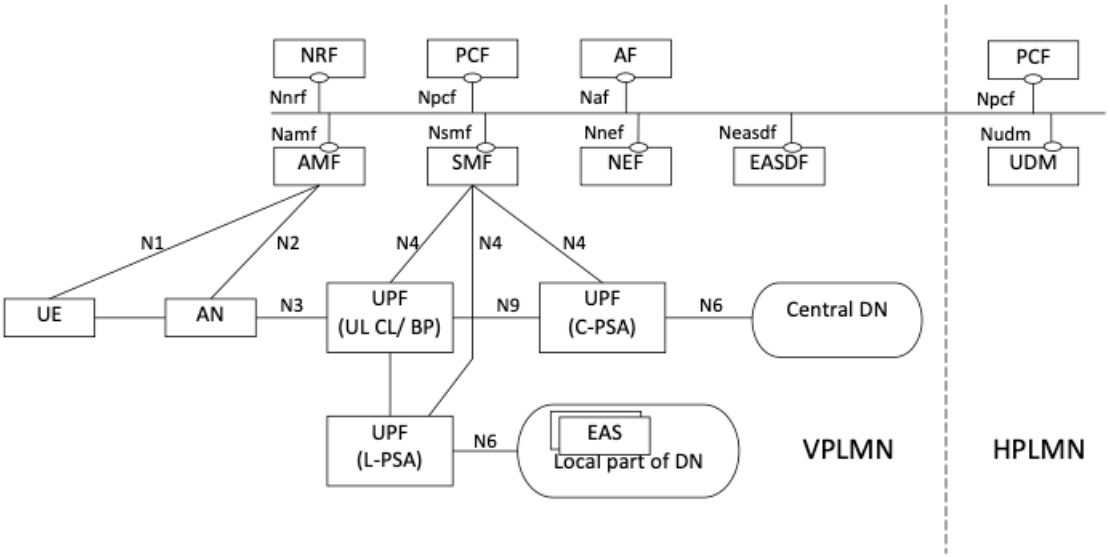


Figure 2 - 3GPP Edge Computing Architecture - Roaming with UL CL/BP

The 5G core per 3GPP standards specifications [14] supports the following connectivity models to enable EC:

- *Distributed anchor point*: For a protocol data unit (PDU) session, the PDU session anchor user plane function (PSA UPF) is in a local site, i.e., close to the UE location. The PSA UPF may be changed due to UE mobility.
- *Session breakout:* A PDU session has a PSA UPF in a central site (C-PSA UPF) and one or more PSA UPF in the local site (L-PSA UPF). The C-PSA UPF provides the IP Anchor Point when UL classifier is used. The EC application traffic is selectively diverted to the L-PSA UPF using UL classifier or multi-homing branching point mechanisms. The L-PSA UPF may be changed due to UE mobility.
- *Multiple PDU sessions:* EC applications use PDU Session(s) with a PSA UPF(s) in local site(s). The rest of applications use PDU session(s) with PSA UPF(s) in the central site(s). Any PSA UPF may be changed due to e.g., UE mobility and using session and service continuity (SSC) mode 3 with multiple PDU Sessions.

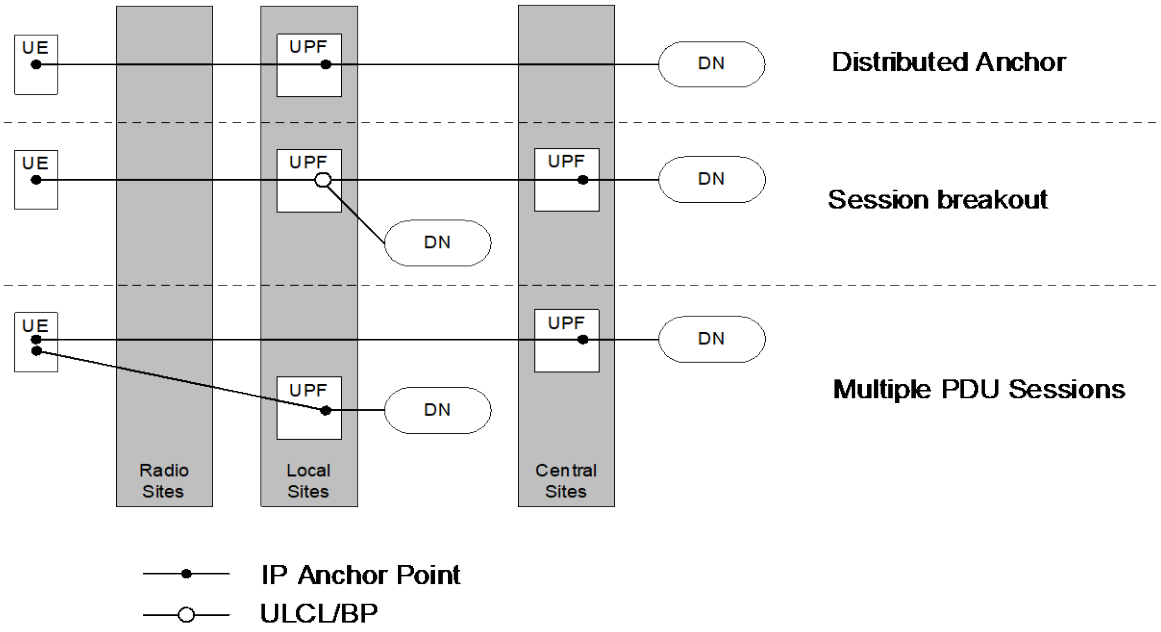


Figure 3 - 3GPP Edge Computing Connectivity Model

The following are some of the salient features of 5GS [14] support for EC:

- Flexible placement of UPFs in the network, which provide IP anchor point or different branch points in the network
- Simultaneous connections to multiple data networks
- Support for multi-homed PDU sessions, using either UL CL or BP
- Enhanced SSC: a connection through a new PDU session anchor point is established before the previous connection is terminated
- The application function (AF) can make a request to influence traffic routing for a given UE

Standardization efforts are currently ongoing for the 3GPP Rel.17 of the specifications [14]. The scope for Rel.17 includes:

- Edge application server (EAS) discovery and re-discovery
- Edge relocation
- Network exposure to EAS
- Support of 3GPP application layer architecture for enabling EC
- Services of EAS discovery function (EASDF) for EAS discovery, DNS etc.

3.2. ETSI

Figure 4 provides the framework of ETSI multi access mobile edge computing (MEC) for the deployment in a network functions virtualization infrastructure and virtualized network functions (NFVI/VNF) environment [15]. MEC offers to application developers and content providers cloud-computing capabilities and an information technology service environment at the edge of the network. MEC is access agnostic, providing flexibility in the operator network; managing different types of sites where the location of the edge will depend on the use case and needed activities to be performed. MEC system incorporates two levels: the MEC host level and system level. The former consists of the MEC host, the platform and the virtualization infrastructure manager, while the latter is composed by the MEC orchestrator, the operations support system.

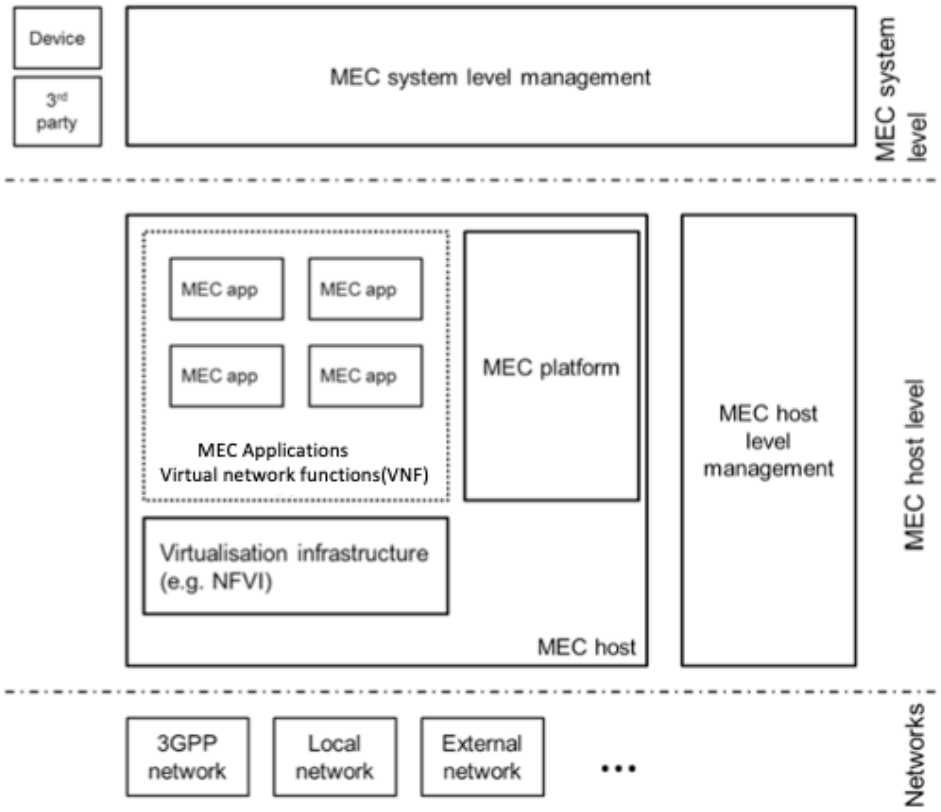


Figure 4 - ETSI Multi Access Edge Computing framework

3.3. IETF

The motto of IETF for beyond edge computing (BEC) methodology is – "Distribute as much as you can, centralized only if you must" [13]. IETF BEC aims to further research and standardize the protocol between multiple BEC gateways, common API across various BEC platforms, user mobility: edge to edge, edge device configuration/management, light-weight virtualization technologies (container/uni-kernel) and local edge security. BEC platform is depicted in Figure 5. The approach here is push the applications and use cases which are latency sensitive towards the edge of the network i.e., edgification of the system. On the other hand, system which scale well, benefit from clustered or centralized processing pushed deeper into the core of the network, i.e., cloudification approach for the system is adapted.

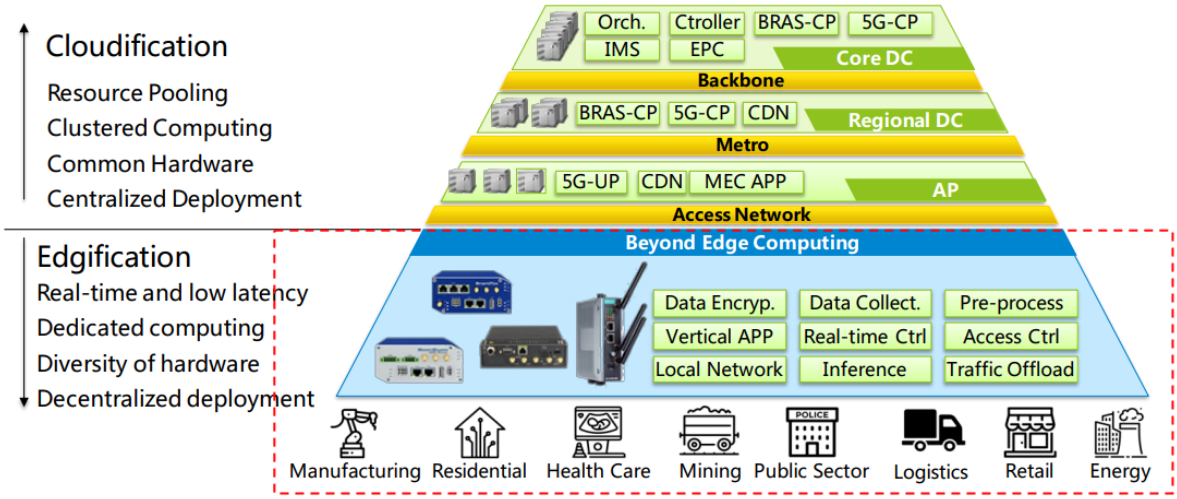


Figure 5 - IETF Beyond Edge Computing

4. Dimensions Of Edge Computing

Figure 6 below (adapted from [10]) depicts different dimensions of EC, the degree of importance these dimensions assume in catering to the needs of different verticals differ. The dimensions of EC are:

- *Time scale/Responsiveness* – real-time, near-real-time, and non-real-time.
- *Communication span* – device-2-device (D2D), device to near edge, device to center of network. How many other entities a given entity interacts with to perform its function, determines the amount of data exchanged and mutual dependence between the systems. Some of the devices may act independently and report their state (e.g., temperature sensors, humidity sensors etc.); while a cooling system may take all this info from multiple sensors and coordinate with coolant, fans etc. to achieve desired temperature control. Similarly, traffic sensors and traffic control systems act in tandem but with different degrees communication span.
- *Degree of control/autonomy* – fully autonomous entities which act independently based on local info such robotics systems, connected vehicles etc.; Systems with some autonomy and local decision making based on predefined policies and rules; and centrally controlled system which collect the data and report to central system and act as per instructions from central system.
- *Location* – immediate vicinity/local/confined to the entity, regional and global. The sphere of influence of a given entity in performing its function.

To this end while connected vehicle, critical healthcare, Industry 4.0 equipment may need real-time responsiveness together with a large degree of autonomy to respond to changes, other applications such as traffic control systems may function well with near real-time response and regionally coordinated control, some enterprise applications may be processed with centralized services as to the response time and processing needs. A given EC deployment for specific requirement need to be tuned based on the processing needs, need for autonomy, sensitivity to response times and specificity for location.

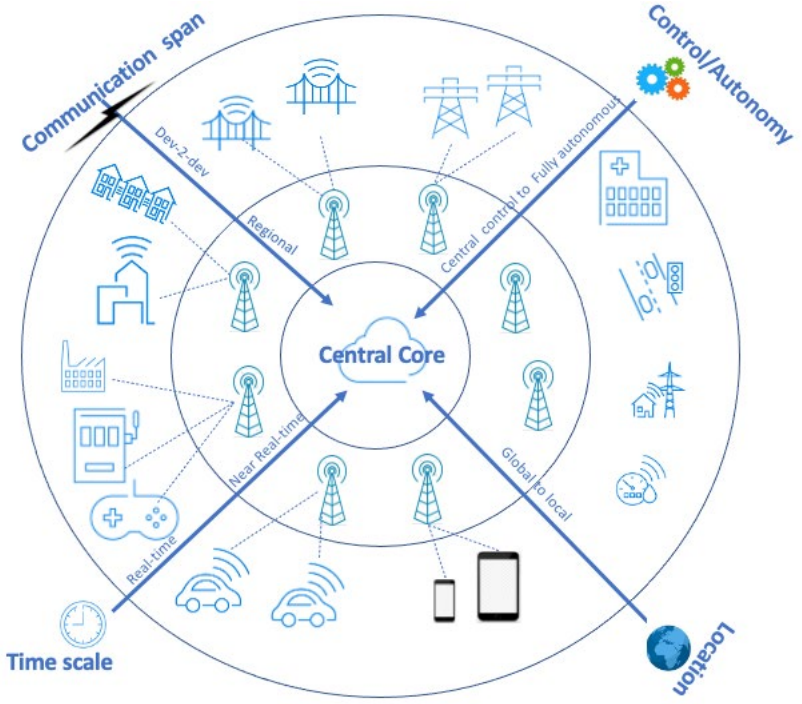


Figure 6 - Verticals that may benefit from Edge Computing

5. Edge Computing Continuum

EC is the key technology, that can support innovative services for a wide ecosystem, stakeholders for EC range from operators, infrastructure providers, application and content providers in the continuum as depicted below. It can span a variety of network locations, form factors, network, and application functions. Centralized computing is performed deeper into the network possibly in the cloud if it suits a given application, and is often used for latency tolerant bulk processing for large number of users. In contrast for latency sensitive tasks requiring some autonomy and local decision-making EC works better. In EC storage and analytics of data is performed closer to end users and devices, in close proximity to where such data is generated. This may differ for different applications, sometimes for the same application there may be multiple points of EC for different types of data in the continuum of the EC depicted above. The 3GPP and ETSI architectures depicted in Figures 2-3-4 enable realization of any combination of this as needed for an application, service or enterprise use case such as industrial control, healthcare, hospitality industry, video analytics, smart city functions, AR/VR/XR applications. Figure 7 describes the EC continuum based on latency and processing needs, assumes ultra reliable low latency communication (URLLC) for access network.

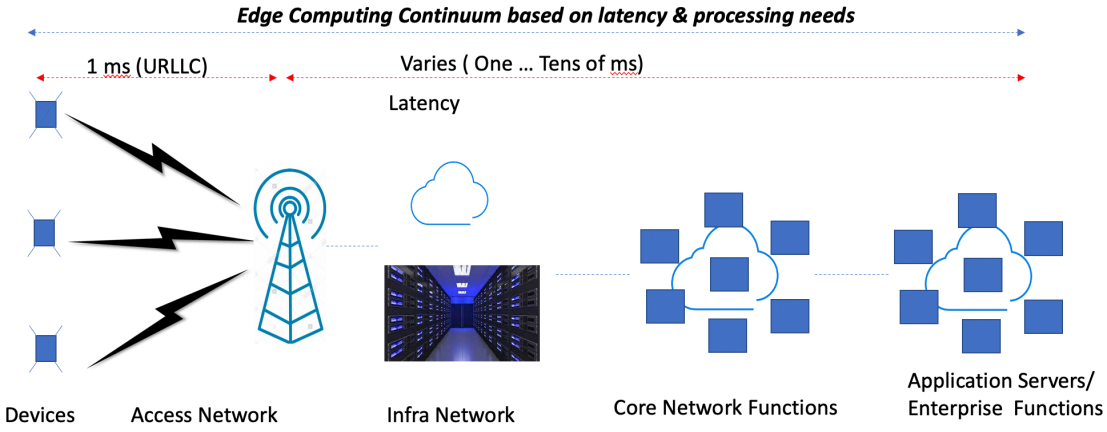


Figure 7 - Edge Computing Continuum

The spatial and temporal proximity needs between devices and application services and systems offering these services is determined by the characteristics of the services under consideration. These include real-time responsiveness, mobility, interactivity, criticality of the function (Industry 4.0, healthcare etc.). These characteristics together with the costs and affordability of a given solution shall largely influence EC deployment in the continuum depicted above. The Figure 8 adapted from [5] depicts continuum of EC from physical end point to central data center, and potential ways to instantiate and scale these. For instance, the edge instances from physical endpoint through compute edge may be singular nodes, the edge deployment from local data center through central cloud could be a cluster of nodes or cloud, based on resource needs and cost considerations.

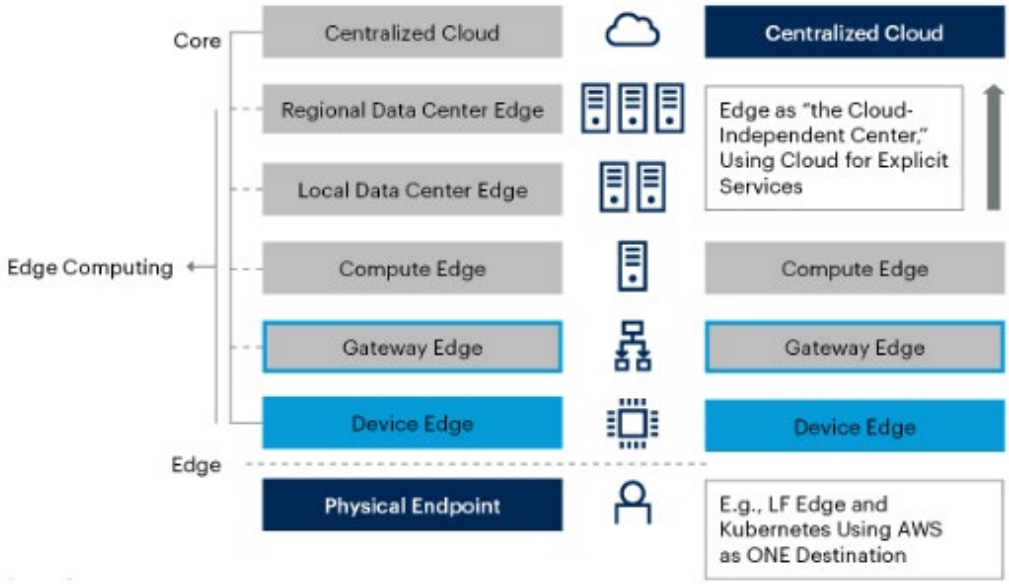


Figure 8 - Edge Deployment Continuum

The original vision for EC is to provide compute and storage resources closer to the user in open standards and in a ubiquitous manner [11,12]. EC is a crucial computing paradigm for multiple verticals IoT, AR/VR/XR, Industry 4.0, and smart cities. The basic characteristics of EC, compute, storage and latency vary widely among these verticals. The specifications from standards bodies like 3GPP and solution offerings from service providers and equipment vendors also have evolved to fulfill these needs. EC specifications from 3GPP currently can support positioning of multiple instances of UPF function at different points network to address latency and data processing needs. Similarly, solutions from vendors provide edge cloud services (e.g., AWS outpost, Azure edge etc.) and standalone servers conducive for EC from many providers.

6. Analytics And Intelligence At The Edge

There is a huge amount of data being generated by the large number of devices connected to the network both through wired and wireless networks. This data is often transported to the central core for processing, analysis and to discern insights and act on them. Often due to the latency of the transport networks and delay in processing, the full potential of this data remains untapped. Therefore, the processing of this data at the edge of the network using various analytics and deep learning (DL) techniques enables deriving insights performing timely actions to realize the value.

Edge computing enables incorporation of DL analytics technology such as computer vision (CV) – in a parking lot to detect expiry of parking duration for a car, to detect availability of parking spot, and natural language processing (NLP) – to provide context sensitive localized information about certain operations. EC can also be beneficial to many applications such as AR/VR/XR, gaming, Industry 4.0 applications, smart cities, health care and hospitality are only to name a few. Incorporation of DL into edge is a huge enabler, can make it possible many previously not feasible applications and use cases, also enhance the quality of experience for existing applications e.g., online gaming.

To realize this potential, there needs to be a match between the capability of the EC nodes, and services, DL processing needs and adequacy of the offered accuracy, latency, energy, memory footprint [9]. The available processing power, energy (battery powered devices/small nodes) and memory in the EC are often the bottleneck. Some of them can be addressed by training the models in a central location in the cloud doing much of the heavy lifting and deploying the trained models at the edge to process the locally generated data and derive insights. The key is tailoring the EC resource and DL models to match one another and meet the needs of the applications. 3GPP has intimated an effort for analytics and machine learning based insights using a federated learning (FL) model, which can potentially fulfill some of these needs.

7. Edge Orchestration And Deployment

EC imposes a few unique challenges to orchestration. Namely:

- *Resource constraints:* resources such as power (e.g., battery powered edge devices/systems), CPU, storage may be severely constrained at the edge
- *Scale:* the number of edge instances may be large, ensuring consistency and coherence across several instances may be difficult
- *Autonomy:* due to resource constraints and loss of communication may necessitate autonomous operation at the edge for extended periods of time

The degree of importance of above-mentioned factors needs to be taken into account in orchestration and deployment of an EC instance. The leading orchestration tools (OpenStack, Kubernetes, ONAP) are

currently focusing on large scale cloud deployments, some these are being adapted for edge orchestration and are still evolving.

ETSI MEC deployment model [16] allows EC to be accessible to a wide range of mobile devices with reduced latency. Figure 9 below depicts, how to make the edge services accessible to large number of devices even when the devices' current access network does not offer the service. While the device during its mobility reaches the target network, which does not offer the application services previously availed by the device, as depicted in Figure 9-II (b) the target MEC host can reach to the application services in prior serving network or alternatively as depicted in Figure9-II (c) the UPF in the target network can reach the application server through the UPF in the prior serving network.

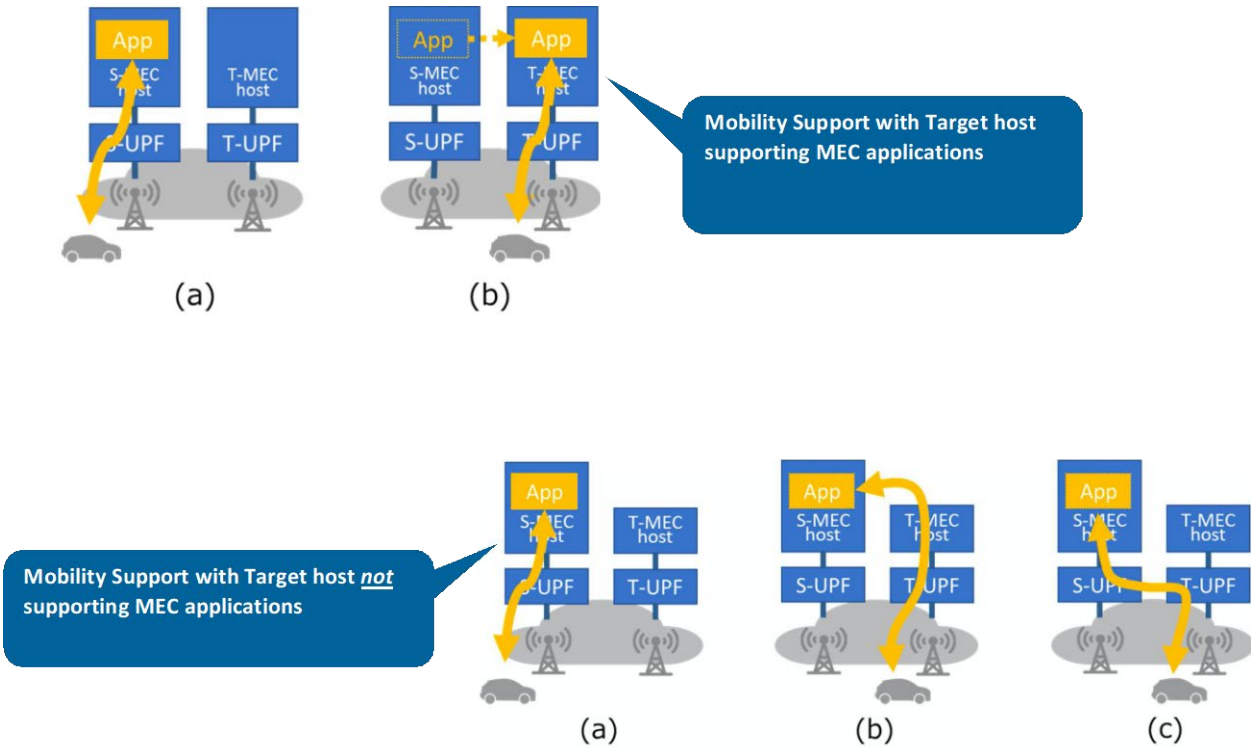


Figure 9 I & II - ETSI MEC Deployment

8. Conclusion

In the last few years significant strides have been made in enhancing the EC architecture, yet there is a need for more improvement. Currently some of EC efforts appear to be in specific silos marked for specific applications, software stacks, sources of data being used, specific cloud and network service providers. This fragmentation across different silos, service providers and multitude of software stacks constrain the stakeholders from realizing the full potential of EC. There is a need for holistic integration of these diverse domains through standardization, industry alliances and market forces. Some of it is being addressed in various standards bodies such as 3GPP, ETSI, IETF and various industry alliances and projects (MANO, ONAP, Kubernetes, etc.). It is imperative on all the stake holders to harmonize and accelerate these efforts across different standards bodies and industry alliances to realize the full potential of EC. EC is already delivering on its promise by significantly optimizing on time-to-insight, time-to-action and cost-of-insight, in the process enabling timely and effective decision making and opening new avenues of opportunities.

Abbreviations

3GPP	3rd Generation Partnership Project
5G-CP	5G control plane
AF	application function
AMF	access management function
AN	access network
AR/VR/XR	augmented reality/virtual reality/extended reality
BEC	beyond edge computing
BRAS-CP	broadband remote access server- control plane
CDN	content delivery network
C-PSA	central PSA
CV	computer vision
D2D	device to device
DL	deep learning
DN	data network
EAS	edge application server
EASDF	edge application server discovery function
EC	edge computing
EPC	enhanced packet core
ETSI	European Telecommunications Standards Institute
FL	federated learning
HPLMN	home PLMN
IETF	Internet Engineering Task Force
L-PSA	local PSA
MANO	management orchestration
MEC	multi access mobile edge computing
NEF	network exposure function
NFV	network functions virtualization
NFVI	network functions virtualization infrastructure
NFVO	NFV orchestrator
NLP	natural language processing
NRF	network repository function
ONAP	open network automation platform
OSS	operations support system
PCF	policy control function
PDU	protocol data unit
PLMN	public land mobile network
PSA	PDU session anchor
SSC	session and service continuity
SMF	session management function
UDM	unified data management
UE	user equipment
UPF	user plane function
UL CL/BP	uplink classifier/branch point

URLLC	ultra reliable low latency communications
VIM	virtualization infrastructure manager
VNF	virtualized network function
VPLMN	visiting PLMN

Bibliography & References

[1] https://www.forbes.com/sites/rkulkarni/2019/02/07/big-data-goes-big/?sh=6a0cc89120d7

[2] https://www.iotacommunications.com/blog/iot-big-data/

[3] https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

[4] https://www.weforum.org/agenda/2018/01/data-is-not-the-new-oil/

[5] Gartner, 2021 Strategic Roadmap for Edge Computing, https://www.gartner.com/doc/reprints?id=1-24JFAZOO&ct=201104&st=sb

[6] https://www.usenix.org/sites/default/files/conference/protected-files/hotedge18_slides_bhardwaj.pdf

[7] https://www.usenix.org/system/files/conference/hotedge18/hotedge18-papers-bhardwaj.pdf

[8] Edge Exchange, Bhardwaj et al. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8790194

[9] Convergence of Edge Computing and Deep Learning: A Comprehensive Survey, Xiaofei Wang, Yiwen Han, Victor C. M. Leung , Dusit Niyato, Xueqiang Yan, and Xu Chen, IEEE

[10] Fog Computing: Principles, Architectures, and Applications, Amir Vahid Dastjerdi, Harshit Gupta, Rodrigo N. Calheiros, Soumya K. Ghosh, and Rajkumar Buyya

[11] All One Needs to Know about Fog Computing and Related Edge Computing Paradigms: A complete Survey, A. Yousefpour, C. Fung, T. Nguyen, K. Kadiyala, F. Jalali, A. Niakanlahiji, J. Kong, and J. P. Jue, *J. Syst. Archit.*, vol. 98, pp. 289–330, Sep. 2019

[12] OpenEdgeConsortium: About - The Who, What, and How, http://openedgecomputing.org/about.html, Technical Report, OpenEdge Computing

[13] IETF: Problem Statement of Edge Computing Beyond Access Network for Industrial IoT, draft-geng-iiot-edge-computing-problem-statement-00

[14] 5G System Enhancements for Edge Computing, Stage 2 (3GPP TS 23.548)

[15] Multi Access Edge Computing (MEC): Framework and Reference Architecture, ETSI GS MEC 003 V2.2.1

[16] Multi Access Edge Computing (MEC): MEC 5G Integration, ETSI GR MEC 031 V2.1.1

Enabling Automation for Mapping Linear Channel Feeds and VOD Files into DASH Structures

A Technical Paper prepared for SCTE by

Yasser F. Syed, PhD.
Comcast Distinguished Architect
Comcast TPX Viper Architecture
1701 JFK Boulevard Philadelphia, PA 19103
215-286-1700
yasser_syed@comcast.com

Alex Giladi
Comcast Fellow
Comcast TPX Viper Architecture
1899 Wynkoop St., Denver CO 80206
215-286-1700
alex_giladi@comcast.com

Ali C. Begen, PhD , Consultant, Comcast

1. Introduction

Established MPEG-2 TS based video workflows have been used for distributing content in traditional consumer delivery ecosystems such as QAM-based cable, satellite, ABR Streaming, as well as over-the-air broadcast. With the advent of adaptive streaming technologies such as MPEG DASH and Apple® HLS, these same content video workflows have been adapted to these new types of video distribution. With that said, manual workarounds are often needed to fit in the new features enabled by adaptive streaming. Thus, the ABR distribution workflows may still depend on MPEG-2 TS source content delivery systems for several years due to existing equipment and network tools, and the need to provide traditional QAM linear services [4]. It is expected that backend systems of MVPDs will change as third-party content contribution and demultiplexed media component delivery becomes more widely used.

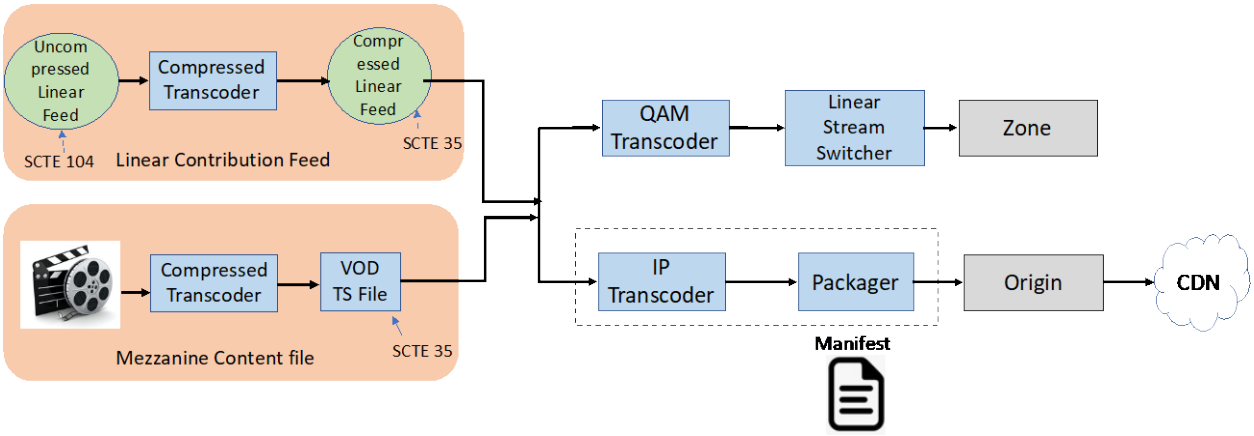


Figure 1 – Source Content workflows for contribution linear feeds and mezzanine files to distribution ABR Streaming and QAM services

This paper describes new mechanisms in existing structures added, proposed or planned in the MPEG-2 Systems (ISO/IEC 13818-1 8th ed) [9], SCTE 214 (DASH constraints) [3]. These additions to existing specifications can then be used in adapting the infrastructure to allow for better automation in mapping of source content into DASH structures for distribution of ABR streaming services.

2. Problem Statement

Between providing MPEG-2 TS source content to the MVPD through traditional source linear channels or VOD mezzanine files and then transcoding/packaging/ and distributing that content for adaptive streaming services, there is no clear mapping from source to distribution that can adjust to source content variations and automate this into adaptive streaming services while providing DASH-enabled, enhanced consumer experiences to the content.

Adaptive Streaming technologies can provide a more individualized customer experience for an asset with choices to match accessibility, language, commentary choices of the viewer. With that said, this is predicated on media content being identifiable and available for these purposes. For example, a selection of English, Spanish, French, and Chinese subtitles assumes that the player is aware of their existence, while a good user experience also depends on a sensible default client behavior. The contribution video ingest infrastructure for video feeds or VOD mezzanine is still operating on an MPEG-2 TS format, be it via fiber, satellite, or IP using TS-over-UDP, SRT, or RIST delivery protocols.

The MPEG-2 TS structure is more geared towards ingestion for QAM-based linear channels and VOD assets where language offerings are more static and limited to 1-2 choices which are determined by hard-coded PID convention ordering (e.g.,101- English, 102-Spanish). Automatically mapping these existing methods over to DASH Role and Accessibility constructs would limit players to static configuration of the channel or the VOD Asset unless a assets specific specialized workaround is done [4]. Additionally, accessibility services such as Audio Description (AD)/ Descriptive Video Services (DVS) are gaining popularity [10]. Initially accessibility tracks were represented using the same hard-coded audio PID construct and often signaling dead or archaic languages (e.g. Middle English) was needed to distinguish between these and the secondary language. Automatically mapping such a setup into DASH constructs is predominantly a manual process because the orthogonality between language and accessibility characteristics of audio streams were determined through PID numbering conventions and overloaded use of the ISO_639 language codes carried in the audio descriptor in the PMT. Often accessibility features would differ program by program on the channel which would require manual per-channel configuration which is usually limited to be a channel configuration instead of a program configuration. Moreover, this approach does not scale.

Lastly, the consumer experience from switching from a main program to an ad needs to be consistent. Achieving consistency becomes harder as the number of player options increase. Similarly, ads also need the same information for smooth playback experience. At the streaming player, another factor needed is the default playback mode set by the customer [10]. Once the player default playback information is known, the switching behavior between the main program and ad, and the playback behavior, can be deterministic and consistent even if the content experience between the main program and the ad do not overlap. But the current mechanisms cannot carry this information dynamically so it can change program to program. An understanding of the content experience offerings in the main program and the default consumer experience of the main program is needed to provide a deterministic behavior for the customer experience as the program moves into an ad or alternate content situation.

3. Latest additions to the MPEG-2 TS language descriptor

The MPEG-2 TS specification contains an ISO_639_language descriptor which has a three-character ISO 639-2 field for the language and an 8-bit audio_type field. [9][15].The descriptor resides in the PMT section. This information can change every PMT occurrence but should be limited to program boundary changes. The descriptor can add multiple table entry pairs of ISO 639-2 languages codes and audio_type codes, but current usage anticipates a single language-audio_type pair. In earlier editions of the MPEG-2 Systems standard, the audio_type codes were limited to accessibility roles -- visually impaired commentary (0x03), hearing impaired (0x02), clean effects (0x01), in addition to undefined (0x00) used for all other audio roles. In past common usage, a single occurrence of the language was used but the audio_type was limited to the undefined values [9][15]. If there was further need to define the audio, the bsmod descriptor values were used but were limited to specific types of Audio formats. These were tolerable in QAM delivery but become more of an issue with multiple different audio formats needing to be carried in adaptive streaming environments [3].

With the 8th Edition of ISO/IEC 13818-1, the structure of the ISO_639_language descriptor was not changed as it is ubiquitously used throughout the ecosystem. However, more values were added to the audio_type table [9]. The resulting efforts provide a clearer approach to map information from the MPEG-2 Stream into DASH values for audio services and accessibility associated services.

Table 1 – Mapping source audio_type values or bsmod descriptors (for (E)AC3) to Role and Accesibility values for audio services

Source Audio_type or bsmod descriptor	Role@value	Accessibility@value
Audio default (audio_type = 0x00 | bsmod [ST] = 000)	Main	N/A
Clean effects (audio_type = 0x01 | bsmod [ST] = 001)	SCTE: Music & Effects	N/A
Primary Audio (audio_type = 0x80)	main[i]	N/A
Native Audio (audio_type = 0x81)	absence of dub	N/A
Emergency (audio_type = 0x82 | bsmod [ST] = 110)	Emergency	N/A
Primary Commentary (audio_type = 0x83 | bsmod [ST] = 101)	main[ii], commentary	N/A
Alternate Commentary (audio_type = 0x84)	alternate, commentary	N/A
Bsmod [ST] = 100 or 111	TBD	N/A

Table 2 – Mapping source audio_type values or bsmod descriptors (for (E)AC3) to Role and Accesibility values for accessibility associated services

Type	Role@value	Accessibility@value
Audio description (audio_type = 0x03 | bsmod [ST] = 010)	alternate	description
Clean audio (audio_type = 0x02 | bsmod [ST] = 011)	alternate	enhanced-audio-intelligibility
Closed Captions[1]	main	captions
Sign language[2]	supplementary	sign

Tables 1 and 2, above, exemplify how one can map the signaling in the MPEG-2 TS mezzanine to the DASH Role or Accessibility values in the DASH MPD for consumer delivery. A similar mapping can be constructed for HLS which uses a different vocabulary for the purpose. It is also possible to map the bsmod values in (E-)AC-3 bitstreams to the DASH constructs if the information is not provided through the ISO_639_language descriptor. AAC provides functionally similar accessibility parameters. Given a basic pair of language code and Audio_type (set to undefined), this would allow the manifest to define a main role for each audio component using a specific language. With the addition of multiple pairs of Language-Audio_types, these same media components could also additionally define the stream as containing AD/ DVS. The tables provide a reference to automatically map what is listed on the transport stream, to role and accessibility assignments in the manifest. To be noted, this mapping in Table 1 and Table 2 is compatible with the role assignments and accessibility assignments of DVB in areas where these configurations overlap [7].

[1] Closed captioning is an accessibility component for a video or text track indicated by the caption service descriptor in the PSI. The equivalent audio_type value would be 0x00

[2] Sign language is identified through the @lang attribute (e.g. “ase” or “bfi”, as defined in ISO 639-3). The proper audio_type value would be 0x00.

In the 8th edition, the audio_type now has additional values of *primary, native, emergency, primary commentary, and alternate commentary*. With these additional values and the use of multiple language-audio_type pairs, there is even more that can be automatically transformed into DASH contexts [9]. For instance, using an audio_type value of "native" can distinguish between original and dubbed audio tracks in a manner that can be automatically processed. With the use of primary and alternate content, a sporting event could have more than one announcer, or just stadium sound, depending on the preference of the listener. In VOD assets, the number of audio experience choices can keep growing but with some direct mapping these can be automatically captured within a single manifest. But even for linear channels, these additional values for audio_type can be beneficial to indicate things like sporting event, native audio, emergency channel audio, and primary audio of the channel. Furthermore, additional information on the audio tracks can benefit DASH constructs by providing a way to have continuous audio media components in the created period, through avoiding overloading of the ISO 639 language value to additionally indicate audio streams with properties like audio description.

Figure 2 and Figure 3 show how the MPEG-2 TS language descriptor values could be set to dynamically provide several playouts, for both linear channels and mezzanine files. These can vary from program to program on a linear channel, or asset to asset on VOD services, so the customer experience offerings can be tailored to the content as well as the experience. For VOD services, there are fewer limitations on customer experiences. The number of languages allowed, for instance, can make a content asset more worldly. Distinguishing between dubbed and native sound, and additional accessibility options, can enable people to hear dialogue more clearly. This benefits both the hard of hearing, and people wanting to watch loud action movies at 2am without disturbing the rest of the family.

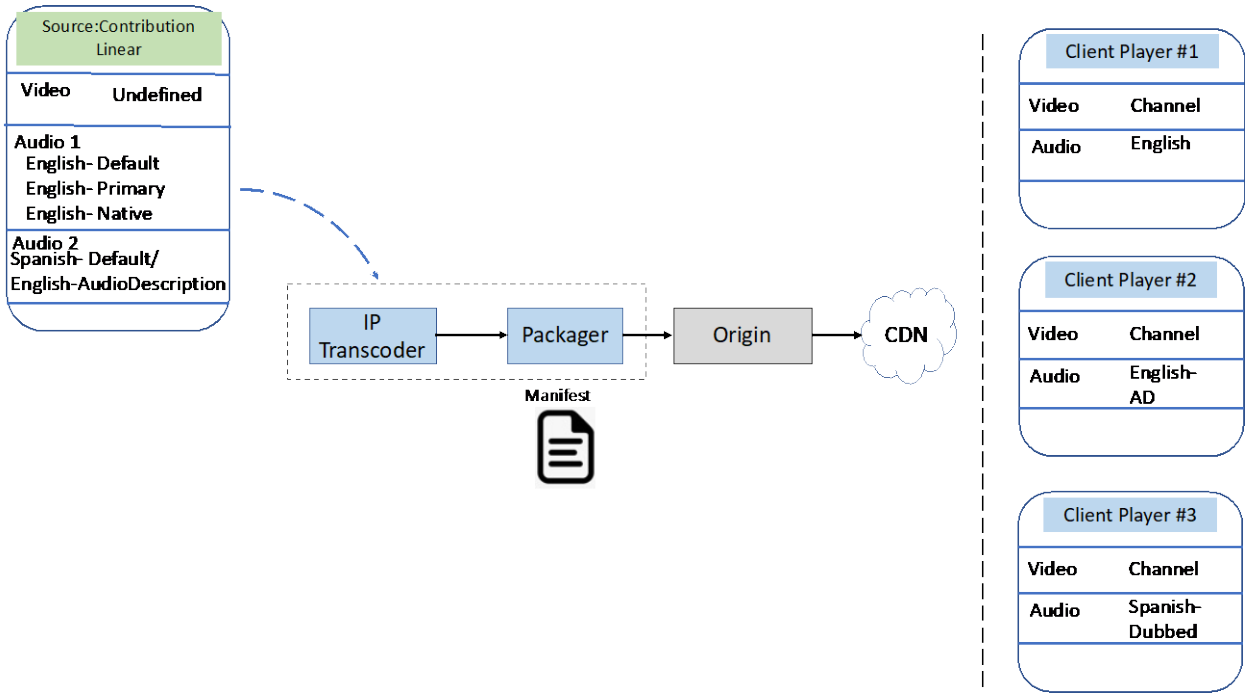


Figure 2 – Example of contribution feed configuration and resulting player options

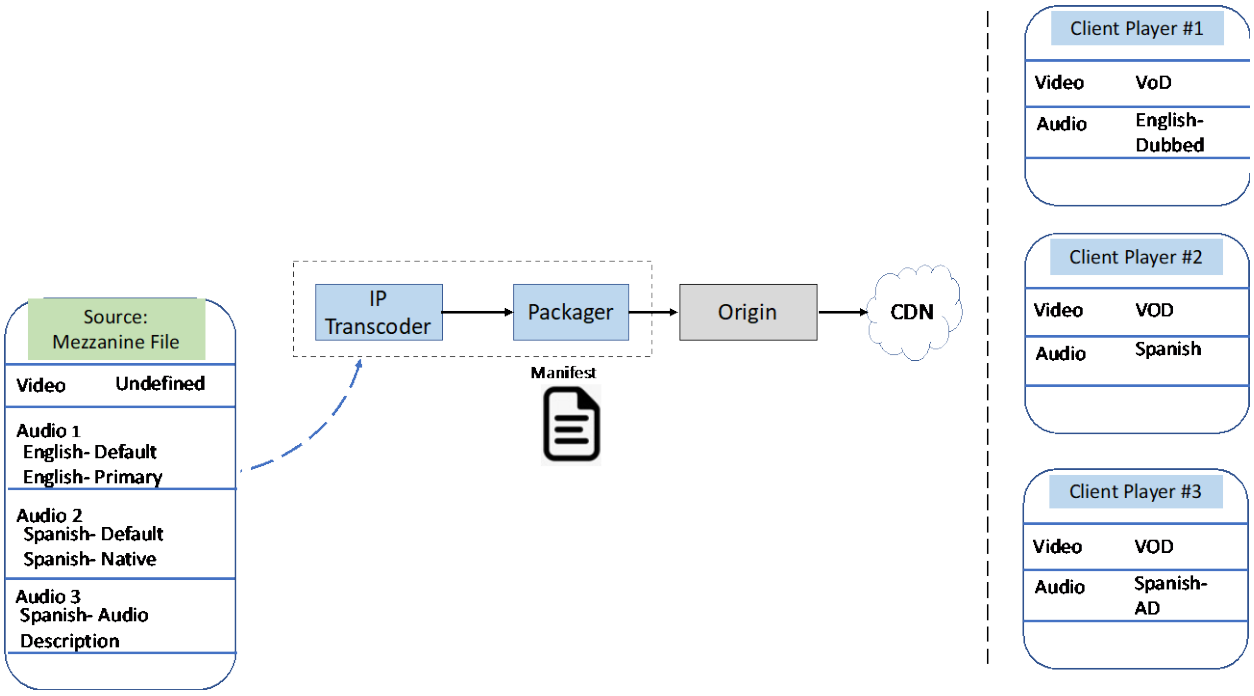


Figure 3 – Example of mezzanine file configuration and resulting player options

Figure 4 shows how the audio_type values can be set to handle different customer experiences for a sporting event. In this case, a sporting event can be offered with two different announcers, or just ambient stadium sound. This can be tailored to use cases like supporting a home announcer and an away announcer (including using different languages), or none at all, or including 3rd party commentary during a sporting event. For example, sourcing comedians for color commentary on sporting events is increasingly popular, as evidenced by the Olympic games.

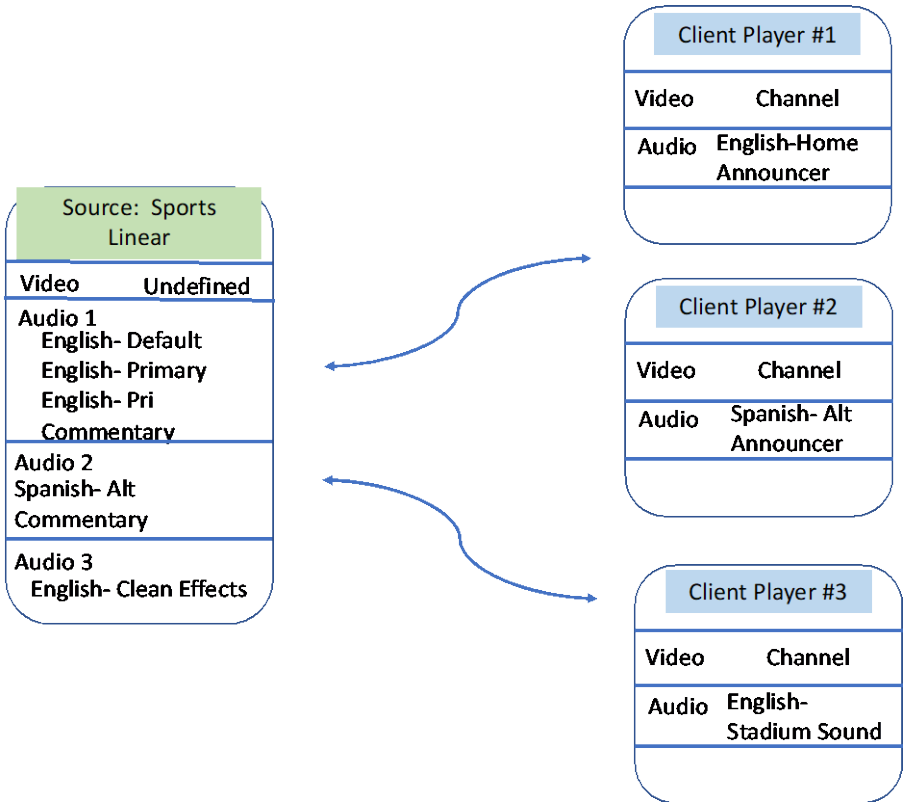


Figure 4 – Sport event contribution feed configuration and resulting player options

4. Using planned modifications to SCTE 214-1

The upcoming SCTE 214-1 revisions have proposed a supplemental property for media language defaults. From a DASH perspective, the player can set consumer preferences for media languages. The NorDig IRD specification provides an example of such logic [16]. From the content side, there was also a media language default that can be indicated in the ISO-639_language descriptor through assigning an audio_type *primary* value, but there is no way to reflect this in the DASH manifest until the media language default supplemental property was proposed. Knowledge of this information helps ad insertion where the language options may not exactly match the main content primary languages. With expanded choices in languages for the content asset, the playout experience can be disruptive between the main program and the 3rd party ad (e.g., switching from French main program to a Spanish ad) and inconsistent depending on implementation. With the primary language of the channel known, this can allow for deterministic ways to map to the customer experience at the client player. Even with an assigned language default language option at the player side, this information is useful especially when no languages overlap -- which can happen during an ad since the ad is independently prepared from the program or asset content. This supplemental property can be placed at the MPD level to indicate language default of the entire content, or provide a language default at the period level, which can override what is set at the MPD level. This supplemental property can also be indicated at a period level as well to indicate in cases of inserted ads what default language to use especially if there is no overlap with the main program [1][2]. From an automation perspective, providing this language default information in the manifest allows an

approach that provides consistent and deterministic playout behavior for the customer experience, even across main program and alternate content such as Ads. Examples of this are shown in Figure 5.

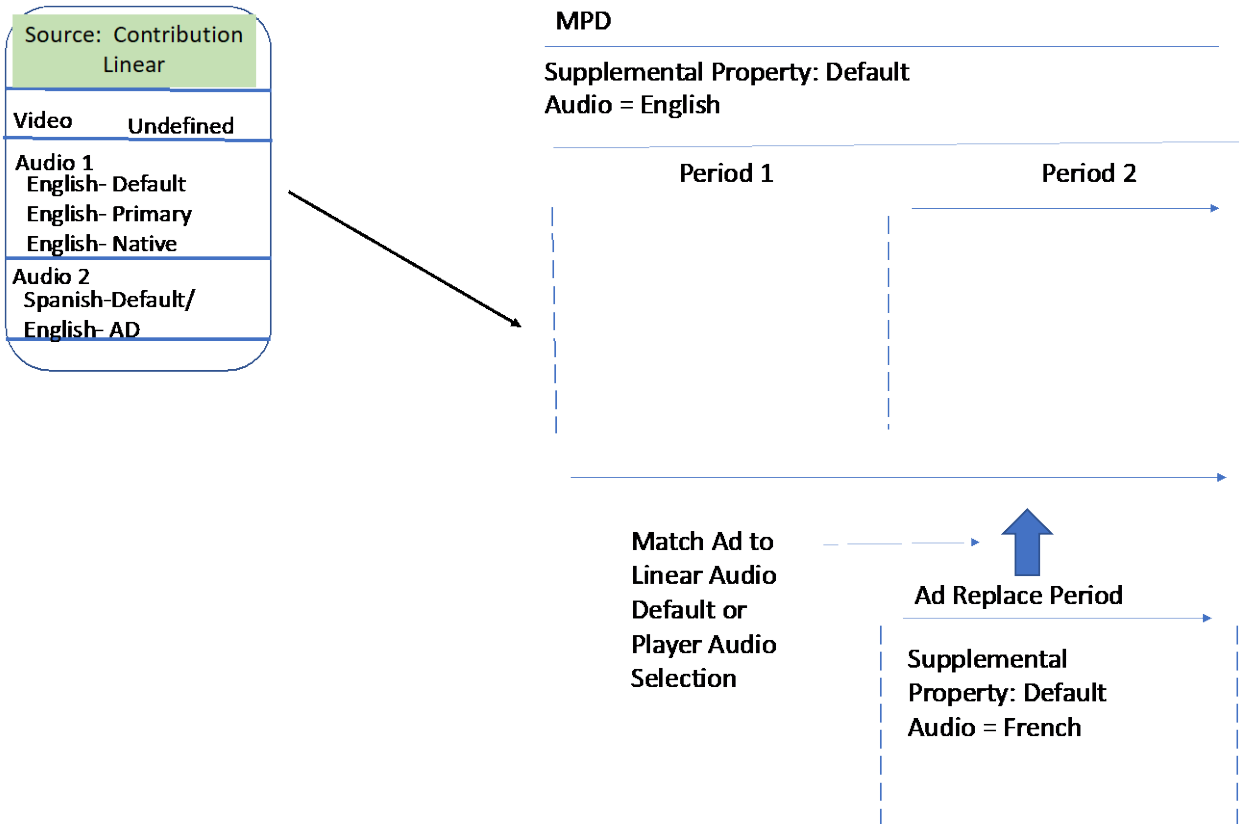


Figure 5 – Use of Media Language Defaults to Align Playout of Main Program and Ads

5. Future and International Extensions

The initial modifications described here worked by only adding additional values to the ISO_639_language structure [11][12]. But the makeup of content assets and linear feeds is expanding and evolving which can go beyond the ISO_639_language Descriptor structure on the audio stream. For instance, more accessibility features such as Video Sign Languages which is placed on the video stream cannot be well described just using the ISO_639_language descriptor. It also introduces the shift from having the content asset built around a single video stream to having several dependent video streams from 3rd parties as part of the content asset. With adaptive streaming there is also a shift to demuxing the media components which may manifest upstream delivery as demuxiplexed MPEG-2 TS streams which may originate from different parties (e.g., multi-language subtitles, dubbed audio, multi video sign languages).

In the next amendment of 13818-1, a new descriptor is proposed called the *Media_service_kind* descriptor that can co-exist with the ISO_639_language descriptor. This allows for a transition strategy from the old descriptor to the new descriptor depending on equipment software modifications. This new descriptor

expands from the ISO_639_language descriptor by applying to video and text streams as well as the audio descriptor and provides a main and dependent relationships between the different media components of the content asset. It also provides more additions to the media_type (similar to audio_type but extended to additional types of media components such as video or text) to address a more complete match between the full set of DASH roles and what is carried in the transport stream which includes values such as forced subtitles, substitution and dialogue. This also accommodates the future trends of having the media components of content being delivered on separate paths by multiple parties.

For the international requirements, the language code will be referencing ISO 639-3 to accommodate multiple sign languages and more dialects with BCP 47 extensions to handle regional dialect languages and scripts [5][6]. Furthermore, it will handle multi-language native audio tracks that may differ from a program-to-program basis but always provide the native audio of the program (ISO 639-3 “mul” code point.) In different parts of the world, this can be a differential playback mode for individuals who know multiple languages and would prefer the program to be played in the native language.

6. Conclusion

Enabling automation needs to define a clear path from mapping content source transport streams to distribution adaptive streaming manifests. With the additional values of audio_type in the ISO_639 language descriptor, adding the supplemental property media defaults in SCTE 214 [3]. We believe these enhancements will provide a clearer path from mapping information carried from the content source into its equivalent DASH roles and accessibility values.

The benefits of having this automation are clear. With DASH Role and Accessibility elements, the content is well described which benefits the customer by providing many more different types of experiences with the same asset and making it more accessible. In the past multiple assets of the same content had to be created separately and displayed differently to capture the different ways of viewing the content asset. But unless there is a way to automatically map this information from MPEG-2 TS structures, the presentation of the content would be limited to avoid the complexities of manually and correctly adding this information to each linear feed or VOD asset. The need to map from a MPEG-2 TS structure is needed due to the existing ingest system that is in place.

Furthermore, there needs to be an approach for this that works for inserting local ads, replacing national ones, or providing alternate content. The content experience can be affected by the ad playout of 3rd party content. Minimal disruption of the experience should be provided as an ad gets played out and when there is no way to avoid disruption of the experience during an ad then the behavior should be deterministic and consistent (e.g. keep the ad in Spanish even if the Spanish audio description is not available).

As content gets more internationalized, the customer experience expands, and these approaches provide a way of including these experiences while still providing a way to put this under the same manifest without changing backend operations.

Lastly in the future content assets and experiences will be expanding to include demuxed media component delivery and things like 3rd party independent signing tracks or new sporting or live event experiences. In anticipation of this, we believe the development of the media service kind descriptor in the next MPEG-2 TS systems standard edition will be needed as well as an approach to have these two descriptors co-exist for some time.

Abbreviations

ABR	Adaptive Bitrate
AD	Audio Description
bps	bits per second
CDN	Content Distribution Network
DASH	[MPEG] Dynamic Adaptive Streaming over HTTP
DVB	Digital Video Broadcasting
DVS	Descriptive Video Services
Ed	Edition
FEC	forward error correction
GOP	Group of Pictures
HD	high definition
HLS	HTTP Live Streaming
IEC	International Electrotechnical Commission
IP	Internet Protocol
ISBE	International Society of Broadband Experts
ISO	International Organization for Standardization
MPD	Media Presentation Description
MPEG-2 TS	MPEG -2 Transport Stream
MVPD	Multichannel Video Program Distributor
PID	Packet Identifier
PMT	Program Map Table
QAM	Quadrature Amplitude Modulation
RIST	Reliable Internet Streaming Transport
SAP	Stream Access Point
SCTE	Society of Cable Telecommunications Engineers
SRT	Secure Reliable Transport
VOD	Video On Demand
UDP	User Datagram Protocol

Bibliography & References

[1] ANSI/SCTE 35, Digital Program Insertion Cueing Message for Cable.

[2] ANSI/SCTE 104, Automation System to Compression System Communications Applications Program Interface (API).

[3] ANSI/SCTE 214-1, MPEG DASH for IP-Based Cable Services Part 1L MPD Constraints and Extensions.

[4] ANSI/SCTE 223, Adaptive Transport Stream

[5] Phillips, A. and M. Davis, "Matching of Language Tags", BCP 47, RFC 4647, September 2006.

[6] Phillips, A., Ed., and M. Davis, Ed., "Tags for Identifying Languages", BCP 47, RFC 5646, September 2009.

[7] ETSI TS 103 285 V1.3.1, Digital Video Broadcasting (DVB); MPEG-DASH Profile for Transport of ISOBMFF Based DVB Services over IP Based Networks.

[8] [HLS I-D] IETF RFC 8216bis, HTTP Live Streaming

[9] ISO/IEC 13818-1, Information technology - Generic coding of moving pictures and associated audio information: Systems

[10] ISO/IEC 23009-1: Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats

[11] ISO 639-2, Codes for the Representation of Names of Languages - Part 2: Alpha-3 Code

[12] ISO 639-3, Codes for the Representation of Names of Languages - Part 3: Alpha-3 Code for Comprehensive Coverage of Languages

[13] [RIST] VSF TR-06-1, Reliable Internet Stream Transport (RIST) Protocol Specification (Simple Profile), October 2018

[14] [SRT] M.A. Sharabayko, J. Dube, J.S. Kim, J.W. Kim, The SRT Protocol, March 2021, available at https://datatracker.ietf.org/doc/html/draft-sharabayko-srt-00

[15] Jan Van Der Meer, Fundamental and Evolution of MPEG-2 Systems: Paving the MPEG Road, Whiley, 2014.

[16] NorDig Unified Requirements for Integrated Receiver Decoders for use in cable, satellite, terrestrial, and managed ITV based networks, 27 October 2018, www.nordig.org

[i] The supplemental property for media language default for audio should be set at the MPD or Period Level
[ii] The supplemental property for media language default for commentary should be set at the MPD or Period Level

Enabling Encryption and Algorithm Revocation for Post-Quantum DOCSIS Certificates

Novel Results in Multi-Key Trust Environments Deployments

A Technical Paper prepared for SCTE by

Dr. Massimiliano Pala

PKI Architectures Team, Director

Cable Television Laboratories, Inc.

858 Coal Creek Cir, Louisville, CO

603.369.9332

m.pala@cablelabs.com

1. Introduction

The cryptography world is going through a revolution. As new computation paradigms emerge and rapidly advance, like quantum computing (QC), the broadband industry needs to start planning how it will address the new security threats that are on the horizon.

Most of the public key cryptosystems like RSA [Rsa16] or ECDSA [Ec05] will not be considered secure when (and if) a large quantum supercomputer is ever built. For the broadband industry this means that, because of the dependency on X.509 [X509] certificates and the RSA algorithm, to provide devices with secure and verifiable identities, the protocols that are used today, e.g. DOCSIS® protocols [Doc31; Doc40], will need to support new algorithms and identities. In fact, network elements like cable modems or Remote PHY (R-PHY) nodes [RPhy18] use, today, their RSA private key and associated certificates chain to prove they are a legitimate and registered entity on the network. To continue to benefit from the security and usability advantages of public-key cryptography (PKC), the broadband industry must provide a mechanism for transitioning to quantum-resistant solutions *in a cost-effective manner*. Although our previous results on Composite Crypto (or Hybrid certificates) provided a promising path forward for the deployment of multiple keys associated with a single identity, our work still left some important questions. For example, an area that was still left to be explored was how to handle complex crypto policies for algorithm validation and deprecation. Because of these limitations, encryption was also left out of scope.

This paper describes our new results in multi-key environments that address the open issues from our previous work and update its technical details [Pala04]. Specifically, in this work we extend the initial proposal and introduce the explicit separation of “AND” and “OR” logic operations across the multi-key signature components. Additionally, our work enables encryption for multi-key certificates (e.g., for S/MIME or document multi-signing purposes) that was, up to now, still an open problem. Together with these important results, this paper also describes our proposal for algorithm revocation and how we leverage the details of X.509 certificates’ public key structures together with extensions in CRLs and/or OCSP responses to provide a dynamic, centrally managed, and easy to deploy algorithm revocation mechanism.

The rest of the paper is organized as follows: Section 2 provides an overview of the current landscape of Post-Quantum (PQ) cryptography and how it addresses the quantum threat. Section 3 describes the composite crypto solution and highlights current limitations of multi-key certificates when it comes to validations or encryption; Section 4 describes the new results that stem from the introduction of Combined Crypto alongside Composite Crypto; Section 5 provides the details on our algorithm revocation mechanism. Section 6 addresses the multi-key encryption conundrum and, finally, Section 7 provides our conclusions and envisioned future work.

2. The Post-Quantum Cryptography Landscape

Although the standardization process that is currently undergoing at NIST has not yet completed, there are interesting trends and practical long-term considerations for PQ algorithms deployment within the broadband industry that we can already highlight.

In order to understand how the new algorithms address quantum resistance, it is important to look at the principles behind solving the Hidden Subgroup Problem (or HSP) and how quantum computers can leverage superposition, entanglement, and interference to efficiently solve HSP for relevant domains.

2.1. The Hidden Subgroup Problem (HSP) and Factoring Keys

When it comes to the link between "classic" cryptography, like RSA or ECDSA, and quantum-based factorization algorithms, such as the one proposed by Schorr, it is not always easy to understand how periodicity comes into play and how post-quantum algorithms address quantum-resistance.

In this section we provide a qualitative explanation of HSP for the classic and post-quantum use-cases by introducing group theory concepts and their intersection with cryptography.

2.1.1. Groups, Cosets, and H-Periodic functions

With the use of modular arithmetic, a cornerstone in modern cryptography, we introduce, de facto, periodicity in the form of cyclic groups. These mathematical constructs consist of a set (e.g., "integers mod N") and a binary operation that takes two inputs and generates outputs in the same set:

$$G \times G \longrightarrow G$$

A group is characterized by three fundamental properties: (a) associativity, (b) a neutral element, and (c) all elements in the group have an inverse. Commutativity is not a core characteristic of a group and this, specifically, is a key differentiator when looking at quantum-resistance as explained in the rest of this section. A commutative group is also called an Abelian group.

When it comes to group theory, there are two definitions that must be well understood: subgroups and cosets. A subgroup H of a group G is defined as a group that still satisfies the group properties and is generated by one or more elements of the group (e.g., "h"). A coset is a similar concept to a subgroup in the sense that it can be seen as "translated" subgroups with respect to an element of the group G. For example, given a subgroup H generated by two elements (h_1 and h_2) and a third element "x" in the group G, the "coset of H with representative x (element of G)" is generated by applying all the permutations starting from the element "x" instead of starting from the neutral element.

The definition of H-periodic functions is strictly connected to the definition of cosets. When a function maps the values of a group to a set, it is said to be H-periodic (H is a subgroup) if, for all cosets xH, the value of the function is the same on all the elements in xH and differs on all elements of the other cosets.

2.1.2. The Hidden Subgroup problem and classic cryptography

The solution to HSP over specific groups can lead to breaking classic and post-quantum cryptography by leveraging the ability of a quantum computer to efficiently find the period of the underlying H-periodic function.

An example of this approach is explained in the famous Schor's paper from 1997 [Shor97]. In his work, Shor teaches us how to use quantum computers to efficiently implement the period finding function which is at the core of the factorization problem. In the case of "classic" algorithms, like the RSA or ECDSA, the underlying groups are commutative and, therefore, easier to deal with. For example, the group definition for the RSA case is $G_{RSA}= (\mathbb{Z}_N, +)$, while $G_{ECDSA}= (\mathbb{Z}_N \times \mathbb{Z}_N, +)$ provides the definition for the ECDSA one. The commutative property of these groups allows us to use the Fourier analysis on Abelian groups by using the Quantum Fourier Transform operation. In the RSA case, for example, given access to the function f that computes exponentiation modulo n, the factorization problem can be reformulated as finding a generator of the subgroup $H = \varphi(n)\mathbb{Z} \in \mathbb{Z}$, where $\varphi(n)$ is the group order and $\mathbb{Z}/n\mathbb{Z}$ is the set of integers modulo n. We can then use the function f as the oracle function for the subgroup H as:

$$f: \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} : x \mapsto a^x mod\ n$$

Once the generator $\varphi(n)$ is found through repeated sampling, computing the factorization of n can be accomplished by using the greatest common divisor (GCD) technique to find the non-trivial factors from the measurements.

2.1.3. The Dihedral Hidden Subgroup problem and Lattices

An interesting group that is relevant for post-quantum cryptography is the symmetry group. This group is defined as the set of all the possible elements permutations π (i.e., N-1 rotations, N reflections, and the neutral element) together with the functional composition operation. The neutral element is, in this case, the permutation that maps everything to itself, i.e., the identity element.

When looking at post-quantum algorithms and their relationship with HSP, we need to start from Regev's 2002 work on HSP. In his paper on Quantum Computation and Lattice Problems [Reg02], Regev shows how a solution to the Unique Shortest Vector Problem (SVP) can be obtained under the assumption that an algorithm that solves the hidden subgroup problem on the dihedral group by coset sampling exists. Regev's work demonstrates the equivalence between solving SVP on lattices and solving HSP on the dihedral group. The main difference with the classic use case is the fact that the group (i.e., domain and operation) for which a solution of the HSP is needed are non-commutative. As a reminder, the dihedral subgroup is a subset of all the permutations that are automorphisms (or symmetries) of the N-cycle (i.e., all the permutations that preserve the structure of the N-sided regular polygon) which include rotations and reflections. The Hidden Subgroup Problem for the dihedral subgroup can be defined as "given an H-periodic function, find H (or find the generators of H)".

Although also in the noncommutative subgroup problem the use of the Fourier analysis is at the center of efficient quantum-based solution, the difficulties of performing it on noncommutative groups makes the noncommutative version of the problem very challenging. Ettinger and Høyer [HeHø04] showed that efficiently solving the HSP for noncommutative groups is possible. More precisely, they show that it is possible to obtain sufficient statistical information about the hidden subgroup with a polynomial number of queries (similarly to the "classic" use case) … However, no known efficient algorithm exists that can leverage this information to find the generator for the subgroup. In their paper, Ettinger and Høyer state:

"Our main result is that there exists a quantum algorithm that solves the dihedral subgroup problem using only a linear number of evaluations of the function which is given as input [...] However, we hasten to add that our algorithm does not run in polynomial time. […] the algorithm applies a certain quantum subroutine a linear number of times […]. We know how to find the subgroup from the data in exponential time, but we do not know if this task can be done efficiently."

The original algorithm from Kuperberg from 2003 to solve HSP on the dihedral group runs in sub-exponential time $\widetilde{O}(3^{\sqrt{2\log_3 N}})$. Known improvements on these constructions are due to Regev [Reg04] and again Kuperberg [Kup13] where the total computation time is estimated to be $\widetilde{O}(2^{\sqrt{2\log_2 N}})$. Table 1 provides the group details (i.e., domain and operation) and specific application of HSP for different groups and well-known applications. For example, solving the HSP for the group of the integer numbers domain ($\mathbb{Z}_N$) with the addition ($+$) operation and (0) as the neutral element can be used in RSA factorization, while ECDSA and El-Algamal algorithms can be broken by solving the HSP for the group identified by the $\mathbb{Z}_N \times \mathbb{Z}_N$ domain with the addition ($+$) operation. In this case, the operation is the component-wise addition, and the neutral element is the pair (0,0).

Table 1 - List of Hidden Subgroup Problem definition and their applications

HSP Group	Operation	Application
$\{0,1\}^n$	XOR	Simon's Algorithm
$\mathbb{Z}_N$	$+ \ mod \ N$	Period Finding Function
$\mathbb{Z}_N$	$+$	Shor's Factoring Algorithm (RSA)
$\mathbb{Z}_N$ x $\mathbb{Z}_N$	$+$	Shor's Discrete Logs (ECDSA, El Algamal)
"Dihedral Group"	Composition of Symmetries (rotations, reflections)	Approximate SVP (and CVP)

2.2. Quantum-Resistant Cryptograpy

As we have seen, lattice-based cryptography does not come, so far, with efficient algorithms, quantum or classic, that can solve the underlying problem efficiently. That is why some of the most promising algorithms that are still present in the final round of the NIST competition are lattice-based. These mathematical objects are, in practice, regular collection of equally spaced vectors or points. In other words, lattices are regular arrays (or grids) of points that are generated by a combination of basis vectors. Lattice-based cryptography properties are rooted in the hardness of solving certain topological problems for which we do not have an efficient algorithm for, like finding the Shortest Vector Problem (SVP) or the Closest Vector Problem (CVP) given a specific basis for the lattice. Algorithms like Falcon [Fa17] or Dilithium-Crystals [Di17] fall in this category and produce the smallest authentication traces overall (i.e., signatures range from 700 bytes to 3300 bytes).

Another class of algorithms to keep an eye on is the Isogenies-based ones. These algorithms use a different structure than lattices and have been proposed for key-exchange algorithms, namely Key

Encapsulation Mechanisms or KEMs. Specifically, isogeny-based cryptography combines morphisms (or isogenies) among elliptic curves to provide Perfect Forward Secrecy (PFS) properties. Although Isogeny-based cryptography is computationally very heavy, it uses the shortest keys in the post-quantum algorithm landscape. SIKE is an example of such a class of algorithms.

Together with these two classes of algorithms, there is another type of algorithm that should be kept in our minds as a possible alternative: hash-based signature schemes. These algorithms rely on very different security property and data structures that are not tractable via HSP. The main issue with stateless hash-based schemes is the size of signatures: the lack of structure in the data comes at the expense of very large cryptographic signatures (although public keys are extremely small). Although the size of signatures hinders, today, their adoption, the security of this class of algorithms is not affected by advancements in HSP solving for non-commutative groups. A well-known hash-based algorithm that will probably be re-included in the NIST standardization process because of its security properties is SPHINCS+ [Sp15].

3. Multi-Keys Trust Environments

X.509 certificates have, so far, been used to link one public key to a single identity. This is true for Trust Anchors (or TAs), Intermediate CAs (or ICAs), and End-Entities (or EE). However, because the encoding of public key data structures inside certificates depends solely on the specific OID used to identify them, the inner BIT STRING that encodes the key value can be re-engineered to accommodate for any data structure. In our original work that was presented at SCTE Tech Expo 2020, we used the algorithm agility feature built in into X.509 certificates and defined a new OID to identify a key structure which implements a `SEQUENCE of SubjectPublicKeyInfo` structures. Each of the structures in the Composite Key encodes a specific public key which encompass the algorithm identifier together with parameters and the key value.

Practically, when a `compositeSignatures` schema is used to encode multiple signatures at once, the value for the algorithm identifier associated with the signature is defined as follows:

```
compositeSignatures OBJECT IDENTIFIER ::= {iso(1) identified-organization(3)
                dod(6) internet(1) private(4) enterprise(1) OpenCA(18227) 11 }
```

The `compositeSignatures` identifier is used to identify the type of signature, and the corresponding value, encoded in the `signatureValue` field, contains multiple signatures and associated parameters. Each of these individual `SignatureInfo` entries carry the information about one of the signatures applied to the certificate in the same order the corresponding public keys appear in the multi-key issuer's certificate.

3.1. Current Limitations

When we first drafted the public release of Composite Crypto, there were still few unresolved issues that were associated with the use of multiple keys in a single certificate. The main issues were related to (a) handling error conditions when only some of the signatures are reported to be bad, (b) the complexity of enabling encryption in multi-key environments, and (c) how to handle ecosystem-wide algorithm revocation.

At the time of publication, specifically, some argued that, although using multiple keys of different type is a valuable feature (not only when it comes to backward compatibility or future-looking deployments), the complications introduced from the use of multiple keys to validate a single object required the deployment of some complex validation policy and additional infrastructure elements. At the same time, a second argument against the standardization of multi-key certificates was related to the impracticality of modifying current crypto libraries to accommodate for new types of error conditions and API changes. A final argument against our idea was based on the difficulty of guaranteeing the correct distribution of validation policies across entire ecosystems, like the broadband one, without the need for deploying additional infrastructure elements.

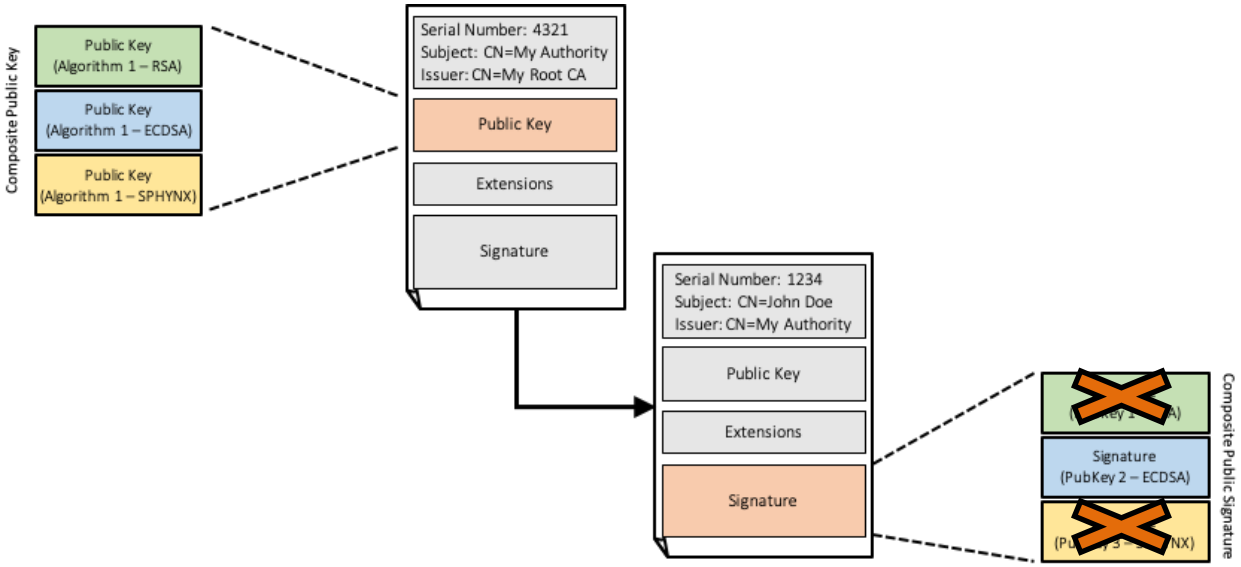


Figure 1 - Example of new error paths introduced with the use of multiple keys

Figure 1 sketches an explanatory error scenario where a composite signature, in this case on a certificate, has only one valid signature component. In our original work, the decision about the overall validity of the signature was left to the crypto library. This ambiguity posed a serious issue for consistency in how applications deal with these new mixed error states. In some scenarios, you want the possibility for the components of a composite signature to be “alternatives” so that a relying party can use the keys they prefer and/or understand (any of the signatures are equivalent). In other situations, instead, you want the possibility to report the signature to be valid only if all the components of signatures verify correctly. In other words, by providing an underspecified behavior, we inadvertently introduced, from an ecosystem perspective, the possibility for unpredictable results.

A problem that did not have a solution until now.

Another aspect that must be considered for signature validation is the level of trust in the algorithm throughout time. In the above example, let’s imagine that two out of three components of the signature are reported to be erroneous. Let’s also imaging that at time t_0, the use of ECDSA alone provides enough security for the identified application and ecosystem. Let’s now move the clock 3 years forward at a time

$t_1 = t_0 + 3$ years. Is the signature still to be considered valid? Without additional indications from a trusted source, applications and users are faced with an impossible task that cannot be easily resolved.

The next section explains how we solved the identified problems by introducing a second data structure that explicitly defines the relationship across the different key and signature components.

4. Composite Crypto vs. Combined Crypto

As described in the previous section, the fact that there were no clear semantics associated with the Composite Crypto was the source of the issues we were facing. This was reflected by the need of requiring crypto libraries to change their current APIs to support different validation policies. Because this might not be practical and might hinder adoption, we defined an alternative mechanism to drive the deterministic behavior when validating multi-key signatures. The core of our solution turned out to be extremely simple: defining, on top of the current ones, a new set of algorithm identifiers for keys and signatures that we call *Combined Crypto*.

With the introduction of these new OIDs (one for `subjectPublicKeyInfo` identifiers and one for `Signature` identifiers), we now have the possibility to explicitly define, via OIDs, the logic operations that crypto libraries must apply when validating the multi-key signatures.

When the *Composite Crypto* identifiers (also referred to as `compositeOr`) are used, the relationship across the different signatures is a logical “OR”. This means that the crypto library can use ANY of the signature components to determine the validity of the entire signature. A simple deterministic algorithm can be defined that goes through the list of signatures and stops at the first correctly validated one. If no signatures are correctly validated (i.e., because the values are corrupted or because the entity does not support the specific algorithm), the overall signature is not valid.

When the *Combined Crypto* identifiers are used instead, the relationship across the different signatures is a logical “AND”. This means that the crypto library must, in this case, positively verify ALL the signature components before being able to declare the whole signature valid. Also in this case, a simple deterministic algorithm can go through the list of signatures and stop at the first incorrectly validated one. If even one signature is not correctly validated (e.g., because of erroneous calculations or unsupported algorithm), the overall signature is not valid.

4.1. Advanced Key Structures

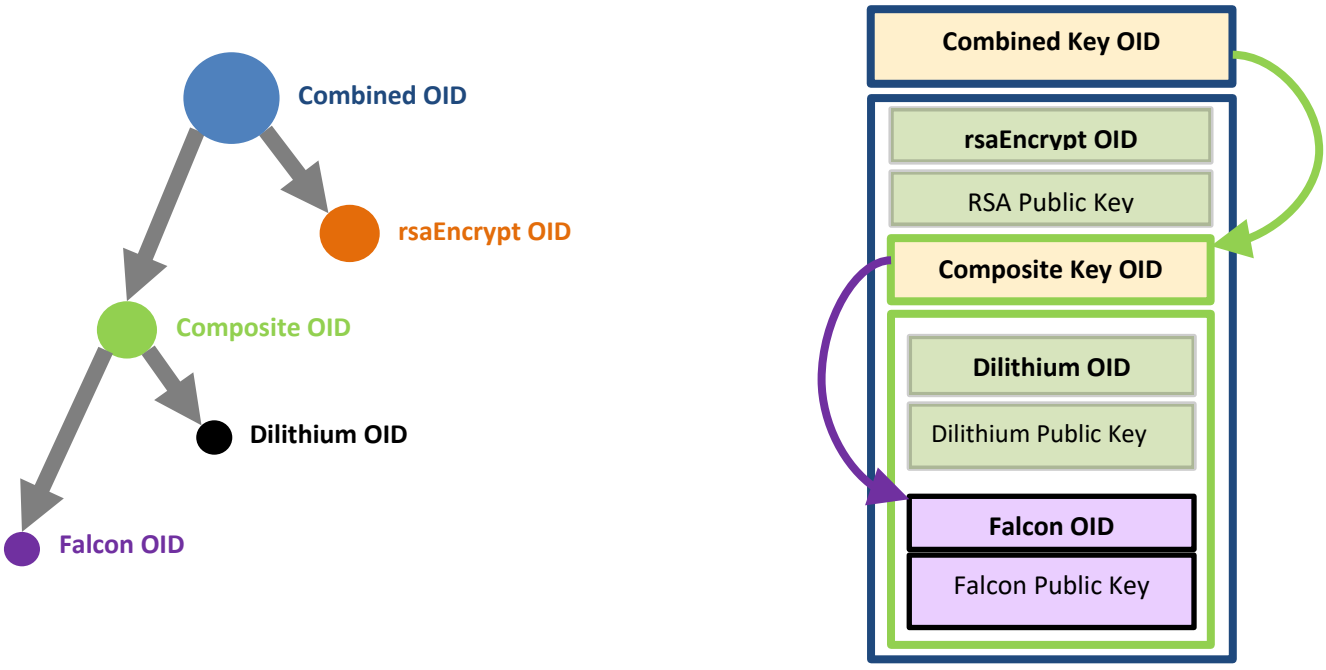


Figure 2 - Tree Representation of a Multi-Key Public Key Info structure

Figure 2 provides a tree representation of a key structure where the use of the “AND” and “OR” logical functions are leveraged. In the provided example, the represented key structure mandates for the use of an RSA component in the Combined Key container together with one of the two components from the Composite Key one. In fact, as previously discussed, when validating the components of a Combined key (as in this example), all of them must be validated correctly and that requires both the RSA and the Composite Key signatures to be valid. Back to the specific example, this translated to the need for the RSA signature to be valid together with, at least, one of the components in the Composite Key, i.e., the Dilithium or the Falcon signatures.

By following the described approach, CAs and PAs can design their certificate profiles with specific key structures for their certificates with deterministic behavior. For example, PKI architects can now decide to use a classic algorithm as the first element in a Composite Key to maximize backward compatibility. Another optimization strategy could be to maximize efficiency by using the fastest algorithm, from a validation standpoint, as the first element in composite keys, while using other FIPS and/or non-FIPS algorithms in combined keys to enforce the use of both types of cryptography together (i.e., classic and post-quantum).

4.2. A Deterministic Algorithm for Multi-Key Signature Validations

When using multi-key certificates, the `subjectPublicKeyInfo` structure of the certificate can have two different types of OIDs. The first type of OIDs is a container OID (i.e., the Composite or Combined ones) while the second type is a "real" algorithm OID such as, for example, rsaEncryption. Equation 1 provides a deterministic algorithm for validating multi-key signatures in pseudo programming language.

```
FUNCTION: Validate Signature Component

-------------------------------------------------------------------------

If Signature OID is Combined:

    For Each Component in Combined:
        If Signature Component is Combined:
            ERROR: Recursion
        End If

        If Validate Signature Component is NOT Valid
            Return False
        End If

    End For
    Return True

Else

    If Signature OID is Composite:

        For Each Component in Composite:
            If Signature OID is Composite:
                ERROR: Recursion
            End If

            If Validate Signature Component is Valid
                Return True
            End If

        End For
        Return False

    Else

        If Validate Signature Component is Valid
            Return True
        Else
            Return False
        End If

    End If
```

Equation 1 - Multi-Key Signature Validation Algorithm

For each of the nested components in combined signatures we evaluate it. We stop the validation process as soon as one component does not verify correctly. In this case the full combined key is invalid. On the other hand, if all components of the combined signature verify correctly, the combined key is considered valid.

For each of the nested components in composite signatures, we also evaluate it. However, differently from the combined key container, in this case we consider the composite key valid if at least one of the components is valid. The algorithm goes through the list of components and considers the composite key valid as soon as one component validates correctly. Conversely, the composite key is invalid if all the components (and not just one as in the combined key case) do not validate correctly.

5. Algorithm Revocation Via CRLs and OCSP Responses

With the possibility of algorithms being completely compromised overnight by quantum computers, PKIs are faced with a new problem: distrusting certificates that use a specific public key algorithm. Independently from the use of multi-key or single-key certificates, the inability, today, to provide such a mass-revocation tool can hinder our ability to effectively revoke the use of an algorithm.

---

*We are missing, today, an important tool in PKIs that is relevant for Post-Quantum algorithms deployment efforts, and that is **Algorithm Revocation**.*

We can easily see the impact of the lack of such tool with the latest example of algorithm deprecation that required a long time to complete (SHA-1). Specifically, when looking at the evolution of the deprecation process, we notice how it has happened very slowly and its resolution used ad-hoc criteria and deployment strategies instead of delivering formal ways to revoke its use across entire ecosystems. As a result, Certification Authorities, although they are there to guide the ecosystem and have the authority to revoke identities as needed, they still lack practical tools and options to enforce algorithm deprecation.

5.1. Policy Authorities as Sources of Trust

Since we introduced the concept of algorithm revocation in conjunction with multi-key environments, some critiques have been directed at the chosen trust model arguing that an external authority should be the one to provide algorithm deprecation information. Because this is an important governance principle, we want to provide additional considerations that can help understanding the principles we rely on when extending existing revocation mechanisms.

The trust model that is usually assumed in PKIs mandates for CAs to keep all participants in the ecosystem behaving according to a common policy. Therefore, CAs are already entities with a clear mandate to protect the integrity of the ecosystem by following verifiable procedures - this includes the possibility to revoke certificates. CAs are, therefore, the entities that, in accordance to defined policies, should provide indications about which type of keys should not be trusted throughout the PKI lifecycle.

In a trust infrastructure, besides the set of CAs that provide their services to the community, it is common practice to deploy a Policy Authority that is responsible for the ecosystem Certificate Policy (CP). When available, the content of the policy document is used to align CAs requirements across the whole ecosystem. In the DOCSIS PKI, the Policy Authority is operated by CableLabs on behalf of the entire ecosystem and is appointed with the task of making sure that the whole PKI is secure. As this governance model has been successfully exported to other ecosystems of interest for the broadband industry, our work can be extended and adopted in other device-centric ecosystems where a common trust infrastructure enables interesting and effective crypto-migration strategies (e.g., Wi-Fi Alliance/Passport 2.0, CBRS-A, etc.)

5.2. Algorithm Revocation vs. Certificate Revocation

When considering revocation and its practical impact over the ecosystem, an important consideration to make is related to the scalability of algorithm revocation vs. certificate revocation. Today, when a crypto algorithm needs to be replaced because of possible security risks or compromises, CAs must revoke every single affected certificate to make sure that the faulty algorithm is not used anymore.

Even when the higher levels of the PKI are protected with quantum safe algorithms (i.e., Root and Intermediate CAs), the option of using traditional revocation mechanism, i.e., via serial numbers, comes with very high costs related to adding many certificates to the revoked lists – both CRLs and OCSP servers are negatively impacted by these massive revocation events and can easily collapse under this added load (e.g., in the DOCSIS PKI the number of active certificates to revoke can be in the hundreds of millions). Conversely, the revocation mechanism described in this invention provides a very efficient way to mass-revoke certificates when and if needed. The mechanism is lightweight both on the Certificate Service Providers (or CSPs) when creating and distributing this information, and on the client when validating certificate chains and signatures. The rest of this Section provides a detailed description of the data structures, procedures, and extensions we defined to enable algorithm revocation.

5.3. Algorithm Revocation and Multi-Key Environments

The lack of standardized secure mechanisms to provide algorithm revocation is not a new problem. However, with the introduction of multiple keys within a single certificate, the problem of algorithm revocation becomes more evident.

In our work we focus on revocation of key structures, rather than a simply focusing on algorithms, to provide the possibility to better manage algorithm trust. For example, there might be the need to revoke a specific key configuration across the whole set of issued certificates (i.e., a specific algorithm or a specific algorithm hierarchy) without having to completely revoke its use in other cases. To address all these use cases, our work allows the ecosystem administrators to revoke, for example, the use of RSA as a primary key type in certificates or within Composite Key containers, but still allow the use of the RSA algorithm when used as a component of a Combined Keys (e.g., RSA + Post-Quantum Algorithm).

An example of the complexity that raises with the introduction of multi-key certificates can be easily shown by looking at the evolution of algorithm deprecation over an extended period. Let's imagine the case where a relying party correctly verifies the signatures on a specific document or certificate and let's also imagine that this process is repeated over and over – while the validation results do not change, the trust in the algorithm itself can change. For example, one of the algorithms used in the PKI might be compromised or there might be a new requirement, during uncertainty, to prevent the use of classic and/or post-quantum algorithms by themselves (e.g., you must use a combined RSA and Falcon signature).

The mechanism described in this work addresses all these cases.

5.4. Key Configuration Revocation

The key configuration revocation mechanism we introduce in this paper focuses on revoking specific key configurations via the use of key configuration revocation lists. To this purpose, we needed to provide a way to identify key configurations that must be distrusted. As discussed in Section 4.1, we can represent

key configurations as binary trees where the Composite and Combined nodes provide the bifurcations in the tree structure, while the individual components represent the end nodes (or leaves) of the tree. Few example configurations are provided in Figure 4.

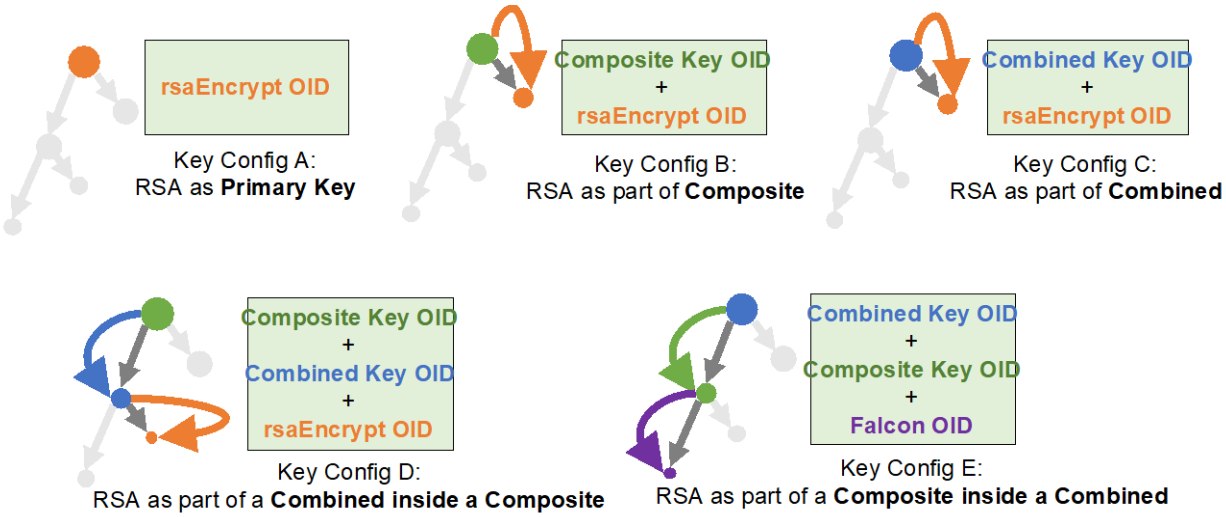


Figure 3 - Example Key Configurations for different types of primary algorithms.

For each of the key configurations that the CA wants to deprecate or revoke (or is instructed to do so by the PA), the CA generates a sequence of OIDs that we refer to as the `KeyConfigRevocationData`. This list of algorithm OIDs is then embedded as the value of an extension in OCSP responses and CRLs that are issued from the CA (or the delegated signer). The value of the `KeyConfigRevocationList` extension is implemented as a `SEQUENCE OF KeyConfigRevocationData`.

More specifically, each of these entries provides information about how to uniquely identify the specific key configuration (e.g., "(Start) → RSA" or "(Start) → CompositeCrypto → RSA"). Additionally, it is possible to specify an optional trust period for the algorithm in the form of `doNotUseBeforeDate` and `doNotTrustAferDate` fields.

The `KeyConfigRevocationList` data structure and associated identifier(s) are defined as follows:

```
keyConfigRevocationList-id OBJECT IDENTIFIER ::=
         {iso(1) identified-organization(3) dod(6) internet(1)
          private(4) enterprise(1) OpenCA(18227) 13 }

KeyConfigId ::= 1..MAX OF OBJECT_IDENTIFIER

KeyConfigRevocationData  ::=  SEQUENCE  {
        keyConfig               KeyConfigId,
            --- Identifier of the specific Key Configuration
            --- identified by this data structure
        doNotUseBeforeDate   [0]  GENERALIZED_TIME      OPTIONAL,
            --- Time before which the key configuration
            --- should not be used
        doNotTrustAferDate   [1] GENERALIZED_TIME    OPTIONAL,
            --- Timestamp after which the key configuration
            --- identified by keyConfig should not be trusted
            --- by the ecosystem clients anymore }

KeyConfigRevocationList  ::= SEQUENCE (1..MAX) OF KeyConfigRevocationData
```

To deprecate a specific algorithm when validating certificates (e.g., RSA), the data structure of the key revocation extension (i.e., the `keyConfigRevocationList`) carries the specific algorithm identifier as the only value in the `keyConfig` field. This configuration would not prevent, however, the use of the identified algorithm inside Composite or Combined keys because the algorithm identifier's list would be different. To deprecate both the use of an algorithm as a primary key in the certificate and as a component of Composite Keys (but leaving the possibility to leverage it in a Combined Key), the CA would generate two entries. The first one carries a sequence that comprises only a single identifier, e.g., the RSA algorithm identifier. This sequence deprecates the use of the algorithm as a primary key. The second one carries the sequence "Composite Crypto OID → RSA algorithm OID". This sequence deprecates the use of the algorithm as a component of Composite Keys (i.e., using RSA inside Composite Crypto keys).

5.5. Deprecating the use of multi-key certificates

CAs might also need a mechanism to deprecate the use of Composite Crypto or Combined Crypto within the ecosystem for when, for example, a transitioning period is over, and infrastructures and devices have fully transitioned to the new algorithms.

In this case, no additional mechanisms are required because the very same approach described in this paper can also be used to deprecate multi-key certificates: the CA generates a `KeyConfigRevocationData` entry where the `keyConfigId` carries only the Composite Crypto or the Combined Crypto object identifier(s) as needed.

6. Solving the Multi-Key Encryption Conundrum

Multi-key environments can provide interesting options to address encryption under today's cryptographic uncertainties. For this discussion, we choose the use case that deals with encrypting a document for a specific recipient as the explanatory relevant use-case. Specifically, the open problem we are focusing on is how to determine which key or set of keys should be used to encrypt a document for a recipient in the presence of multiple certificates and algorithms.

Similarly to the algorithm revocation case, linking multiple keys to the same identity is not a new problem and still we have no standardized solutions for it. In fact, there is no accepted procedure, today, to securely link together identities contained in different certificates that might even be issued from different CAs or different PKIs.

6.1. Encryption, Certificates, and Multiple Algorithm Support

To better explain the issues that crypto libraries and applications need to address when supporting multiple algorithms to encrypt data, let's go back to our example and describe the process of encrypting a document that is to be shared with a single recipient. In our example, let's assume that multiple algorithm support (e.g., RSA and Dilithium-Crystals) is required but only single-key certificates are deployed. This can happen, for example, when encrypting an e-mail for a recipient that might have multiple certificates, i.e., one with an RSA key and another with a Dilithium-Crystals one. For brevity and clarity, in the rest of the discussion we omit the description of the procedures for encrypting the data via a symmetric algorithm (not relevant for our discussion) and focus on the differences, when considering the encryption process, between single-key and multi-key certificates.

Before encrypting, applications must validate the revocation status of the recipient's certificate by accessing the certificates' revocation information (i.e., CRLs or OCSP responses) from the appropriate URL for all the certificates in the validation chain of the recipient. This is an essential step that prevents the leakage of the encrypted information for compromised certificates or keys. Without any indication of what the status of the algorithm (or key configuration) is or might be in the future, applications will happily encrypt the data for each of the certificates and possibly leak the encrypted content if one of the algorithms is broken.

This simple example shows the two main issues that the industry faces under the current crypto uncertainty when single-key certificates are used: dealing with the inefficiency of using multiple certificates connected to a single identity (i.e., need to interact with multiple infrastructures/services for a single encryption/validation operation) and the inability of efficiently communicating how to leverage the security of multiple algorithms together (i.e., "AND" or "OR" operations).

When looking at the first issue, multi-key certificates provide a distinct advantage: the need for less queries to the infrastructure. Specifically, because applications have to validate only one certificate chain per recipient, the number of requests to OCSP and CRL servers is greatly reduced. For example, in a three-tier infrastructure (i.e., Root CA, Intermediate CA, End-Entities) with three different algorithms deployed via single-key certificates, applications might need to perform up to six different OCSP or CRL queries and securely store 3 different Root CAs, while when multi-certificates are used, applications might need up to only 2 different queries and securely store a single Root CA. When looking at the second issue, the application that is performing the encryption is faced with the same uncertainty we noticed in the first formulation of our composite cryptography proposal (i.e., lack of deterministic behavior) because there is no possibility to dictate if the keys in the different certificates are equivalent or if they must be used together.

Ultimately, this one-to-one paradigm (i.e., one key for one identity) is also reflected everywhere in X.509 trust infrastructures where the assumption is that different certificates are associated with possibly different identities. Multi-key certificates solve the underlying conundrum by using a single identity, thus enabling the use of multiple algorithms across the board: from network functions to document signing.

6.2. More Efficient Encryption Process with Multi-Key Certificates

As described earlier, the ambiguity that was introduced with the initial proposal for multi-key certificates is completely resolved in this work by using explicit logic operations across keys and signatures that are completely defined by the specific OID used (Composite or Combined). Also in the encryption case, we leverage the separation of "OR" and "AND" logic operations to provide crypto libraries with deterministic encryption and decryption behavior. Table 2 provides a summary of the differences between Composite and Combined crypto when it comes to encryption options. Specifically, a Composite Key is enabled for encryption if at least one of the components algorithms supports encryption while a Combined Key is enabled for encryption if all the components' algorithms support encryption.

Back to our example, by providing algorithm deprecation information together with certificate revocation information, the encryption process can be performed even more securely than we do today and increase flexibility by supporting forward-looking or backward-compatible key structures. Even

outside the multi-certificate use-case, the availability and use of key configuration deprecation enhances the security of the whole ecosystem and help to prevent possible data breaches.

Table 2 - Encryption Operations for Composite and Combined Crypto

Composite Crypto	Combined Crypto
When Encrypting for a Composite Key, the encryption is performed with ***all the public keys SEPARATELY***	When Encrypting for a Combined Key, the encryption is performed with all the ***keys in a COMBINED way***
When Decrypting with a Composite Key, the decryption ***can be performed with ANY of the private keys*** related to the single public key components (OR)	When Decrypting with a Combined Key, the ***decryption must be performed with ALL the private keys*** related to the single Public Key components (AND)

7. Conclusions and Future Work

In this work we provide a description of the latest results when it comes to Composite Crypto and deployment of post-quantum algorithms. Specifically, we extend our original proposal to address the origin of the processing uncertainty that affected our original proposal: an incomplete design.

By adding a new set of OIDs, we can now express what the relationship across signatures (or keys) should be, thus providing a deterministic validation and encryption process. This simple enhancement unlocks deterministic behavior for crypto libraries without the need for deploying complex validation policies as it was initially envisioned. In other words, with the discussed new additions to our framework, the key structure of multi-key certificates itself provides clear validation, encryption, and decryption processing rules for crypto libraries.

On top of these important results, we identified CRLs and OCSP responses as the preferred mechanism to carry sequences of OIDs (and validity periods) to deprecate individual key configurations. This mechanism for algorithm revocation can be used in conjunction with both single key and multi key certificate environments.

Ultimately, the considerations contained throughout the paper show that the use of multi-key certificates can lower the cost of multiple algorithm deployment and provide the possibility to better manage, at the ecosystem level, the risks related to cryptographic failures. As we continue to evolve tools and specifications for multi-key environments, we envision that their deployment might become a common mechanism for delivering dynamic crypto-agile ecosystems in the future and, at the same time, simplifying new algorithm deployments and support algorithm migrations processes.

Abbreviations

CA	certification authority
CBRS-A	citizens broadband radio service alliance
CRL	certificate revocation list

CSP	certificate service provider
CVP	closest vector problem
DER	Distinguished Encoding Rules
DOCSIS	Data Over Cable Service Interface Specifications
DH	Diffie-Hellman
EC	Elliptic-Curves
ECDSA	Elliptic-Curves Digital Signing Algorithm
EE	end entity
FIPS	Federal information processing standard
HSP	hidden subgroup problem
ICA	intermediate certification authority
I-D	internet draft
IETF	Internet Engineering Task Force Standards Organization
KEM	key encapsulation mechanism
KEX	key exchange (algorithm)
NIST	National Institute of Standards and Technologies
PA	Policy Authority
PKC	public-key cryptography
PKI	public-key infrastructure
OCSP	online certificate status protocol
OID	object identifier
PFS	perfect forward secrecy
PQ	Post quantum
PQA	post-quantum algorithm
QC	quantum computing
R-PHY	Remote RF Layer (PHY)
RSA	Rivest-Shamir-Adleman (cryptosystem)
SHA-1	Secure Hash Algorithm (160 bits)
SCTE	Society of Cable Telecommunications Engineers
SVP	Shortest vector problem
TA	trust anchor
TLS	Transport Layer Security
SCTE	Society of Cable Telecommunications Engineers
S/MIME	secure e-mail message format
Wi-Fi	wireless
X.509	standard format for digital certificates
XOR	exclusive OR operator

Bibliography & References

[Ec05] American National Standards Institute, *Public Key Cryptography for the Financial Services Industry: the Elliptic Curve Digital Signature Algorithm (ECDSA)*, ANSI X9.62, November 2005.

[Rsa16] The Internet Engineering Task Force (IETF) – IETF RFC 8017. PKCS #1: RSA Cryptography Specifications Version 2.2, edited by K. Moriarty et al., November 2016. Also available at https://datatracker.ietf.org/doc/rfc8017/

[Doc40] *Data-Over-Cable Service Interface Specifications, DOCSIS 4.0, Security Specifications.* CableLabs Publication, 2019. Available as CM-SP-SECv4.0-IO1-190815.

[Doc31] *Data-Over-Cable Service Interface Specifications, DOCSIS 3.1, Security Specifications.* CableLabs Publication, 2020. Available as CM-SP-SECv3.1-IO9-200407.

[X509] ITU-T Recommendation X.509 (2005) | ISO/IEC 9594-8:2005, *Information Technology - Open Systems Interconnection – The Directory: Public-key and attribute certificate frameworks*.

[RPhy18] *Data-Over-Cable Service Interface Specifications, DCA – MHAv2*. Remote PHY Specification. Available as CM-SP-R-PHY-I10-180509.

[Pala04] The Internet Engineering Task Force (IETF) – I-D draft-ounsworth-pq-composite-sigs-04 - Composite Keys and Signatures For Use In Internet PKI, edited by M. Ounsworth and M.Pala, Jan 2021. Also available at https://datatracker.ietf.org/doc/draft-ounsworth-pq-composite-sigs/

[Shor97] Peter W. Shor. “Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer”. In: SIAM Journal on Computing 26.5 (Okt. 1997), S. 1484–1509. ISSN: 0097-5397, 1095-7111. DOI: 10.1137 / S0097539795293172. arXiv: quant-ph/9508027.

[Reg02] Regev, O. “Quantum computation and lattice problems.” The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings. (2002): 520-529.

[HeHø04] Mark Ettinger, Peter Høyer, Emanuel Knill, The quantum query complexity of the hidden subgroup problem is polynomial, Information Processing Letters 91 (1) (2004) 43–48.

[Reg04] *Regev, O.. “A Subexponential Time Algorithm for the Dihedral Hidden Subgroup Problem with Polynomial Space.” arXiv: Quantum Physics (2004).*

[Kup13] Kuperberg, G.. “Another Subexponential-time Quantum Algorithm for the Dihedral Hidden Subgroup Problem.” TQC (2013).

[Fa17] Falcon - Fast Fourier Lattice-based Compact Signatures over NTRU. https://falcon-sign.info.

[Di17] Dilithium-Crystals - Dilithium digital signature scheme. https://pq-crystals.org/dilithium/.

[Sp15] SPHINCS+ Stateless hash-based signature algorithm website. https://sphincs.org.

End to End Telecom for Healthcare Architecture

A Cable Industry Perspective

A Technical Paper prepared for SCTE by

Dr. Sudheer Dharanikota
Managing Director
Duke Tech Solutions Inc.
111 Fieldbrook Ct. Cary, NC 27519
+1-919 961 6175
sudheer@duketechsolutions.com

Clarke Stevens
Principal Architect, Emerging Technologies
Shaw Communications
1401 Lawrence St., Suite 1550, Denver, CO 80202
+1-720-723-2316
clarke.stevens@sjrb.ca

1. Executive summary

The COVID-19 pandemic forced people to consider new ways to manage healthcare and wellness care. One of the first places people turned was to online solutions. Video conferencing provided a way for physicians and other healthcare workers to provide remote care for patients even though opportunities for meeting in-person were limited. This provided a basic form of Telehealth. Many people had to manage their personal health and wellness conditions while primarily remaining at home. Again, network connectivity became essential for these rudimentary Aging in Place (AIP) use cases. While people relied on networking services, most did not really take advantage of the true possibilities of networked services. There are numerous networked sensors available that can measure many important parameters. There are smart home devices that can automate common tasks. More importantly, networks of family, friends and caregivers can be provided with information and contacted under common management. AIP and Telehealth has not nearly met its potential. This is an opportunity for cable operators. While there are piece parts that are available from specialists for smart home components, health management and communication, nobody has yet really integrated these parts into a cohesive service. Cable operators provide networking services to the home, work with existing familiar interfaces (like television) and aggregate services for consumers as core competencies. Therefore, cable operators are well-positioned to work with various partners to provide complete integrated AIP and Telehealth packages. This can increase convenience, reduce costs and improve care for all stakeholders in the healthcare value chain.

2. Introduction

The Healthcare industry is going through a major transformation to modernize the infrastructure, reduce the cost and increase the quality of care. In a series of articles, we have suggested how the Telecom industry can assist the Healthcare industry [1][2][3]. We call this inter-industry collaboration Telecom for Healthcare (T4H). Even though the T4H opportunity is not limited to these two major intersection points, we focus on Aging in Place (AIP) and Telehealth use cases to illustrate our thoughts on the end-to-end T4H architecture. (Refer to [4] for six different opportunities that a Telecom operator can address through the T4H architecture covered in this paper.) The SCTE Data Standards Subcommittee, in which the authors are members, is actively working on T4H solutions for the AIP and Telehealth areas in working groups three [5] and four [6].

Figure 1 provides a quick summary of the T4H opportunity and challenges from AIP and Telehealth points of view. Many of the needs, challenges, and Telecom opportunities of both markets are similar (refer to the SCTE working group analysis at [5][6]). Some of the high-level use cases that need to be supported for these two markets include:

A. Providing basic communication between the subscribers (users) and the providers/caregivers
B. Providing seamless communication between the users and the stakeholders
C. Monitoring the subscribers (users) for health, mobility, fall detection, etc.
D. Analyzing the data collected from the subscribers (users) and properly notifying the stakeholders
E. Assisting the T4H service providers with claims by documenting accountability
F. Offering managed services to support installations, product support, and other services to improve adoption and retain customers

The goal of the paper is not to elaborate on the use cases but to use them to motivate the end-to-end architecture. For additional information refer to the working group documents.

In the next sections, we elaborate on the T4H architectural needs, provide a framework, discuss details on individual components, summarize the findings and propose next steps.

3. End-to-end high-level T4H architecture

Figure 2 provides a high-level end-to-end architecture proposed by Duke Tech Solutions (DTS) in their market analysis [4] based on different T4H market opportunities. The framework is further elaborated in this paper with a second level of architectural details.

To understand the end-to-end T4H architecture, first, we need to understand the users, the service providers, and the other stakeholders (refer to Figure 1).

	Aging in Place	Telehealth
Subscribers (Users)	Older adults (65+), caregivers	Individuals, providers
Stakeholders	Family members, care givers, doctors, service personnel etc.	All family members, providers, (payors)
Needs	Communicating, monitoring, service, support, integration	Communicating, monitoring, integrating with provider systems
Challenges	Ease of use, provider network integration, problem solving	Ease of use, device and EMR integration, remote monitoring,
Telecom opportunity	End to end solution, managed services, provider integration	End to end solution, managed services, provider integration

Figure 1 - Telecom for Healthcare opportunity and challenges summary

Subscriber or Users: These are the folks who are the primary subjects of the T4H platform. For AIP, the elders who are aging at home are the primary users. The service infrastructure revolves around the elder's needs in the AIP use cases. For Telehealth, the users are typically the family members who are using the T4H platform at home or the patients who use the platform in the care centers (such as Veterans Administration (VA) satellite healthcare facilities [7]).

- Service providers: In the traditional healthcare industry physicians and nurses in offices and hospital systems are primary service providers. In the proposed T4H emerging system, the intent is to go beyond healthcare to wellness. Here it is important to realize that doctors are not the sole provider of wellness services. Hence, we introduce the concept of T4H service providers. For AIP the service providers include the caregivers (both healthcare and non-healthcare related), network providers and technicians. For Telehealth use cases, doctors are still included. This distinction is important for understanding the relationship between the users and the service providers from different architectural points of view.
- Other stakeholders: In the T4H environment, other stakeholders are also interested in the wellness of the users. These include, for example, family members, friends, and payors.

Understanding the classification of different stakeholders, we proceed with an explanation of the components in the T4H solution. We adopt the architectural framework provided in [4], DTS's Telecom for Healthcare Environment Framework (DTEF), to evaluate the end-to-end solution components proposed in this paper. We will use AIP and Telehealth use cases (as provided in section 2) in oder to do this.

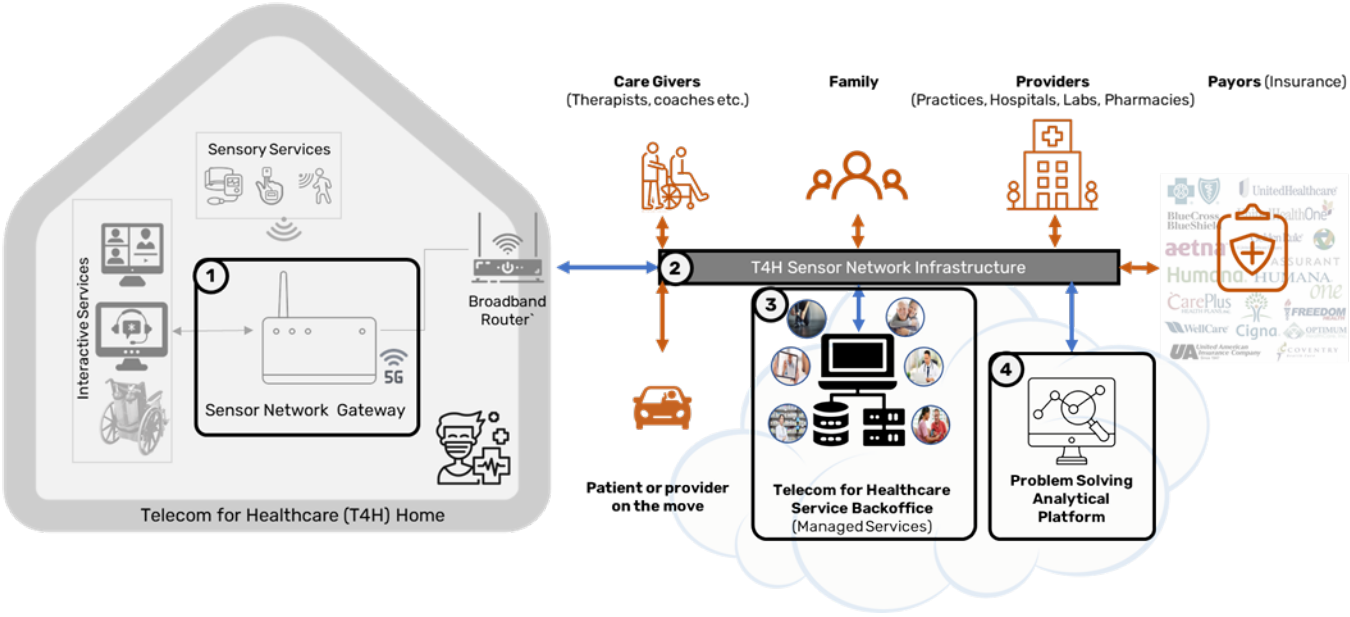


Figure 2 - DTS's Telecom for Healthcare Environment Framework (DTEF) based components

1. In-home healthcare/wellness aware gateway: From use cases A, B, C, it is clear that there needs to be a gateway in the T4H home. This gateway, as shown in Figure 2, acts as an integration point for monitoring the sensor devices (e.g., motion sensors, remote patient monitoring equipment) and integrating with the interactive services endpoints (such as unified communication services). This **Sensor Network Gateway** can be a standalone device or integrated with other vendor equipment such as the set-top box or residential gateway. In this paper, we treat it as a logically separate device.
2. T4H aware network infrastructure: Again from the use cases A, B, C it is clear that the T4H requires to connections between the users, the providers, and the other stakeholders. This requires not only reusing the exiting telecom infrastructure, but will also need to meet reliability, security and privacy requirements specified by the T4H architecture. The communications infrastructure will have to meet the needs of the sensor network traffic, unified communications traffic, and notifications to the different stakeholders. To differentiate (or to keep the focus on) the T4H needs, we call this the **T4H Sensor Network Infrastructure**.
3. T4H aware service back office: The cable operators have all the required infrastructure for managing end-to-end services. As mentioned in use case F, it is essential to turn the fragmented, gadget-oriented point solutions into a well-oiled managed service. This can only be accomplished by Telecom operators who have access to such infrastructure and have been managing communications infrastructure for 90+% of the households in the US. We call such infrastructure as **T4H Service Backoffice**.

4. T4H aware problem solving analytical platform: Finally, as mentioned in use cases D and E, this infrastructure attempts to solve the problems stakeholders are facing. These problems and related algorithms may be unique to the healthcare/wellness industry, but the infrastructure is similar to infrastructure the telecom operators use today. We call this repurposed analytical platform the **T4H Problem Solving Analytical Platform**.

These solution components are explained in the following sections. Note that in this paper we provide a block diagram level architecture. The detailed architectural specifications will be worked out in more detail at the SCTE DSS WG3 [5] and WG4 [6].

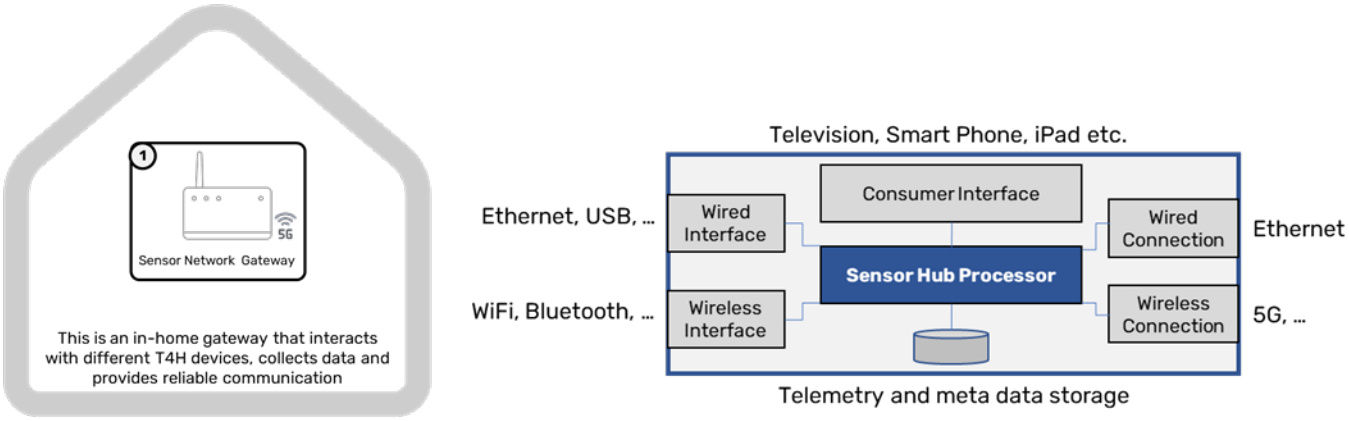


Figure 3 - Sensor Network Gateway architecture

4. In-home T4H architectural components

Figure 2 shows different tasks that need to be done in a T4H capable home. These tasks include:

- Support for different data streams: The AIP and Telehealth infrastructure needs to support typical data streams generated in a T4H home. These include sensor and actuator data streams, streams to record events and real-time streams such as video and audio communication between T4H stakeholders.
- Communication with existing in-home broadband devices: These might include consumer consoles (such as TVs or smart speakers), and smart home devices (such as smart locks, lights or video doorbells).

To increase the adoption of T4H solutions and for ease of use, the T4H in-home components need to be on the same logical network.

- The T4H physical networking can be dependent upon the use case for any particular device. Most components are likely best connected with an in-home broadband network, but certain devices (such as a locator device) may need to be connected even if the user is beyond the limits of the in-home network. The important thing is that the networked devices can communicate with each other on a secure logical network. Critical components may require a secondary backup network connection in case the primary network fails. Cloud-based services also protect against exclusive dependency on the in-home network.

    - The in-home solutions shall integrate remote patient monitoring devices, sensor devices (such as fall detection, motion sensors, etc.), and other IoT devices that are used for wellness needs.
    - Additionally (to increase the utility and ease of use of the system), the T4H solutions shall be integrated with the frequently used consumer consoles (such as Television for the elderly), smartphones, and other handheld devices. Again, a cloud-based solution simplifies and experience that can be duplicated on whatever console is convenient.
- Provide installation and support services: The operator shall also streamline the installation and support services to improve the ease of use of the integrated solution.

Based on the above high-level needs, we propose that a Sensor Network Gateway functionality be developed for supporting T4H solutions. The block diagram of such a gateway is presented in Figure 3. This gateway will have wired (Ethernet, USB, etc.) and wireless (WiFi, Bluetooth, BLE, etc.) interfaces to integrate the T4H devices and other IoT devices (such as turning on light, placing a phone call to the family member). They will need limited internal storage for the temporary storage of sensor data and to perform local analytics on time-critical events. The gateway will also need to integrate the consumer access interfaces (TVs, smartphones, iPads, etc.). The gateway shall provide redundant Internet connectivity with an Ethernet interface to wired broadband and a 5G or other wireless connection as a backup. The gateway shall offer an easy installation process and support self-install where possible.

5. T4H communication architecture

Although cable operators already own a capable end-to-end Telecom infrastructure, it may need to be adapted to meet T4H needs. That is the focus of this paper. The T4H sensor network infrastructure is used to provide communication between stakeholders and users, to collect in-home sensor information, to provide intelligent notifications, and to offer T4H managed services. This infrastructure, as shown in Figure 4, will be used in the T4H case for the following:

- Unified communications infrastructure: This is the collaborative software (similar to [8]) that is used to create a communication environment for different players in the T4H ecosystem. The challenge of offering UCC software is the ability to scale it to the individual consumer level and

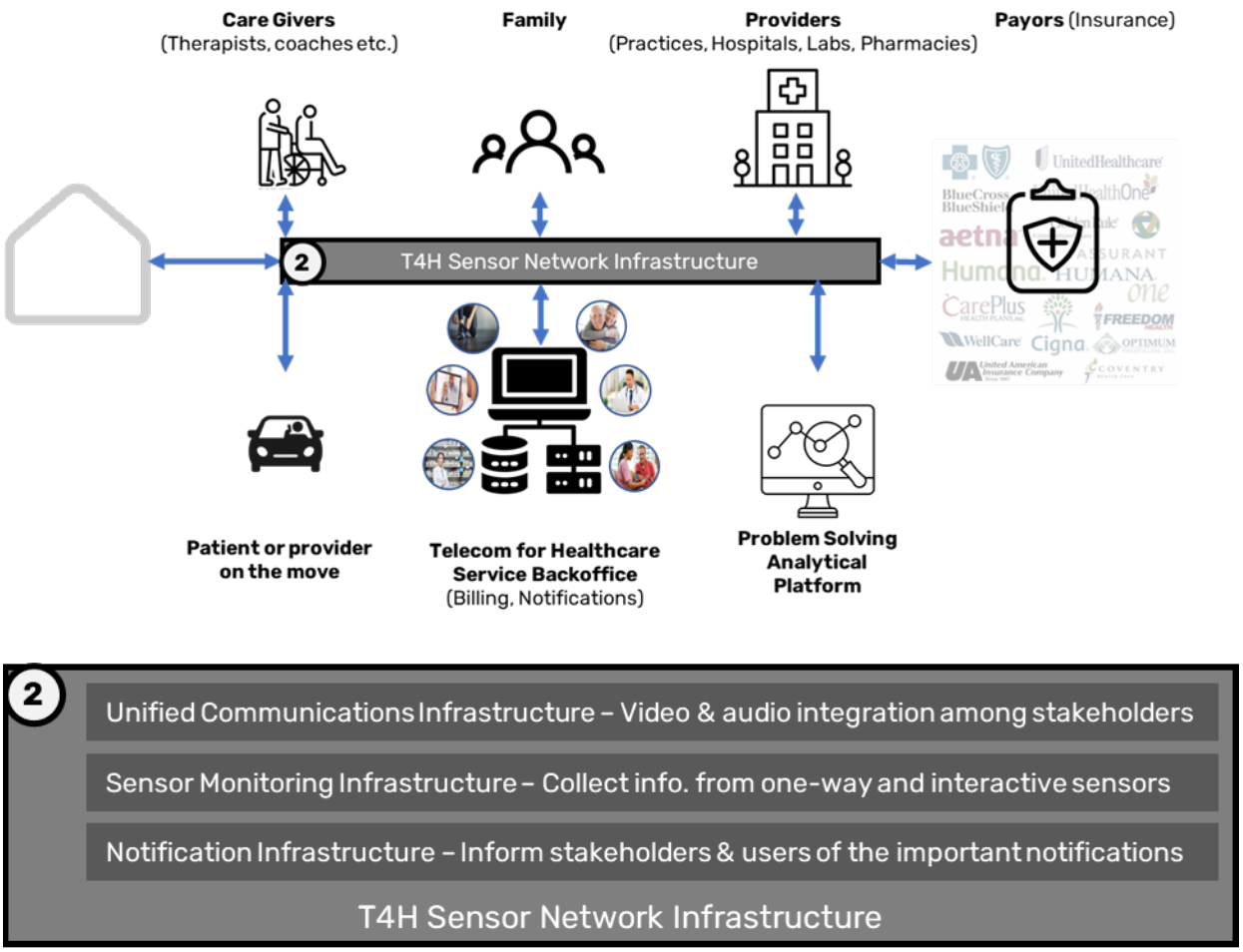


Figure 4 - UCC, sensor monitoring, and notification traffic on T4H Network Infrastructure

still host it for a large volume of deployments. Current UCC software may need to be enhanced for the healthcare market in order to be compliant with HIPAA. UCC software must seamlessly offer secure communication between parties in the communication session and still provide privacy to protect the shared and recorded content. Launching the UCC solution on different consoles (such as TVs, smart devices, etc.) is also an important requirement.

- Sensor monitoring infrastructure: This is a stream of information coming from the T4H from sensor devices. The traffic can be from remote patient monitoring (RPM) devices (typically medical devices), and other sensory devices (such as motion sensors, door openers, etc.) used for both AIP and Telehealth. The communication infrastructure will need to assist in bringing the devices online, providing secure communication of the information to the analytical platform, and potentially track metadata from the sensor devices. The platform must implement HIPAA requirements and offer privacy similar to what operators provide for PII. In healthcare, privacy extends to PHI (Personally identifiable Healthcare Information). Note that the overall QoE (Quality of Experience) for such streams is essential to understand. We refer you to [9] for our preliminary thoughts on consumption patterns for the T4H data.
- Notification infrastructure: The other major data streams included in the T4H communication infrastructure are the notifications. To enable this notification infrastructure, the endpoints (stakeholder, software, etc.) shall be registered to receive specific events. The infrastructure will

need to provide secure communication between the communication platform and the stakeholders.

6. T4H service back-office architecture

The T4H service back office, as shown in Figure 5, is used for providing managed T4H services. This includes installation, support, troubleshooting, connectivity management, billing, and more importantly business rule management. Note: The Cable operators already have much of the T4H infrastructure in place for managing their existing broadband and other in-home services. In the remainder of this section (as shown in Figure 6) we provide a discussion on what tasks need to be completed in order to implement the above-mentioned service components. We also discuss why cable operators are suitable for providing these services and what additional capabilities they need to develop.

- Connectivity management: The operator needs to manage basic connectivity services. These include the UCC, in-home sensor devices, notifications, etc. Cable operators have already been

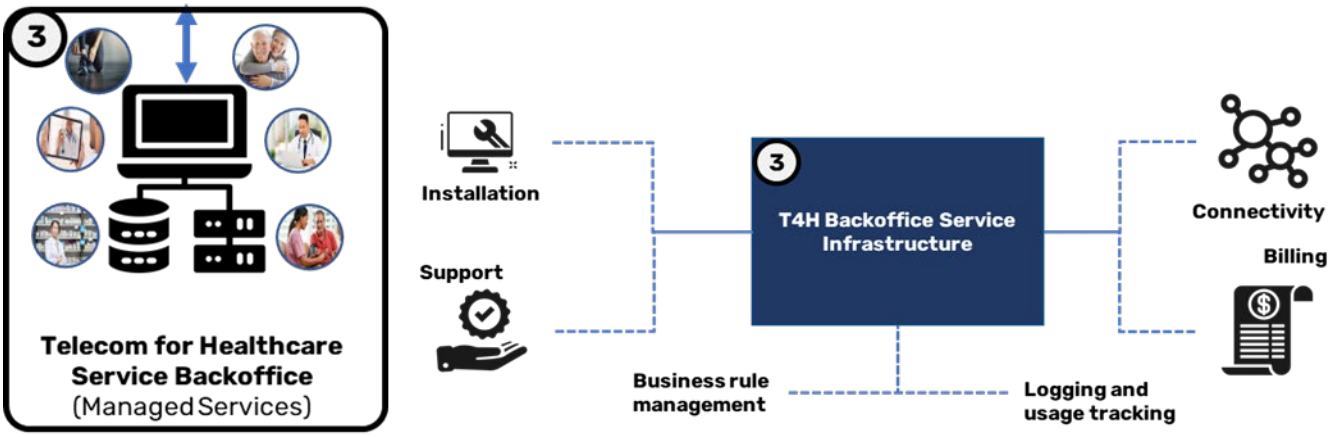


Figure 5 T4H service back-office infrastructure components

- managing such in-home services in many cases. Lately, operators are assuming the management of more extended services such as video surveillance, IoT-based services, etc. for a large customer base. The connectivity services that are offered to regular customers need to be extended for T4H services.
- Business rule management: Managing T4H customers, stakeholders, and service providers is done through business rules. This requirement will add new rules-based management. These include the user service enablement, the stakeholder notification management, and other per user management. This is similar to the Cable operator subscriber management. The cable operators have been managing such business/subscriber rules for more than 70M subscribers in the US. Extending similar concepts to T4H healthcare/wellness needs is the new challenge that cable operators face.
- Logging capabilities: Metadata and telemetry data collection are essential for T4H services. This data is required to enable new problem-solving capabilities for AIP and telehealth use cases. Additional information on such data gathering requirements can be found in [10]. Cable operators are used to gathering and analyzing gigabytes of customer and network-specific daily data. Extending the same capability for T4H services is a new opportunity and challenge for cable operators.

- Installation and support services: One major issue facing the T4H industry is their fragmented way of handling the market space. No single company is responsible for the quality of care. We recommend that Telecom companies extend their installation and support services to include T4H. Refer to [2][3] for the business case behind such an offering. Cable operator's experience in providing end-to-end installation services, their service assurance tools, and their "boots on the ground" to manage on-premise services can be effectively leveraged to support T4H services. This will require cable operators to expand the use of some tools and train additional teams to handle T4H services.
- Billing services: One of the challenges facing the next generation healthcare and wellness industry is billing capabilities for next-generation products. The Telecom industry is experienced in creating products for in-home services, billing for them, and collecting payments from customers. Extending the same capabilities for T4H services is required.

7. T4H analytical service architecture

It is necessary to tune the analytical platform to meet the needs of the T4H problem space by collecting the right data [10], providing appropriate analytics to solve problems, and offering a flexible notification engine for stakeholders. Such a platform is essential to add value to the raw data.

Some of the reference AIP and Telehealth problems that need attention include:

- Trend analysis: Providing trend analysis for basic time-series information gathered from an IoT

	Connectivity	Rule Management	Logging	Installation	Support	Billing
Tasks to perform	UCC, In home sensor devices, end to end services	Service, notification, PHI, and other per sub rules	Different T4H related meta data and Telemetry info.	UCC, IoT (Healthcare and non-healthcare devices)	T4H healthcare and non-healthcare services	User and stakeholder service billing and collection
Why Cable operators?	Extensive experience with in-home service mgmt.	Have been managing 70M+ customers	Used to managing tera bytes of customer specific info.	Highly experienced with in home, e2e service installs	Boots on the ground, service management tools and org.	Elaborate systems to offer and manage service models
Capability development for Cable operators	Need to tune the connectivity focus to T4H	Healthcare/wellness related rules	Collect T4H specific data and address the right problems	Repurpose to T4H vertical (RPM, monitoring installs, ...)	Repurpose to manage T4H	Repurpose to manage T4H

Figure 6 - Why MSOs are suited for T4H services, and what capabilities do they need to develop?

- sensor or telemetry information gathered from different RPM systems.

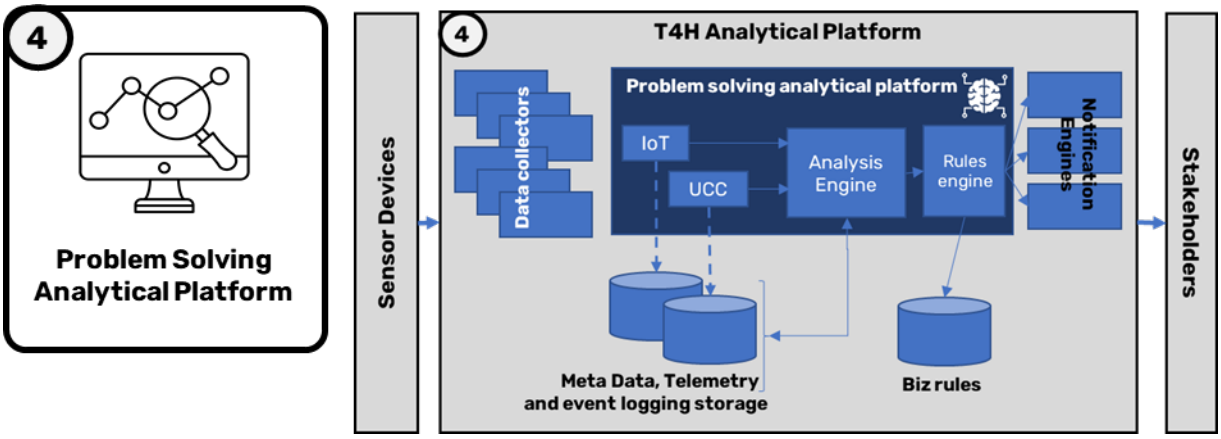


Figure 7 - T4H analytical plaform that manages problem-solving notification infrastructure

- Anamoly detection: Identifying different anomalies such as fall detection, etc. based on the collected sensor data.
- Correlation: Correlating different health and wellness-related conditions with indicative data.
- Resolution analysis: Assisting with different analytical algorithms and the data collected, then tying that data to T4H problem resolution.
- Quality of care forecasting: Providing different success metrics to quantify the quality of care provided by the cable operator's T4H solution.

To meet the above needs the cable operator will have to create the following architectural components, as shown in Figure 7:

- Metadata collectors: Collectors for different types of metadata for interactive systems (such as UCC), sensor networks (such as IoT), and telemetry data per customer. The operators need to adjust their existing monitoring capabilities to meet T4H needs.
- Analytical engine: An analytical engine is required to solve the problems highlighted earlier. This is accomplished using data collected from the various sources. These analytical tools need to consider the Protected Health Information (PHI) [11] restrictions and offer solutions for the classes of problems considered previously.
- Rules engine: Different business rules must be programmed for classes of services offered to the T4H customer for different classes of problems the platform needs to solve.
- Notification engines: The relevant observations from analysis must be communicated to the stakeholders through this infrastructure. Such a notification infrastructure is generally present and tuned for the SLA (Service Level Agreement) needs of the customers.

8. Conclusion and next steps

In order to realize the promise of Aging in Place and Telehealth, all parts of the ecosystem need to be integrated in a way that is convenient for all stakeholders. Cable operators are well-positioned to be this integrator because they already provide network and television services to customers and they have the underlying infrastructure that can install, support and manages such services. However, there is still a lot of work to be done to realize this vision.

In this paper, we propose and architecture capable of supporting AIP and Telehealth use cases and providing key features to users, caregivers, payors and other stakeholders. Some companies will undertake this integration opportunity eventually, but there is currently an opportunity for cable operators to step up and assume this role.

The challenges are not insignificant, but the cable industry has already shown an aptitude for this sort of work as they have created widely adopted standards, aggregated services, created an army of installation and support personnel, and formed trusted billing relationships with millions of consumers and businesses.

SCTE is already working on AIP and Telehealth within their standards development organization. With serious engagement and active participation, cable operators can agree on the problems that need to be solved and come up with solutions and strategies to create a new and potentially lucrative revenue source in AIP and Telehealth for operators around the world.

9. References

[1] Sudheer Dharanikota, Ayarah Dharanikota, *Why are cable operators natural fit to support Telehealth – An inter-industry perspective*, 2020 SCTE Expo, available here

[2] Sudheer Dharanikota, Ayarah Dharanikota, Dennis Edens, Bruce McLeod, *Aging in Place business case for cable operators*, SCTE Journal, June 2021, available here

[3] Sudheer Dharanikota, Ayarah Dharanikota, Dennis Edens, Bruce McLeod, *Telehealth business case for cable operators*, SCTE Journal, September 2021, available here

[4] Duke Tech Solutions market Research, *Telehealth market report – A Telecom based opportunity analysis*, available here

[5] Data Standards Subcommittee, Working Group 3, Aging in Place, available here

[6] Data Standards Subcommittee, Working Group 4, Telemedicine, available here

[7] US Department of Veterans Affairs, *South Texas Veterans Health Care System (STVHCS) – Satellite clinic division*, Available here

[8] Cisco, Unified Communications and Collaboration, available here

[9] Sudheer Dharanikota, *What are the impacts of changing consumption patterns on bandwidth usage?* DTS white paper, available here

[10] Sudheer Dharanikota, Jason Page, Metadata/Telemetry support from Cable Operators to address Telecom for Healthcare opportunity, 2021 SCTE Expo, available here

[11] HHS.gov, *Health Information Privacy*, available here

Ensuring HFC Network Resiliency During Extended Utility Outages

A Technical Paper prepared for SCTE by

Toby Peck
Sr. Director of Broadband Product Management
EnerSys
3767 Alpha Way, Bellingham, WA 98229
360-392-2247
Tobias.peck@enersys.com

Jay Frankhouser
Senior Director Product Marketing & Product Management for Energy Storage
EnerSys
2366 Bernville Rd, Reading, PA 19605
610-659-7755
Jay.Frankhouser@enersys.com

1. Introduction

High-speed Internet is ingrained in our culture. From email and social media to videos and entertainment, the online experience is part of everyday society. One primary enabler of high-speed Internet is the cable broadband network, a collection of Hybrid Fiber Coax (HFC) networks connecting residential and business users to the Internet. In the US alone, the cable broadband industry has over 78 million high-speed Internet connections.

The Covid-19 pandemic made these broadband connections even more essential. As much of the workforce transitioned to work-from-home status, businesses became reliant on stable Internet connections to perform core functions. Parents depended on these connections to support remote learning for their children. The line between business services and residential customers blurred, transforming high-speed data connections into a universal requirement.

Because the traffic flowing across broadband connections are vital for business, school, and health in general, the need for dependable service has become even more critical. Yet, at the time when we seem to need it most, our utility grid was exposed as a potential weak link for network resiliency. Across the United States, there have been waves of multiple day power outages, with reasons stemming from tropical storms, snowstorms, fires and freezing temperatures. Due to the susceptibility of the grid to severe weather, natural disasters, and planned outages, customers are purchasing home generation solutions at record rates. However, electricity at the customer's premises exacerbates the broadband problem, since TVs and laptops are functional, but without a connection to the Internet.

As performance of the grid brings to question the reliability of the HFC network, government entities have begun to legislate network backup time. High-speed Internet connections are so important that Multiple-System Operators (MSOs) are being mandated to keep their networks online, regardless of the utility grid. These mandates present a fundamental question for network reliability: How can we ensure power availability for critical HFC services during extended outages?

There are several potential solutions that continue to be explored, however, in many situations the simplest and cost-effective answer is adding batteries to existing HFC network elements to extend run time. However straightforward the concept may be, these batteries require space, environmental protection, thermal management, ongoing maintenance, budget, and end-of-life management. Solutions are heavily dependent on run time requirements, battery selection and, in many cases space limitations at existing broadband power system locations. Which battery chemistry provides the best run time? Which meets the budget? Which can be maintained by MSO technicians?

This paper explores various extended run time solutions using both Lithium Ion (Li-ion) and Thin Plate Pure Lead (TPPL) batteries. For each battery chemistry, comparisons are presented for run time, space requirements, thermal management, relative cost, and perspective comparison to higher total cost of ownership (TCO) for on-site generation solutions.

2. Extended Run Time

For discussion purposes, this paper will consider an "extended outage" as a utility outage that lasts longer than the average run time of the installed backup batteries plus run time of an average portable generator. This is assumed to be approximately 12 hours. For the last two decades, a three to four hour backup run time was typical for many cable outside plant (OSP) sites. This amount was determined to be sufficient for the operator to roll a truck with technicians, replacement equipment, or a generator. But with severe

weather and natural disasters causing longer grid outages across broad swaths of the network, standard run times have been questioned.

The California Public Utility Commission (CPUC) has been at the forefront of investigating these requirements. Since the state had experienced several instances of prolonged outages driven by utilities de-energizing sections of the grid to prevent possible wildfires, the CPUC created a new extended backup requirement. This was specifically developed for wireline carriers to enact "comprehensive resiliency strategies to prepare for catastrophic disasters and power outages." The new requirements adopted a 72-hour backup power requirement for the wireline providers' facilities in Tier 2 and Tier 3 High Fire Threat Districts[i]. This new requirement is far from the three to four-hour historic backup standard for which HFC sites were initially designed.

Not all regions will follow the CPUC's lead, but many areas are ideal candidates for extended run time. Texas was in the news for outages caused by freezing conditions, the Midwest and Southeast US are vulnerable to hurricanes, tornadoes or winter freezes and Northeastern US and Canada experience frequent extended outages due to annual winter storms. With the HFC network becoming increasingly important for society, it is likely some form of run time extension will be adopted in these regions as well.

While adding energy storage seems like a straightforward method of compliance, there are many challenges to overcome. Some reasons existing locations do not always lend themselves to easy upgrades include:

- Currently most sites have been designed specifically around the form-factor of three or six batteries in case size 27 (306 x 173 x 225mm) or case size 31 (330 x 173 x 240mm)
- Sites are often in easements in front of residences or on utility poles, so there is usually little or no space for additional battery cabinets to extend run times
- Many areas have height restrictions for cabinets making it impractical to add battery extensions to existing locations
- In locations where space does exist for additional battery enclosures, re-permitting can cost $10K or more per site and could add up to 6 months to install
- When upgrading existing sites with established utility service connections, it is often a requirement to maintain that connection to reduce costs and minimize logistical issues with local electrical utilities.

These hurdles are not show-stoppers but do eliminate the opportunity for a one-size-fits-all solution. The result is to look at the application, the location, and the installation to determine the right fit. The analysis starts with diverse options for energy storage and backup. After establishing different backup alternatives, the paper looks at the key factors for determining the best fit per application.

Figure 1 - Example cable broadband powering sites. The variability in available space, loading, and local regulations makes a one-size-fits-all solution for extended run time nearly impossible.

3. Backup Power Methods

There are a variety of possibilities available when considering how to provide backup power to the outside plant broadband network, including batteries, generators, fuel cells, and flywheels. Solar power is also a consideration when combined with an energy storage solution. While solar power, fuel cells, or flywheels could fill a niche, they do not currently present a viable solution for the broad deployment required to meet the new governmental mandates. These options must continue to be explored as their technologies mature, however they currently present significant issues regarding deployment and scalability. For this reason, this paper focuses on more prevalent alternatives – generators and batteries.

3.1. Generator Overview

Generators convert fuel (diesel, propane, natural gas, etc.) via internal combustion engine into electricity. The output can be either alternating current (AC) or direct current (DC), depending on the type of generator. A major advantage for generators is the ability to deliver power and keep the network running

as long as they have fuel. This becomes a key factor in the viability of generators as a solution for extended run time scenarios.

There are two broad categories of generators that can be used for extended run time scenarios in the cable broadband network: portable generators and curbside or stationary generators. Portable generators are designed for deployment during extended outages and provide flexible run time augmentation, then retrieved and stored until the next outage. Curbside generators are permanently stationed at the site to provide additional backup when needed. Most major operators have deployed both types of generators for extending run time in outage situations and have strategies for when and where to deploy based on unique site requirements.

3.2. Portable Generators

Portable generators have been broadly used in OSP for years as a flexible option to add backup time only when and where it is needed. Nearly all MSOs have an operational strategy for deploying portable generators to extend plant backup time during an outage. These strategies usually start by assuming a baseline backup time that is resident at each site based on the batteries installed. From there, remote status monitoring tracks the duration of the utility outage. Once an outage duration reaches a critical point set by the operator the site is at risk of dropping plant power and a network management system will trigger a technician to deploy a portable generator to the site to augment run time.

This strategy has proven effective over the years but has its limitations when dealing with outages of an extended nature. Portable generators are designed to be relatively small to be easily transported and deployed, and most suitcase style generators are optimized to completely enclose the engine to reduce noise. A standard 2 or 3kW suitcase style AC portable generator ranges from 50 to 100lbs. and will provide six to ten hours of backup per tank of gasoline for an average plant load of 600W. Due to this limitation, extended outages often require portable generators to be refueled multiple times, adding significant operational cost and environmental risk of technicians spilling gasoline in the field. Additionally, the flexibility and portability of generators make them an easy target for theft, often requiring additional cost and effort to secure. Another drawback to portable generators is that their output voltage quality is typically less stable than utility power. Variations in frequency and fluctuations in voltage can interact with the broadband UPS, causing the UPS to switch between utility line mode and battery backup mode. To avoid this undesirable behavior, modern UPS systems employ an AC input desensitizing feature to overcome poor input voltage quality typical of portable AC generators.

Portable generators can be staged from an operators' regional facilities to be within physical proximity of UPS locations that may require extended backup power. A typical portable generator staging approach utilizes trailers with multiple generators that are ready to deploy when needed. An example of a portable generator trailer is shown in Figure 2.

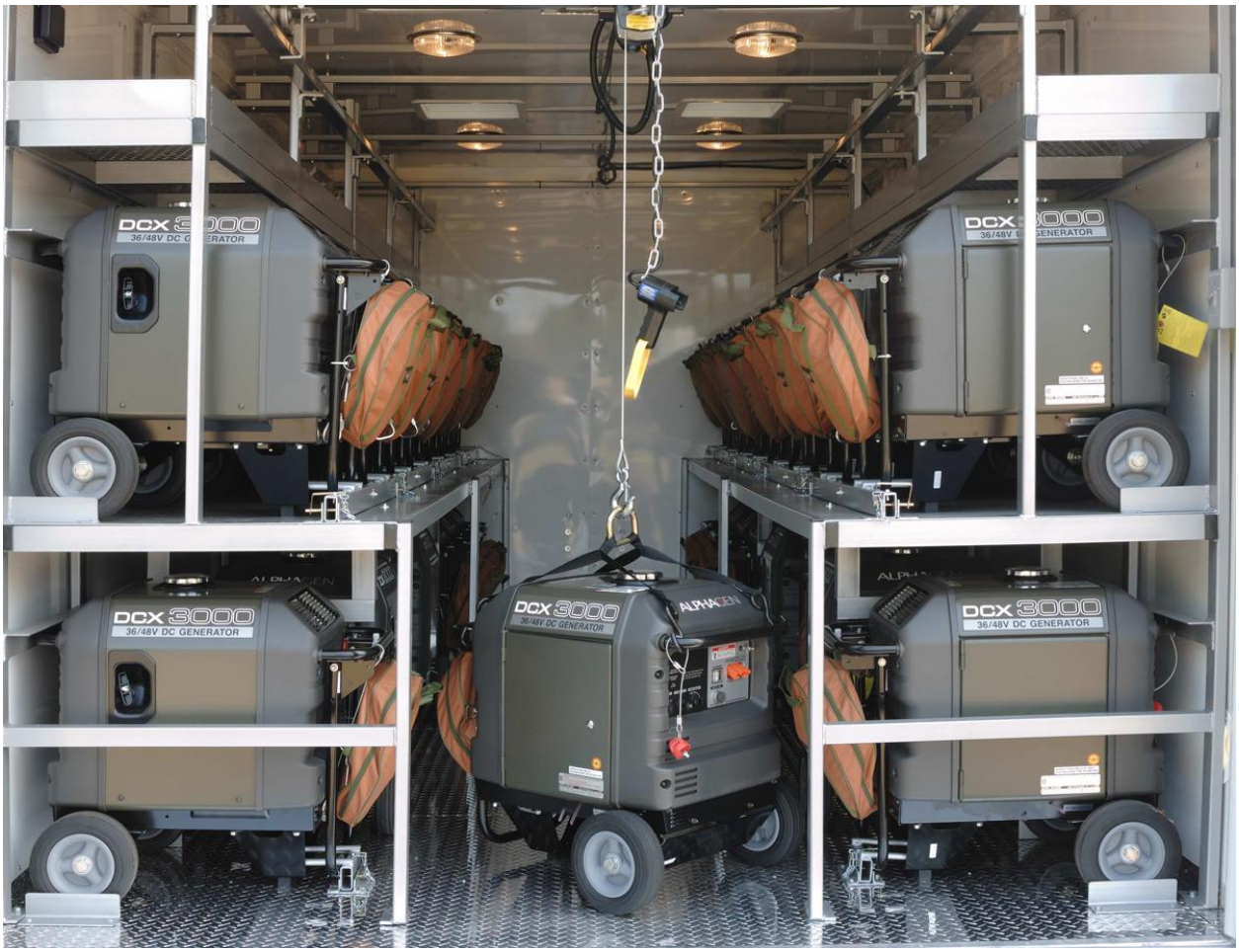


Figure 2- Portable Generators Staged for Rapid Deployment During an Outage

Run time limitation and theft risk of small portable generators can also be mitigated by using larger, tow-behind portable generators designed for construction sites. These larger generators can be capable of storing more than 50 gallons of gasoline allowing them to provide significantly more backup time, and due to their size, there is a reduced risk of theft. Larger generators do, however, have some drawbacks. Once again, due to their size and design, they can only be deployed one-at-a-time, meaning activation in a widespread outage can take a small army of technicians or making dozens of trips to and from the operator's warehouse. In addition, the amount of space required to store a fleet of large generators can be prohibitive and carry higher costs to maintain. In conclusion, larger portable generators are sized for larger loads than what's typically needed for cable, making them highly inefficient in this application.

3.3. Curbside (Stationary) Generators

Curbside or stationary generators are permanently stationed next to an existing cable powering site with critical loads to provide consistent support for extended outages. Historically, curbside generators were deployed in centralized powering architectures, where multiple power supplies were installed in a single location which fed power to the plant in multiple directions like the hub of a wheel. This allowed the significant cost to purchase and install the generator, and connect fuel to the site, to be distributed across multiple power supplies, making the cost per supply more reasonable. Stationary generators used in OSP powering applications provide DC voltage and connect directly to the system's DC bus with the battery

string and power supply inverter. This allows the generator to avoid frequent stops and starts from momentary outages by only starting when DC voltage drops below a certain level. A typical curbside generator installation included multiple enclosures: one enclosure houses the power supplies and limited battery backup. A second enclosure houses the generator. Some installations include a third enclosure to house propane fuel tanks for locations where natural gas is not available. A typical curbside generator installation is shown in Figure 3.

Figure 3 - Typical Curbside Generator Installation Using Natural Gas

Stationary generators are likely fueled with plumbed natural gas or with liquid propane cannisters. Of these two, generators that operate on natural gas are better candidates for delivering extended run time. When generators are fed natural gas via a pipeline it is less likely to be affected by a power outage. The drawback is many sites are not in locations with access to natural gas, and most sites that could have access are not currently equipped with natural gas due to upfront costs of digging a trench, running a new gas line to a site, installing and certifying a metered service, and permitting. Combined this process can add months of delays and thousands of dollars to site turnup and eliminate financial viability to add natural gas connection to meet the extended run time requirements. Liquid propane is used when natural gas is not available and the inherent limit to the amount of fuel that can generally be stored on-site for backup makes propane a less desirable fuel for extended outages of more than 24 hours.

Another factor to consider when looking at curbside generators as an option is the amount of maintenance required for proper functionality during extended outages. Since generators run on an internal combustion

engine, which is a complex system of moving parts, it is necessary to perform maintenance at least twice annually. An increased number of extended outages may also necessitate more frequent maintenance visits. Because of this additional maintenance requirement, TCO models should always bear this in mind.

One other key factor that cannot be overlooked when deploying portable or curbside generators is the noise level of the engine and its impact on overall customer satisfaction. As shown in Figure 1, many OSP powering sites are adjacent to customer residences or within utility easements on customer property. This alone can be a point of frustration, but when generator noise is added, it can often lead to customer complaints and, in some cases, customers lobbying local governments for restrictive ordinances to be placed on cable installations. For this reason, potential impact to customers is a primary factor when considering generator-based solutions.

4. Battery Overview

Batteries, the most common source of backup power, provide DC power to a standby power supply to be converted to AC plant power in times of utility outage. Batteries come in many shapes, sizes, capacities, and chemistries. Lead-acid chemistry has been a fundamental stalwart of energy storage for many decades. Traditionally, indoor applications were supported with flooded lead-acid batteries. These batteries have a liquid electrolyte, generate oxygen at their positive electrodes and hydrogen at their negative electrodes. Over time, water loss occurs, which requires the electrolyte to be topped off with water on a regular basis. A byproduct is the production of hydrogen and oxygen gas which must be ventilated to avoid hydrogen accumulation and the potential of combustion.

Both outdoor and indoor applications use valve regulated lead-acid (VRLA) batteries. The electrolyte is not liquid, but instead is immobilized with either a gel or absorbent glass mat (AGM). The result is less water loss, eliminating the need for topping off with water. Gas emissions are much lower in this solution, therefore ventilation requirements are far less than flooded lead acid batteries. Additionally, VRLA AGM batteries are generally rated non-spillable per UN 2800, which reduces regulations required for safe shipping.

The arrival of TPPL technology, a type of AGM VRLA battery, transformed lead-acid battery performance. TPPL batteries have thinner, high purity grids meaning more of them can be stacked into the battery. The result is increased surface area contact between the grid and active material boosting power densities. Therefore, TPPL batteries can deal with much higher current peaks and have faster charging capability. Standard TPPL batteries in case sizes designed to fit in most standard OSP cable enclosures, generally have higher capacities than other equivalent AGM VRLA batteries. For Extended Run Time applications where space is available for additional ground-mount cabinets, larger, high capacity TPPL batteries can be an additional high-value option.

Positive attributes for VRLA batteries include ease of transport, ease of installation, reduced maintenance, and higher energy densities compared to their flooded counterparts. A 98% recycling rate at end of life is also beneficial to the environment. Technicians in general like working with VRLA batteries, as they are typically non-spillable, include handles for improved installation and handling, and require limited maintenance. TPPL versions include greater temperature tolerance, higher energy density and longer life.

Two key downsides of VRLA batteries are slightly higher costs measured against flooded batteries and heavier weights associated with increased energy density. The high purity of TPPL models also increase initial cost, but the TCO is favorable due to the improvements in energy density and battery life.

4.1. Lithium Ion Batteries

Lithium Ion (Li-ion) batteries were originally conceived in the early 1970's and first commercialized by Sony 30 years ago[ii], but even as a result of continued research, development, and investment, they are still in infancy. Although there is still a positive and negative electrode, separator, and electrolyte similar to traditional lead-acid batteries, the key method of operation is the shuttling of Lithium Ions between electrodes. Temperature, charge voltage, end of discharge voltage and impurities limit the shuttling ability over time resulting in the loss of battery capacity.

The positive and negative electrodes of Lithium Ion batteries vary by technology and application, with this variance defining cell voltage and energy density. More research has been conducted on the positive electrode with Nickel-Manganese-Cobalt (NMC) and Iron Phosphate (LFP) being the most common, however research to improve the negative electrode is equally promising for the future.

Since multiple Lithium Ion chemistries exist, selecting the right one for an application is important. In Figure 4, the selection of chemistry will determine the voltage operating window and can impact the system cell count. LFP for instance operates at a lower voltage, therefore, to meet a higher system voltage, more LFP cells will be needed over other chemistries, such as NCM or NCA to meet the same system voltage window.

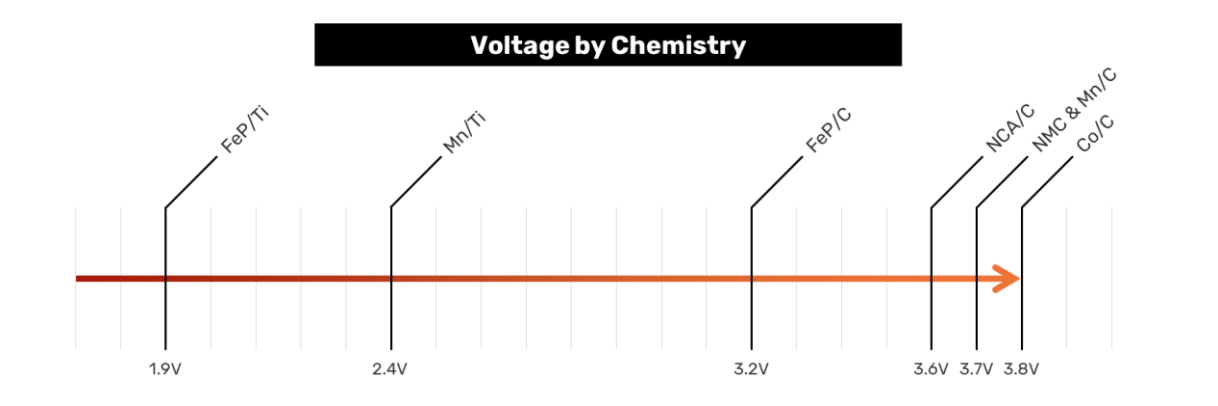


Figure 4- Standard Voltages of Lithium Ion Battery Cells based on Chemistry

Some key attributes for LFP and NMC chemistry batteries are depicted in the spider chart shown below. Since the CPUC project requires the providers to maximize run time and local jurisdictions would prefer as little physical presence as possible, NMC is a desirable choice to meet these goals. It also offers both cold and warm temperature tolerance and suggests long life in this type of application.

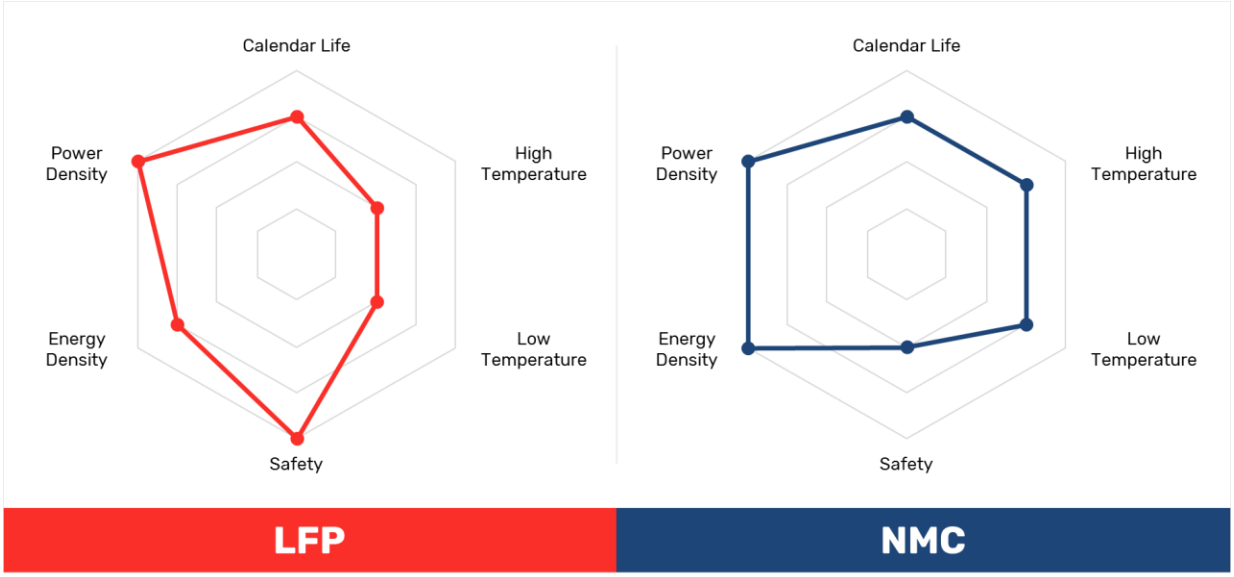


Figure 5- Spider Chart Showing Key Attributes of LFP & NMC Lithium Ion Batteries

4.2. Safe Lithium Ion Battery Design

When building Lithium Ion battery modules and systems, product developers must incorporate varied and numerous safety layers into the design as the battery needs to have optimize life, performance, and most importantly, safety. A graphic representation of these safety layers is shown in Figure 6. The Western Automotive-Grade Cell is the high-quality center of the layered approach. This is the optimal starting point for any designer as they develop an integrated system. Mechanical Safety System layers are designed to fit the battery housing, while providing physical protection for the critical internal elements. Operating the system in a safe way, while monitoring the key diagnostic parameters adds the Application-Level Software and Functional Safety Layers to the design.

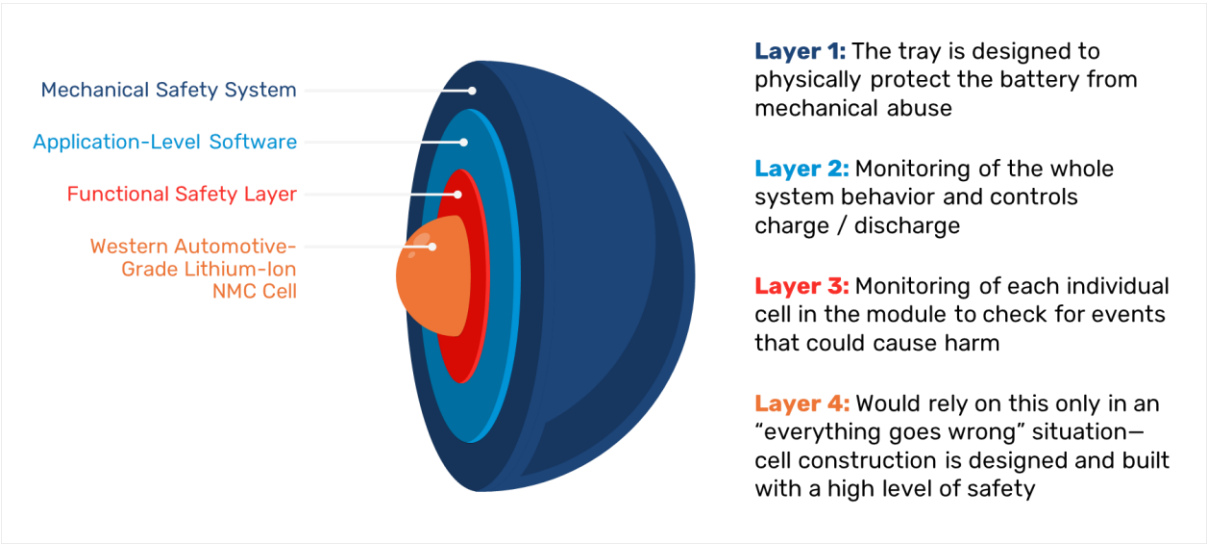


Figure 6- Graphic Representation of Layered Safety Design for Lithium Ion Batteries

Lithium Ion batteries require a Battery Management System (BMS) since, unlike lead-acid batteries, individual cells in a Lithium Ion system do not inherently work with each other to maintain system harmony and balance. The BMS can be active or passive, but it is critical to safe operation and application of the battery for successful system design. As a BMS gathers data, it raises alarms and makes critical decisions. With the proper communication protocols and interaction, the BMS can input information to the network operation center to allow assessment of battery health and overall system readiness prior to, during, or after an outage.

4.3. System Design for Lithium Ion Batteries

When designing any power system that integrates energy storage, it is always recommended to consider how battery performance, life and safety can be maximized by the broader system. This concept becomes significantly more important when considering lithium battery deployment in the OSP, where there is limited environmental control and most power supplies have been designed to manage unintelligent lead-acid batteries with brute force charging methods. Elements such as the system enclosure, power supply charger, internal infrastructure, and remote monitoring systems should all be optimized for integration with Lithium Ion batteries. Simplicity and safety for Lithium Ion products comes through a fully-integrated, engineered system.

First, it is vital to see if there is ample space for proper airflow between lithium batteries, and any necessary passive or active heating and cooling to maximize battery life and performance for outdoor systems. Next, the intelligence inherent in the BMS, required for safe deployment of a Lithium Ion battery, presents an opportunity to use the power supply charger as an additional safety layer. By creating a coordinated communication path between the Lithium Ion BMS and power supply, charge currents can be safely managed and optimized for the battery modules. Additionally, a Lithium Ion BMS should have the ability to provide valuable information that can be leveraged by remote status monitoring such as state of charge and state of health of the battery modules. Conversely, adding lithium batteries to an OSP power supply that has not been designed and tested for interoperability and communication can be potentially dangerous as the potential exists for the power supply charger to provide improper charge current to the BMS and harm Lithium Ion batteries. For these reasons, it is of the utmost importance that Lithium Ion batteries be designed into, and thoroughly tested with their intended OSP power system before being deployed in the field.

4.4. Comparison of TPPL and Lithium Ion Batteries

TPPL and Lithium Ion batteries have unique attributes, yet many applications will see the combination of both chemistries drive value for the application:

Physical Attributes:

Chemistry	Energy Density	Form Factors	Weight	Physical Orientations	Include BMS Electronics
TPPL	High	Many	Heavier	Most	Not required
Lithium Ion	Higher	Limited	Lighter	All	Yes, required

Operational Attributes:

Chemistry	Ventilation	Cut Off Voltage	Cycling Capable	Partial SoC Operation	Recycling
TPPL	Limited	Variable	Limited	Limited	98%
Lithium Ion	None	Fixed	Significant	Excellent	Limited

Figure 7- Comparison of the benefits of TPPL Lead-Acid and Lithium Ion Batteries

With noted differences above, each chemistry has physical and operational advantages depending on the application. Every day more applications arise where a Lithium Ion solution is the most viable solution because of energy density, longer life and greater cycling capability. However, in many situations TPPL remains the most cost effective and flexible solution available.

5. Requirements for Extended Run time Solutions

Generators and/or batteries (TPPL or Lithium Ion) can be used to meet extended run time requirements, but the right selection depends on analysis of several factors:

1. **Required backup time**: *The CPUC requires 72 hrs., however this may vary as other states develop backup recommendations around their unique needs.*
2. **Real estate:** *How much space is available at the site for additional enclosures or generators. Many sites have limited space and would require significant redesigns to create additional space.*
3. **Existing cabinet space:** *Using the existing cabinet where utility connection has been established is often paramount to avoid significant costs for running a new service. Maximizing energy density within an existing cabinet is the best strategy in these situations.*
4. **Location of the existing plant (aerial/underground)***: Generally, this determines the location of the established cabinet.*
5. **Power system load***: OSP power system loads can vary substantially where two sites with virtually identical systems can require drastically different upgrades for extended run time capability.*
6. **Fuel availability***: Primarily, is natural gas readily available at or near the site? This is one of the biggest questions when determining curbside generator viability.*
7. **Local regulations***: Local laws can restrict everything from the height, width, or weight of an enclosure to the type of fuel you are allowed to use and the permits required to do so.*
8. **Security concerns***: Sites at risk for theft can have unique challenges securing batteries, portable generators, or propane tanks. Additional security measures can drastically change TCO.*

9. **Total initial system cost**: *Upfront spend for the system cabinet, equipment and turn-up. This is often the biggest limiting factor for traditional extended run time solutions like curbside generators.*

5.1. Scenario Solution Comparisons

There are many different scenarios to consider when defining solutions. For brevity, this paper addresses several popular scenarios using common loads and varying available space, fuel, and security needed at each system location. Some assumptions were made for scenario comparisons:

- All loads are critical for support as driven by customer agreement or government regulation, therefore allowing the network to go down has significant financial consequence
- At least one extended outage greater than 12 hours occurs annually
- Costs for all equipment are based on currently deployed products plus known installation expenses
- Cost per truck-roll for generator maintenance or refueling is around $200 USD
- Impact of security concerns will be higher for lead batteries and portable generators having greater potential for reuse or value from recycling
- Lead-acid battery weight and energy density calculated from known TPPL configurations
- Lithium Ion battery weight and energy density is based on known NMC cell configurations
- Lithium Ion battery useful life assumed to be approximately two times that of TPPL

Scenario 1:

First, let's look at the run time capabilities for lead-acid batteries, lithium-ion batteries, and portable generators when the load is 4 Amps and there is only enough space for one pole mounted cabinet with a weight limit of approximately 600 lbs.

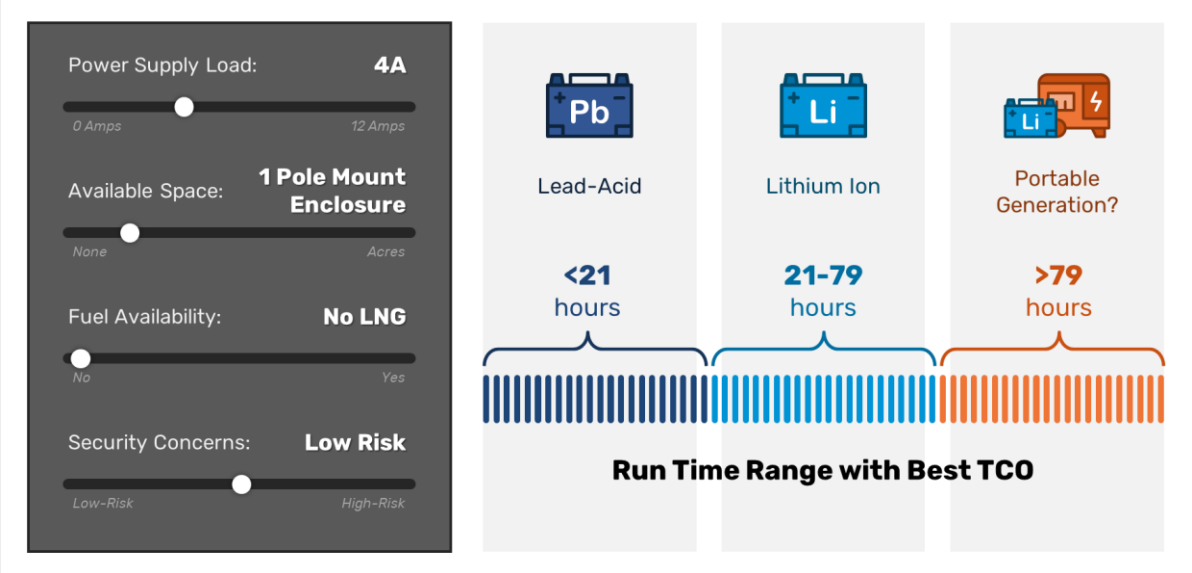


Figure 8- Run Time Ranges for Solutions in Extended Run Time Scenario 1

Due to space and weight restrictions of many pole mount installations, Lithium Ion batteries can play a key role in meeting extended run time requirements in this application and avoid the significant operational costs of portable generator deployment and refueling. However, as with any situation where

space is limited, after a certain run time threshold, the ability to continually bring new fuel on site to keep a portable generator running will be the only available tactic.

Scenario 2:

The next hypothetical scenario involves a ground mount site with an 8A load allowing space for one additional cabinet and an available natural gas hookup for a curbside generator. It is assumed this is a roadside deployment with little potential for noise to impact customer perception.

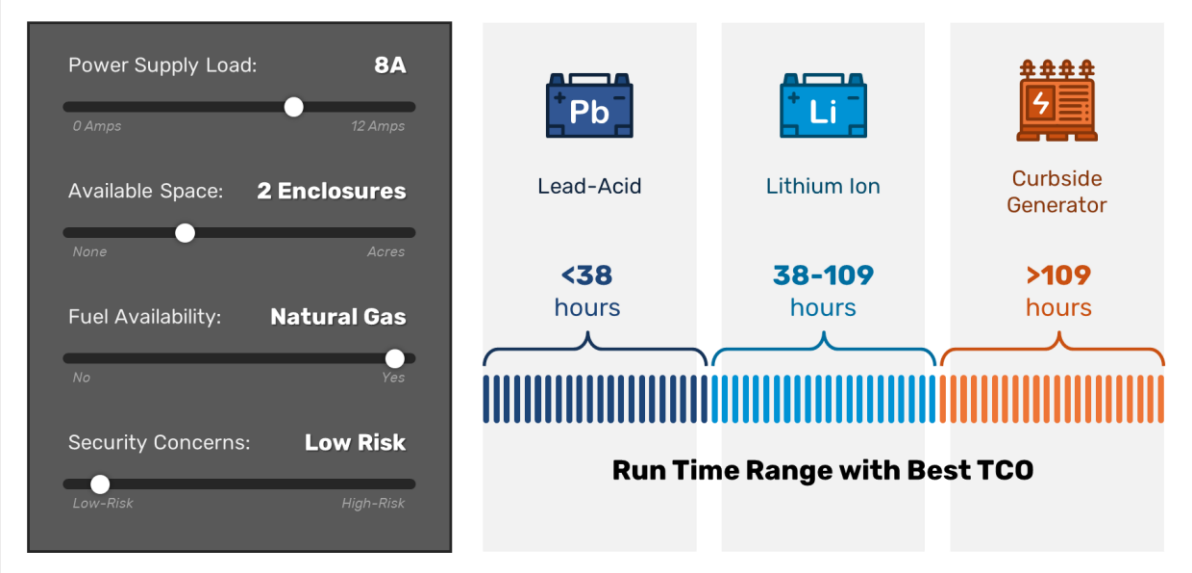


Figure 9- Run Time Ranges for Solutions in Extended Run Time Scenario 2

In this scenario, additional space afforded has an impact on useful range of battery solutions, as the amount of energy that can be stored on site is significantly increased. All three solutions can now be considered with using a second ground-mount enclosure on site. The TPPL solution no longer needs to be limited to case size 31 batteries and can now leverage an additional enclosure with High-Capacity TPPL batteries that may have been too heavy or taken up too much space for a pole mount application. The availability of a natural gas connection supports an on-site generator solution for extreme outages and does not require an army of technicians to refuel portable generators. Once again, Lithium Ion batteries show great promise as an alternative to a curbside generators in this extended run time scenario.

Scenario 3:

The first two scenarios considered common loads on the lower end and middle end of the load spectrum. This scenario has a higher site load of 12A, there is some restriction in available real estate, and an available natural gas line near the site. For this scenario we will assume this site has a history of frequent battery theft.

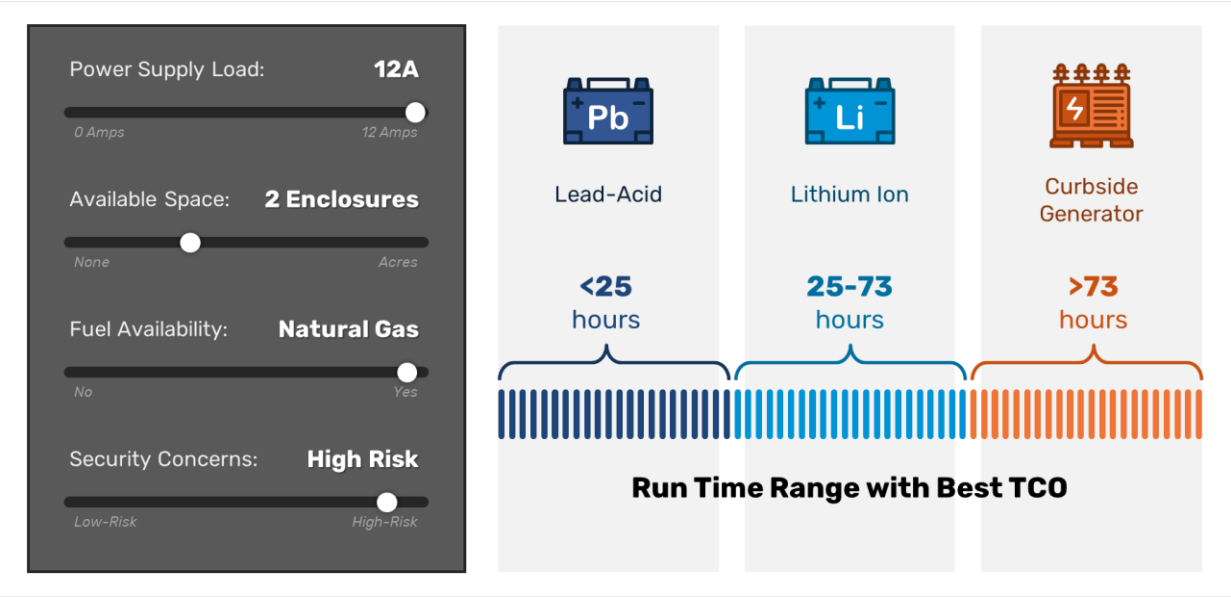


Figure 10 – Run Time Ranges for Solutions in Extended Run Time Scenario 3

The increased load on the power supply has an obvious effect on the available run time from battery-based options where Lithium Ion continues to show promise as the best TCO option to protect against extended outages. A subtle point to be understood here is the higher threat of theft could change the viability of lead acid, by building significant cost into the model for site security options. It could also mean planning multiple replacement cycles for the TPPL option, which also brings risk of the site being without backup during a critical outage.

5.2. Scenario summary

The number of variables in the explored scenarios further emphasizes the fact that there is no one-size-fits all solution for extended run time outages. Each variable needs to be carefully considered in order determine the best solution at each site. The key to having a solution for every site is to have a variety of tools to adjust for variability in site dynamics and run time requirements.

6. Conclusion

Power availability for critical HFC service during extended outages has become increasingly vital for operators because of new government regulations and greater global reliance on the broadband network. Simultaneously, the problem of extended outages does not seem to be going away any time soon. Data from the U.S. Energy Information Administration shows from 2018-2020 there were 164 extended outages in the U.S. alone that impacted more than 50,000 customers at a time, the largest of which impacted more than 1.4 million homes. [iii]These outages have begun to garner the attention of state and federal lawmakers and, in the case of California have driven strict wireline resiliency regulations.

Current available solutions addressing the challenge of extended outages have their own unique set of benefits and limitations, so having a suite of solutions to address the nuances of varied existing HFC powering deployments is imperative. High capacity TPPL lead-acid batteries will allow for adequate run time extension in a number of scenarios. And, while portable and stationary generators will continue to be useful in limited situations, advancements in Lithium Ion battery technology, and economies of scale driven by electric vehicles, have broadened the applications where lithium has a better TCO. Due to its

extreme energy density, light weight, cycling capability and advanced intelligent diagnostics, Lithium Ion technology as a part of an integrated, engineered system will become the keystone of many operators extended run time network resiliency programs.

Abbreviations

AC	Alternating Current
AGM	Absorbent Glass Mat
BMS	Battery Management System
CPUC	California Public Utility Commission
DC	Direct Current
HFC	Hybrid Fiber Coax
LFP	Iron Phosphate
Li-Ion	Lithium Ion
MSO	Multiple-System Operator
NMC	Nickel-Manganese-Cobalt
OSP	Outside Plant
TCO	Total Cost of Ownership
TPPL	Thin Plate Pure Lead
UPS	Uninterruptible Power Supply
VRLA	Valve Regulated Lead Acid

Bibliography & References

[i] *"Decision Adopting Wireline Provider Resiliency Strategies," Public Utilities Commission of California, February 11, 2021.*

[ii] *https://www.energy.gov/science/articles/charging-development-lithium-ion-batteries*

[iii] *Data analyzed for 2018-2020 from the US Energy Information Administration website: https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=table_b_2*

Evolved MVNO Architectures for Converged Wireless Deployments

A Technical Paper prepared for SCTE by

Omkar Dharmadhikari
Lead Wireless Architect
CableLabs
Louisville CO
303.661.3875
o.dharmadhikari@cablelabs.com

Ojas Choksi
Executive Technical Advisor
CableLabs
Louisville CO
o.choksi-contractor@cablelabs.com

John Kim
Distinguished Technologist and Director of Connectivity
CableLabs
Louisville CO
303.661.3473
j.kim@cablelabs.com

1. Introduction

With the advent of smartphones and widespread deployment of mobile technologies like 4G LTE (and now 5G NR), wireless connectivity is becoming an integral part of a connectivity service offering for multiple system operators (MSOs). Several MSOs that lack mobile network infrastructure have relied on the mobile virtual network operator (MVNO) model to supplement their connectivity offerings with wireless. New entrants like DISH in the United States and Rakuten[1] in Japan (who are deploying their own mobile infrastructure) are also taking advantage of MVNO arrangements to supplement their network coverage.

As the name suggests, an MVNO is a wireless service provider that does not own the end-to-end mobile network; instead, it leverages a portion of a mobile network operator (MNO) network via a business agreement. Typically, MVNOs focus on the marketing, billing, and customer facing aspects and rely on the MNO network and/or services infrastructure to deliver the connectivity and/or services.

The MVNO's payments to the MNO are typically based on usage by the MVNO's customers. Although voice usage has stabilized, data usage continues to grow at a compound annual growth rate (CAGR) of more than 25%.[2] This makes the offloading of data for MVNOs even more important to remain profitable in the long run. The continued success of MVNOs is also important to MNOs in terms of sustaining their wholesale revenues (resulting from the MVNO data usage).

Until now, MVNO offerings from MSOs have relied on Wi-Fi to offload data usage from the MNO's network. Shared spectrum such as the Citizens Broadband Radio Service (CBRS) is significantly reducing the barriers for new entrants and small operators to further offload data usage and improve the economics of an MVNO offering. Deployment of shared spectrum will require MSOs to build their own cellular infrastructure to enable offload using CBRS in areas of high usage.

As MSOs look to deploy their own wireless infrastructure, they will have to contend with three disparate sets of wireless infrastructures—the MSO's community Wi-Fi network, the MNO's 4G/5G network, and the MSO's own 4G/5G network. Maximizing offload via the MSO's own wireless assets, ensuring a consistent user experience, and enforcing uniform and personalized policies as users move in and out of coverage of these three networks will require deployment of new converged network architecture and related capabilities.

This paper is focused on describing options for these evolutionary converged architectures that can be utilized by the MSOs as they progress in building out their own 4G/5G networks.

2. Overview of MVNO Models

Globally, several different MVNO models have been implemented and can be broadly classified as depicted in Figure 1. The classification is based on the amount of mobile network infrastructure owned by the MVNO and the degree of control over the management of different aspects of MSO subscriptions and their service offerings.

[1] "Rakuten Mobile Completes Connection of MMEs with KDDI Roaming Areas Through S10 Interface," press release, April 2020, Rakuten Mobile
[2] Figure 13 of "Ericsson Mobility Report," November 2020, Ericsson

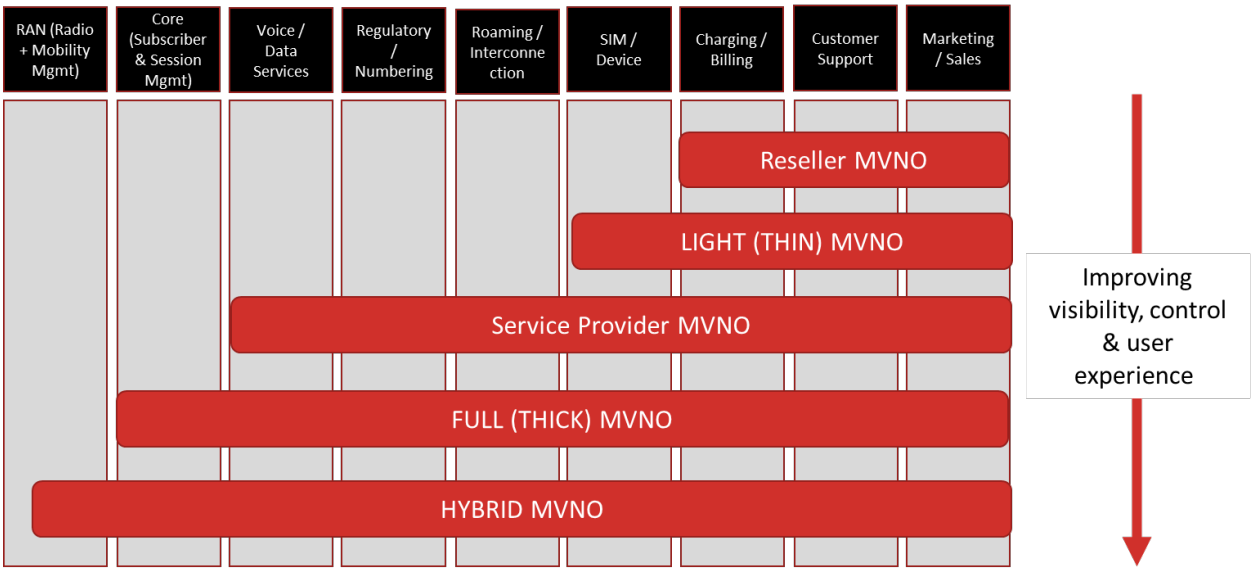


Figure 1: MVNO Types and Degree of Control

Reseller MVNO

A reseller MVNO manages the billing and customer network support functionality while utilizing an MNO for subscriber identity module (SIM) management, device management, core network, and radio infrastructure. The MVNO has no control over subscription, policy, and mobility management and lacks visibility into a subscriber's usage pattern to customize the service offering.

Light (Thin) MVNO

The light MVNO model is like the reseller model, except the SIM and handset are managed by the MVNO; all other aspects remain unchanged. The MVNO still has limited to no control on network policies, subscription, and mobility management of the subscriber on the cellular network and lacks visibility into data usage by its customers.

Service Provider MVNO

In comparison to the first two types of MVNOs, a service provider MVNO is responsible for deploying, operating, and managing its own service platform, thereby enabling the MVNO to differentiate its service offering from that of the MNO. However, the radio and core infrastructures still belong to the MNO, and the MVNO still has limited control on network policies, subscription, and mobility management of the subscriber on the mobile network.

Full (Thick) MVNO

In addition to the service platform, a full MVNO deploys certain mobile core network nodes, such as packet gateways and policy controllers, to have more control over the policy and session management of its customers while allowing the MVNO to leverage the MNO radio and mobility management infrastructure. The data traffic is routed back to the MVNO's mobile core network.

The MVNO manages billing, customer network support, SIM credentials, handset functionality, subscriptions, and policies and has full visibility into data usage. However, the MVNO has limited to no control over mobility management of the subscriber on the cellular network. This model is like the traditional home routed (HR) roaming model specified as part of 3GPP standards.

Hybrid MVNO

A hybrid MVNO (H-MVNO) is a relatively new model wherein the MVNO owns a mobile radio network deployed in specific geographic areas. These networks could be small cell hotspot deployments or traditional regional mobile deployments.

Like the full MVNO model, billing, customer network support, SIM credentials, handset functionality, subscriptions, and policies are managed by the MVNO. Additionally, the user device is prioritized to access and utilize the MVNO radio network when available and to use the MNO radio network only when it is outside the coverage of the MVNO radio network. This model can be of particular interest to many MSOs who may have or are planning to have hotspot and/or regional mobile deployments.

To deliver a seamless (converged) experience across the two wireless networks, a varying degree of convergence (interoperability) between the two networks is required. The degree of convergence and interoperability between the networks will depend on the type of applications used by the end users, types of services to be provided by the H-MVNO network, and the desired level of visibility into the subscriber usage to enable customized service plans, as well as the amount of operational coordination that is acceptable to both the H-MVNO and the MNO.

The hybrid MVNO model and associated converged architecture options are the focus of this paper. The next section covers the converged architecture options for evolving a traditional reseller or light MVNO into a hybrid MVNO.

3. Architecture Options for a Hybrid MVNO (H-MVNO)

The architecture options outlined below are categorized into two broad categories based on device capability regarding SIM support.

1. Architecture options for devices with dual SIM support
2. Architecture options extending support to single SIM devices

3.1. Architecture Options for Dual SIM Devices

In this section, we analyze the architecture options available to H-MVNOs for dual SIM devices.

3.1.1. Architecture Option 1: Independent Mobile Core Networks

Dual SIM devices have been around for some time. Traditionally, dual SIM devices were used to manage two separate phone lines with a single device (e.g., one for personal use and the other for business use or while traveling internationally). With the advent of embedded SIM (eSIM) technology, most smartphone manufacturers now support at least two SIMs—a physical SIM and an eSIM. With two SIMs in a device, the device can be configured to connect to two different networks from two different operators, making it a logical option for H-MVNOs to consider.

Most devices that operate in a dual SIM dual standby (DSDS) mode have two radio receivers but a single transmitter (although this may change with the proliferation of 5G), implying that the user can actively transmit data on only a single network at any given time. While on an active data session on one network, the DSDS device remains in standby mode on the second network, continuing to listen to paging messages. Upon receiving a page, based on service settings, the device either provides an indication of the incoming page to the user or automatically connects to the standby network while transitioning to standby on the first network.

Using DSDS-capable devices is the simplest way for an H-MVNO to maximize the use of its own network and utilize the MNO's network only when the user is outside the coverage of the H-MVNO's own mobile network. Figure 2 shows the network architecture for an H-MVNO utilizing DSDS, where both the MNO and H-MVNO have independent mobile cores and subscriptions. The MNO SIM is configured in the MNO Home Subscriber Server (HSS), while the H-MVNO is configured in the H-MVNO's combined unified data management (UDM) + HSS system. The transition between the H-MVNO and MNO networks in this architecture is controlled and managed by the intelligence in the device.

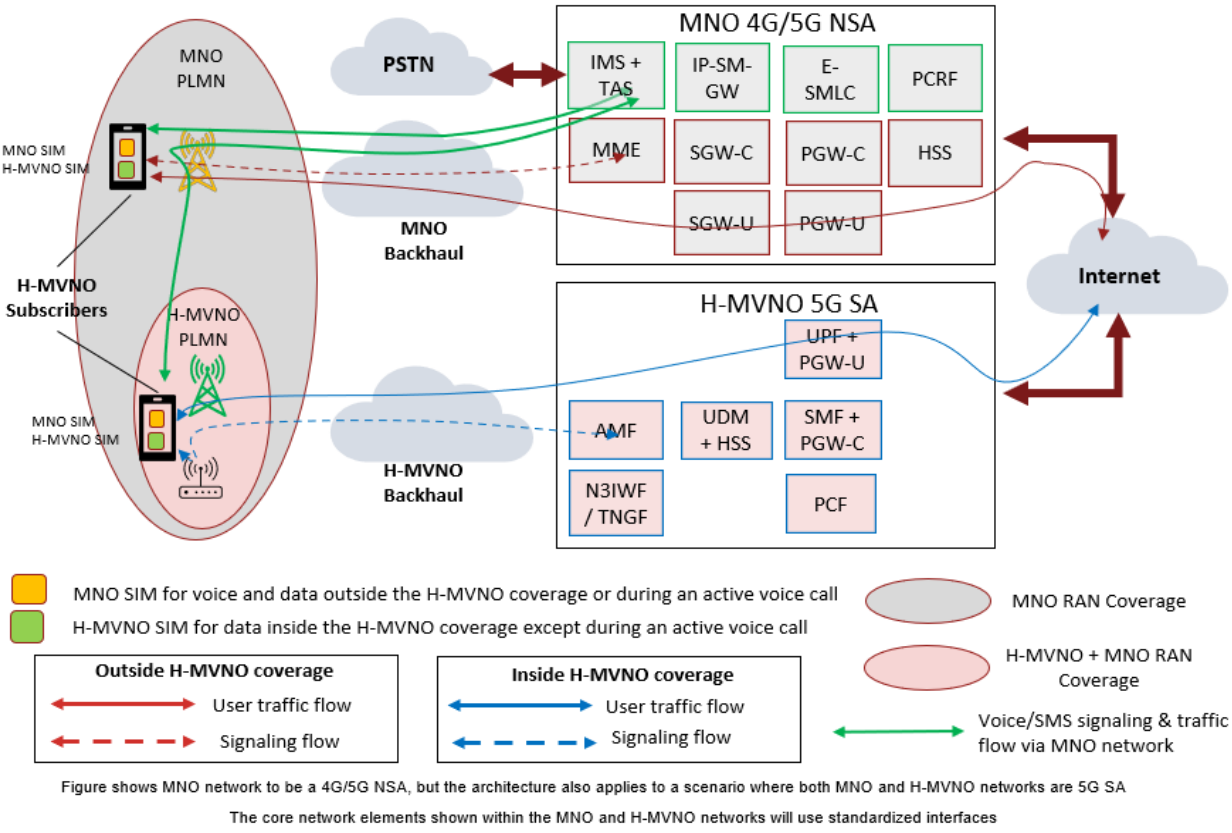


Figure 2: Hybrid MVNO Architecture Option 1—DSDS-Based Architecture

When the H-MVNO subscriber is inside the H-MVNO network coverage area, the user device connects to the H-MVNO network via the H-MVNO SIM (active SIM) to access the Internet while the MNO SIM is in standby, significantly reducing the time spent on the MNO network. On the other hand, when the H-MVNO subscriber is outside the H-MVNO coverage footprint, the H-MVNO SIM is not connected, and the device connects to the MNO network via the MNO SIM (active SIM) to access the Internet. Because of

the limited coverage of an H-MVNO network and the lack of a seamless handover between the two networks in this architecture, voice service will be over the MNO network.

The transition between the two SIMs (and the associated networks) while a user is active on one network is critical. It occurs in two cases.

- Paging-based—the device inside H-MVNO coverage receives a paging message from the MNO network for voice related service, and the user accepts it.
- Coverage-based—the device active on the H-MVNO network moves outside H-MVNO coverage or the device active on the MNO network moves inside H-MVNO coverage.

In addition to having to depend on dual SIM devices, this H-MVNO architecture option faces the challenge of the DSDS switching between networks. The custom intelligence built into all dual SIM devices before the finalization of the 3GPP Release 17 (Rel-17) standard will have vendor-specific implementations, which can result in a variety of user experiences. In June 2019, CableLabs conducted testing[3] on a subset of commercially available dual SIM phones to analyze the behavior and impact on user experience with regard to network connectivity, data-session transition, and paging on a non-active network. The test observations and results indicated that user experience varies significantly depending on the vendors' implementations on the dual SIM devices. The implementations supported paging-based scenarios but were not efficient for coverage-based scenarios, requiring user intervention and impacting user experience when switching between the networks corresponding to the two SIMs .This is because the device implementation assumes overlapping coverage from both the networks.

An H-MVNO needs to work closely with device vendors to customize the devices for effectively managing the transition between the two networks, specifically for hand-in/hand-out of H-MVNO network coverage and paging-based scenarios.

The H-MVNO network also needs to be able to gracefully become aware of the DSDS switching events (responding to either a mobility event or the user accepting a voice call through the MNO network) so as to avoid degrading the key performance indicators (KPIs), perform internal context cleanup (discard buffered download data, delete bearer context, etc.), and avoid wasting system resources. 3GPP is trying to address some of these issues related to the use of multiple SIMs. However, because of the ongoing pandemic, finalization of 3GPP Rel-17 has been delayed, and the earliest availability of these standards-based features in devices and infrastructure will likely be in 2023–2024.[4]

The voice and messaging services, including emergency calling/texting, will be provided via the MNO network utilizing the MNO SIM. While in H-MVNO coverage, the data services are via the H-MVNO access network except when active on a voice call. While on the voice call, the data services will be via the MNO access network utilizing the MNO SIM. While outside the coverage of the H-MVNO network, the data services will be via the MNO's network.

Finally, H-MVNOs lacks real time visibility into their subscribers' data usage statistics and patterns over the MNO network and have no control over policy, subscriptions, mobility, and user experience management when its subscribers are outside the H-MVNO network coverage. To overcome these data visibility and policy challenges resulting from having two separate anchor points with their own policies, user traffic needs to be routed back to the H-MVNO network. One way this can be achieved is via an over the top (OTT) VPN-like solution comprising a custom device-side application (e.g., connection manager) and a server-side application located in H-MVNO's cloud platform (private or public). Another is a

[3] "Dual SIM—An Alternative to 3GPP-Based Roaming Models," technical brief, June 2019, CableLabs R&D Wireless
[4] 3GPP Release 17, Study on System Enablers for Devices Having Multiple Universal Subscriber Identity Modules (USIM), TR 23.761, v1.5.0 (June 2, 2021)

standards-based approach that does not require a custom device client. There are two different standards-based ways (that do not require a client-side application) to achieve this and are described in sections below.

3.1.2. Architecture Option 2: Enabling S8 Interface Between Networks

One evolutionary approach to overcome some of the limitations of the above architecture option is shown in Figure 3.

This architecture evolution relies on implementation of standards-based interfaces between the two networks. The H-MVNO subscriber traffic across both the H-MVNO and MNO SIM connections is always anchored within the H-MVNO mobile core (routed from the MNO to the H-MVNO via the S8[5]-based interface). This ensures full visibility into subscriber data usage, irrespective of which network is used. The H-MVNO and MNO SIMs continue to be configured in their subscriber databases (HSS for MNO and UDM+HSS for H-MVNO) within their own respective mobile cores. The mobility management entity (MME) within the MNO mobile core selects the H-MVNO packet gateway (SMF+PGW-C) based on the default access point name (APN) value specified in the connection request from the device and the APN subscription data from HSS. The DNS query from MME will include both the tracking area code (TAC) and the APN to pick the nearest SMF+PGW-C to the end user. The SMF+PGW-C then selects a corresponding UPF+PGW-U network element nearest to the end user (from a latency perspective). Termination of S6a and S8 in different administrative domains has not been contemplated within 3GPP and will have to be agreed to and coordinated between the MNO and the H-MVNO.

[5] Although MNO network is depicted as a 4G network, this can easily be extended in a similar way for a 5G System deployment

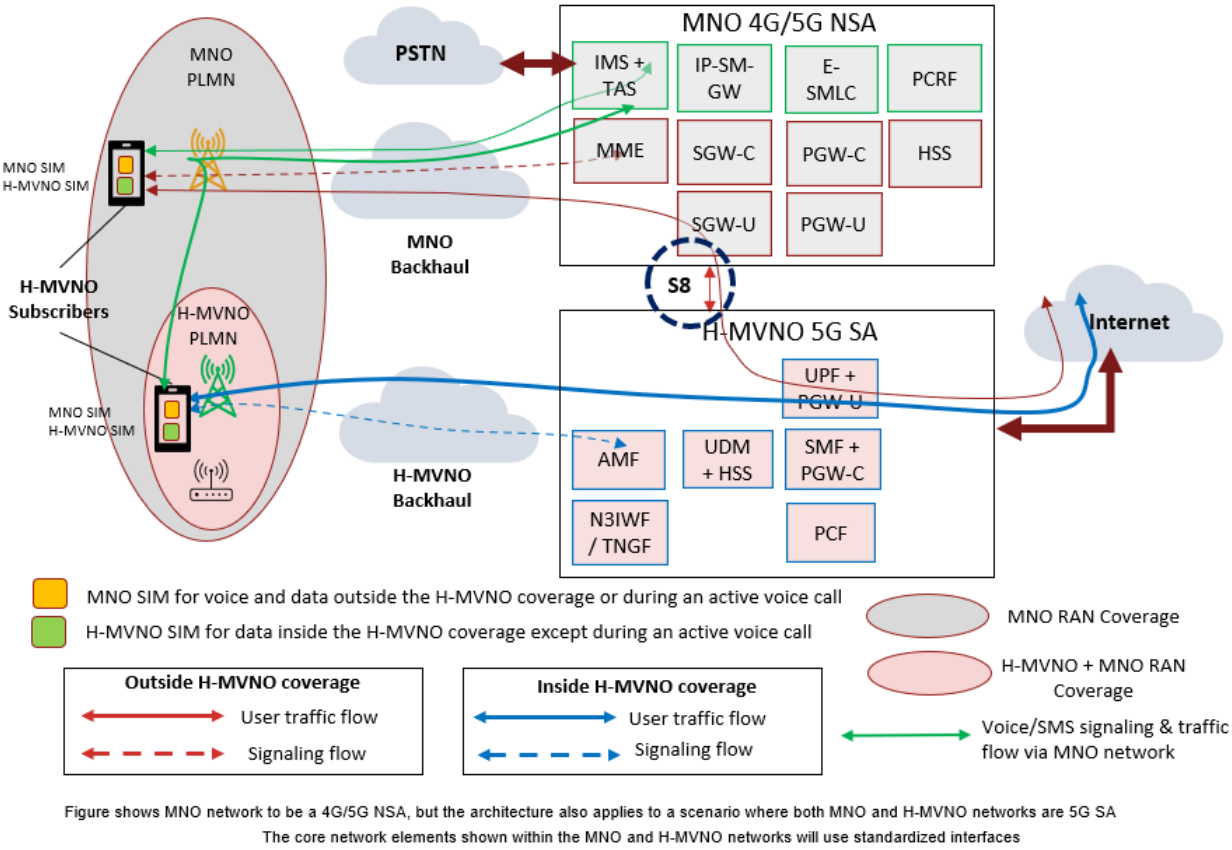


Figure shows MNO network to be a 4G/5G NSA, but the architecture also applies to a scenario where both MNO and H-MVNO networks are 5G SA
The core network elements shown within the MNO and H-MVNO networks will use standardized interfaces

Figure 3: Hybrid MVNO Architecture Option 2—Evolved DSDS Architecture

In this architecture option, like Option 1, the MNO SIM continues to be provisioned in the MNO HSS. Like Option 1, the voice and messaging services, including emergency calling/texting, continue to be provided via the MNO network utilizing the MNO SIM. While in H-MVNO coverage, the data services are via the H-MVNO access network except when active on a voice call. While on the voice call, the data services will be via the MNO access network utilizing the MNO SIM. While outside the coverage of the H-MVNO network, the data services will be via the MNO's network. However, unlike option 1, this architecture option uses a common data anchor point located within the H-MVNO network, thereby providing full visibility into the data usage patterns and statistics.

Some coordination to set up the connectivity between the MNO and H-MVNO networks will be required. Additionally, in this architecture, the MNO serving gateway (SGW-C) may have to generate the charging records for H-MVNO subscribers.

A key difference between the traditional 3GPP specified roaming architecture (home-routed) and this converged architecture (Option 2) is that while the roaming interface (S8) is utilized to interconnect the two domains for user data, the user subscription continues to be provisioned in the MNO HSS, resulting in continued use of intra-domain S6a interface for the control plane. The inter-domain interfaces between the MNO and the H-MVNO can be secured in the same way as they are done today for roaming, by using a secured connection.

Because the H-MVNO has control over the data traffic irrespective of which access network the subscriber is on (MNO or H-MVNO), it can implement uniform policies and functionalities in the H-

MVNO core network to manage steering/switching traffic, not only between the two 3GPP access networks, but also between the MNO's 3GPP network and the H-MVNO's non-3GPP (Wi-Fi) network.

In this architecture, it is possible to support access traffic steering, switching, and splitting (ATSSS) functionality across the H-MVNO's non-3GPP (untrusted, trusted Wi-Fi) network and the MNO's 3GPP network (even when the MNO 3GPP network is 4G[6]). While on the MNO's 4G network, the ATSSS traffic rules on the devices will be updated through the non-3GPP leg of connection. This will require some additional functionality within the H-MVNO infrastructure, as the SMF+PGW-C will have to retrieve information about the serving SMF+PGW-C and UPF+PGW-U and ensure that the same ATSSS anchor is assigned. The ATSSS capability will have to be enabled across both SIMs. However, unlike architecture option 1, no client-side application will be required.

By enabling ATSSS functionality across the H-MVNO's wireless (cellular and Wi-Fi) assets and the MNO's cellular network, a fully converged standards-based architecture could be realized, giving H-MVNOs tremendous flexibility in utilizing all available wireless access networks for user data transmission.

Additionally, capability to transfer the 3GPP leg of ATSSS-compliant Multi-Access Protocol Data Unit (MA-PDU) session from the H-MVNO's 5G network to the MNO's 4G network can also be contemplated as a custom capability within the H-MVNO mobile core; no enhancements are required in the MNO's 4G network. Depending on the implementation of the device IP stack, device side enhancements may not be necessary.

The next architectural option depicts a solution to enable ATSSS between the H-MVNO's Wi-Fi and the MNO's 4G network without requiring enhancements within the H-MVNO's core network or requiring ATSSS to be enabled across both SIMs. It also facilitates use of Wi-Fi connectivity when available, irrespective of whether the user is in the H-MVNO or MNO coverage.

3.1.3. Architecture Option 3: Enabling S6a and S8 Interface Between Networks

Figure 4 depicts further enhancements to architecture option 2. In this architecture option (option 3), the H-MVNO SIM is configured to roam onto the MNO's network when outside the coverage of its home network (H-MVNO network). Therefore, in addition to the S8 interface, this option requires support of inter-domain roaming S6a interface between the MNO and the H-MVNO networks for the H-MVNO SIM.

In this option, the data sessions are established using the H-MVNO SIM irrespective of whether the device is inside or outside the H-MVNO's network coverage area. The MNO SIM is utilized only for data sessions during an ongoing voice session. As in previous architecture options, the voice sessions continue to be established using the MNO SIM.

The roaming S6a interface is used by the MNO's core network to authenticate access via the H-MVNO's SIM. In the event an MA-PDU session was previously established, the anchor SMF+PGW-C is retrieved by the MNO's MME from the H-MVNO's network via the S6a interface. Otherwise, the PGW selection is done by the MME in the same way as described in option 1. The only time ATSSS MA-PDU will not be utilized is when the user is accessing data services during an ongoing voice call and Wi-Fi coverage is unavailable.

[6] 5G defined ATSSS capability is transparent to the 4G from a signaling perspective and does not require any modifications to the MNO deployment

In this option, since data is via the H-MVNO SIM across both networks, this option requires ATSSS configuration only for the H-MVNO SIM. Additionally, no customization for ATSSS across multiple SIMs is required in the H-MVNO's core network to utilize Wi-Fi connection irrespective of whether the user is in the MNO or the H-MVNO coverage. Furthermore, since a single SIM is used for data across the two networks, data session management can be more gracefully managed as the device transitions in and out of H-MVNO coverage. However, in option 3, when UE switches between the two SIMs (based on network availability), UE has to de-register from the source network and register on the target network and may experience a longer interruption to reestablish the session as opposed to Option 2, where separate SIMs are used for data operation on each network and both SIMs are registered to the networks simultaneously.

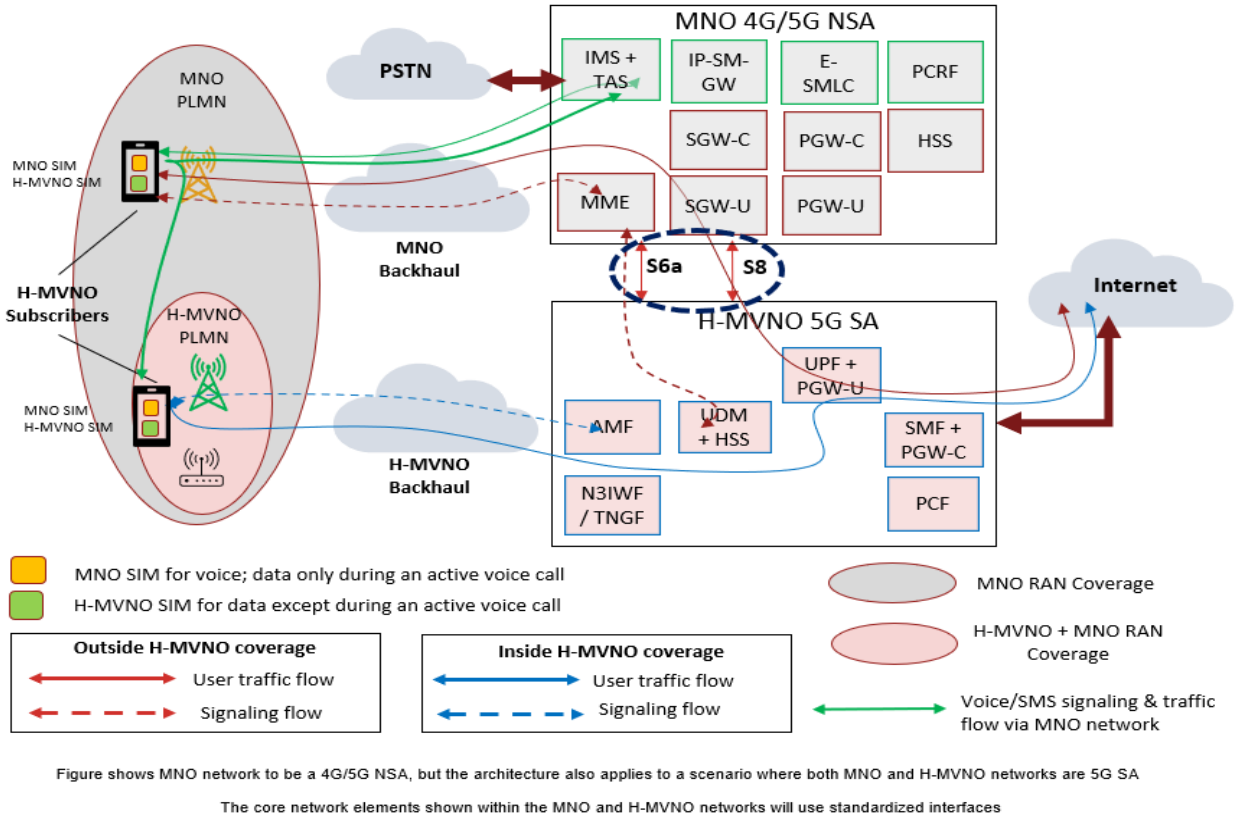


Figure shows MNO network to be a 4G/5G NSA, but the architecture also applies to a scenario where both MNO and H-MVNO networks are 5G SA

The core network elements shown within the MNO and H-MVNO networks will use standardized interfaces

Figure 4: Hybrid MVNO Architecture Option 3—Evolved DSDS Architecture

Table 1 compares the key salient features of the DSDS and evolved DSDS architecture options.

Table 1: Comparison of H-MVNO Architecture Options 1, 2, and 3

Attributes	H-MVNO Dual SIM Architectures		
	Option 1 (DSDS)	Option 2 (Evolved DSDS)	Option 3 (Evolved DSDS)
Enforcement of uniform policies irrespective of the serving network without device/network customization	×	✓	✓
Uniform traffic policy management regardless of the SIM or associated network without device/network customization	×	✓	✓
Full visibility into data usage statistics and pattern irrespective of the network used without device/network customization	×	✓	✓
Use of standardized ATSSS feature to improve session continuity across all accesses	×	×	✓
No need of coordination of roaming interfaces between MNO and H-MVNO	✓	×	×

Though the evolved DSDS architectures solve the issues of converged policy, full visibility into subscriber usage statistics and patterns and leveraging Wi-Fi connection when available, the following challenges still remain.

- Support for dual SIM is needed across the H-MVNO's device portfolio. There are also related concerns (i.e., efficiently managing transitions between the two SIMs as the subscriber moves in and out of H-MVNO access network coverage, especially if two separate SIMs are used for data sessions across the H-MVNO and MNO networks).

- The device on the MNO network is unable to immediately switch back to the H-MVNO network as soon as it moves inside the H-MVNO access network coverage without customization on the device side. One approach to facilitate this transition in the evolved DSDS architecture is to build intelligence in the anchor point (PGW/SMF/UPF), either through custom signaling or further enhancements to standards based ATSSS signaling.

The next section presents additional standards-based architecture options to overcome these challenges.

3.2. Architecture Options Extending Support for Single SIM Devices

In this section, we analyze the architecture options available to H-MVNOs to enable low-latency handovers between their networks and that of MNO networks. Even though the architecture options presented below are focused on single SIM devices, they are backwards compatible with above DSDS architecture options and will also be able to support dual SIM devices.

3.2.1. Direct N26 Interface Between H-MVNO and MNO Networks

One way to facilitate seamless handover of devices between MNO and H-MVNO access networks is through the implementation of the standards-based N26 interface. This architecture option is depicted in Figure 5. In this architecture, the H-MVNO SIM is provisioned within the H-MVNO network. Through the roaming interfaces, S6a and S8, the device can obtain the service by connecting through either the MNO or the H-MVNO access network. What separates this architecture option from the previous options is the introduction of the mobility interface, N26, between the two networks. With this interface, the device will be able to seamlessly move between the two networks through the execution of inter-PLMN (public land mobile network) handover procedures. The networks will be able to control the mobility aspects (connected mode handovers and idle mode cell reselection) as the device moves in and out of the H-MVNO coverage footprint rather than relying on the device to switch the data sessions between the two SIMs, as was the case for the dual SIM architecture options 1 to 3.

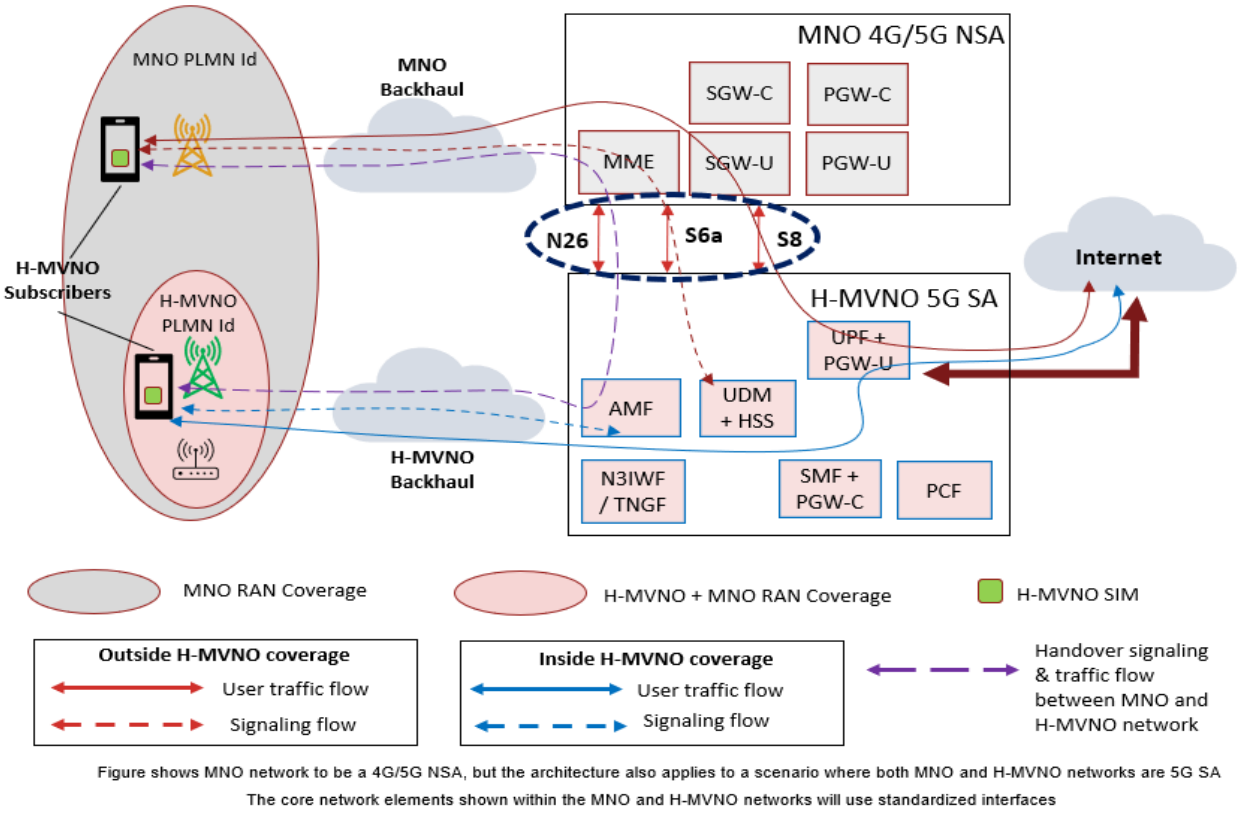


Figure 5: Architecture Option 4—Handover Interface Between MNO and H-MVNO

To enable option 4, several potential operational issues have to be addressed.

1. Securing the handover interface between the two networks
2. Having the appropriate mobility configuration and parameters (e.g., connected-mode and idle-mode configurations, event thresholds) for H-MVNO devices in the MNO network, taking into consideration the differences in the spectrum bands in the two access networks and the overlapping network coverage
3. Ensuring minimal impact to the MNO users from the additional signaling traffic in the core network caused by potential ping-ponging between the two access networks as the H-MVNO device frequently transitions in and out of the H-MVNO footprint
4. In case of multiple H-MVNO partners, the MNO core serving its subscribers having to interface with multiple core networks, creating significant operational challenges and risks

The first issue is similar to option 3 with regards to enabling secured connection with inter-domain S6a and S8 interfaces between the two networks. The same techniques used to interoperate and secure those interfaces also can be utilized for the N26 interface. The Rakuten network is based on this hybrid MVNO architecture option. Rakuten has an MVNO arrangement with KDDI and has enabled the S10 interface between its 4G core and KDDI's 4G mobile core.

The second issue can be addressed with custom configurations within the radio access networks for access barring, handovers, and redirection using RAT Frequency Selection Priority ID (RFSP ID) that has been specified in the standards.

Regarding the third issue, the actual amount of increase in excessive handovers and ping-ponging in the MNO core network will depend on the degree of contiguous deployment in the region by the H-MVNO (i.e., the number of handover boundaries between the two networks). Indoor/outdoor transitions could also result in many handover boundaries depending on the type of spectrum (low- or mid-band).

One way to address this issue of ping-ponging and corresponding signaling overload on the MNO's core network is by using a dedicated core for the H-MVNO traffic. This use of dedicated core also allows us to address the fourth issue. This architecture option utilizing a dedicated core is described below.

3.2.2. Dedicated Core for H-MVNOs

The dedicated core architecture option depicted in

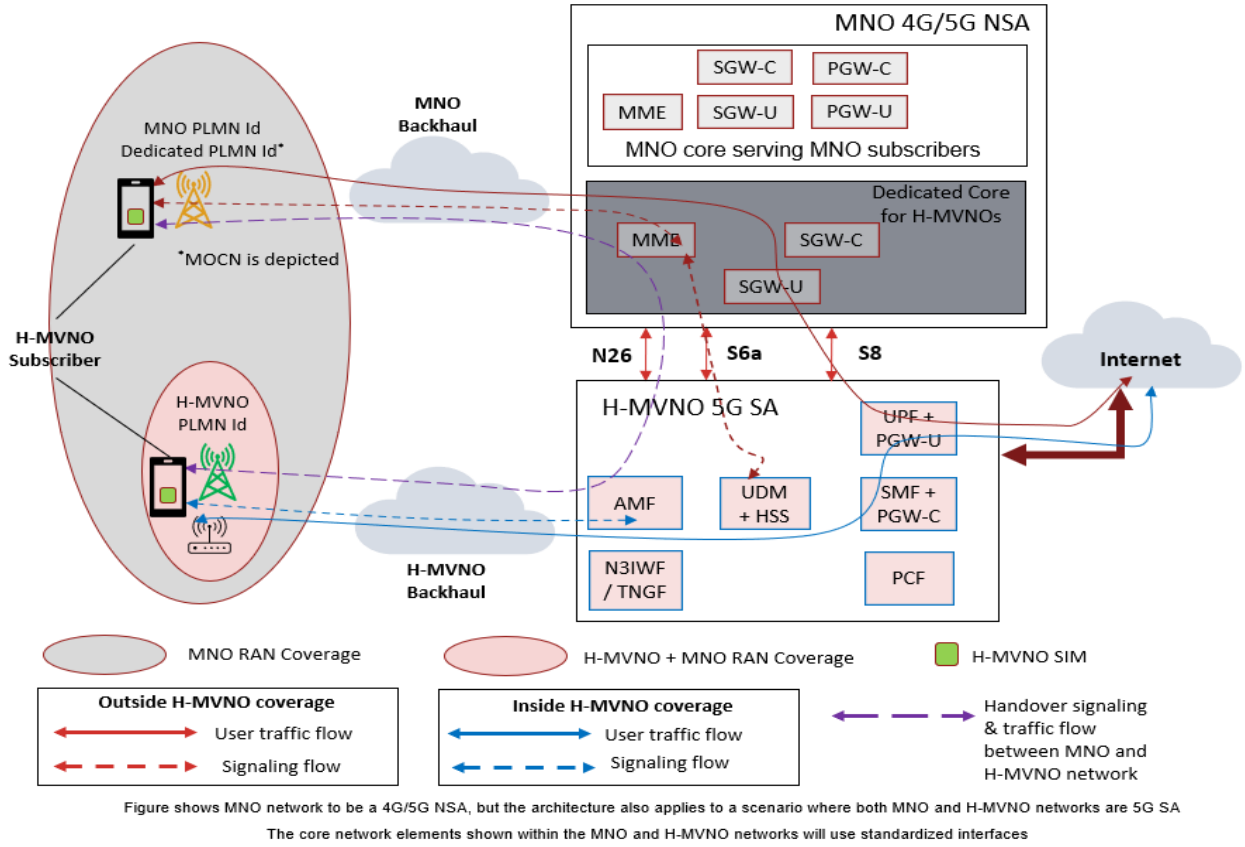


Figure shows MNO network to be a 4G/5G NSA, but the architecture also applies to a scenario where both MNO and H-MVNO networks are 5G SA

The core network elements shown within the MNO and H-MVNO networks will use standardized interfaces

Figure 6 helps alleviate the impact of ping-pong handovers and the risk to MNO core operation. The H-MVNO users' signaling, and user traffic streams will be processed within the dedicated core rather than the MNO's core, thereby isolating signaling load generated from the mobility of the H-MVNO's user devices from that of the MNO's user devices. The dedicated core comprises a dedicated mobility management entity (MME) and, optionally, an SGW (SGW-C + SGW-U).

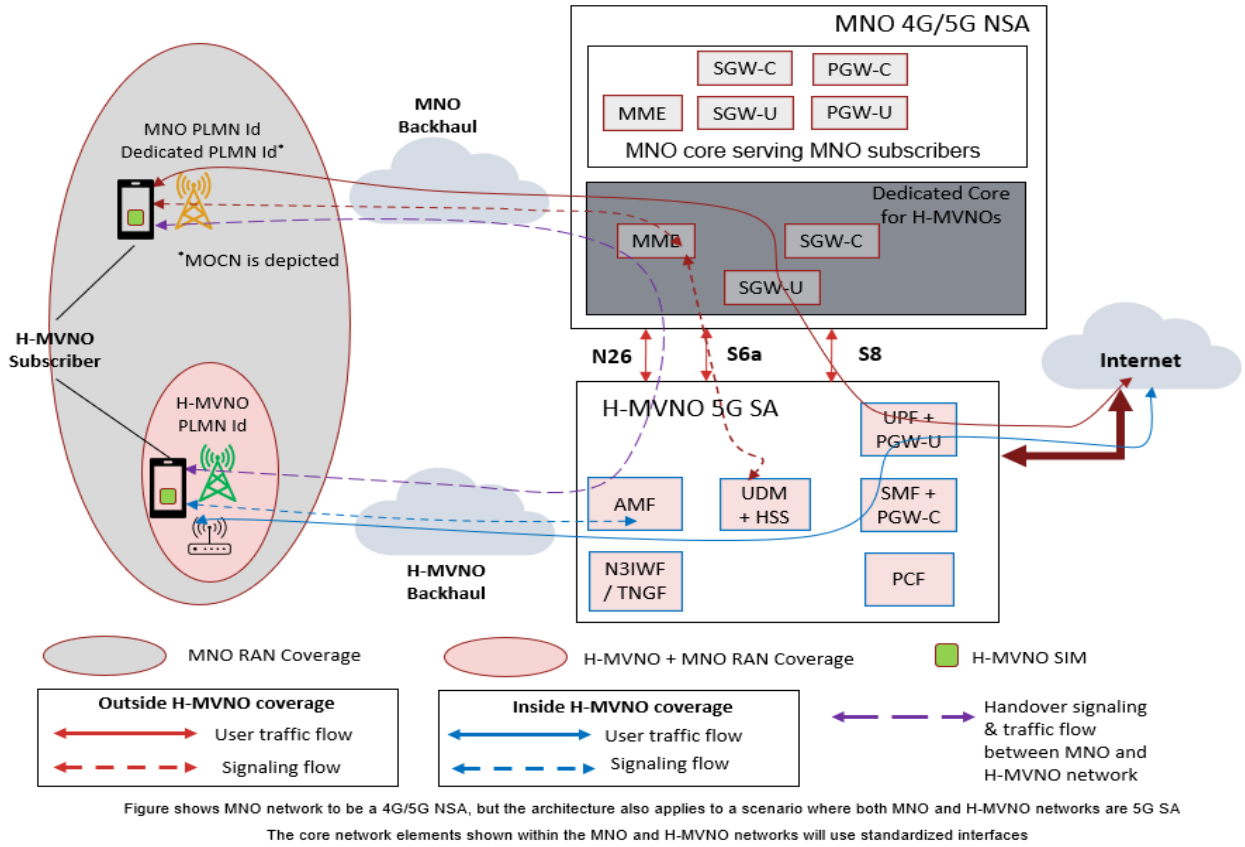


Figure 6: Architecture Option 5—Dedicated Core for H-MVNO

The dedicated core can be deployed and managed within the MNO network (as shown in Figure 6) or externally, depending on the MNO's operational policies. A key aspect of this architecture is that a single dedicated core can be shared across several H-MVNOs having an agreement with the MNO. The dedicated core would separate the traffic of each H-MVNO through the PLMN identity (part of the international mobile subscriber identity, IMSI) and route it to the appropriate anchor points in the H-MVNO home networks. The MME hosted within the dedicated core will perform a DNS query to select the SMF+PGW-C in individual H-MVNO networks, giving them full control over their subscribers' traffic.

Depending on the capabilities of the H-MVNO devices and the MNO access network, the dedicated core can be implemented using one of the following standards-based options.

Multi Operator Core Network (MOCN) specifications: In this option, the MNO access network broadcasts two PLMN IDs—one for its core network and one for the dedicated core. The dedicated core PLMN ID broadcast by the MNO is distinct from that used by the H-MVNOs in their home networks. The H-MVNO user devices are programmed to access the dedicated core PLMN ID when outside of H-MVNO access network coverage. As described previously, the dedicated core uses the home PLMN embedded in the IMSI to route the traffic to each H-MVNO network. The MOCN feature has been available since 3GPP Release 8 and is currently supported by most access vendors. Several advanced MOCN features, such as PLMN-specific configurations, parameters for access barring, handovers, and redirection, are likely to be available from access vendors, allowing for distinct handover settings for MNO and H-MVNO user devices. This can enable handover parameter configurations to be customized,

facilitating handovers between MNO and H-MVNO networks that do not impact the handover operation and performance for MNO user devices.

- **Standardized 3GPP feature—Dedicated Core (DECOR) or Enhanced Dedicated Core (eDECOR):** Implementing the DECOR feature in the MNO core and access network redirects traffic to the dedicated core based on information received in the subscription profile from the H-MVNO HSS. eDECOR-aware user equipment provides the dedicated core network ID (DCN-ID) when it is accessing the MNO network, which uses it to route traffic to the dedicated core. A key advantage of this approach is that the MNO does not have to broadcast multiple PLMN IDs. However, MNO access and core networks need to support the DECOR redirection/routing capabilities to have H-MVNO users serviced by the core dedicated to the H-MVNOs. In addition, to isolate the handover configurations for MNO and H-MVNO devices, additional functionality will be required in the MNO access network. One standards-based approach is to tie custom handover configurations by using a standardized index called the RAT Frequency Selection Priority ID (RFSP ID). One limitation of DECOR, compared to MOCN, is that the MNO will be unable to configure separate access barring settings for H-MVNO user devices. It may not be critical, however, given that MNOs do not currently require separate access barring configurations for their MVNO user devices.

The selection of one of these options will be based on the capabilities of the MNO access and core infrastructure and the H-MVNO devices.

Note: These options can be viewed as precursors of the network slicing concept specified as part of 5G. In the MOCN option, the network slice identifier is the PLMN ID. In the DECOR/eDECOR option, the slice identifier is either the DECOR parameter specified in the user profile, or the dedicated core network identity (DCN-ID) provided by the device.

Table 2 compares the salient features of the inter-network handover and dedicated core architecture options.

TABLE 2: COMPARISON OF H-MVNO ARCHITECTURE OPTIONS 4 AND 5

Attributes	H-MVNO Single SIM Architectures	
	Option 4 (Inter-Network HO)	Option 5 (Dedicated Core)
Seamless handovers	✓	✓
Minimizes signaling traffic impact on MNO core serving MNO subscribers due to handovers	✗	✓
No additional complexity within the MNO network to support multiple H-MVNO networks	✗	✓

Though the dedicated core architecture solves issues related to seamless handovers, it does introduce the following operational overheads in terms of network planning and configuration.

- The MNO needs to enable either MOCN or DECOR to facilitate isolation of the signaling traffic and minimize signaling load (caused by potential ping-ponging between the MNO and H-MVNO networks) on the MNO core control functions.

- The MNO needs to verify interoperability between the MNO access network and the core dedicated for the H-MVNOs.[7]
- The H-MVNO needs to ensure interoperability of the N26 between the dedicated and the H-MVNO core networks.
- The MNO needs to configure H-MVNO-specific mobility parameters in its access network; the H-MVNO needs to configure MNO-specific mobility parameters.
- Support for additional interfaces for supporting voice calls and SMS (described in the following section).

3.2.3. Voice and SMS Services Implications

Voice and SMS services are expected to be an integral part of all mobile service plans; therefore, it is critical to evaluate the ability to support them in the context of architecture options 4 and 5. Voice and SMS services require additional architectural components in the mobile core network that are not required for data services. This additional functionality is required to support i) voice and SMS applications, ii) mobile number portability (MNP), iii) interconnection/ interwork with PSTN, and iv) emergency services as required by the local regulator.

Given the widespread deployment of voice over LTE (VoLTE), 3GPP has defined several ways to leverage it for enabling voice service over 5G NR. Voice services likely will be enabled using VoLTE until 5G deployments become more ubiquitous. 3GPP has specified redirection from the 5G network to the 4G network for voice sessions to facilitate use of VoLTE. One way to redirect is through handover at the time of the voice media setup procedure, and another way is through radio channel redirection. The method of deployment is left up to the MNO operators and will depend on their network capabilities and configurations.

Described below are several ways to facilitate voice and SMS services as part of architecture options 4 and 5 using MNO's LTE deployment.

3.2.3.1. Voice/SMS Services Architecture Option A

One option is for an H-MVNO to support voice and messaging services via its own or a partner's voice/messaging platforms, as shown in Figure 7. In addition to showing the architectural impact, Figure 7 also depicts voice and SMS data flows when the device is in the coverage region of H-MVNO+MNO and MNO-only networks.

i) The voice traffic flow after the redirection to the MNO network (i.e., when the user is inside the H-MVNO coverage) and/or when camped/connected to the MNO RAN (i.e., when the user is outside the H-MVNO network coverage) is depicted by the solid green line.

ii) The voice handover/redirection signaling flow when camped on the H-MVNO is depicted by the dashed purple line.

[7] Problems surrounding interoperability can be minimized by ensuring that the MME and the MNO radio access network is from the same vendor.

iii) The SMS traffic flow after the redirection to the MNO network (i.e., when the user is inside the H-MVNO coverage) and/or when camped/connected to the MNO RAN (i.e., when the user is outside the H-MVNO network coverage) is depicted by the dashed yellow line.

iv) The SMS traffic flow when camped/connected to the H-MVNO RAN is depicted by the solid yellow line.

In voice/messaging architecture option A, the mobile connection for voice and SMS is always anchored at the H-MVNO's UPF, irrespective of whether the UE is located within the coverage area of MNO or H-MVNO's RAN. One perceived disadvantage of this option is that H-MVNO will have to deploy the voice and SMS infrastructure and manage the additional functionalities around MNP, emergency calling and interconnection to the PSTN. Using a voice service partner capable of supporting mobile voice and messaging can overcome these drawbacks.

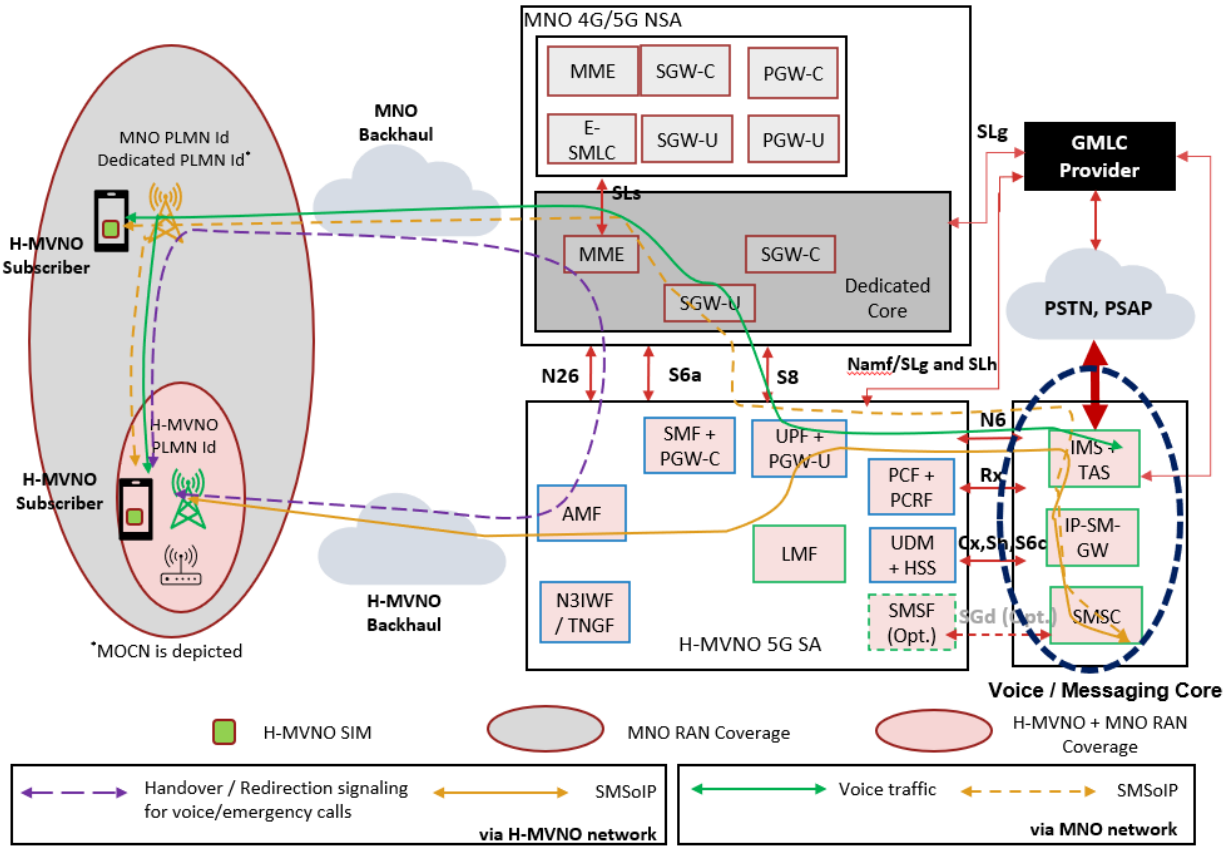


Figure depicts Single SIM Architecture (Option 5), but the voice option described is also applicable to Option 4*

Figure 7: Voice Architecture Option A for Single SIM Devices

An alternative to option A is to leverage MNOs' voice and messaging platforms (i.e., effectively, the MNO becomes their voice service partner). Given that all MNOs have a robust voice and SMS platform, this could be an attractive option for H-MVNOs, especially those who do not have the scale to deploy their own. As described below, there are two ways to leverage MNOs' infrastructure—voice/messaging architecture options B and C.

3.2.3.2. *Voice/SMS Services Architecture Option B*

As shown in Figure 8, in voice/messaging architecture option B, MNO's voice/SMS services platforms are used while the voice subscription and related credentials are configured in the H-MVNO UDM+HSS. In this architecture, additional 3GPP interfaces between the MNO's IMS, IP-SM-GW, and SMSC and the H-MVNO's UDM+HSS, PCF+PCRF, and SMSF are also configured as depicted in Figure 8. The MNO's voice and SMS platforms will use these interfaces to authenticate/authorize the user and retrieve the subscription and registration status/info (using the MCC+MNC from the IMSI of user device). This will ensure successful registration for voice and SMS services within the MNO's IMS and IP-SM-GW. It will also ensure proper forwarding of the incoming calls and text messages, irrespective of whether the user is camped on H-MVNO or MNO radio access network (RAN). Like in option A, the voice and SMS data connection will always be anchored in H-MVNO's UPF. The voice traffic and SMSs will be transferred from an H-MVNO's UPF to the MNO's IMS voice and SMS platforms through a secure data connection. The MNO's IMS will interface with the H-MVNO's PCF+PCRF (either directly or via a local MNO PCRF) through the Rx interface to set up dedicated bearers for the voice media traffic.

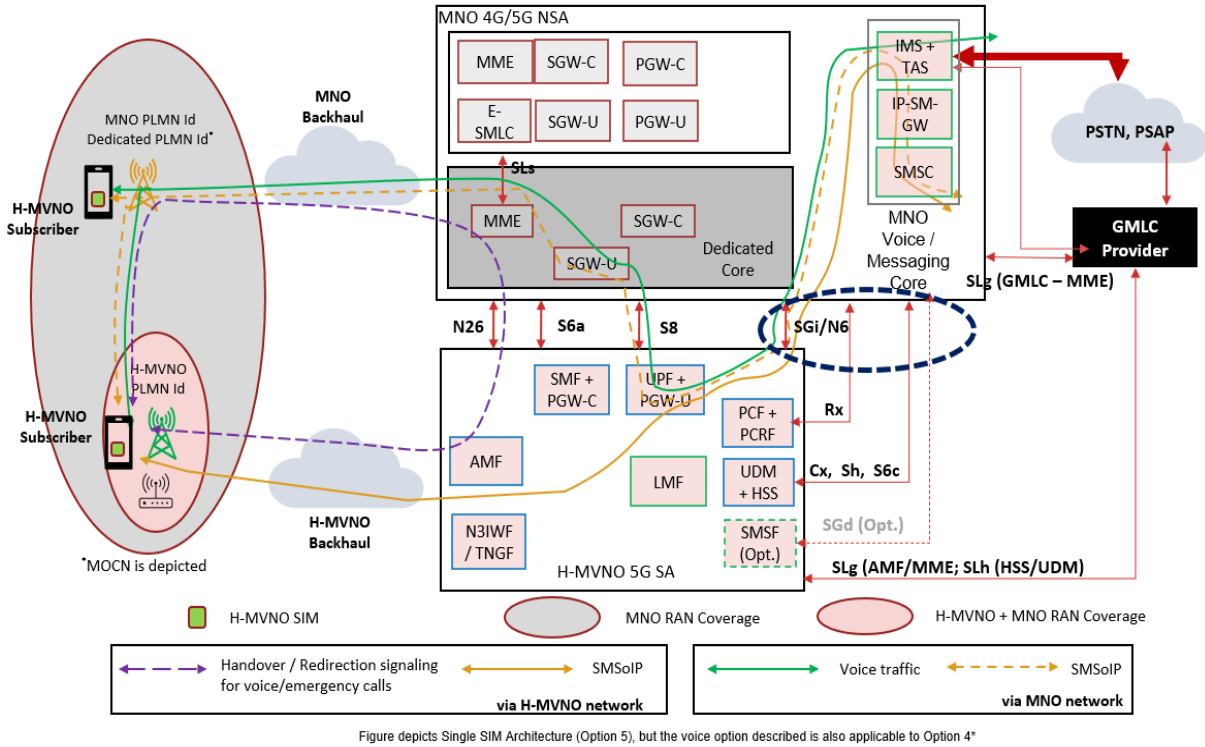


Figure depicts Single SIM Architecture (Option 5), but the voice option described is also applicable to Option 4*

Figure 8: Voice Architecture Option B for Single SIM Devices

Like option A, Figure 8 depicts the voice and traffic flows related to voice handover/redirection signaling (dashed purple), voiced traffic (solid green), and SMS flows (solid and dashed yellow).

User Connected to H-MVNO RAN

The voice call will be initiated, then redirected or handed off to the MNO's RAN. If an emergency call is initiated by the user, it will be redirected to be initiated via the MNO's network. The data anchor will continue to be the H-MVNO's UPF via the dedicated core. The voice signaling and media will be routed from the H-MVNO's UPF to the MNO's IMS.

The text messages will not result in redirection or handover to the MNO's RAN. The text messages will be forwarded via the H-MVNO's UPF to the MNO's IP-SM-GW via the IMS, which will then forward it to the SMSC for delivery to the target user. If a 5G control plane (NAS) is used to transfer the SMS (in the event IP-SM-GW is unavailable), then an additional interface between the MNO's SMSC and the H-MVNO's SMSF will have to be configured.

One aspect of voice and text services is the ability to accurately determine the user location during emergency calls and texting. The Gateway Mobile Location Centre (GMLC) must be able to query the serving control node (AMF/ dedicated MME) from the UDM+HSS to request the location of the user device from the serving control node. The MNO's GMLC partner must be capable of routing the query to the H-MVNO's network. Alternatively, the MNO network must be able to forward or redirect the query from the GMLC provider to the H-MVNO's UDM+HSS.

For emergency voice calls, the dedicated MME will retrieve the location from the MNO E-SMLC to deliver it to the GMLC. However, for text to 911, the H-MVNO will have to deploy its own LMF in its core as the user will not redirected to the dedicated MME for text service while in the H-MVNO coverage.

User Connected to MNO RAN

Even when the user is in the coverage of the MNO's RAN, it will continue to be anchored in the H-MVNO's UPF via the dedicated core (architecture option 5) or through a direct interface (architecture option 4). The voice call and text will be routed back to the IMS via the H-MVNO's UPF to the MNO's IMS platform. IMS will continue to interact with the H-MVNO's PCF+PCRF to set up dedicated bearers for the user traffic. Similarly, SMSC will continue to query the UDM+HSS to identify the serving node (either SMSF or the IP-SM-GW) so that the incoming text can be sent to the serving control node for delivery to the user's device.

The location retrieval by GMLC for emergency calls will be performed in the same way as described above. The text to 911 will be delivered via the dedicated core and the location retrieved using dedicated MME and MNO's E-SMLC

3.2.3.3. Voice Option C

If the H-MVNO operator is using a separate ISIM instead of deriving the IMPU from the USIM, one could also consider provisioning the ISIM in a separate HSS located within the MNO network or in the dedicated core as shown in voice architecture option C in Figure 9. This will allow the MNO's IMS functions to authenticate locally without having to interact with the H-MVNO HSS. Furthermore, if the IMS APN is anchored in the MNO's PGW, it will also eliminate interaction between the MNO's PCRF and the H-MVNO's UPF for the setup of the dedicated bearer.

Like options A and B, Figure 9 depicts the voice/SMS signaling and traffic flows when the device is in the coverage of MNO and H-MVNO networks. The color scheme used for the traffic flows is identical to that for options A and B.

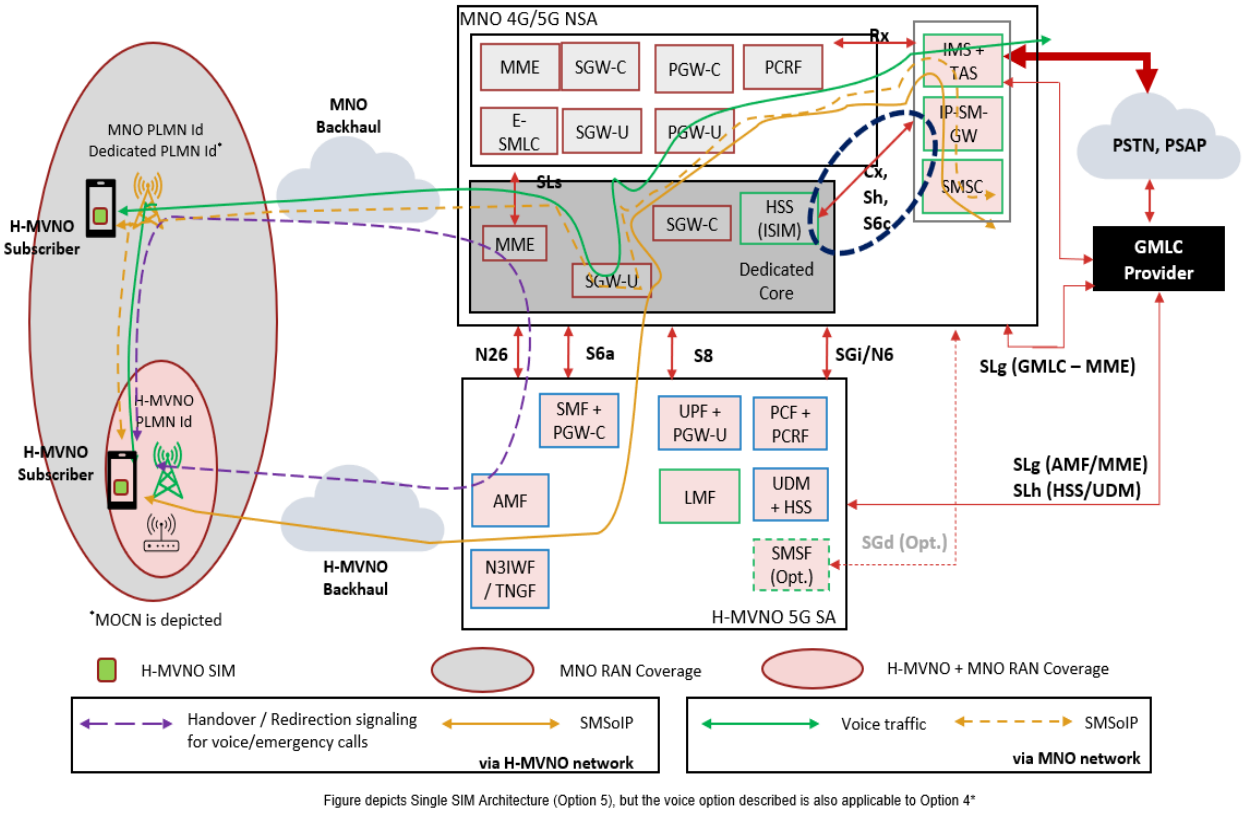


Figure depicts Single SIM Architecture (Option 5), but the voice option described is also applicable to Option 4*

Figure 9: Voice Architecture Option C for Single SIM Devices

One drawback of anchoring IMS APN in the MNO's PGW is that if the MNO has deployed 4G EPC core only, then the dedicated core needs to incorporate an interworking function to translate HTTP interface based 5G session control signaling messages received from the H-MVNO AMF into GTP-C based 4G EPC signaling messages to successfully set up the MNO's PGW as the anchor. To circumvent the need for this new interworking function, a dedicated UPF for voice could be deployed as part of the dedicated core.

In this architecture, the SMSC will still have to interact with the H-MVNO's UDM+HSS to determine the SMS serving node(s) (SMSF or IP-SM-GW). In addition, the GMLC provider will still have to interact with the H-MVNO's UDM+HSS to query the serving node (AMF/MME) to retrieve the location of the user device during text to 911 sessions while the UE is connected via the H-MVNO's RAN. Like option B, H-MVNO will have to deploy the LMF in its core to determine user location during text to 911.

Table 3 below summarizes the pros and cons for the above voice architecture options.

Table 3: Benefits and Impacts of Various Voice Architecture Options

Voice Options	Benefits	Impacts
Option A	• No coordination required with the MNO	• Requires H-MVNO to deploy IMS, TAS and SMSC (or partner with a voice/SMS service provider). • H-MVNO (or its voice partner) will have to support MNP, interconnectivity with PSTN, emergency calling, and associated location requirements. • H-MVNO (or its voice partner) needs to enable interconnection with PSTN and for outbound roamers.
Option B	• H-MVNO does not have to deploy voice and SMS platforms—IMS, TAS, IP-SM-GW, and SMSC. (SMSF may be required to facilitate delivery of SMSs when IP-SM-GW is unavailable.)	• Coordination is required with the MNO to enable interfaces between MNOs' IMS and SMS platforms and H-MVNOs' 5G core systems. • H-MVNO will still have to deploy location management function (LMF) (i.e., location server) to deliver location for text to 911 initiated by the user over H-MVNO's network. • MNO needs to enable forwarding/redirection of queries from GMLC provider to the appropriate function in the H-MVNO network and the dedicated core network.
Option C	• H-MVNO does not have to deploy voice and SMS platforms—IMS, TAS, IP-SM-GW, and SMSC • Fewer interfaces required between the MNO and H-MVNO networks compared to option B	• Dual provisioning required (USIM to be provisioned in H-MVNO network, while ISIM to be provisioned in MNO network). • If MNO has not deployed a combined 4G/5G core, an interworking function is required to translate session management messages. • Interface between SMSC and HSS+UDM will still be required to retrieve the serving SMSF if it is deployed in the H-MVNO network and not statically configured. • H-MVNO will still have to deploy LMF (i.e., location server) to deliver location for text to 911 initiated by the user over H-MVNO's network. • MNO needs to enable forwarding/redirection of queries from GMLC provider to the appropriate function in the H-MVNO network and the dedicated core network.

4. Conclusion

CableLabs recognizes the evolving mobile industry landscape driven by the introduction of 5G and the availability of new and innovative spectrum options. Many of our members are either MNOs supporting MVNOs or MVNOs looking to deploy their mobile networks. Globally, MVNOs have been around for some time, with varying degrees of control being made available to them. As regulators are also looking to facilitate increased competition, MVNOs are expected to play an integral part in that effort.

In a data-centric connectivity environment, because of continually growing usage and a lack of control over subscribers, MVNO arrangements have typically constrained the MVNOs' mobile service plans. As the data usage continues to grow, it is generally recognized within the industry that data offloading onto Wi-Fi and/or own mobile network is necessary for the success of an MVNO.

This paper presents evolutionary converged MVNO architectural blueprints that will allow CableLabs members who are MVNOs or are considering becoming one to converge their wireless connectivity service and maximize the offload onto their own mobile and Wi-Fi deployments while improving MVNO user experience. They will also allow an MNO member to differentiate itself from its competitor by offering greater flexibility to its MVNO customers and creating a win-win strategy for itself and its wholesale partners (MVNO). **Error! Not a valid bookmark self-reference.** and Table 5 summarize a comparative assessment of the different architecture options.

Table 4: Comparative Assessment of Architecture Options

Attributes	Dual SIM Architectures			Single SIM Architectures	
	DSDS	Evolved DSDS		Inter-Network HO	Dedicated Core
	Option 1	Option 2	Option 3	Option 4	Option 5
Support for dual SIM devices	✓	✓	✓	✓	✓
Support for single SIM devices	✗	✗	✗	✓	✓
No need for customization on the device to manage network transition	✗	✗	✗	✓	✓
Enhanced user experience through full data visibility, uniform policy, and subscription management	✗	✓	✓	✓	✓
Maximized use of H-MVNO access network when available without a connection manager client in the device	✗	✓	✓	✓	✓
Seamless low-latency mobility	✗	✗	✗	✓	✓
No interfaces needed between the two networks	✓	✗	✗	✗	✗
Minimizes signaling traffic impact on MNO core serving MNO subscribers due to handovers	✓	✓	✓	✗	✓
Use of standardized ATSSS solution for improved session continuity across all accesses	✗	✗	✓	✓	✓
No custom configuration of mobility related parameters in MNO's access network	✓	✓	✓	✗	✗

Table 5: Comparative Assessment of Single SIM Architecture Voice Options

Attributes	H-MVNO Single SIM Architectures (Voice)		
	Option A	Option B	Option C
No coordination required for enabling interfaces between MNO IMS and H-MVNO network	✓	✗	✗
No need for H-MVNO to host voice and SMS platforms	✗	✓	✓
No need for additional interface to authenticate H-MVNO subscribers	✓	✗	✓
No need for dual provisioning (USIM and ISIM separately provisioned across the two networks)	✓	✓	✗

The blueprints presented in this paper are standards based and evolutionary in nature. Members could start with the evolved DSDS option, then evolve to implement full mobility between their and MNO deployments by using the dedicated core option. The dedicated core architecture allows MNOs to isolate their infrastructure and customers from any excessive MVNO-related signaling or user traffic resulting from mobility between the MNO and MVNO access networks.

For dual SIM architecture options, voice is always carried over the MNO SIM utilizing the MNO network. This dual SIM approach can also be considered as an alternative to voice options A-C for architecture options 4 and 5 whereby N26 is used to enable low-latency data-centric applications (e.g., VR/AR applications) while the voice/SMS is supported using the second SIM via the MNO network. For single SIM devices requiring voice/SMS support, one of the voice architecture options will have to be considered.

Members could leverage the architecture options discussed in the paper to converge policy and subscription infrastructure across all access networks (MNO, MVNO, Wi-Fi) and enable a seamless user experience to their customers irrespective of the underlying wireless access technology.

Abbreviations

3GPP	3rd Generation Partnership Project
AMF	access and mobility management function
APN	access point name
AS	access stratum
ATSSS	access traffic steering, switching, and splitting
DCN-ID	dedicated core network identity
DECOR	dedicated core
DNS	domain name server
DSDA	dual SIM dual active
DSDS	dual SIM dual standby
eDECOR	enhanced dedicated core
EPC	evolved packet core
eSIM	embedded SIM
GMLC	gateway mobile location center
GTP-C	general packet radio service tunnelling protocol control plane
GTP-U	general packet radio service tunnelling protocol user plane
GW	gateway
H-MVNO	hybrid mobile virtual network operator
HR	home routed
HSS	home subscriber server
IMS	Internet protocol multimedia subsystem
IMSI	international mobile subscriber identity
IP-SM-GW	internet protocol-short message-gateway
KPI	key performance indicator
LMF	location management function
MCC	mobile country code
MME	mobility management entity
MNC	mobile network code
MNO	mobile network operator
MNP	mobile number portability
MOCN	multi-operator core network
MVNO	mobile virtual network operator
NAS	non-access stratum
NSA	non-standalone
OTT	over-the-top
PGW-C	packet gateway control plane
PGW-U	packet gateway user plane
PLMN	public land mobile network
PSAP	public safety answering point

PSTN	public switched telephone network
RAN	radio access network
RAT	radio access technology
RFSP	radio access technology/frequency selection priority
RRM	radio resource management
SA	standalone
SCTE	Society of Cable Telecommunications Engineers
SIM	subscriber identity module
SMF	session management function
SMS	short message service
SMSC	short message service center
SMSF	short message service function
SPID	selection priority identity
TAC	tracking area code
TCC	text control center
UDM	unified data management
UE	user equipment
UPF	user plane function

Bibliography & References

3GPP TS 23.002, “LTE; Network architecture” (Release 16), v16.0.0, July 2020.

3GPP TS 23.122, “Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode” (Release 16), v16.7.0, October 2020.

3GPP TS 23.167, “IP Multimedia Subsystem (IMS) emergency sessions” (Release 16), v16.2.0, July 2020.

3GPP TS 23.203, “Policy and charging control architecture” (Release 16), v16.2.0, November 2020.

3GPP TS 23.204, “Support of Short Message Service (SMS) over generic 3GPP Internet Protocol (IP) access” (Release 16), v16.0.0, November 2020.

3GPP TS 23.228, “IP Multimedia Subsystem” (Release 16), v16.4.0, October 2020.

3GPP TS 23.234, “3GPP system to Wireless Local Area Network (WLAN) interworking” (Release 13), v13.1.0, March 2017.

3GPP TS 23.271, “LTE Location Services (LCS)” (Release 16), v16.0.0, July 2020.

3GPP TS 23.273, “5G System (5GS) Location Services (LCS)” (Release 16), v16.4.0, July 2020.

3GPP TS 23.401, “General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access” (Release 16), v16.10.0, March 2021.

3GPP TS 23.501, "System Architecture for 5G System" (Release 16) , v16.6.0, October 2020.

3GPP TS 23.502, “Procedures for the 5G System" (Release 16), v16.5.0, July 2020.

3GPP TS 23.503, “Policy and charging control framework for the 5G System” (Release 16), v16.5.0, July 2020.

3GPP TS 24.008, “Mobile radio interface Layer 3 specification; Core network protocols” (Release 16), v16.6.0, October 2020.

3GPP TS 24.193, “Access Traffic Steering, Switching and Splitting (ATSSS)” (Release 16), v16.0.0, July 2020.

3GPP TS 24.171, “Control Plane Location Services (LCS) procedures in the Evolved Packet System (EPS)” (Release 16), v16.0.0, July 2020.

3GPP TS 24.228, “Signalling flows for the IP multimedia call control based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP)” (Release 5), v5.15.0, September 2006.

3GPP TS 24.229, “IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP)” (Release 16), v16.8.0, January 2021.

3GPP TS 24.301, “Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS)” (Release 16), v16.6.0, October 2020.

3GPP TS 24.341, “Support of SMS over IP networks” (Release 16), v16.0.0, November 2020.

3GPP TS 24.501, “Non-Access-Stratum (NAS) protocol for 5G System (5GS)” (Release 16), v16.5.0, August 2020.

3GPP TS 24.571, “Control plane Location Services (LCS) procedures” (Release 16), v16.2.0, November 2020.

3GPP TS 29.171, “LCS Application Protocol (LCS-AP) between the Mobile Management Entity (MME) and Evolved Serving Mobile Location Centre (E-SMLC)” (Release 16), v16.1.0, November 2020.

3GPP TS 29.172, “LCS Protocol (ELP) between the Gateway Mobile Location Centre (GMLC) and the Mobile Management Entity (MME)” (Release 15), v15.0.0, July 2018.

3GPP TS 29.173, “Diameter-based SLh interface for Control Plane LCS” (Release 16), v16.0.0, July 2020.

3GPP TS 29.211, “Rx Interface and Rx/Gx signaling flows” (Release 6), v6.4.0, June 2007.

3GPP TS 29.212, “Policy and Charging Control (PCC); Reference Points” (Release 16), v16.4.0, November 2020.

3GPP TS 29.214, “Policy and charging control over Rx reference point” (Release 16), v16.4.0, November 2020.

3GPP TS 29.228, “IP Multimedia (IM) Subsystem Cx and Dx Interfaces; Signaling flows and message contents” (Release 16), v16.1.0, November 2020.

3GPP TS 29.229, “Cx and Dx interfaces based on the Diameter protocol” (Release 16), v16.2.0, June 2020.

3GPP TS 29.244, “Interface between the Control Plane and the User Plane nodes” (Release 16), v16.5.0, November 2020.

3GPP TS 29.272, “Mobility Management Entity (MME) and Serving GPRS Support Node (SGSN) related interfaces based on Diameter protocol” (Release 16), v16.6.0, March 2021.

3GPP TS 29.274, “Tunnelling Protocol for Control plane (GTPv2-C)” (Release 16), v16.5.0, November 2020.

3GPP TS 29.303, “Domain Name System Procedures” (Release 16) v16.3.0, November 2020.

3GPP TS 29.328, “IP Multimedia (IM) Subsystem Sh interface; Signaling flows and message contents” (Release 16), v16.1.0, November 2020.

3GPP TS 29.329, “Sh interface based on the Diameter protocol” (Release 16), v16.0.0, July 2020.

3GPP TS 29.338, “Diameter based protocols to support Short Message Service (SMS) capable Mobile Management Entities (MMEs)” (Release 16), v16.0.0, July 2020.

3GPP TS 36.331, “Radio Resource Control (RRC)” (Release 16), v16.1.1, July 2020.

3GPP TS 36.413, “S1 Application Protocol (S1AP)” (Release 16), v16.5.0, April 2021.

3GPP TS 38.331, “Radio Resource Control (RRC); Protocol specification” (Release 16), v16.3.1, January 2021.

3GPP TS 38.413, “NG Application Protocol (NGAP)” (Release 16), v16.2.0, July 2020.

Execute The Upstream Makeover Without Leaving Scars

A Technical Paper prepared for SCTE by

Dr. Robert Howald
Fellow
Comcast
1800 Arch Street, Philadelphi PA
robert_howald@comcast.com

Larry Wolcott
Fellow
Comcast
1401 Wynkoop Street, Denver CO
larry_wolcott@comcast.com

Leslie Ellis
President
EllisEdits
1401 Wynkoop Street, Denver CO
leslie@ellisedits.com

1. Introduction

For decades, and within the halls of this very event (Cable-Tec Expo), technologists, being a practical bunch, described the substantial task that is widening the 5-42/54 MHz upstream signal path, as the kind of monumental event that would happen but once in a lifetime. This was usually expressed as "not in my lifetime," or variants.

The title of this paper, and the breadth of technical literature happening concurrent with this paper, is, first, an acknowledgement that widening the reverse path is very much going to happen in our lifetimes. It's also an assurance that while going to a Mid-Split (85 MHz) or High-Split (204 MHz) upper spectral boundary for the upstream, home-outwards signal path is a network makeover, it is not a network rebuild. There are ways to accomplish a larger upstream signal path that are precise, reasonably swift, and forgiving – all vital elements to a "makeover without scars."

Informed by substantial lab, field and live/production environment experiences, this paper aims to illuminate why a roomier upstream path is happening now. It will describe the major things that matter, when preparing for and enacting a systemic widening of that narrow sliver of upstream spectrum at the low end of the frequency band, between 5-42 MHz. A spectral area renowned for its many signal-squelching quirks, like impulse noise. The intent is to share what works and what doesn't, when it comes to accomplishing an upper spectral boundary of 85 MHz or 200 MHz.

Because, unquestionably, the upstream path is intrinsic to all two-way applications: It is one of the two ways.

Mid- and High-split upstream configurations coincide with increasingly powerful DOCSIS 3.1 options, even as DOCSIS 4.0 is emerging. The optimal near-term expansion and long term DOCSIS 4.0 migration varies by operator. Each operator must necessarily consider its network starting point, given the interdependency of bandwidth initiatives such as node splits, Distributed Access Architectures (DAA), upper spectral boundaries, and fiber-deeper topologies. It's also worth noting that DOCSIS Annex-A (conversationally known as "Euro DOCSIS") reflects the fact that our colleagues "on the other side of the pond" send signals upstream in the spectrum between 6-65 MHz and do so very successfully.

The mathematics of Compounded Annual Growth Rate (CAGR) provide a straightforward way to quantify capacity growth and network lifespan. "Billboard speeds" must also inform the upgrade roadmap, but generally the "What" of traffic engineering and lifespan management is tractable analyses.

The "How-to" of spectrum migration is where it gets complicated. Operators understand investments in node and amplifier upgrades from previous cycles. However, these cycles didn't address upstream spectrum, largely because usage patterns didn't warrant it. The 42/54 MHz split has been in place for decades, and devices that adhere solely to it, particularly set-top boxes (STBs), are in many millions of homes. Production-scale tools, techniques, and processes must be developed to ensure that a new, wider upstream path can be efficiently operationalized, while being transparent to customers.

This paper will describe the analysis, tools, techniques, and processes to enable this upstream bandwidth transformation, focusing on production operationalization of an 85 MHz Mid-Split including:

- How homes may be impacted by a mix of device spectrum capabilities
- Mid-Split activation using SC-QAM and OFDMA
- Existing metrics and tools to assess home health

- New automation techniques to enable a seamless transition for customers with the activation of new upstream spectrum
- Cross-functional tools and processes for Tech Ops, Care, and Serviceability
- Identify and discuss some of the differences between Mid-Split and High Split (204 MHz), and of DOCSIS 4.0

Widening the upstream to stay ahead of heavy bidirectional consumption is a multi-dimensional topic. Readers will learn about new tools and operational practices that can smooth this transformation.

2. A Brief History of Cable's Upstream Path

The term "upstream path" is synonymous with the "reverse path" and the "return path" because it came second, after the "forward" signal path, from Headends to homes. For the first few decades of cable television's evolution, from the late 1950s to the late 1970s, the upstream signal path wasn't necessary. Television signals were broadcast downstream, through the plant, to homes; subscribers turned on their TVs, and watched. Nothing was "clicked," and none of those clicks moved upstream, from homes to Headends, because nothing was clickable.

In the late 1970s, some operators experimented with televisions and rudimentary data services that encouraged consumers to interact. Coincident with that, attention started to focus on building a two-way path to augment the existing one-way, downstream plant. That involved installing modules into existing amplifiers that fed a signal upstream, to the headend, then balancing that two-way signal path. From the late 1970s until the mid-90s, in fact, operators expressed their two-way-readiness in terms of what percentage of amplifiers were "two-way-capable." This meant that the amplifier housing had an empty slot for the reverse module.

Spectrally, the 5-42/54 MHz reverse path is an inhospitable zone, highly susceptible to signal ingress and impulse noise. What makes it worse is that most noise – upwards of 70%, by some estimates – originates inside homes. Because the upstream signal path is a multipoint-to-point architecture (the exact opposite of the downstream signal path), any noise generated in a home is funneled upstream, through taps to nodes, getting amplified as it moves to the headend. This effect is called noise funneling. Noise funneling is bad.

The harsh conditions of the upstream path required a sturdy modulation type, relative to the QAM-styled modulation used to carry signals downstream, towards homes; QPSK was an early workhorse. Using a lower-order modulation, like QPSK, is not unlike slowing down when driving on a road with deep potholes: It's the only way to get to the destination, without gaining any unplanned "adventure badges" on your vehicle.

Over time, as fiber reached deeper into neighborhoods, which shortened amplifier cascades, it became possible to move to higher and higher orders of modulation in the 5-42 MHz upstream: 16-QAM and 32-QAM and 64-QAM via DOCSIS 3.0 SC-QAM. Today using DOCSIS 3.1 OFDMA, up to 1024-QAM will be viable, especially in DAA systems. Use of 2048-QAM may also be achieved, and 4096-QAM is within the standard. These increasingly bandwidth efficient formats allow ever-increasing amounts of data to be carried from homes outwards, to the Internet or cloud.

3. Differences Between the Upstream and Downstream Signal Paths

There are a few notable differences between the forward/downstream signal path, and the reverse/upstream signal path. They are briefly noted here.

One is channel widths. Because the upstream path was never envisioned (or designed) to carry video, its channel widths aren't a static 6 MHz, as they are in the downstream (home-facing) path. Upstream channel widths typically use one of three sizes: 1.6 MHz, 3.2 MHz, and 6.4 MHz.

The upstream signal path was envisioned as a way to move small amounts of information, such as a click to order a movie, or, later, a click of a mouse to request a web page. When voice-over-IP entered the service mix, audio signals began moving upstream. All are negligible, relative to the "carrying size" of broadcast video. So, until recently (hello, webcams!), traffic type was also a differentiator between what moved upstream vs. downstream.

Modulation is a third difference between the downstream and upstream signal paths. The width differences are to accommodate multiple modulation rates for sending traffic: QPSK, 16-QAM, and 64-QAM.

A fourth difference – and an omnipresent conversation – is the matter of noise and ingress funneling in the upstream direction, which makes upstream more susceptible to over-the-air (OTA) signals. As the upstream bandwidth grows, some OTAs flip from downstream phenomenon to upstream, and in doing so become more troublesome. In the lower spectral regions, it used to be the off-air analog channels, which vacated the band coincident with digital. There's the FM band, which sits between 88-108 MHz. Potential issue: Interference. There's also the Aeronautical Mobile and Radio Navigation, between 108-137 MHz. Potential issue: Signal leakage. And let us not forget legacy out-of-band signaling, used by some set-tops and modems to move things like guide data, and command-and-control information.

Some readers may remember the big-growth days of high-definition TV, and the concerns about having enough downstream capacity to carry them all. Suddenly, we needed to add capacity, adjust channel lineups, advance another leap in video compression (at the time, to MPEG-4), and roll out things like Digital Terminal Adaptors, or, for some operators, Switched Digital Video.

These days, downstream capacity is reasonably under control (even as 8K TVs started rolling into retail this summer). It's keeping ahead of the growth in upstream demand that drives a larger part of our plant augmentations. As it turns out, after DOCSIS 3.1, the only viable tool in the non-fiber-deep playbook, besides constantly splitting nodes (which is increasingly inefficient) is to add spectrum.

Ironically or not, while this paper was being written (summer 2021), the author was participating in nightly overnight maintenance window sessions aimed at bringing to production the new "scar detector" tool on live nodes to activate on the Mid-split band. It was about as good as it gets for upstream geeks!

4. HFC Spectrum Relationships: Mid-Split, High Split and DOCSIS 4.0

The HFC network in North America has been limited to 42 MHz or less in the upstream direction since there has been HFC plant. The launch of HSD services increased the focus on the upstream because of the central role it now plays in providing a quality Internet experience for the ever-increasing range of demanding real-time applications. Fortunately, the growth of Internet traffic per year has been generally quite predictable, although slightly less so in the upstream than the downstream. In the upstream, year-on-year fluctuations historically experienced periods where traffic has been more dynamic, and periods where it has been flat. It all nets out to an average usage per year that is predictable enough to let capacity planners do their jobs effectively.

There has been over 20 years of growth managed almost exclusively by a fixed amount of upstream spectrum between 5-42 MHz. The amount and type of traffic moving upstream largely populates the quality spectrum available and managing new growth has transitioned from new QAM carriers and node splitting to node splitting and more node splitting. The COVID-19 pandemic accelerated this activity [2]. DOCSIS 3.1 can be used below 42 MHz, but better QAM bandwidth efficiency is no match for spectrum when it comes to adding capacity – according to that Shannon guy (http://www.inf.fu-berlin.de/lehre/WS01/19548-U/shannon.html). For high SNR (Signal-to-Noise Ratio) cable networks, Capacity ~ [BW/3] * SNR [dB]. The key thing to note is "dB." Capacity increases directly proportionally to bandwidth, and only logarithmically proportional to SNR.

As nodes get split smaller and smaller, it tends to become less efficient to continue to split. It is less likely to yield a 50/50 split, so the full benefit of the split is not realized. Whereas a 50/50 split buys ~3 years of growth at a CAGR of 25%, if the node is split 60/40 or 70/30, it is less. It is not unusual for one port of a node to be naturally more heavily loaded with traffic than another, since these ports feed different neighborhoods, and one, for example, may include a student housing complex or have a high density of business customers, while another may service less online-active customers.

Figure 1 shows a set of upstream expansion options available for MSOs. Many are active or imminent. They can be viewed as sequential in time, with some overlap and market-based criteria informing the timing and path to 10G. The architectures are described further below.

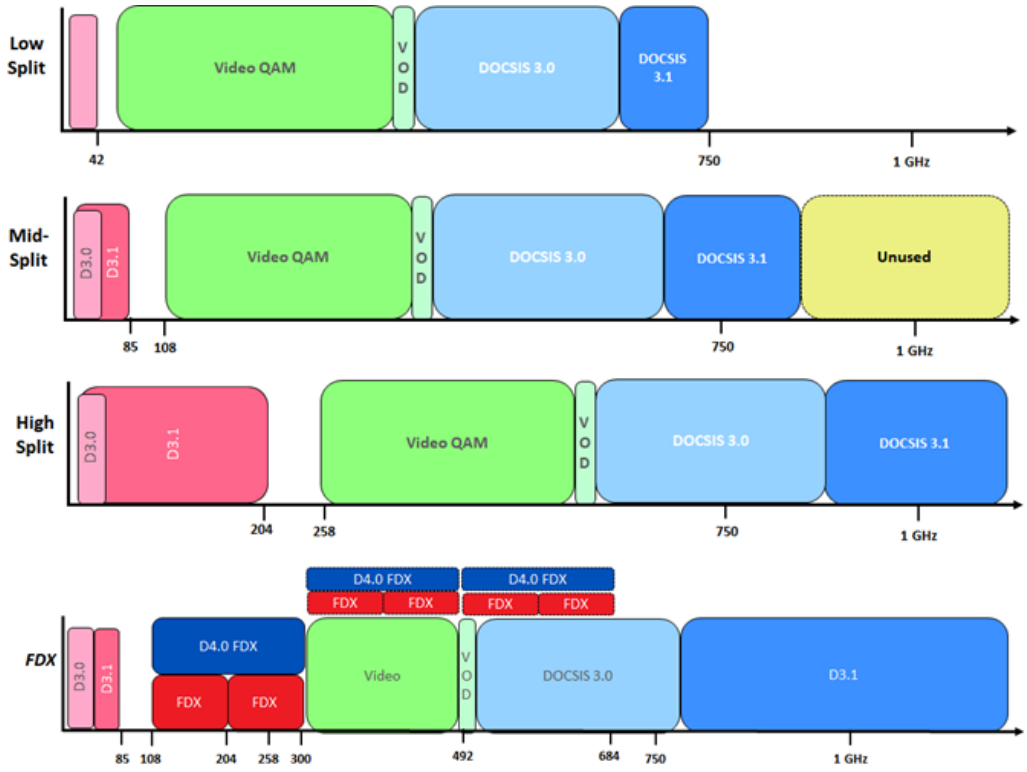


Figure 2 - Spectrum Migration Options through DOCSIS 4.0 (FDX only)

4.1. The Mid-Split: 5-85 MHz (defined initially in DOCSIS 3.0)

The Mid-Split is viewed as a practical steppingstone and with a relatively light touch because it resolves the upstream capacity challenge as we mathematically know it today. When combined with a node split, it defers additional augments to address congestion for at least 5 years, typically more (depending on D3.1 vs D3.0 assumptions). Furthermore, with an all-OFDMA channel, it can support around 600 Mbps. With a 4xSC-QAM DOCSIS 3.0 payload in the 5-42 MHz portion, about 450 Mbps is expected. Speeds up to 300 Mbps are expected in scale, under some traffic engineering guidelines tied to new utilization patterns.

An 85 MHz payload consistent with most MSO DOCSIS 3.0 usage today is 4x64-QAM DOCSIS 3.0 carriers, and a single OFDMA block from approximately 40 MHz to 85 MHz. This configuration is shown in **Figure 2.**

At this time (summer 2021), the Time and Frequency Division Multiplexing (TaFDM) feature, which allows spectrum to be DOCSIS 3.0 in some time slots and DOCSIS 3.1 in other time slots, has not been enabled. This remains an option, depending on the penetration mix of DOCSIS 3.0 and DOCSIS 3.1 modems and the net efficiency provided.

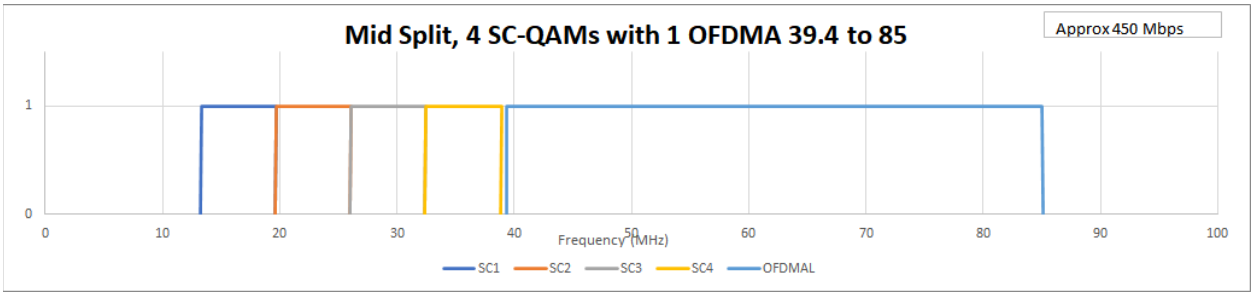


Figure 2 – Mid-Split DOCSIS 3.0 + DOCSIS 3.1 Loading Configuration

4.2. The High-Split: 5-204 MHz (efined initially in DOCSIS 3.1)

The High-Split is another popular option, as it stretches the speeds possible in the upstream to 1 Gbps or slightly more, as was demonstrated in the fall of 2020 [3]. Since the Mid-Split is such a powerful solution itself for capacity, going to a High-Split is a potentially very long-term solution with respect to capacity. However, practical capacity benefits are driven ultimately by the number of High Split-capable devices that can access that spectrum. There are many more Low-Split and Mid-Split modems deployed today compared to High Split, although this could change over time, and in particular for those that deploy with High Split.

4.3. DOCSIS 4.0

Like the High-Split, the primary value of DOCSIS 4.0, FDX or FDD, is upstream speeds. DOCSIS 4.0 fully attacks the historical asymmetry of downstream and upstream capacity, bringing multi-Gigabit symmetric capability to HFC. As it is defined today, the upstream will achieve 5-6 Gbps when fully activated. A path to 10 Gbps upstream is available by extending the bandwidth in FDX or FDD above today's 684 MHz limit with two more OFDMA blocks, to 1068 MHz. While this is easy to draw on a diagram, it creates challenges like upstream transmit power from a cable modem to overcome high coaxial losses.

Figure 3 summarizes the speeds associated with the options shown in **Figure 1**.

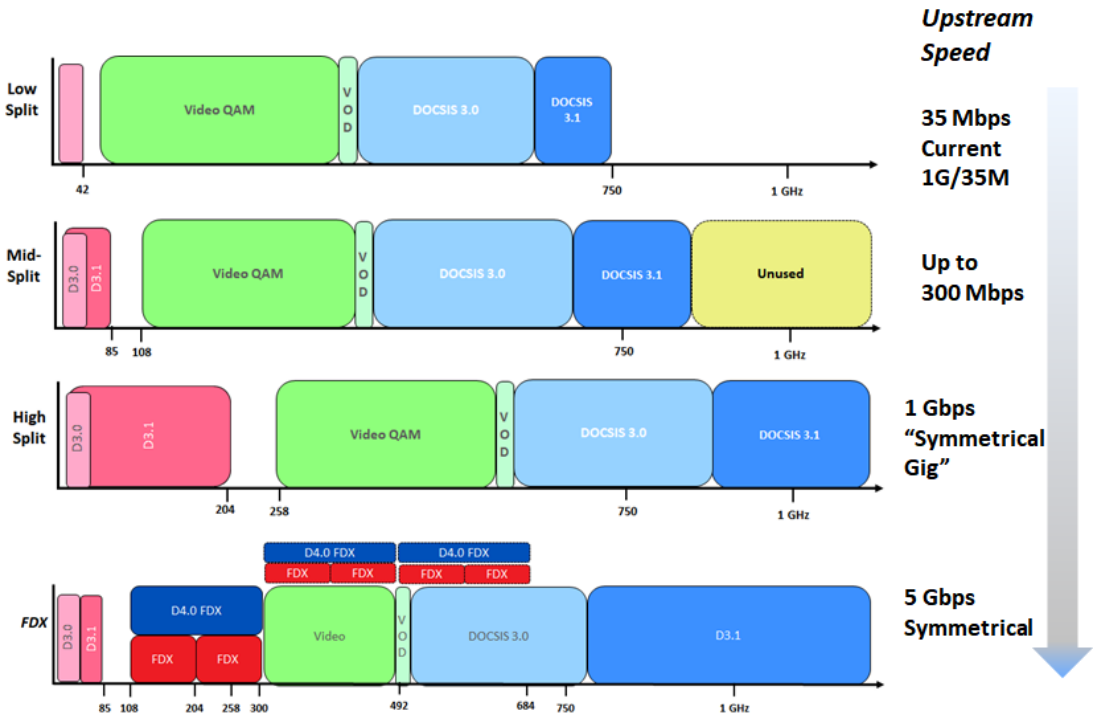


Figure 3 – Spectrum Migration and Implications to HSD Speed Tiers

We will get into the nitty-gritty details of the paper from this point on by mostly examining the Mid-Split scenario. Many of the same concepts are applicable to High Split, although there are some important differences that we will call out in the next section. There are deeper details, software and tool development, and mature processes that can be explained more readily using the Mid-Split case study due to its longevity, so we will lean on that for the bulk of the deep dives.

4.4. The Math

Mid-Split expansion takes the available upstream bandwidth from 37 MHz to a limit of 80 MHz. It was defined in DOCSIS 3.0, with the upper limit selected in part to fall just below the FM radio band in the US, while preserving the important downstream video out-of-band (OOB) signals widely used by legacy QAM set-top boxes (STBs). Per the earlier discussion, it is typically the upstream that drives network upgrade activity.

Because of the average per-user peak-busy-hour (pbh) upstream is still in the hundreds of kbps range, the upstream payload generally grows more slowly than downstream. Plus, because the new upstream spectrum is much cleaner, the Mid-Split impact on network lifespan is extremely powerful.

Figure 4 shows the time runway generated by three options – node split, node split plus upgrade to Mid-Split, and finally N+0 with Mid-Split. While N+0, with smaller service group size, offers the longest runway of the three, an N+x migration tied to a node split is also a very effective way to extend HFC lifespan to nearly 7 years in this analysis.

A key benefit of N+x with spectrum migration is its ability to add capacity quickly when compared to N+0. With the Covid-19 spike eliminating months of CAGR lifespan, N+x upgrades bring more US bandwidth to the network quickly to reset the lifespan timeline. The naturally slower pace of deeper fiber construction

will leave too many areas without an augment for too long of a period of time. With the capacity growth "time" erased due to the pandemic, alternatives such as drop-in HFC upgrades that are both fast and effective make a sensible augmentation step. Having a diverse strategy, not one-size-fits-all, adds important flexibility to deal effectively with adjustments for situations like Covid-19.

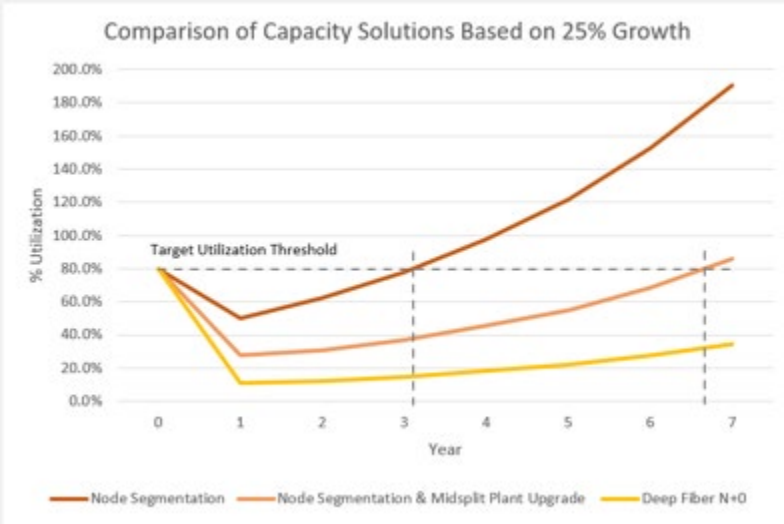


Figure 4 – Upstream Lifespan Expansion Options [2]

Lastly, looking ahead to future capacity and speed demand, and coupled with the objective to push fiber deeper into the network whenever possible, adjustments are being made to the architecture where it makes sense. For example, adding fiber in an underground network without the benefit of conduit is an inherently slow process. However, by providing the flexibility to allow a strategically placed amplifier (e.g., to allow an N+1 network) or two, there will be less construction, increased node size, and decreasing cost per household passed (HHP.) Combined, all speed the pace of the network upgrade and deliver the added bandwidth to more HHPs/year.

4.5. How About Some Real Visuals, Larry?

Comcast and other MSOs, such as Shaw Communications, have been building and activating Mid-Split spectrum for about 5 years. As much as can be gleaned from the crisp PowerPoint visualizations that led up to this moment is not nearly as exciting as displaying the real thing! **Figure 5** shows an activated Mid-Split upstream of 4xSC-QAM and the rest OFDMA. Upstream traffic is bursty, so this spectrum tool, the Yeti upstream spectrum analyzer application, has been placed in Max Hold mode to show the full 85 MHz band over time being utilized by the CM traffic. Most of the traffic is still in the DOCSIS 3.0 SC-QAM – the "Heatmap" view would show this – because most of the CMs in the system are DOCSIS 3.0. This balance is changing rapidly but DOCSIS 3.0 CMs are still the majority.

Figure 5 – Activated Mid-Split of 4xSC-QAM + OFDMA

5. New Spectrum, New Challenges

It is a sizeable project to upgrade the access network to support a new spectrum split. While some of the equipment is hosted in Hubs and Headends, where it is more easily accessible and centralized, upgrading outside plant (OSP) is more challenging. It requires going into the field, to every active device which has a diplexer – which is to say, every active device – and upgrading it to the new split. In some cases, this is something that can be done by changing a plug-in filter inside of the housing (not a live housing, removed from network) but in most cases it is not this simple. Many amplifiers in the field are decades old, made by vendors who no long support the product line, requiring swapping of devices altogether. Regardless, many operators have made the decision that the time is now for a frequency split upgrade and are committed to executing it.

Most of the above applies directly to an upgrade to a High-Split when it comes to upgrading actives in the field. However, the nature of the upgrade to Mid-Split is a lighter touch, with respect to other important variables. These are important to understand as part of a decision criteria on the upstream plan.

The seemingly intuitive “if more spectrum for Mid-Split is good, then even more spectrum for High-Split must be better” runs up against some significant new complications, outlined below.

5.1. Legacy QAM STB OOB Carrier

QAM video STBs that do not have a DOCSIS STB Gateway (DSG) control channel utilize the SCTE 55-1 or SCTE 55-2 protocol to get the necessary information to the STB. Signals returned from the STB over the upstream path in the 8-12 MHz range, typically. For fellow upstream geeks – yes, that part of the spectrum is terrible, but the protocol calls for very simple, robust modulation that is inefficient, but the traffic requirements of this modem are very low by today's HSD standards.

In the downstream, however, the "out-of-band" (OOB) carrier must be in the band 70-130 MHz according to the standard. When the upstream stops at 85 MHz, there is plenty of spectrum to place the OOB signal. When the spectrum extends to 204 MHz, legacy QAM STBs are stranded unless they can receive this OOB channel some other way. There are creative ways to do that, however it is some version of one-off solution.

Some of the thinking at the time of the DOCSIS 3.1 standard was that legacy QAM STBs were on the decline and would largely be out of the network by the time the High-Split was deployed. In addition, there was a move towards all-IP video delivery, which is still true today - although with somewhat less urgency based on changing business dynamics of the 10 years since the specification was being developed. One such class of box (General Instrument / Motorola DCT2000) is a model old enough that its tuner is at a fixed frequency and will not tune up or down from this frequency, which is about 72 MHz. Because it cannot tune up, it is incompatible even with Mid-Split. As a result, where Mid-Split is deployed, a pre-requisite is to swap these STBs out of the network. Due to the age of these STBs, the number of these STBs are very small, and the burden thus relatively low.

This is one of the very important aspects of High-Split compared to Mid-Split. Eliminating the OOB can make it a more invasive procedure for customers, as the best operator option is to extract these non-DSG-capable STBs, and this is more likely to "leave scars." The other import aspect attributable to High-Split is the Neighbor Interference (NI) phenomenon, which we will discussed later in this paper.

5.2. Aeronautical Leakage Band

The Aeronautical band, 108-137 MHz, is one in which there are requirements on operators to ensure there is not egress above a certain amount that could interfere with over-the-air (OTA) users in that band. Operators have mature processes and equipment to monitor this, placing "leakage carriers" in and around this band (and others) to measure leakage systemically to ensure compliance. It acts inherently as plant hygiene, so there is substantial benefit to operators - because where there is egress, there is the possibility of ingress. Ingress, of course, has been haunting the upstream for many years, and especially as the DOCSIS HSD upstream has grown more critical.

In practice, tones are placed close to the 108-137 MHz band in the downstream spectrum line-up and measured by specialized equipment. Of course, this works fine for Mid-Split. Mid-Split ends at 85MHz and the leakage band begins above that. However, this band, for High-Split is now in the upstream. To the letter of the FCC requirements, the upstream transmit power of a specification-compliant DOCSIS 3.1 modem cannot exceed the FCC limit, even at maximum transmit power. However, rather than do away with this aspect of plant maintenance altogether and lose the value it brings to operators and regulators alike, leakage measurements in this band will continue. Techniques which can measure leakage coming upstream from the modem are required, however, which is much different than today. Since these transmissions are bursty and short, the probability of catching them using today's auto-pilot drive-by method is not sufficient. Instead, techniques that use probe signals sent from CMs that are scheduled, and such that burst detecting equipment – also new for these meters – can capture the burst and assess the

leakage performance is needed. It is a more complex and intricate solution than is needed on the downstream, but early proof-of-concepts have shown it viable.

5.3. Cable Modem Maximum Upstream Transmit Power

The maximum Total Composite Power (TCP) of a DOCSIS 3.1 cable modem is 65 dBmV. An "average" upstream transmitter in our footprint launches at about 43 dBmV/6.4 MHz. Extrapolating to 4x SC-QAM carriers, this becomes a TCP of 49 dBmV, still plenty of headroom to 65 dBmV. Extrapolating a uniform Power Spectral Density (PSD) over Mid-Split, this becomes roughly 53 dBmV using the Mid-Split configuration of Figure 2 – again, still plenty of headroom, but keeping in mind that 43 dBmV was the average.

The 90% point for upstream TCP is about 51 dBmV/6.4 MHz, meaning 90% of cable modems transmit 51 dBmV/6.4 MHz or lower. Extrapolated to Mid-Split, this is a TCP of about 61 dBmV. Our headroom is disappearing! Indeed, the 55 dBmV/6.4 MHz case, which would be the limit of the TCP that the modem can transmit, when extrapolated to Mid-Split, is about a 99% point on the cable modem upstream, Tx power cumulative distribution function (CDF). Of course, 1% represents a very small relative likelihood of running out of gas, but it is certainly not negligible for a large DOCSIS footprint when the population of DOCSIS 3.1 devices (all Comcast Mid-Split capable devices are DOCSIS 3.1) in the field is growing.

Now consider these extrapolations for the High-Split case:

Average US Tx: 43 dBmV → TCP (High-Split) = 58 dBmV

90% Point US Tx: 51 dBmV → TCP (High-Split) = 66 dBmV

This suggests that there could be a significant increase in the number of CMs that will be transmitting at their maximum and more, reaching the amplifier or node port at lower than designed levels, all else the same. This may impact network performance (lower MER) and throughput, in addition to creating challenges for operations and maintenance in aligning the network. It is unlikely this would be noticeable to a customer during normal use of Internet applications. But it could increase slightly the probability of a speed test failure for a 1 Gbps upstream service, which does not have a lot of capacity headroom above 1 Gbps.

5.4. FM Band

One of the benefits of the Mid-Split spectrum stopping at 85 MHz is that the FM radio band begins at 88 MHz. Indeed, by the time discussions about an expanded DOCSIS 3.0 spectrum were happening, the role of plant ingress and impact on the upstream was beginning to be felt and understood. FM radios broadcast from 88-108 MHz and can be very powerful signals when nearby transmitting antennas on major stations with the most powerful signals. It is expected that this band will suffer in terms of guaranteed MER across part of the network.

There is no better option for working effectively through an FM band with residual radio noise ingressing onto the cable than using DOCSIS 3.1 OFDMA. However, if FM radio signals create high interference, as reflected by a low MER, there is only so much that can be done by OFDMA. The 88-108 MHz span is a modest chunk of spectrum – about 12% of the newly added capacity for High-Split, so the impact also is expected to be modest.

As with the Total Composite Power (TCP) case, this scenario could also make it more difficult for the High-Split solution to achieve the 1 Gbps or greater target, given the small amount of headroom that exists.

6. Tool Time !

A suite of existing Comcast tools is essential to Mid-Split activation, which will become apparent as we describe the new tool development and tech ops processes supporting the initiative. This section is an introduction to the essential tools that we will reference along the way.

6.1. Premise Health Test (PHT)

Premise Health Test is a tool used to assist technicians in the diagnosis of any customer premise. PHT has evolved over nearly a decade, beginning as Home Integrity Check (HIC), starting with DOCSIS-only measurement values. It has been expanded to be more comprehensive, including many Proactive Network Maintenance (PNM), Wi-Fi, MoCA, EPON and other measurements. The test is usually invoked before and after installs and repairs to provide outlet-level readings from the installed equipment, where available. In addition to pass-or-fail, PHT also provides details about the service to facilitate the troubleshooting process. **Table 1** has a complete list of pass-or-fail criteria.

Table 1 – PHT Pass-or-Fail Criteria

Metric	Upper Fail	Lower Fail	Single Threshold Fail
Actual US TX	> 54 dBmV	< 25 dBmV	
Partial Bonding			Registration state <>4 Downstream state <>1
FLUX ICFR			>= 3 dB ICFR (ICFR Indicators)
DS RX	>13 dBmV	<-13 dBmV	
DS SNR			< 33 dB
SpectraCM Impairments*	* Full Spectrum Devices Only		Any Individual Impairment <>ACP
MoCA PHY Rate * MoCA Network *	* MoCA Capable Devices Only * MoCA Segmented Network Devices * MoCA Unexpected/Foreign Devices		< 200 Mbps (XG to XI, XB to Xi, XG to XB devices only; Xi to Xi or RNG150 to RNG150 not in scope)
FM Ingress	* All DOCSIS Devices		Severe Ingress Condition
EPON	>-8.0 dBm >-4.0 dBm	<-28.5 dBm Downstream <-28.0 dBm Upstream	Please Refer to Market RTM Process for any Failing Light Level Conditions
All Out			All Devices are Unresponsive
Wi-Fi – RSSI	* Re-Launch 10-29-19; None Pass/Fail		< -70 dBm RSSI Range at Xi5/Xi6

6.2. Yeti

The Yeti tool comes from the PNM suite, providing real-time upstream spectrum capture information to users. Historically, operators have relied on hardware-based spectrum analyzers to perform this function. Since the advent of PNM, spectrum capture capabilities are now available in the cable modem, from both downstream and upstream burst receivers. Upstream spectrum capture implemented within the burst receiver provides several distinct advantages over external, hardware-based solutions. First, it eliminates

the need for additional hardware, which typically occupies valuable headend space and requires facilities power and cooling. It also allows operators to take advantage of the powerful burst receiver demodulators and CMTS core scheduling information. For example, in-channel demodulator performance metrics can be displayed along with the spectrum capture traces, with thresholds and colorization to aid in human interpretation. Another powerful feature is the CMTS scheduler's "quiet time" mode, which captures spectrum traces when no modems are transmitting. This provides users with a "noise-only" view of the upstream spectrum, simplifying the troubleshooting process by removing the cable modem bursts. **Figure 6** shows an example of the Yeti display, including SC-QAM and OFDMA bursts, in-channel demodulator statistics and threshold-based colorization.

Figure 6 – Yeti Upstream Spectrum Capture Display with SC-QAM and OFDMA

6.3. SpectraCM

This downstream spectrum capture tool provides a cable modem-oriented view of the RF spectrum. In the PNM suite, it's referred to as Full Band Capture (FBC) and provides the modems with downstream receiver spectrum capture. It's especially useful when upgrading from Low-Split to Mid-Split operation because of the switchable diplex filters in the cable modems. With the diplex filter operating in Low-Split mode, the cable modem is very effective at capturing the noise environment from within the home, which, as mentioned earlier, is often the source of RF ingress. Then when upgrading to Mid-Split operation, the frequency spectrum up to 85 MHz changes direction, creating a notorious funnel effect, which complicates troubleshooting. Having the ability to switch between diplexer modes of operation allows operators to automate this traditionally difficult troubleshooting process. **Figure 7** illustrates an example of SpectraCM being used with Yeti to match and locate ingress of VHF television signal ingress.

Figure 7 – SpectraCM (bottom) Locates Ingress with Yeti (top)

7. Activate Spectrum So No One Will Notice

It is an axiom of good network strategy and upgrade practices to think of success like a baseball home plate umpire does: You've had a good game when nobody notices that you were there. This is absolutely the case for migrating spectrum, with the caveat that they will notice in a good way that you were there, as the products enabled by these network upgrades – in particular upstream HSD speeds – become available. Until then, though, the internal satisfaction that is flat trouble call metrics and meeting schedule and budget targets will have to do.

We describe below the potential issues to manage and share practices that help to achieve a seamless customer experience.

7.1. Activate Spectrum So No One Will Notice

With the decades of spectrum split in North America being set at 42 MHz/54 MHz, all QAM video STBs deployed are configured this way. They were built to receive video channels beginning at 54 MHz. When the network is NOT configured this way, and instead is upgraded to enable the diplex split to expand the upstream and activate new spectrum above 42 MHz, the QAM STB's point of view for Mid-Split or High-Split changes. This is shown in **Figure 8**.

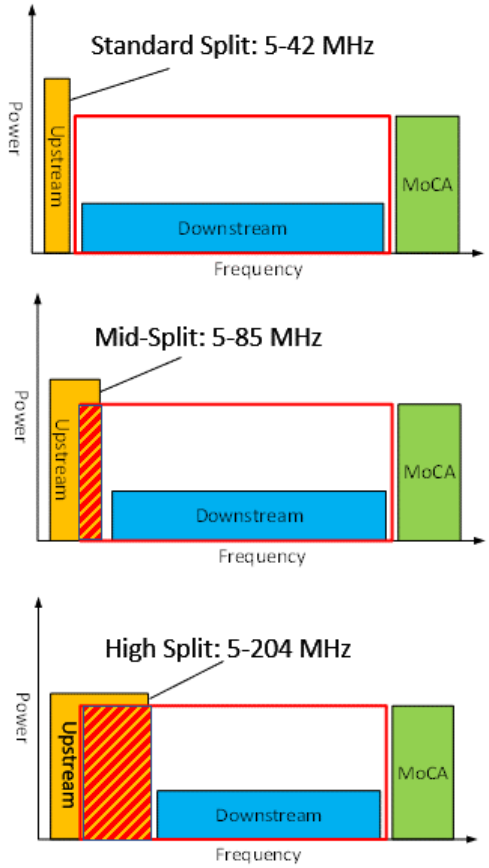


Figure 8 – New Spectrum Splits vs. Standard Deployed Equipment

The red cross-hatched areas in **Figure 8** represent the spectral overlap imposed on a QAM STB by a Mid-Split capable cable modem when utilizing that band. Any signal energy that appears above 54 MHz can be seen by the STB downstream receiver, because it is built expecting to operate on downstream signals that begin at 54 MHz. Unfortunately for the STB, in a home that also contains a Mid-Split capable cable modem (CM), the CM sees that band as "eligible" for placing carriers when the CMTS is configured to allow CMs to use it. In a CMTS that is properly load balancing, with much of the existing traffic volume generated by devices with a 42 MHz limit, and OFDMA turned on above 42 MHz, the Mid-Split capable devices would expect to be utilize the 40-85 MHz spectrum for transmissions.

Note that the RF processing front-end of the STB is not acting on any specific signal type – it is simply adapting its front-end gain to deliver the ideal level to the A/D converter using its Automatic Gain Control (AGC) function. AGC measures the total energy in the downstream band and doesn't care about its origin. It adds or subtracts gain to deliver the signal intact and at the optimized level – a balance between Signal-to-Quantization-Noise Ratio (SQNR) and clipping distortion – to the A/D converter, on its way to the digital processing. Thus, if new Mid-Split upstream energy on the STB receiver is very high, the STB receiver will add attenuation to keep the right operating point on the A/D converter. When this happens, the *desired* video channels will inadvertently be pushed into the noise floor through a phenomenon called "Adjacent Channel Interference" or ACI. If it attenuates too much, then the QAM video signals can be pushed down enough to cause low SNR in these channels, and video pixelization could ensue. Potential makeover scarring alert!

Note the above description is a "static" or time-fixed snapshot view of upstream energy and spectral overlap with signals moving downstream to a STB. Actual upstream traffic is called "bursty" because it bursts on and off. This is important because the AGC function has dynamic characteristics, but they tend to be slow (levels don't change very much, typically). As a result, the duty cycle (off/on ratio) and burst duration of upstream signals is a factor that can impact the AGC implementation of different STBs.

Note that, as represented in **Figure 8**, the nature of the levels is not favorable – the downstream receive level is low (DS Rx), while the upstream transmit level (US Tx) is high. Until now, there was a diplex filter to separate them, but now, between 54-85 MHz, this is no longer the case. As discussed, the US Tx average level is about 43 dBmV/6.4 MHz, with the vast majority at 51 dBmV/6.4 MHz or less. By contrast, downstream receive levels typically target 0 dBmV/6 MHz, but range widely by design from a minimum of about -12 dBmV/6 MHz to a maximum of +10 dBmV/6 MHz.

Fortunately, between a CM and a QAM STB there will be an RF splitter, the design of which will inherently isolate the port of a CM from the port of a STB by some amount. This scenario is illustrated in **Figure 9**, showing just one isolation path between modem (MTA) and a STB's RF inputs.

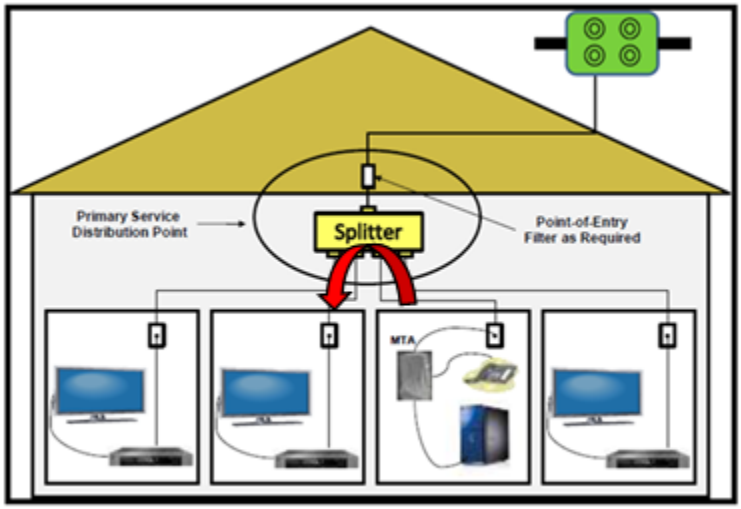


Figure 9 – Mid-Split Band Energy Isolation Path Across and RF Splitter

How much energy leaks through to the STB? That question depends directly on the splitter and home wiring shown in **Figure 9**. Home wiring – RG6 cable – has a very predictable dB/loss per foot and is easily modeled, from which some basic assumptions can be made about the range of run lengths in non-celebrity, which is to say "reasonably sized" homes. The most important factor with respect to the ACI phenomenon is the splitter(s) used to distribute RF to devices for video and data services. **Table 2** shows the specification for the isolation parameter (paragraph 13.0) for approved Comcast splitters.

The parameter used to evaluate the relative risk of video interference is called Carrier-to-Adjacent-Channel-Interference Ratio (CACIR). The threshold at which video degradation can be observed varies by STB model, and these have been individually characterized for every model still in use. Automated tools can discover the model type directly or through information in the billing systems and can apply a CACIR threshold according to the model type. For purposes of this paper, we will use the worst case empirically observed lab test value of -22 dB CACIR as the threshold value for deriving statistics. It seems like a 20 dB higher signal should completely blow up an RF front-end, right? However, note that the -22 dB is a

single carrier-to-single carrier comparison. There is, of course, more total bandwidth in the downstream than the upstream, so the total difference in power of the interfering energy to total downstream energy is less than this.

Testing was done as both "always on" and over a range of fixed duty cycles meant to emulate the "on/off" bursts of real upstream traffic. It is only when the upstream is bursting on, of course, that there are signals that can interfere with a downstream STB.

Table 2 – Port-to-Port Isolation of Comcast Approved RF Splitter

COMCAST

Field Equipment Specification: Mid-split Passive RF Drop Splitters

ID	Specification	Requirements				Test Method / Other References
12.0	Return Loss	**Minimum RF Port Return Loss, All Other Ports Terminated (MHz)**				ANSI/SCTE 144 2007
		5-1002	**1003-1218**	**1219-1675**		
		≥ 18 dB	≥ 15 dB	≥ 5 dB		
13.0	Isolation	**Minimum Output Port to Output Port Isolation (MHz)**				ANSI/SCTE 144 2007
		5-10	**11-85**	**86-1002**	**1003-1125**	
		≥ 25 dB	≥ 35 dB	≥ 25 dB	≥ 20 dB	
14.0	Maximum Output Port to Output Port Isolation 1,125 MHz-1,675 MHz	**Splitter Device**		**1126-1,675**		ANSI/SCTE 144 2007
		2-way		≤ 28 dB		
		3-way (unbalanced)		≤ 30 dB		
		4-way		≤ 30 dB		

Comcast specifies a minimum isolation through the Mid-Split band of 35 dB. That says that a lot of the upstream transmit power is going to be attenuated on the way to the STB. (Yay!) Is it enough? Usually. And, fortunately, when it is not, it is discoverable and easily remediated. For a scar-free upstream makeover, the key statement is "discoverable."

Because of the importance of quantifying the potential for ACI, an existing model of Comcast-approved splitters was measured for actual performance. Practical performance of these splitters is shown in **Figure 10**. Note that the band between 54-85 MHz is, in the case of this model, the "sweet spot" of RF isolation, with 45 dB being more characteristic of splitter performance. This is an extremely valuable 10 dB with respect to quantifying the potential to impact video.

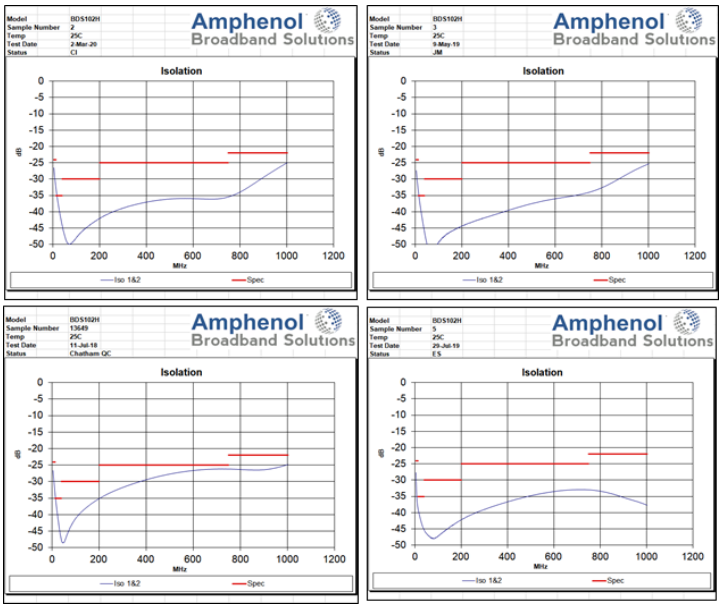


Figure 10 – Measured Port-to-Port Isolation of Comcast Approved RF Splitter

For every "sweet spot," there is of course a counter example. In the ACI scenario, the most concerning counter example, with respect to video service impact, is the use of a low-cost, off-the-shelf splitter that might be found in the cable TV accessories section of a home improvement store, such as Lowe's or Home Depot. (For old timers, this is where we used to say, "Radio Shack," but you may be hard pressed to find anything related to radio on the shelves there anymore ... if you can even find a storefront.)

Figure 11 shows the isolation performance of this type of splitter.

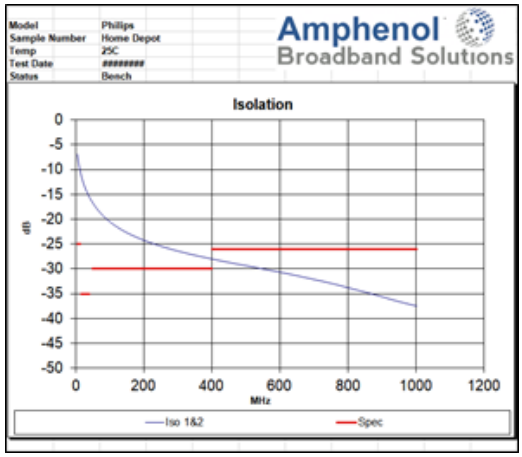


Figure 11 – Measured Port-to-Port Isolation of an Off-the-Shelf RF Splitter

We believe these retail splitter scenarios create the most likely risk to STB video degradation, as they perform with 15-20 dB worse isolation than the Comcast requirements. This in-home scenario is easy to envision occurring in practice and is generally out of the operator's control. Through the automated diagnosis of homes slated for a wider upstream, we can find such scenarios and take proactive measures prior to activating the Mid-Split spectrum.

Eliminating the possibility of video interference attributable to the activation of Mid-Split spectrum is a primary criterion for smooth spectral transition. Using empirical data of measured Comcast DS Rx and US Tx from production CMs, and typical home network assumptions for splitter and coaxial LAN runs, **Figure 12** shows the probability of ACI reaching the threshold for video interference to be 1.29% for the worst-case sensitivity among all STBs tested. This number looks small, and it is. However, when measured against the total number of broadband subscribers with video service, it is not negligible, and needs to be managed with proper tools and processes.

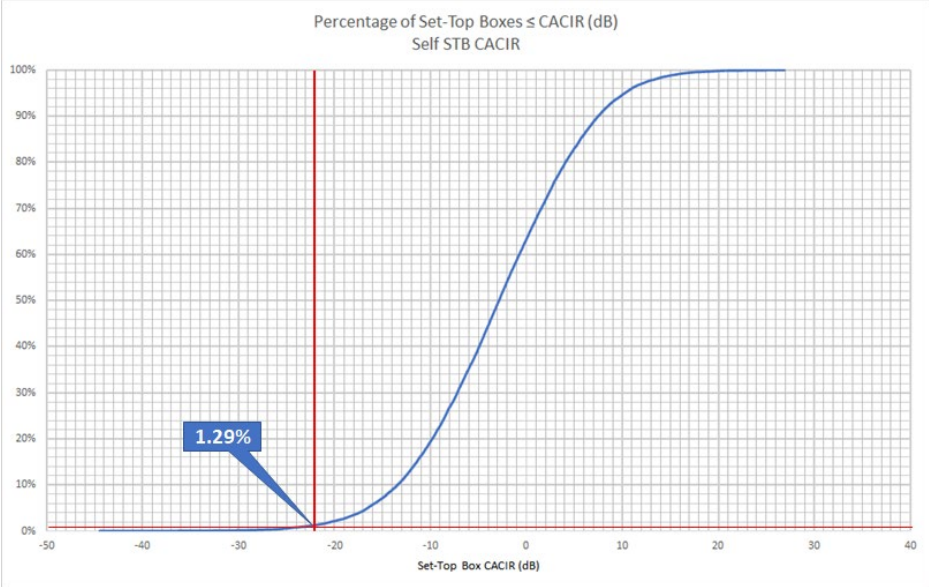


Figure 12 – Probability of Interference ≥ ACI Threshold of STB

One important note is that **Figure 12** shows the correlation of measured dB relationships with a lab-observed video impairment, under a set of fixed upstream transmission patterns. In real life, the upstream duty cycle is low, and the transmissions are relatively random in both size and duration. It is difficult to precisely correlate RF impairments measured in dB, to customer-impacting video degradation from real traffic, and further, to degrade it enough to generate a trouble call (versus the so-called silent sufferer – a worse scenario). While it is straightforward to create an impaired condition in the lab, how this translates to field exposure will be something that will be continually learned over the course of trials and the scaling-up of Mid-Split activation. That is the genuine way to test the hypothesis and lab measurements when encountering real traffic.

7.2. In-Home Amplifiers

In addition to OSP and traditional CPE devices that provide residential services that only know the 5-42 MHz Low-Split, many homes also use drop amplifiers to overcome losses across the in-home coaxial network. As you might expect, these are also built with a Low-Split diplexer.

These devices must come out…eventually…. but because they may or may not be customer-impacting, they do not *necessarily* have to be tackled coincident with the activation of Mid-Split spectrum. From a capacity perspective, every drop amplifier that can be removed so that it doesn't block Mid-Split energy from exiting the home is good for capacity. The operations perspective depends on the percentage of homes that include a drop amplifier – estimated at 15-20% but can cluster depending on geography and practices.

Methodically removing in-home drop amps over a period of time may make more sense than dealing with amplifiers transactionally, meaning only as part of a service call or product upgrade. A proactive plan to address drop amps will eliminate the perpetual limbo state that is mixed-mode devices working in mixed-mode spectrum.

With capacity and product in mind, we can itemize home amplifier management into two buckets:

1) Capacity-driven – Referring again to **Figure 4**, the coaxial lifespan when doing digital node splits and upgrading the network to Mid-Split is shown to be almost 7 years. This includes new capacity made available by Mid-Split (450 Mbps used), which depends on the DOCSIS 3.0 QAM bandwidth consumed and no considerations for TaFDM. Again, this only works if modems can access the Mid-Split bandwidth, which only happens if the CM is capable (new enough) of doing so. All DOCSIS 3.1 CMs we use are Mid-Split-capable, and our DOCSIS 3.0-only CMs are not. CMs are migrating to DOCSIS 3.1 status steadily, so over the 7-year period it is safe to assume the vast majority will be installed and capable of Mid-Split upstream connectivity. However, even if a Mid-Split-capable CM is present, any home that cannot allow the spectrum to pass out of the home is one that cancels the capacity gains. The "real" penetration of DOCSIS 3.1 OFDMA is decreased accordingly, and the 7-year lifespan is compromised. Thus, as mentioned, over time, these amplifiers must be removed, so that the capacity plan can deliver on its lifespan promise. How quickly this must be done is a mathematical analysis of utilization vs the "real" penetration trajectory.

2) Product-driven – One of the key benefits of the Mid-Split is the ability to deliver HSD speeds in the upstream such as 100 Mbps, 200 Mbps, 300 Mbps, even higher, as OFDMA begins to replace DOCSIS 3.0 QAMs in the upstream. Traffic engineering rules developed for these speeds account for utilization and total capacity and are considered reasonable expectations for potential product offerings. Once such products are made available, customers with home amplifiers will be (self)-blocked from receiving them. Interest in speeds that require Mid-Split would trigger immediate action, to remove the blocking amplifier. The challenge is how to manage this efficiently and, more importantly, in a way that is impact-free to the customer. The good news is these blocking devices (amplifiers or any filtering within the band that may have been installed inline) can be discovered remotely and in real-time. While the customer cannot get the new upstream speed immediately, a rapid and transaction-based process can serve to notify the customer that additional steps are required to support the speed upgrade. That we have detected the need for additional steps is a communications decision that is out-of-scope for the purposes of this paper. Either way, appointment scheduling can commence to eliminate the problem and get the new product speed to the customer.

   The product case is a different kind of operational task than the "capacity driven" case and is difficult to pre-plan because it is impossible to predict exactly which customers seeking the new upstream speed will also have in-home amplifiers. This adds to the long and growing list of reasons to proactively replace drop amplifiers during, for example, scheduled truck rolls to homes that also have been discovered as having a drop amp. Or, as mentioned above, build a program to recover in-home amplifiers proactively, over time, in a way that spreads out the cost of the effort, controls it better, and unifies the network for all homes -- rather than during product requests, which introduce a reactionary mode.

In a nutshell, the problem statement for Mid-Split activation caused by the 54-85 MHz spectrum overlap identified in **Figure 8**, is to develop a way to unobtrusively discover the state of a home with respect to these two criteria:

1) Potential for video interference
2) Ability to support DOCSIS upstream pass-through in the Mid-Split band

To enable this home-by-home assessment in scale, we developed an automated in-Home Assessment Test – aka iHAT – to enable a seamless migration of capable CMs to utilize Mid-Split when conditions 1 and 2 above are satisfied, as shown in **Figure 13**.

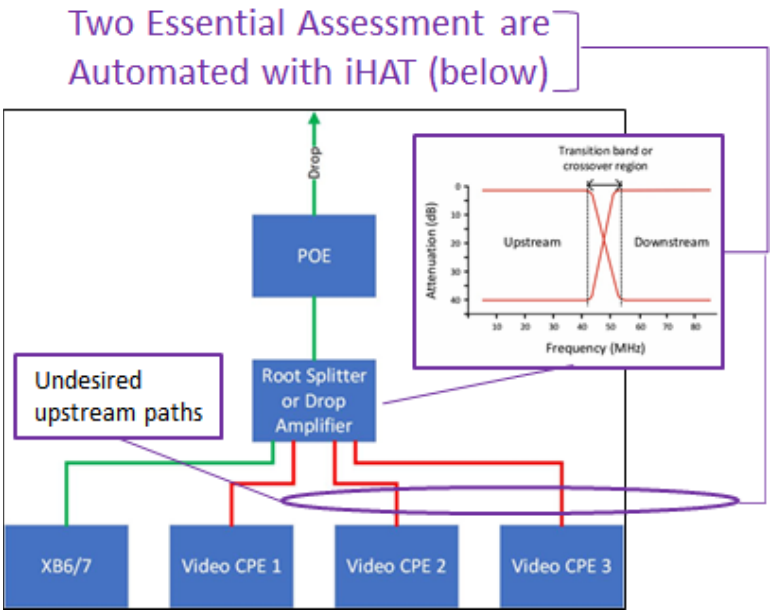


Figure 13 – The Two Basic RF Assessments Evaluated by iHAT

7.3. What's Under the HAT?

The "Black Box" view of iHAT that includes its functional core is shown in **Figure 14**.

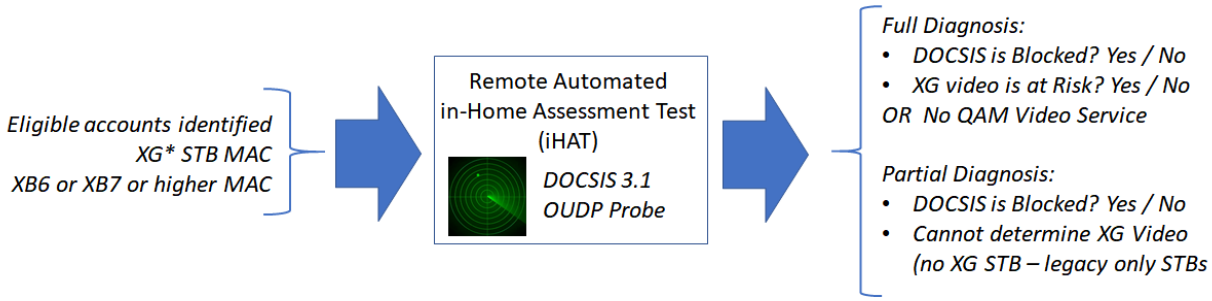


Figure 14 – The iHAT "Black Box" – Method and I/O

7.3.1. Incoming !

As shown in **Figure 14**, iHAT retrieves a list of devices, by account, on a particular Mid-Split-enabled RemotePHY Device (RPD) node, after the node is cutover, activated, and services restored that meet a pre-cutover state of performance (we are activating Mid-Split with OFDMA-only on DAA platforms). When an account is identified as having a Mid-Split-capable CM – for us, this includes the DOCSIS3.1 gateway family of XB6, XB7, and XB8 – it is deemed eligible for an iHAT test. With one of those devices present, it will be possible to place the CM in Mid-Split mode to determine whether its upstream transmissions in the Mid-Split band are able to be seen and received by a Mid-Split enabled vCMTS and DAA node, or if they are blocked.

When an account also includes the "XG" class of QAM STB, the iHAT evaluation will look both for DOCSIS Mid-Split pass-through and the potential for video interference. This XG family, the majority of QAM STBs in the Comcast network, supports the proactive network maintenance (PNM) and SpectraCM functionality needed to capture RF measurements that are the basis for iHAT scoring of video interference potential. Older QAM STBs do not support this capability. In a home that includes an XG class STB, that measurement taken is a reasonable proxy for the expectation for other non-XG STBs with respect to their isolation from Mid-Split spectrum energy.

If there is no XG-class STB present at all, but "legacy" QAM STBs are present, then no iHAT assessment can be made with respect to the potential for video degradation. At the outset, these homes will default to Mid-Split activation as data is accumulated. After some scale of statistical significance is built up, the policy will be revisited to determine if a course correction is needed. Also, as we observed in **Figure 12**, the risk of video impairment is very small. By NOT defaulting to activating in these homes, the alternative being committed to is to roll a truck to each home that only has QAM STB and take "iHAT" style isolation measurements manually when only a very small fraction may be impacted.

7.3.2. Start Your Engines

The method iHAT uses to make its determination is based on the DOCSIS 3.1 OFDMA Upstream Data Profile (OUDP) feature, which allows a pre-defined "probe" signal to be scheduled by a CMTS and generated as a test signal. The probe can be defined by center frequency, bandwidth, and duration. When iHAT runs, it schedules this probe signal, home by home, to be burst into a portion of the Mid-Split spectrum.

Figure 15 shows the probe signal centered at about 80 MHz. It is 1.6 MHz wide (a common reference bandwidth for OFDMA bandwidth used in the DOCSIS 3.1 requirements), has a PSD at the ranged OFDMA power, and lasts 3-5 seconds. These are empirically-derived values through trial-and-error testing and optimization in the lab.

Figure 15 – Probe Signal Used in iHAT via DOCSIS 3.1 OUDP Feature

When the probe is fired, the time stamp is used to instruct the XG STB when to execute a Full Band Capture (FBC), and with that capture, samples are returned to that include levels of the OUDP probe and the first few downstream QAM channels. By determining the relative levels of these components and comparing them to an interference threshold value, making offset adjustments that account for the test probe not occupying the Mid-Split band completely, the home can be classified as to whether it needs remediation.

The OUDP method provides three major advantages:

1) It is part of the DOCSIS 3.1 specification, so a required featured to be compliant to the specification (when asked for!)
2) It can be a scheduled event within a system's normal operation, and therefore is very non-intrusive, happening without a customer's awareness or service interruption
3) As a scaled down (in total power) representation of an actual upstream signal, it does not actually create enough interference to impact video. Instead, it emulates what a small portion of the filled spectrum would look like and extrapolates mathematically to draw the proper pass/fail conclusion.

Iterative optimization of the parameters yielded a repeatable, reliable result that correlates well as a mathematical extrapolation with the video threshold testing that forms the foundation of ACI analysis.

The probe signal can also be used to evaluate blocking of the Mid-Split upstream by a drop amplifier, because if this is so, the CMTS will not be able to observe the probe. However, as part of the iHAT test, Mid-Split becomes active on a modem prior to an OUDP probe being launched, once the CMTS has a configuration that supports it. Ranging information of the OFDMA band (DOCSIS 3.1 ranging) is available to determine if the upstream was successfully sounded. If not, this is typically sufficient cause to identify a home with a drop amp issue, at which point the CM can remain in partial service or reverted to Low Split mode. In either case, the home state is logged as "remediation required."

In future iterations of iHAT, a time-out on partial services re-tries will be used to force an auto-revert of the modem into the Low Split band. However, it is anticipated that, rather than do this with filter switching in and out, the vCMTS will support multiple bonding group (BG) operations that include both a 4-channel BG of all-DOCSIS 3.0 QAMs, and a 5-channel BG that is the former plus one OFDMA block of 40-85 MHz. This will simplify iHAT testing and shorten time consumed by avoiding modem reboots that force a diplex filter switch to Mid-Split, in order to execute the test. Instead, CMs will arrive on the scene in Mid-Split mode, by default, and if conditions such as drop amps block Mid-Split signal passage, then the lower 4-Channel BG will be deployed on that modem. In the case of a product need for that device (100Mbps upstream, for example), there will, of course, still need to be a rapidly-executed action scheduled for a good customer experience, to eliminate the blocking amplifier or filter, and to provide the service speed requested.

Figure 16 shows two sample outputs from iHAT. In these screen captures, the test was launched locally. To support deployment in volume production, these tests will run from the cloud.

In the top screen capture in **Figure 16**, we see the measurements being made by the iHAT tool. Notice the 2nd to the last value, which is what is compared to the 22 dB threshold "US OUDP Power minus DS Rx Power." This is well below the threshold and thus this measurement is a "pass." These values are stored for trend analysis and optimization. Another one of the values of particular importance for this is the last row, "Isolation." With iHAT, we now have the game-changing tool of being able to see the RF isolation between a gateway and XG STB in every home. Note also that 35.84 dB is very close to the Comcast isolation spec observed in Table 4.

In the bottom capture of **Figure 16**, we see the explicit result: "video interference was above threshold for at least one set top box," and also the isolation value, in this case called out by its variable name in the actual code as "dbCDelta" of 24.16 dB – above threshold.

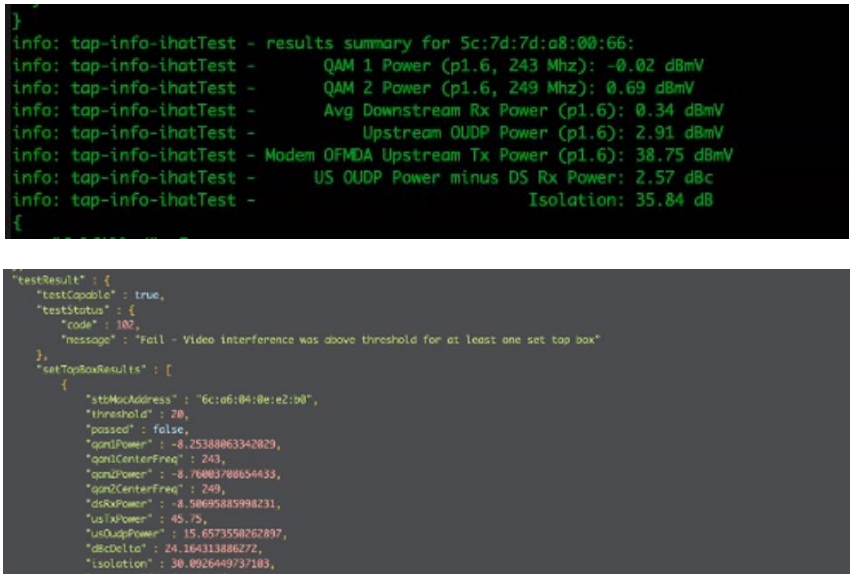


Figure 16 – iHAT Screen Captures of Video Assessment: Pass (Top), Fail (Bottom)

7.3.3. The Answer is…..

As shown in **Figure 14**, the output of iHAT is straightforward:

- DOCSIS Mid-Split pass-through (pass/fail)
- Potential for video degradation without intervention (pass/fail)
- Partial diagnosis – DOCSIS Mid-Split compatibility only; this is the case of a customer that has video services but no XG STB to support the telemetry needed to run the video assessment of iHAT

For a home to be declared ready to be activated, both DOCSIS and video tests must pass. If either does not pass, the home is left in Low-Split mode, and the home is dispositioned for remediation with the associated reason code (DOCSIS or video). How to optimally process the remediation queue itself is a discussion among many stakeholders. In addition to the pass/fail "answer" at the heart of iHAT, the RF measurements taken in the home, such as the isolation measured from the CM to the STB, are recorded and stored for purposes of trend analysis and iHAT optimization.

A fourth "state" that iHAT technically discovers is that the home is simply "ineligible" for Mid-Split because it has a CM which is only capable of Low-Split. In this case, the RF test engine does not run at all. This discovery occurs on the front-end, during the filtering of accounts connected to the RPD to only those that include a Mid-Split-capable CM.

Note that for a fully automated solution, iHAT receives input account/device data and develops output conclusions and an accompanying set of numerical parameters associated with the result for use elsewhere. In this sense, iHAT is the test function, with appropriate interfaces into other key back-office tools and subsystems, to operationalize the completely automated solution into the end-to-end ecosystem. In this context, iHAT is the "engine" of the overarching *Mid-Split Spectrum Upstream Launch* (MUSL) method, which we shall discuss next.

8. The Muscular Frame Supporting the iHAT Engine

8.1. MUSL-Up: End-to-End Device Activation Overview

As described above, the innovative iHAT tool provides a relatively non-intrusive view into a customer's home, and, on a home-by-home basis, will make a Go-No Go declaration with respect to readiness for activating spectrum in the Mid-Split band. iHAT scores a home's DOCSIS readiness for passing spectrum to 85 MHz, and its likelihood of creating video interference.

Referring to **Figure 17**, we can show how iHAT fits within the broader operational perspective, going from the trigger of a DAA Mid-Split node cutover on the far left, to the completion of activation on the far right. Note that "Tier 1" and "Tier 2" refer to different categories of markets which inform the upgrade strategy used at Comcast, however they have no real bearing on the flow otherwise.

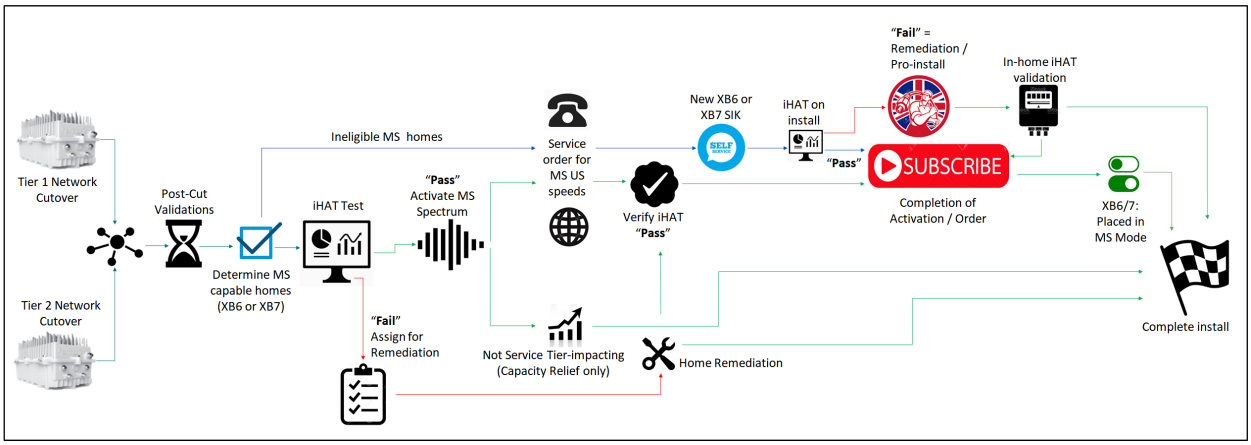


Figure 17 – Simplified Mid-Split Activation Flow: Cutover Through Activation

Beginning on the left, when a Mid-Split network upgrade occurs, internal tools will notify systems when construction is complete and officially closed out. This triggers the spectrum activation process, notifying other tools that the network is now be able to take this step. Two things must happen prior to letting iHAT sweep across the node and validate homes where Mid-Split can be turned on. They are:

1) *Post-Cut validation* – Ensures that the network has resumed to BAU metrics after the cutover. It is not uncommon to have a short period of elevated network activity shortly after a cutover, and it is desirable to resolve any residual cutover issues prior to moving to Mid-Split. This can be time-based, or it can be directly associated with, for example, observation of trouble TC metrics, pre-cut vs post-cut.

2) *Determine which homes are eligible for activation* – This boils down to whether the DOCSIS CPE is capable of Mid-Split. At Comcast, all DOCSIS 3.1 Gateways are Mid-Split-capable.

On item 2) above, if a home is ineligible, then iHAT does not run. Following this arrow to the top path in Figure 2, there is no immediate required step to get that customer a Mid-Split-capable modem. There is an effective loss of capacity for every CM that cannot access the DOCSIS 3.1 spectrum, because it forces utilization in the Low-Split band, rather than accessing the faster and wider OFDMA spectrum.

There is guidance in the in the field on what triggers a DOCSIS 3.1 upgrade for a customer – a particular speed tier for example. Over time, DOCSIS 3.0 CMs will organically disappear from the field, and it is likely at some point there will need to be a proactive effort to remove the older CMs in the network to maximize the DOCSIS 3.1 capacity.

Now, as shown in **Figure 17**, when a customer decides to upgrade their speed tier to one that requires Mid-Split, then getting them a gateway capable of that becomes a priority. Also, this customer's home needs to be evaluated for its ability to be placed in Mid-Split mode. So, as a new Mid-Split-capable gateway is brought onboard, one of the first things it needs to do is call upon iHAT to determine the state of the home for Mid-Split. If the iHAT "pass" is recorded, then the activation process continues, and iHAT sets the device into Mid-Split mode. It then becomes capable of using the OFDMA spectrum between 40-85 MHz. If iHAT records a "fail," then the customer is notified that a technician must come to the home to complete their install, and that their new speed tier will not be available until this "Pro Install" step happens. When the remediation is complete, the technician will validate onsite with iHAT, in this case triggered locally from the Performance Health Test (PHT) application.

If the eligibility conditions are in place – Mid-Split capable CM, and a STB model with the necessary telemetry capability – we move to the right of the blue checkmark of **Figure 17**: "iHAT Test." Let's now follow the lower path under the iHAT test icon – "iHAT fail."

8.1.1. The Remediation Queue

As noted, unless there is a speed upgrade required by a customer, there is not necessarily an immediate need to provide them with a Mid-Split capable gateway. However, it is still important that the iHAT score be logged. The fact that the home needs to be remediated is documented and populated into tools used by agents and technicians. Homes in this category are placed into a "Remediation Queue." iHAT will identify the specific failure mode, so that technicians know what needs to be done when they arrive. In general, remediation tasks are well-understood and known to technicians, and include changing out home amplifiers for alternative devices, checking home splitter configuration, models, and wiring, to bring the home to Comcast compliance standards. After remediation is performed, the iHAT test is run to validate readiness for Mid-Split spectrum, and the activation then completed.

When to schedule a home for remediation, assuming there is no speed tier motivation, is a business decision. They can remain in Low Split mode until that time, with some impacts on the network side. There are multiple variables to consider that have to do with capacity, efficiency, and proactive expense. Ultimately, however, all homes in the remediation queue will need to get serviced to extract the full DOCSIS 3.1 capacity and maximize the upstream runway these architectures are made to deliver.

Also note that a customer's iHAT "score" is not necessarily static. This is a very important point – the home has never been static, but now there is quantifiable information that can be used and leveraged to account for this, to improve the customer experience. Changes to the coaxial network in the home made by the customer, or new CPE brought into the home, can both affect the iHAT score. These events are "on demand triggers" that will call on iHAT to run off-cycle even after the initial iHAT sweep of the node at cutover.

8.1.2. iHAT Pass

The most straightforward flow in **Figure 17** is down the center, left to right. Both branches are logical and easily understood. An iHAT "Pass" means that the DOCSIS signals up to 85 MHz are able to be received by the vCMTS receiver, indicating that there is no home amplifier or filter blocking this transmission. AND it means that the home has been checked for RF isolation between the CM and the STB and determined not to be of concern.

Going to the lower green flow down the center of **Figure 17**, this is the case where there is no speed upgrade involved. The spectrum is being turned on to maximize efficient use of upstream capacity. The Tier 1 and Tier 2 plans are counting on use of this capacity to defer any future network augmentation by at least 5 years. So, while it may not be noticeably service-impacting to a customer, it is network- impacting. It may be indirectly service-impacting by lowering the congestion on that node overall (a good thing) and providing lower utilization spectrum for that customer to take advantage of for their current services.

The upper green flow is the case when a speed tier upgrade request is made, and there is already a Mid-Split-capable device present. Because a home's iHAT score is not static, a new iHAT score is taken prior to upgrading the customer. The customer expectations for the new service will be higher, and the awareness will be acute to service-impacting issues, so it is prudent to be certain that the home is still in a "ready" condition. In addition, because the customer now has, for example, a speed tier of 200 Mbps, they will have bursts of energy more likely to utilize a wide chunk of the Mid-split band at once, a condition that more aggressively exposes the STB to energy that can cause video degradation. If this "updated" iHAT result is still "pass," then activation is completed. If not (this is not shown), this home reverts to a

Remediation state, and because of the desire for a new service tier, it is a Remediation Queue with a higher priority.

8.2. MUSL-Building Logic

As powerful as the iHAT engine is, to use it in a production flow such as **Figure 17**, and, as importantly, seamlessly in production scale, it cannot be done on a home-by-home basis via human interaction. The information iHAT needs to run and the information needed by other systems to act on the iHAT outcome must be automated, and the interfaces to these other functions built for production scale. A logical flow diagram for the overarching MUSL ecosystem is shown in **Figure 18**. As shown, within the MUSL framework, like its role in **Figure 17**, iHAT is the engine. **Figure 18** speaks more to the software logic and definition of the adjacent subsystem interfaces that are implied by the flow of **Figure 17**.

The interfaces for iHAT for use in production are highlighted in the red box at the bottom of **Figure 18** and briefly described below. These represent interfaces for MUSL to distribute this important information to stakeholders, for the end-to-end operational success of Mid-Split activation.

Customer Accounts – Serviceability: When there are new upstream speeds that only a Mid-Split upstream can provide, it is important that the systems to upgrade a customer, whether online or through a service call, recognize the home's readiness state, as identified by iHAT. Alternatively, these tools can trigger an instant iHAT test for an updated result.

Biller – new CPE: When a customer changes CPE, possible iHAT variables that are affected are the device DOCSIS capabilities, the sensitivity to interference of a new video CPE, and the possibility of a wiring change in the home. It is prudent, given these potential changes to the iHAT state, to test (or re-test) the home.

XOC – Job Scheduler: When a home "fails" iHAT, it goes into a remediation queue, with a flag for what needs to be remediated (video or DOCSIS). For a Tech Ops plan based on proactive remediations, occurring routinely and not waiting for a house call to take care of iHAT-known issues, iHAT can report its findings per account to the local XOC tools that queue, prioritize, and schedule jobs.

Sales – Serviceability: Similar to Customer accounts, sales representatives should be able to quickly assess whether a customer is eligible for Mid-Split speeds by accessing iHAT status in existing sales tools.

Care – iHAT status, ITG Updates: When a care agent takes a customer call, after some amount of Interactive Troubleshooting Guide (ITG)-led questioning, the possibility of the issue being Mid-Split-related will be considered. A check on the iHAT status of that home, or an instantaneous iHAT test, can help the triage process.

Tech Tools – Tech360: Similar to care agents, when technicians are enroute or onsite at a customer home, part of the awareness they can have is the home readiness state, as determined by iHAT. More deeply in the tools, the sequence of steps to diagnose and fix a MS-related issue should also be available.

Inventory Management: As remediations are made at relatively large scale to remove old drop amps, procurement awareness to the deployment of alternative solutions can ensure that the supply pipeline is tracked and nurtured. This is even more important for proactive amplifier replacement plans, to ensure supply alignment with the plan for the alternative solutions – passive or active.

Data Sciences: As iHAT data is accumulated, new information about the home RF environment – isolation performance, trends over time, and correlations across neighborhoods – can be stored and processed for future optimizations, and to inform future process implications and costs.

Note that the MUSL flow leading into iHAT is from the perspective of a cutover to a Mid-Split-capable DAA node and network, from which flows notifications to activate the spectrum. This includes the instantiation of iHAT, to figure out who can use it at the RPD level. Once this cycle is complete and the Mid-Split node is in operation using the extra spectrum, several reasons were identified to check home readiness status via iHAT on an individual account basis. Noted at the bottom of **Figure 18,** these are referred to as "Asynchronous iHAT Triggers." They are options that can become part of new Mid-Split operations and maintenance practice, and include:

- New CPE device: XG STB or XB HSD gateway triggers new state-of-home update
- Buy-flow for new product offering that requires Mid-Split: Serviceability tools trigger up-to-date home state
- Premise Health Test (PHT): Tech in the home can trigger from available tools to assess state-of-home locally, as well as to assure that remediation work is completed properly
- Care (E360 tool) – Agent ability to see home state in real-time triage and possibly have access to reverting to Low Split and scheduling remediation
- (New) Home Metrics: Indications that imply a high likelihood that the equipment in the home has changed location or wiring has been changed, such as a persistent DS level change

Lastly, note the hourglass in the middle-right of **Figure 18**. As Mid-Split rolls out in scale, a determination will be made on whether a periodic update of all homes is warranted, and if so, how often. It will be based on empirical data that will reflect findings of just how dynamic the home environment is, and whether it is enough to warrant sweeping all devices on the RPD periodically, or spot checking if the asynchronous triggers do not provide enough off-cycle visibility. The tool will allow for periodic revisiting of each RPD. This eventually could place a lot of additional compute and overhead traffic on the network, storage (and associated cost), and state management resources. For an issue of small enough scale, this might not be warranted – experience and scale will tell.

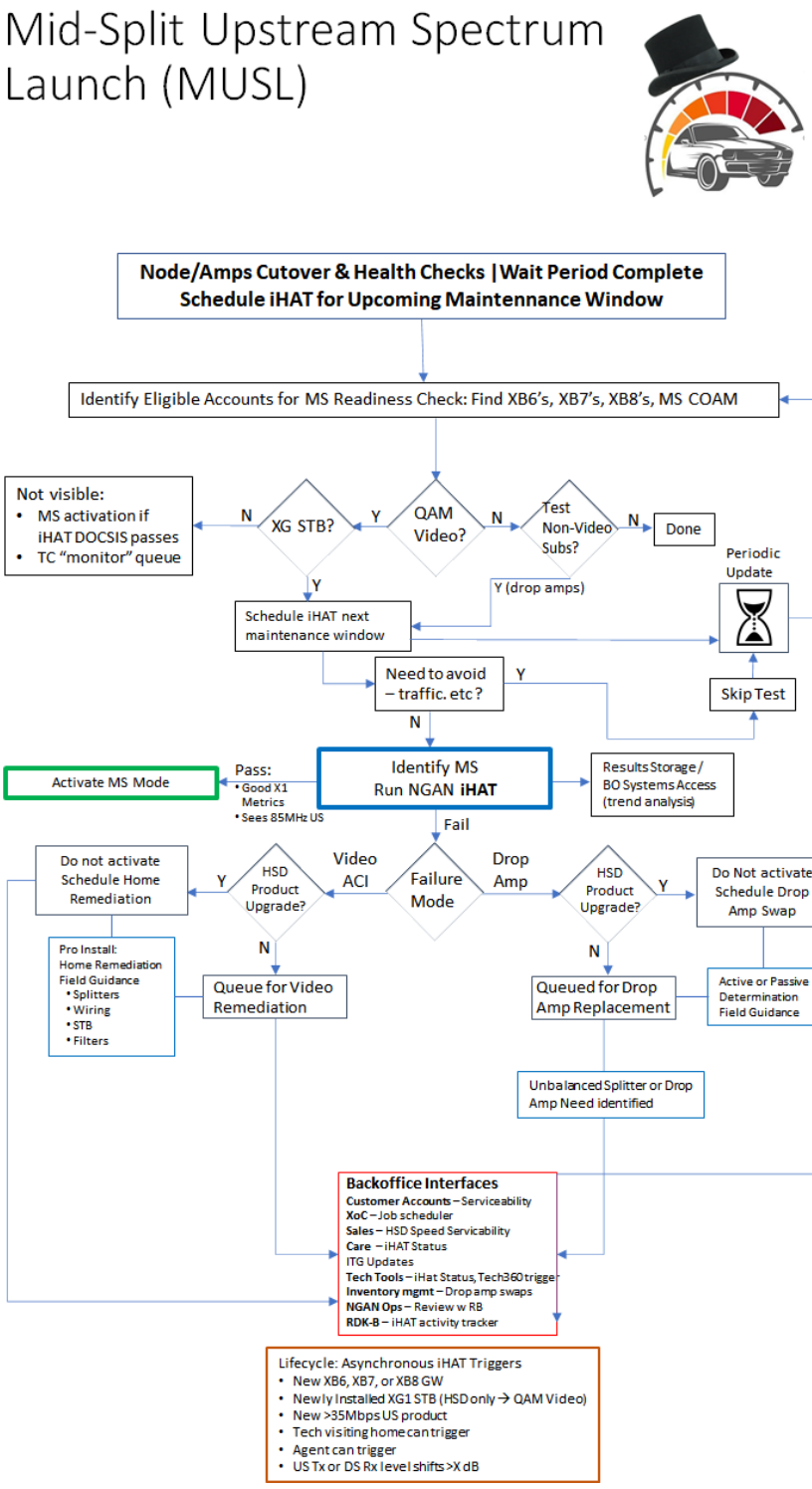


Figure 18 – iHAT as the Engine of the MUSL Framework

8.3. The Care and Feeding of Mid-Split

The iHAT tool described above prepares nodes and homes for the launch of Mid-Split. However, because the home environment can change, and is not under control of the operator, a particular state declared by iHAT is impermanent. New devices can and do come into the home, and customers do change wiring and add in-home passives and actives. Service changes, in particular a broadband speed upgrade for the upstream, increases the likelihood that video service interference could be observed, and exposes any blocking drop amplifiers.

In addition, the video interference potential is statistical in nature: There is an inherent (but small / <1.5%) probability that conditions in the end-to-end system, including OSP and home, shift in such a way that the threshold of interference for visual impairment observed in lab testing is breached.

Finally, the software tool itself will take time to mature, scale, and optimize, and should not be expected to operate perfectly to every potential negative use case and error condition it could encounter at scale.

Because there is a finite probability that a customer could experience video or HSD issues with newly activated OFDMA spectrum, the introduction of Mid-Split spectrum could lead to new inbound call types. As a result, there will need to be process updates such as for ITGs and Line-of-Questioning (LoQ) scripts to diagnose whether Mid-Split is the cause of these issues in the home.

By contrast, a DOCSIS failure is not a statistical phenomenon – either the upstream signal path is blocked by a drop amp, or it is not. As such, there is not very much nuance required around Care processes. The most intricate part of the practice of remediating a DOCSIS failure is two-fold:

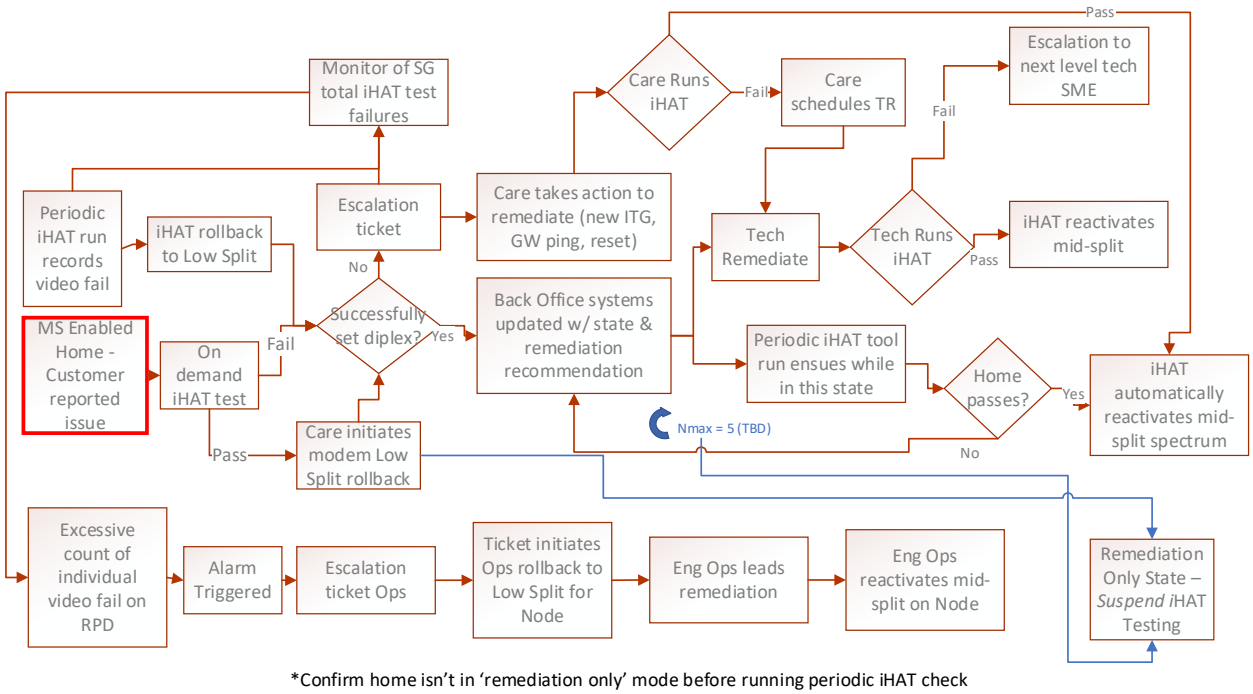
1) A home drop amp is replaced with what?
   The knee-jerk answer is another drop amp that supports the expanded frequency split. However, this places a new frequency barrier in place that is likely to be an obstacle in the future, such as for 10G FDX. Best practices and training are being built around a methodology that prioritizes a passive termination at the point of entry, if is not a DOCSIS termination itself (such as for DOCSIS 4.0, in an all-IP home configuration). If a typical splitter implementation is inadequate, a specialized unbalanced splitter may be possible to assure healthy levels at each.

2) Proactive Remediation
   Until HSD products are launched that require the Mid-Split spectrum (upstream speeds of 50+Mbps), the additional upstream spectrum is unlikely to affect the customer experience, one way or the other. To the extent that the added capacity reduces the average upstream utilization and provides a more uniform HSD experience, the customer experience should generally improve with the use of Mid-Split.

   Of course, as mentioned, one of the key benefits of Mid-Split spectrum is the launch of higher upstream speed tiers. As these become available, interested customers who live in a household with a Low-Split upstream limitation will require an additional step. Because faster upstream products are an inevitability, and their penetration will likely grow over time, it makes sense to consider a proactive plan to remove the amplifiers with transactional house calls made for other reasons, and eventually for the specific purpose of pulling the Low-Split drop amplifiers out of the system. This is a business balancing act of operations investment at the right time, to stay ahead of the trajectory of these speed tiers. The alternative is that a percentage of customers who want these speeds will have to await a scheduled truck roll to receive them. However, much can be

known in advance with the iHAT tool, so that messaging on the buy-flow front-end can be developed to make that outcome as smooth and efficient as possible for the customer.

For the potential video impact scenario, there is additional nuance and more options to ensure that a quality video experience is maintained. **Figure 19** charts this nuance.

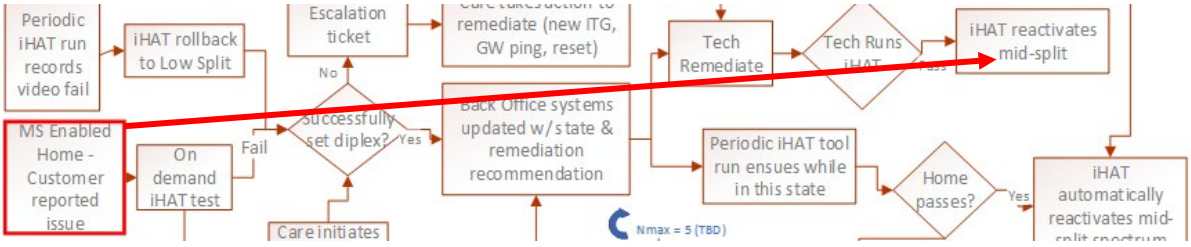


*Confirm home isn't in 'remediation only' mode before running periodic iHAT check

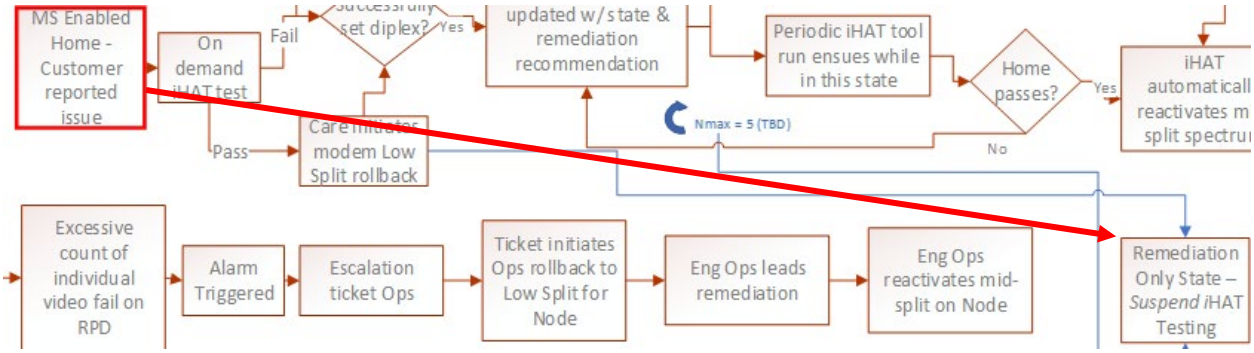
Figure 19 – Care Flow for Support of Mid-Split Related Service Impacts

The flow details are self-explanatory. The diagram assumes that the issue has been diagnosed as likely due to Mid-Split. All other possible causes more probable than Mid-Split have been checked, as they typically would have before reaching this branch of an overall triage flow. In summary form, the fundamental sequence of events in **Figure 19,** after a call to an agent leads to a potential MS diagnosis, are as follows:

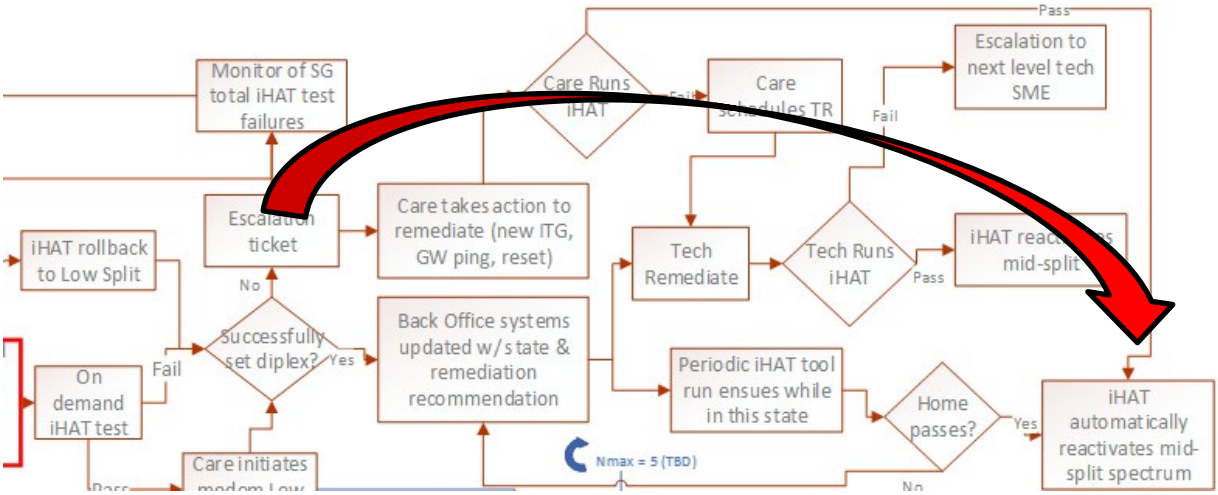
1) Run iHAT for an up-to-date state check and disposition of home, compared to prior state
2) Check if iHAT *before|after* state aligns with the TC (iHAT "fail" and device goes from Mid to Low Split). In Low-Split, the video issue will be eliminated, if it was indeed a Mid-Split-related issue. If so, leave the customer in Low-Split mode, schedule remediation, determine the root cause, and reactivate the Mid-Split region for that home, as charted below:



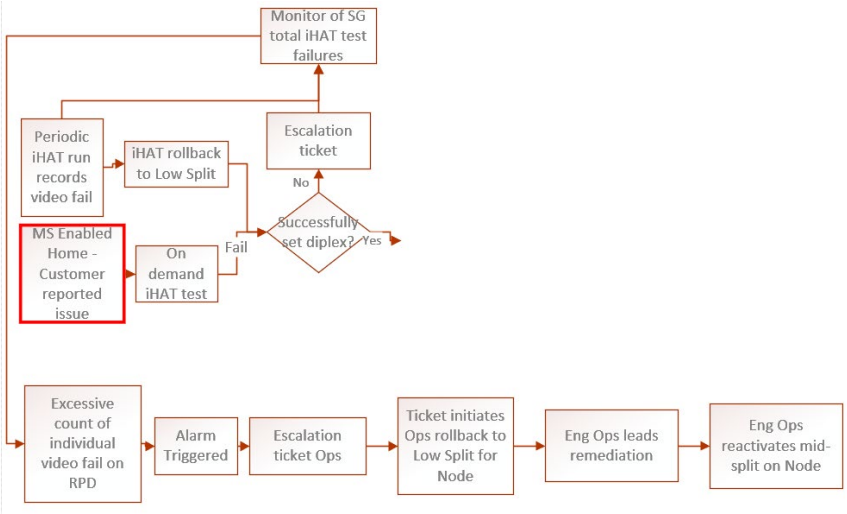
3) If iHAT *before|after* is Mid-Split → Mid-Split, then iHAT believes there should not be a video impairment and yet this is what the customer is experiencing. This does not mean it *is* definitely Mid-Split related. but making this determination on-site is the next step. First, however, the home is manually (via Care directly or via escalation to Operations) reverted to the Low-Split to eliminate the video issue. Again, if this does NOT eliminate the video issue, it is not related to Mid-Split. Assuming this step eliminates the video issue, the account is scheduled for remediation. It is also removed from any further iHAT updates – referred to as the "Remediation Only" queue – that would place it back in Mid-Split mode (because iHAT is giving an erroneous result of "pass" to begin with).



4) A next branch of the flow deals with the case where iHAT does diagnose that the device should be in Low-Split, but it does not properly revert the device to that state. Over time, with code maturity and optimization, we expect this scenario to get diminishingly small. The path first has Care or Eng Ops try to set the modem manually as above, or even factory resetting the modem. If unsuccessful, the path escalates into on-site remediation.



5) A final branch of interest is the scenario where, knowing "typical" metrics for iHAT DOCSIS and video test "fail," a certain threshold of "too many" is set that indicates the issues are probably not on a home-by-home basis, but more systemic. In this case, Engineering Operations is brought in to triage the situation immediately:



8.4. Scars? What Scars?

In summary, the seamless migration to Mid-Split, from a customer experience perspective, looks like this:

- The full 85 MHz upstream signal can exit the home intact and get onto the network – not blocked by drop amps, filters, or otherwise poor frequency response
- The full 85 MHz signal can be activated, and it causes no QAM video artifacts on any STB in the home
- Any video issues that are encountered by a customer and diagnosed as being caused by Mid-Split can be eliminated remotely and immediately
- Products (speeds) that need Mid-Split spectrum can be delivered to a customer who wants the speeds simply and effectively by any buy-flow means available

We have not detailed further, but these additional components fill out this list:

- In-Home filters are not required to launch Mid-Split spectrum
- Mid-Split can be activated using a self-install kit (SIK) model most of the time

9. Mid-Split → High Split and DOCSIS 4.0

9.1. High Split (5 MHz – 204 MHz)

While this paper focused on a Mid-Split migration scenario, operators are looking also to High-Split and, further out into the future, DOCSIS 4.0. Earlier in the paper we described some of the incremental challenges of operationalizing High-Split, compared to Mid-Split. We left one of these challenges out of that discussion until we were able to go into the Mid-Split details of RF isolation management. Of course, High Split has the overlapping band phenomenon with STBs, only worse. There is much more bandwidth for High-Split that extends into and thus overlaps the forward band. Because of this additional energy that will be launched into the spectrum between 54 MHz and 204 MHz, which currently overlaps the input bandwidth of a STB, there is an even greater chance of ACI interference. Therefore, a migration to a High-Split also includes a migration of the home that is receiving High Split-enabled HSD services (e.g., 1 Gbps symmetric service) to an all-IP configuration – i.e., no QAM video. Without a QAM STB, we can ensure that the video service in the home is not affected by the new service.

However, the relatively loud and wide upstream to 204 MHz has enough energy that this is not necessarily the end of the interference story. A phenomenon that has undergone much study is that of "Neighbor Interference," (NI) whereby the cable "Tap" neighbor of a High Split services user may be "close" enough in the dB sense to have services impacted on the adjacent STB or CM. This scenario is shown in **Figure 20**.

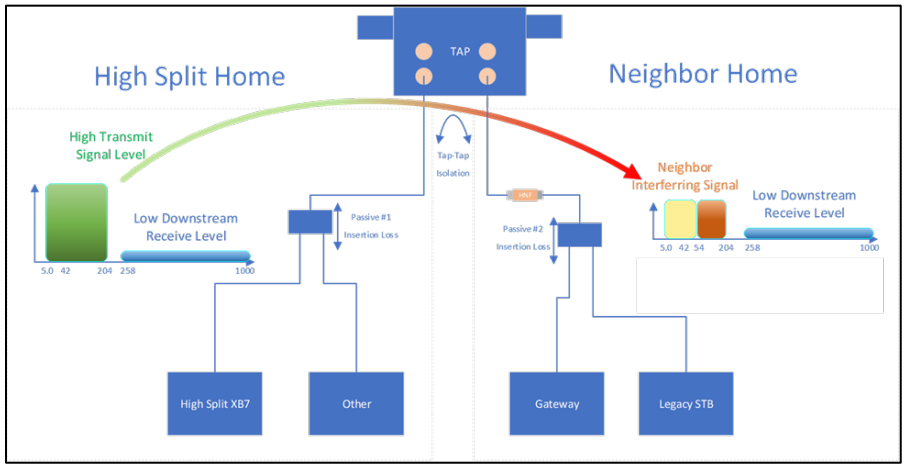


Figure 20 – The Neighbor Interference Phenomenon of High Split

As with the in-home scenario described for Mid-Split, it is also an RF port-to-port isolation and ACI phenomenon with Neighbor Interference, but it is instead the Tap ports that are of consequence. There has been substantial characterization of Tap isolation performance and STB and CM ACI sensitivity recently for this extended upstream band, and the likelihood of this issue has become very well quantified. The MUSL and iHAT tool kit can be applied to a High-Split upstream with these major differences:

1) The OUDP probe signal of interest needs to be modified to one that can be consistently correlated to an extrapolated equivalent of *High-Split* signal energy
2) The potentially at-risk" STB is in a neighboring home, and physical addresses and the relationship of CMs in the field, to Tap ports, is generally not easily known in an automatable way
3) The "victim" device can also now be a non-High-Split CM

Mitigation of NI involves different processes. Visiting a neighborhood home for remediation because a different neighbor on the block upgraded their HSD would be an awkward process on every conceivable level. Thus, the bias for NI would be towards blocking filters in the OSP at the "guilty" Tap port. Documenting filters installed would be an important way to simplify new customer adds going forward. Complete quantification of this phenomenon with respect to the customer experience is difficult until some correlation of High-Split transmissions to video and non-High Split HSD performance can be documented.

9.2. DOCSIS 4.0

Assuming the all-IP home policy that is anticipated for DOCSIS 4.0 homes, coexisting with QAM STBs and legacy CMs in the plant using DOCSIS 4.0 Extended Spectrum technology would take on the equivalent model as High-Split NI, but with the Ultra High-Split bandwidth as the spectrum limit, and STB sensitivity characterization work to be done.

It's worth noting that in DOCSIS 4.0, the very nature of the FDX technology at its core is an overlapping upstream and downstream. As such, it is inherent in the protocol to introduce new technology to manage this overlap. There are two ways this is done, as shown in **Figure 21**:

Echo Cancellation: "Self" cancellation at the PHY level, leveraging knowledge of the transmitted signal at the co-located receiver.

Sounding/Scheduling: Avoids having an FDX CM transmit upstream when it is known that a neighbor could be affected in the downstream that is receiving packets.

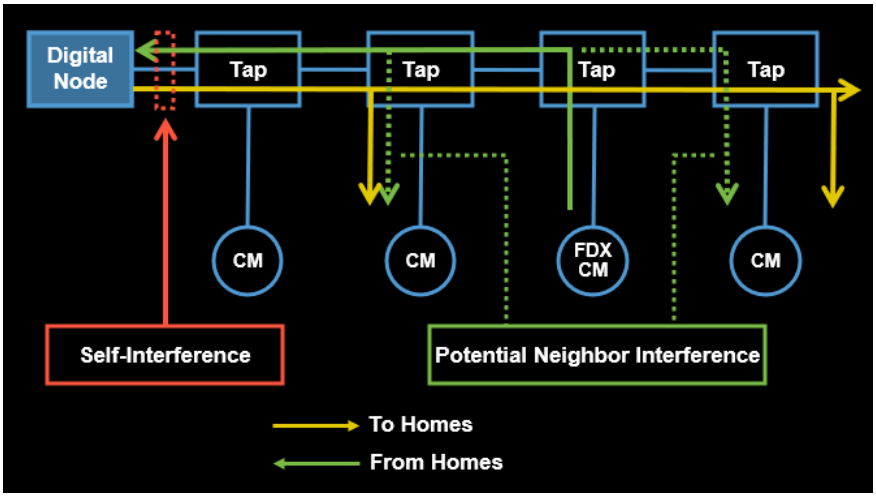


Figure 21 – The Two New Technology Features of DOCSIS 4.0 FDX

For the latter, FDX determines the RF dB isolation relationships among modems to form Transmission Groups (TGs). FDX "sounding" is effectively the built-in, standardized version of NI for the FDX band (108-684 MHz) in FDX systems.

For FDX, the limitation that arises is that these relationships can only be discovered in DOCSIS 4.0 CMs and DOCSIS 3.1 CMs with an FDX-L SW upgrade. FDX-L is a way to make DOCSIS 3.1 CMs aware that they are connected to a DOCSIS 4.0 system, and thereby have their traffic scheduled within the context of the TG assignments. DOCSIS 3.0 CMs cannot participate in sounding at all. The number of DOCSIS 3.0 CMs continues to decline rapidly in the field, but they will not be completely removed from the network before FDX is deployed.

There will be more to come on DOCSIS 4.0 migration challenges and solutions as the 10G technologies continue to be developed [1].

10. Conclusions

Operators are recognizing that, as good as the upstream has been to them since the launch of HSD services, it has given nearly all that it can at this point and needs a spectrum boost to continue to deliver value and support continually growing HSD services, capacity, and speeds. The next step is to add spectrum and launch the next long runway of capacity, with new speed expectations in mind. With the commitment to spectrum comes a commitment to managing it through an HFC lifetime of legacy equipment that is simply not built for it. Building the technology, tools, processes, and practices to enable this transition is a

challenge all operators are working through, with a seamless and non-disruptive experience for the customer as the top priority.

In addition, as in any network evolution that touches the outside plant, making sure that enough is done to the network for the longer term, once the commitment has been made to go out and touch it, is an important part of the upgrade. For access network engineers, the billiards analogy is that, as you are lining up the 6-ball at the side pocket, it is important you make sure that after sinking it you've left the cue ball lined up neatly behind the 12-ball at the corner pocket. With the right series of deft maneuvers, the 10G-ball will be lined up to finish out the game.

Acknowledgements

Many thanks to my esteemed peers Leslie Ellis and Larry Wolcott for joining me for this paper, which due to unique circumstances required a full-court press effort to meet the deadline.

Abbreviations

ACI	Adjacent Channel Interference
AGC	Automatic Gain Control
BAU	Business-As-Usual
BG	Bonding Group
CACIR	Carrier-to-Adjacent Channel Interference Ratio
CAGR	Compound Annual Growth Rate
CDF	Cumulative Distribution Function
DAA	Distributed Access Architecture
DSG	DOCSIS Settop Gateway
FDD	Frequency Division Duplex
FDX	Full Duplex DOCSIS
FTTH	Fiber-to-the-Home
HHP	Households Passed
iHAT	In-Home Assessment test
LoQ	Line-of-Questioning
MER	Modulation Error Ratio
MTA	Media Terminal Adaptor
MUSL	Mid-Split Spectrum Upstream Launch
NI	Neighbor Interference
OFDMA	Orthogonal Freqeuncy Division Multiple Access
OOB	Out-of-Band
OUDP	OFDMA Upstream Data Profile
OSP	Outside Plant
OTA	Over-the-Air
PHT	Performance Health Test
QAM	Quadrature Amplitude Modulation
SNR	Signal-to-Noise Ratio
STB	Settop Box
TaFDM	Time and Freqeuncy Division Multiple Access
TCP	Total Composite Power

Bibliography & References

[1] Baumgartner, Jeff, "Comcast Full Duplex DOCSIS trial pumps out 4-Gig symmetrical speeds," LightReading, 4/2/2021.

[2] Howald, Robert, Repair the Ides of March: COVID-19 Induced Adaption of Access Network Strategies, 2021 SCTE Expo, Oct 11-14, Atlanta, GA.

[3] Robuck, Mike, "Comcast notches 10G Milestone with a trial of 1.25-Gig symmetrical speeds in Florida," Fierce Telecom, Oct. 8 2020, https://www.fiercetelecom.com/telecom/comcast-notches-10g-milestone-trial-1-25-gig-symmetrical-speeds.

[4] Thompson, Robert, and Rob Howald, Dan Rice, John Chrostowski, Ronini Vugumudi, Amarildo Vieira, and Zhen Lu, Rapid and Automated Production Scale Activation of Expanded Upstream Bandwidth, 2021 SCTE Expo, Oct 11-14, Atlanta, GA.

Exploring Multi-Access Edge Compute in Converging Access Networks

A Technical Paper prepared for SCTE by

Andrii Vladyka
Technical Product Manager, Cable Access
Harmonic Inc.
2590 Orchard Parkway, San Jose, CA 95131
+1 408 542 2559
Andrii.Vladyka@harmonicinc.com

Asaf Matatyaou
Vice President, Solutions and Product Management, Cable Access
Harmonic Inc.
2590 Orchard Parkway, San Jose, CA 95131
+1 408 542 2559
Asaf.Matatyaou@harmonicinc.com

Howard Abramson
Principal Architect, Cable Access
Harmonic Inc.
2590 Orchard Parkway, San Jose, CA 95131
+1 408 542 2559
Howard.Abramson@harmonicinc.com

1. Introduction

Cable access networks and equipment has changed dramatically since the inception of high-speed data over hybrid fiber coaxial (HFC) networks. This evolution has enabled cable to become the dominant supplier of broadband access worldwide. The ongoing need for scale and operational efficiencies to maintain this lead is anticipated and will enable Multi-System Operators (MSOs) to meet subscribers' voracious appetite for bandwidth and internet-enabled devices while lowering the overall cost of ownership so that MSOs can stay competitive with alternate access providers.

A common element to cable access network evolution has been a focus on purpose-built hardware appliances, employing advances to information encoding (e.g., DOCSIS® 1.0, 1.1, 2.0, 3.0, 3.1, and 4.0), that operate on-premises at *carrier scale for a single service medium, DOCSIS*. These advances were complemented by the modularization and distribution of functionality to various portions of the operator's access network (inside and outside plant) using tightly coupled (industry standard) hardware appliances continuing to operate at carrier scale and including other service medium such as passive optical networks (PON). In recent years, the industry has acknowledged the utility and applicability of public (and private) cloud operator approaches to organizing, operating, and deploying massively scalable computational networks for the access network. These *cloud-native* architectures started by partitioning functions into software-based applications running on commercial off-the-shelf (COTS) servers and Ethernet switches to *virtualize* typical hardware-based network functions (VNFs). This approach to VNFs has evolved into the orchestration of *containerized* cloud-native functions (CNFs) that comprise a collection of disaggregated loosely coupled microservices that can operate in public, private and distributed portions of an operator's network to advance *carrier scale toward cloud scale*. Further evolution of these networks is now occurring along two dimensions to i) leverage elasticity of hyperscale cloud service providers' (CSPs) compute resources and ii) deep edge computing infrastructure to converge deployment and operations for a diverse collection of access technologies and services into a single multi-access, multi-service edge computing (MEC) platform. A paper presented last year focused on exploring the former aspect of access network evolution. This current paper focuses on the latter point — access service convergence.

This paper summarizes the role of CNFs and edge compute in access network convergence based on the practical experience of deploying distributed access architecture (DAA) DOCSIS networks and PON. This convergence is in a form of CNFs on a common cloud-native platform leveraging general purpose x86-based compute resources. Motivation and strategies for convergence, as well as its impact on network topology change and daily operations is considered. The concepts of cluster resource elasticity and horizontal scaling, along with the benefits of infrastructure resource sharing between access media (e.g., DOCSIS and PON) workloads in MEC deployments, are also analyzed. Finally, furthering DOCSIS and PON network convergence with other access technologies that provides a path for building CNF-based service agnostic networks will be described.

2. Terminology

The following interpretation of key industry terms is assumed throughout the paper.

Multi-access Edge Computing (MEC): applies to a system which provides an IT service environment and cloud-computing capabilities at the edge of an access network and contains one or more types of access technology that is within close proximity to its users [1].

Cloud-native Network Function (CNF): network functionality delivered in software via cloud-native development and delivery practices [2].

Cluster: a set of compute nodes that can be viewed as a single system. Cluster nodes are connected via a converged network (e.g., Ethernet-based) and managed by platform management software.

Multi-tenant cluster: sharing compute resources for infrastructure pods and workload pods serving different types of applications (e.g., DOCSIS, PON, wireless), while providing required levels of resources and components isolation.

Access network convergence: providing wireline and mobile services from a single flexible, programmable connectivity platform whose hardware, software and data storage resources spanning multiple geographic locations are shared across multiple access technologies. For operators, network convergence reduces the complexity and cost of providing multi-service offerings [3].

Pod: a group of one or more software containers, orchestrated by Kubernetes [4].

Container: an application that has its own file system, CPU, memory and process space. It is similar, but more lightweight compared to a virtual machine (VM), as containers have relaxed isolation properties to share the Operating System (OS) among the applications.

3. Historical Overview of Access Network Convergence and Virtualization in Cable

Historically, the telecommunication equipment market was dominated by hardware-based appliances tailored for a specific access layer technology, implementing one or more layers of the Open Systems Interconnection (OSI) model, from L1 to L7 [5]. Although such appliances employ industry standard protocols (e.g., SSH, Telnet, SNMP, NETCONF, etc.) each is built and managed using proprietary methods. The result has been economic and operational overhead that inevitably comes with the proliferation of disparate networking appliances on the same network.

The evolution of telecommunication equipment was (and remains) focused on the following.

1. Increased performance and capacity (e.g., throughput, network latency, and port density) facilitated by innovations in purpose-built, highly integrated application-specific integrated circuits (ASICs) and systems on a chip (SoCs).
2. Reduced construction cost and operations maintenance.
3. Increased availability, reliability, and security.
4. Improved network management and network operations.

Each of these advances has been mostly focused on individual network elements that each operate at *carrier scale and availability* rather than focus on the network as a converged system. Examples of these network improvements include the original Modular Cable Modem Termination System (M-CMTS) followed by the Converged Cable Access Platform (CCAP), later the Modular Headend and Distributed Access Architecture (MHAv2 and DAA), Remote PHY (R-PHY), and Flexible MAC Architecture (FMA) specifications; all of which were closely tied to advances in physical layer optimizations via the DOCSIS 2.0, 3.0, and 3.1 standards. Today, the cable industry recognizes challenges for taking networks to the next level, as described by CableLabs' vision of the 10G platform — a combination of technologies that will deliver symmetric multigigabit internet speeds [6].

As an alternative to optimizing individual network appliances, a new approach to meeting these challenges has emerged. This can be described by the following two points.

1. Network and service convergence.

2. Network and service virtualization.

Network convergence and network virtualization has presented a convenient approach to solving these challenges given exponential growth of computational efficiency available in general-purpose processors using common design patterns and leveraging open source software libraries, tools, and network operating systems. The increased performance of general-purpose compute has enabled telecommunication equipment vendors to develop solutions that run on generic, compute platforms that can implement virtual network functions (VNFs) in software. By using COTS server hardware, operators are reducing variations and consolidating vendor proprietary hardware appliances that were deployed in their network. The transition from purpose designed hardware appliances to similarly capable virtualized functions running on COTS servers is illustrated below.

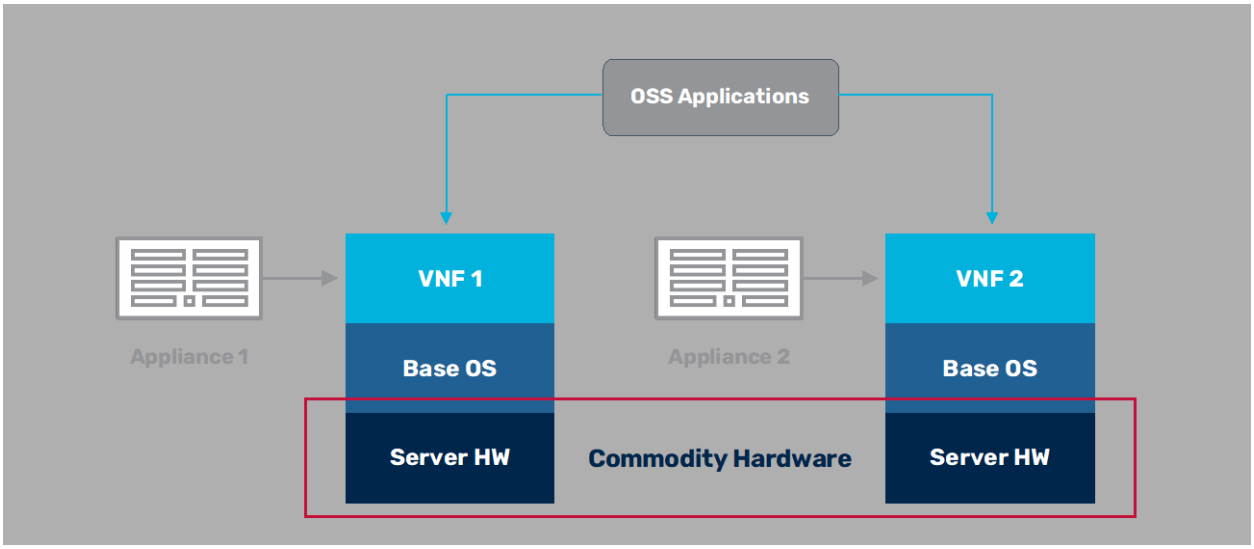


Figure 1 – Transition from Hardware-based Appliances to VNFs on Commodity Hardware

This first transition from hardware to VNFs remained *carrier scale and availability* of individual elements as it has been mostly focused on substituting network appliances with a better mouse trap [11].

Based on advances by the major hyperscalers, the next natural step in network convergence and virtualization has been the adoption of a new paradigm that operates at *cloud scale and availability*: the shift from VNFs to CNFs operating under the management and control of a container orchestration system (such as Kubernetes). This container orchestration system provides a time tested, scalable, de-facto standard for managing cluster resources (such as central processing unit (CPU), networking, memory, and storage), network function partitioning & abstraction, and operation of the network. Rather than managing individual appliances at carrier scale, this *cloud-native* paradigm creates an open environment where homogeneous general-purpose compute can be shared by networking applications from any vendor design deployed in the form of CNFs that is scalable and isolated from other vendor CNFs. The following diagram illustrates independent network functions (from one or more vendors or the operator themselves) operating as CNFs to provide an end-to-end network solution.

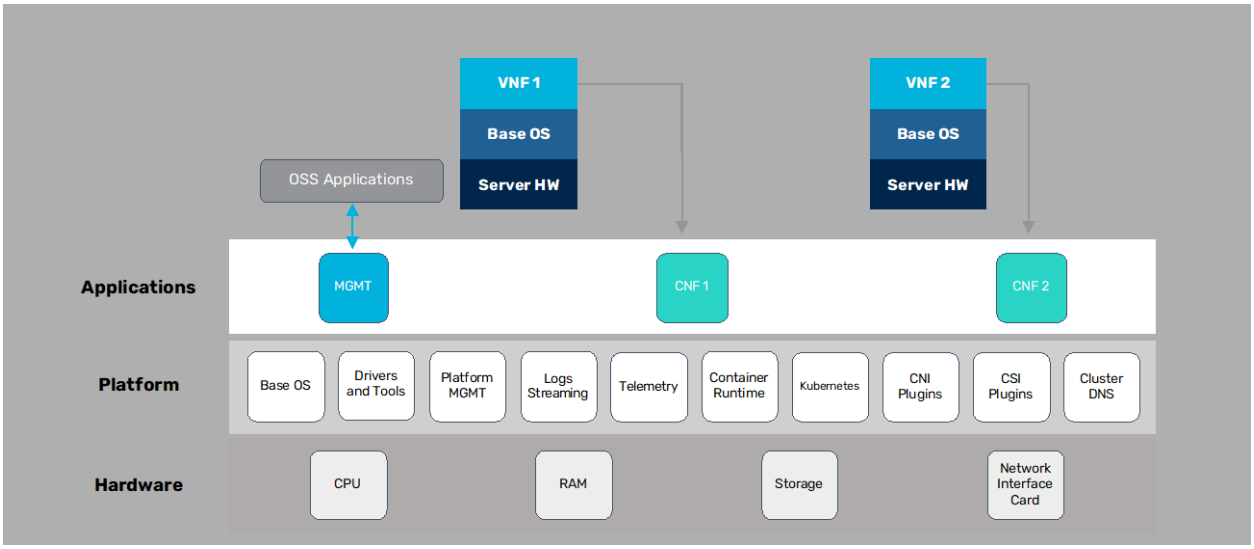


Figure 2 – Transition from VNFs to CNFs

To date, access network convergence and virtualization has helped cable operators to address the following challenges.

1. Sustaining ever-growing network performance requirements by leveraging a short and predictable cycle of increasing performance gains available in general-purpose compute.
2. Reducing network capex by riding economies of scale provided by commodity hardware.
3. Optimizing network opex by reducing the number of disparate devices deployed and maintained in the network.
4. Improving network reliability by employing native high availability (HA) facilities that come with the flexibility of network functions implementation in software, in general, and in particular with the use of Kubernetes-orchestrated CNFs.
5. Simplifying and unifying network management and operations by reducing the number of different systems to manage and leveraging service-based telemetry and logging for receiving and storing the information from different CNFs in a uniform way.
6. Advancing *carrier-scaled* applications that can grow or shrink both in terms of performance/capacity of an individual CNF (vertical scaling) and in terms of the number of deployed CNFs (horizontal scaling) at *cloud scale* one CNF and one server at a time.

In subsequent sections of the paper, we describe different aspects of the current state of cable access networks convergence and virtualization.

4. Access Networks Convergence in Cable: The Current State

Network convergence can be viewed by considering three different access technologies employed in cable. Looking at cable's HFC, it becomes clear that the same fiber carrying analog and digital wavelengths for video and data can be used for Ethernet and PON. In fact, DOCSIS, Ethernet, and PON are three edge technologies that have been used in cable networks for years. These access technologies are typically applied for different subscriber categories and managed by different personnel using unique operations. For example, it is often the case that DOCSIS is reserved for residential and small medium businesses (SMB), Ethernet is used for enterprise and mobile xHaul, and PON is used for managed property and multidwelling unit (MDU) subscribers. Although provisions were made for designing

hardware appliances with support for multiple edge access technologies in the CCAP specification, it is typically the case that DOCSIS, Ethernet, and PON are provided by different vendors using quite different approaches to managing and operating the network. The impact is a non-uniform and disparate collection of point products that must be managed by cable operators. The result is a lack of convergence and economies of scale that come from modern approaches to deploying and managing cloud-native networks. Alternatively, when considering access solutions based on virtualization, it becomes possible to consider workloads for two or more of these access technologies within the same converged platform.

Key factors that contribute to access networks virtualization include the following.

1. Growth in general-purpose CPU performance capable of software implementation of network functions in the form of CNFs with the same or better performance as of field-programmable gate array (FPGA)/ASIC-based hardware appliances.
2. Maturity of software frameworks that accelerate packet processing workloads running on a wide variety of CPU architectures [12].

The trigger for virtualization of the CMTS (vCMTS) was the release of the DOCSIS Remote PHY Specification [10], that partitions the DOCSIS PHY from components responsible for MAC and upper layer protocol functions. This separation of the lowest layer of the DOCSIS protocol enables the repartitioning of upper layer software-only implementations. In this way, virtualization of the CMTS was achieved and brought into operator's facilities [11]. Using general-purpose compute resources for running DOCSIS vCMTS workloads makes it possible to consider workloads for other access technologies on the same cluster of compute resources.

The following diagram illustrates an example of how access workloads, operating on general-purpose COTS servers, can be applied to DOCSIS and PON.

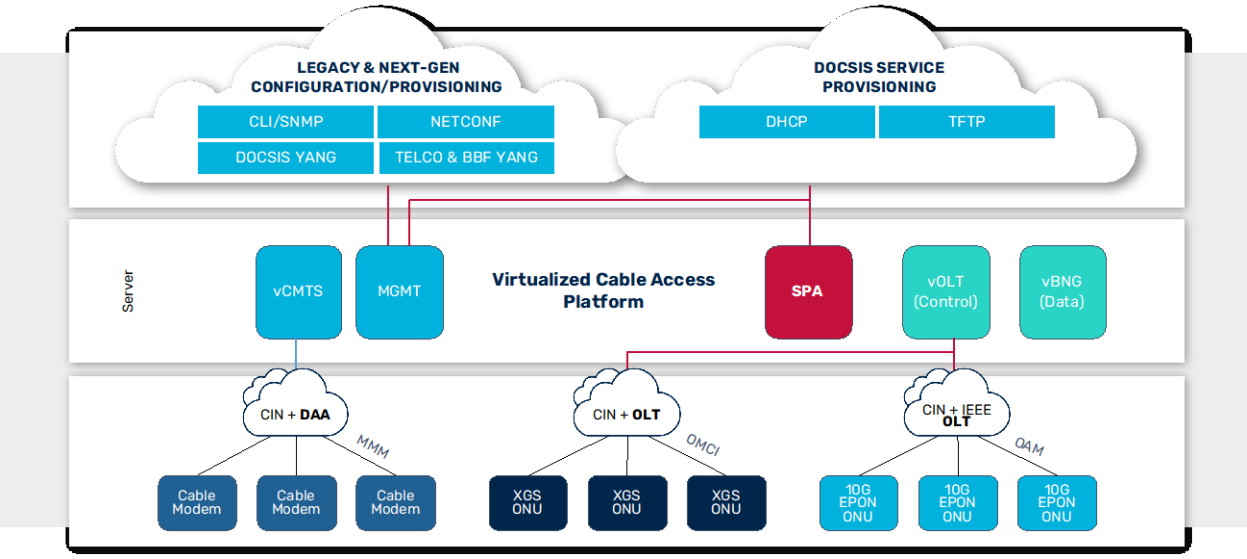


Figure 3 – Access Network Convergence and Uniform Service Provisioning for Any Access

This is explained by the following.

- vCMTS — DOCSIS workloads are provided by CNFs that virtualize the CMTS upper layers serving DAA-based PHY devices.
- vOLT— PON workloads are provided by CNFs that virtualize the control of external physical OLTs
- vBNG — upper layer functions typically provided by hardware appliances for PON are provided by CNFs performing subscriber management and user plane operations
- MGMT and SPA — OSS/BSS, fault, configuration, accounting, performance, security (FCAPS) is provided by Management and Service Provisioning Application CNFs to complete the convergence

In this way, not only are different access technologies and protocols converged on common COTS servers, but they are presented to the operator's network with a uniform and scalable platform that can share and load balance available resources.

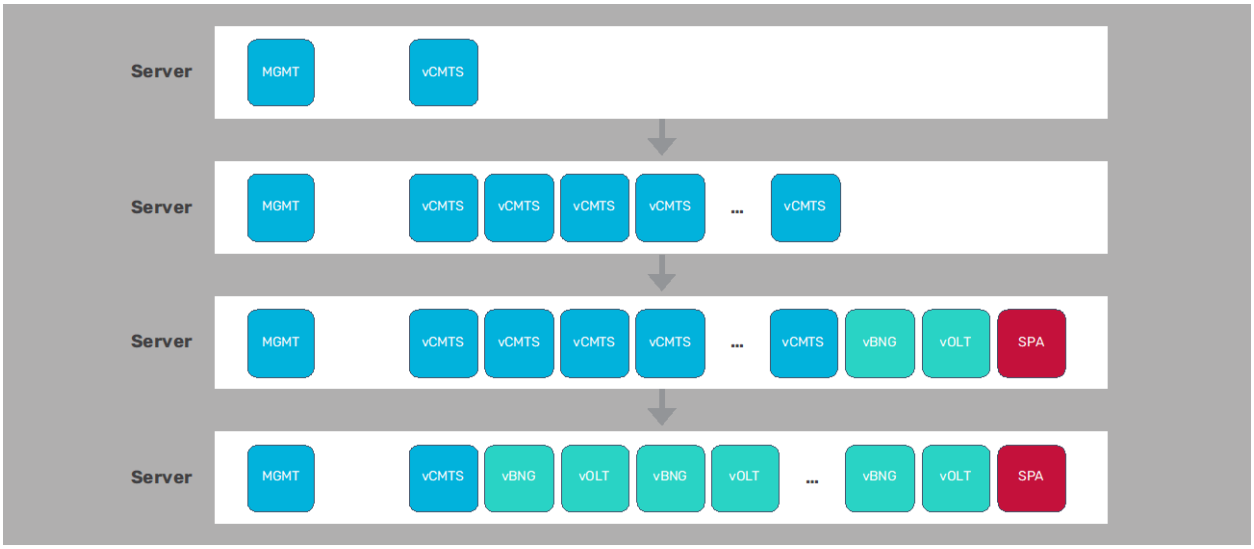


Figure 4 – Horizontal Scaling of CNFs on a Converged Access Platform

Figure 4 illustrates the example of horizontal workloads scaling on a converged access platform.

1. The deployment starts with a relatively small number of DOCSIS service groups (SGs) connected to one CNF instance that implements vCMTS functions.
2. As the number of connected DOCSIS SGs grows, the number of vCMTS workloads increases proportionally.
3. With the introduction of a new type of access technology (PON), CNFs implementing vBNG, vOLT, and SPA functions are instantiated on the same cluster resources.
4. Over time, the number of CNFs implementing different types of access technologies may change. For example, some of the DOCSIS SGs may be converted to PON SGs, which is reflected by the proportional change in the number of CNFs of a certain type running on the platform.

As a converged access platform, the solution now offers common interfaces for managing and operating the entire network. One example of this convergence is the presentation of traditional DOCSIS service provisioning.

Cable operators have long benefited from a simple and standard method of provisioning individual subscriber services. The utility of this basic approach for defining individual subscriber connections and

service level agreements (SLAs) was recognized and applied in the DOCSIS Provisioning of EPON (DPoE) standards [17]. While originally defined and qualified for 1 Gbps EPON, this same DOCSIS-based method for defining subscriber connections, when partitioned as a collection of CNFs, can be applied to any media, including ITU-T (GPON and xGPON), IEEE (10G, 25G, and 50G EPON), and Ethernet. Further, by separating provisioning and management CNFs from the media-specific CNFs, adoption of other methods becomes far more flexible and scalable. As an example, by applying cloud-native APIs within Management and Service Provisioning Applications, use of CLI, SNMP, NETCONF, and of course virtualized cable modem (vCM) methods specified in DPoE, are all available and of little consequence to the vCMTS, vOLT, and vBNG CNFs that they serve.

While this is one way of converging an access network, it is but one example. The point is that the impact of access network convergence in the cable industry results in the following benefits.

1. Increased service agility.
2. Reduced operational cost.
3. Simplified and uniform network operations.
4. Increased network reliability, availability and security.

Next, we will explore the connection of cable access network convergence with the MEC concept.

5. Multi-access Edge Computing in Cable

The ETSI definition for Multi-access Edge Computing enables the implementation of applications as software-only entities that run on top of a virtualization infrastructure, which is located in or close to the network edge [7]. The MEC framework represented in the diagram below shows the general entities involved. These can be grouped into system level, host level and network level entities.

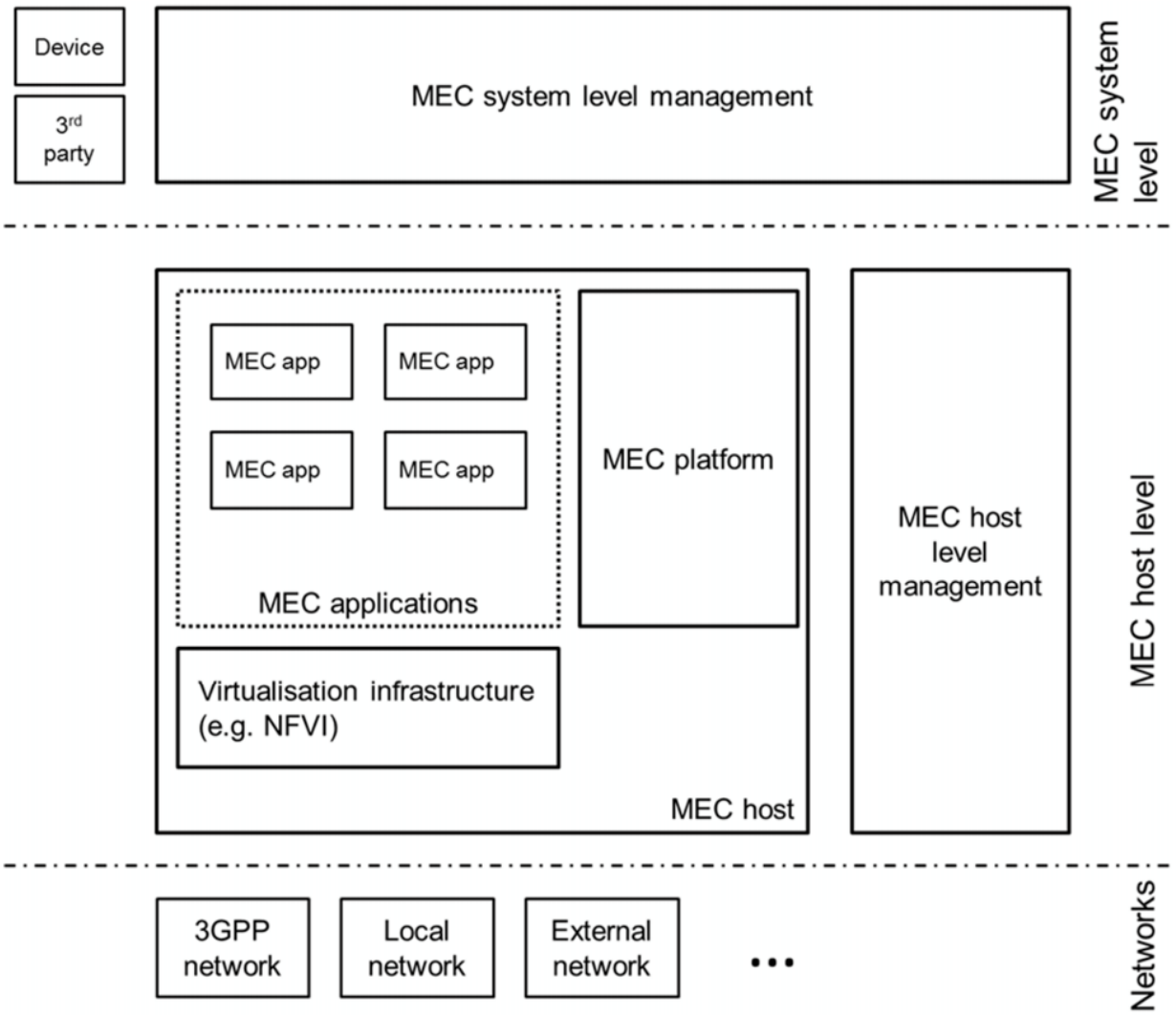


Figure 5 – ETSI Definition of Multi-access Edge Computing Framework [7]

The framework for Multi-access Edge Computing consists of the following entities.

1. **MEC host** – an entity that contains a MEC platform and a virtualization infrastructure that provides compute, storage, and network resources for the purpose of running MEC applications.
    a. **MEC platform** – the collection of essential functionalities required to run MEC applications on a particular **virtualization infrastructure** and enable them to provide and consume MEC services.
    b. **MEC applications,** instantiated on the virtualization infrastructure of the MEC host based on configuration or requests validated by the MEC management.
2. **MEC system level management** that includes the multi-access edge orchestrator as its core component. The orchestrator is responsible for the following functions.
    a. Maintaining an overall view of the MEC system.
    b. Application packages on-boarding.
    c. Selecting appropriate MEC host(s) for application instantiation based on constraints, such as latency, available resources, and available services.
    d. Triggering application instantiation, termination, and relocation.

3. **MEC host level management** which handles the management of the MEC-specific functionality of a particular MEC host and the applications running on it.

While the MEC concept originated in the context of mobile/wireless networks, the general definition of the MEC concept and framework has a lot in common with the real-world architecture of the current generation of the virtualized cable access platforms deployed today [8].

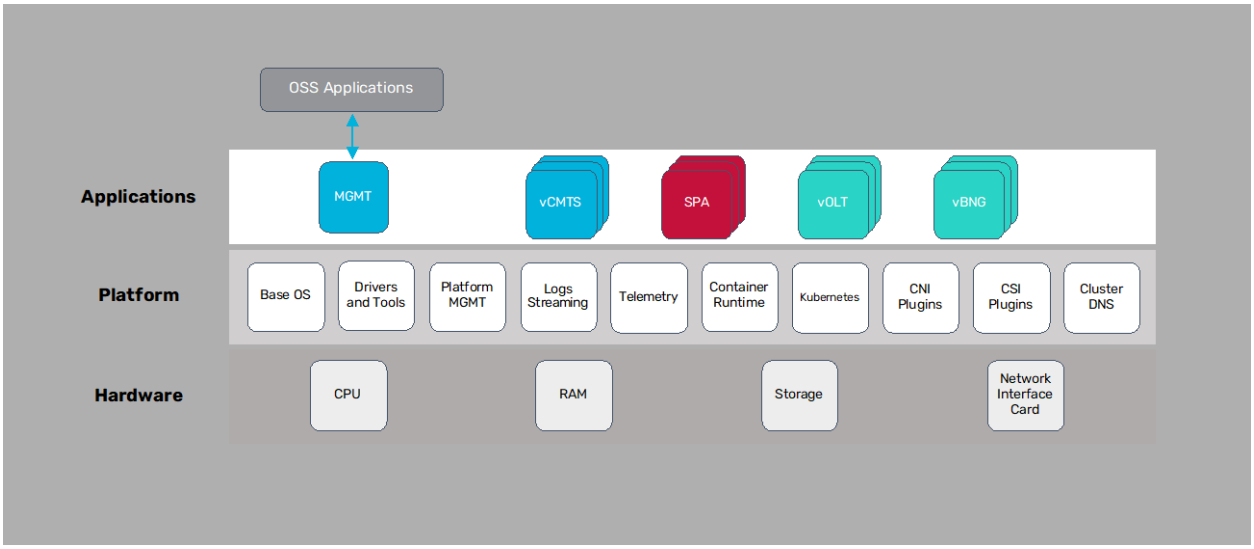


Figure 6 – A Host of a Virtualized Cable Access Platform

As can be seen from the figure above, the implementation of a virtualized cable access platform is like the generic ETSI definition of MEC compute framework. This includes the following similarities.

1. The role of the platform layer of the virtualized cable access platform is equivalent to those of the MEC platform and virtualization infrastructure entities.
2. Applications running on a virtualized cable access platform are equivalent to MEC applications. Examples of applications running on top of the virtualized cable access platform are:
    a. Virtual cable modem termination system (vCMTS).
    b. Virtual optical line terminal (vOLT).
    c. Virtual broadband network gateway (vBNG).
    d. Service provisioning application (SPA).
    e. Management (MGMT) applications implementing “northbound” interfaces toward operations support systems (OSS) as well as “southbound” API calls toward platform and CNFs.
3. MEC system level management in a virtualized cable access platform is partially covered by the OSS applications, and partially implemented by a set of tools for deployment automation and monitoring.

In fact, the cable industry has been leveraging the MEC concept for years [9] without calling it “MEC.” Future directions for the application of the MEC framework within cable may include the following.

1. Further network convergence and proliferation of the applications running on top of the virtualized cable access platform:

a. Adding new types of broadband access applications on the virtualized cable access platform.
b. Adding applications specific for the hospitality industry (e.g., digital signage and targeted advertising applications).
c. Converging applications implementing 5G CU/DU/cCore components with other broadband access applications.
d. Exploring new business models with healthcare applications. This use case also covers more generic list of applications dealing with sensitive customer data, where application data is not allowed to leave the perimeter of the organization.

2. Integration with hyperscalers and offloading certain types of workloads traditionally deployed in CSP infrastructure to the network edge.
3. Adopting CI/CD practices on the organizational level to accelerate the development and deployment of edge compute applications.

6. Dealing with Transport Matters in a MEC Era

As shown on Figure 5, the MEC framework is generally agnostic to the networks connecting MEC hosts. Luckily, the current state of the virtual access networks convergence doesn't leave many variations for networking technologies connecting applications running on the edge compute hosts with their consumers. To explain the concept and why it's perceived to be a positive way of network evolution, we need to take a step back and state the challenges attributed to the networks of the past.

Consider the following diagram showing a typical legacy DOCSIS network with analog fiber going from the hub to the field.

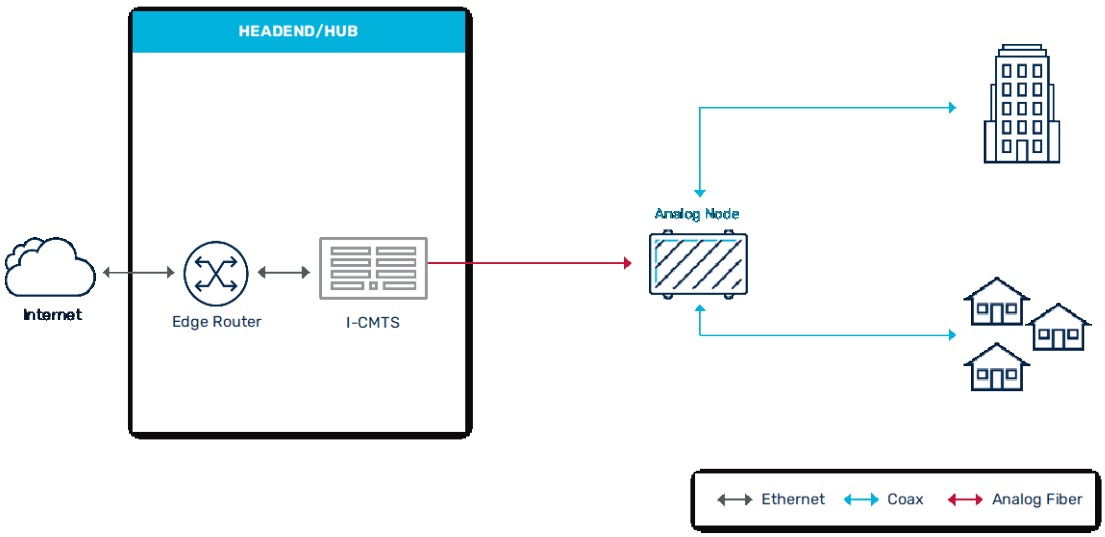


Figure 7 – Legacy DOCSIS Network

Analog connectivity between headend/hub and fiber node doesn't leave options for choosing the connectivity type for the end subscribers. If, for example, a business subscriber is looking for a speed tier that cannot be provided by a DOCSIS network, then it would require the cable operator to lay a dedicated optical fiber line from the headend/hub to the customer's facility. The cost of such work and the time it would take to do the construction makes the business case not economically viable. Similarly, delivering high-speed internet connectivity using Ethernet as a last mile technology would require adding more

physical appliances (e.g., access switch, BNG) to the operator's facility and integrating them with the service provisioning systems. The alternative, of course, is to use the same fiber with different digital wavelengths that can carry a variety of access protocols. This is a key concept of the DAA.

With DAA, Ethernet transport becomes a universal means of transporting broadband access applications running on top of the compute resources (wherever they are) to the devices implementing physical layer connectivity. The result is a converged multi-access network as shown below.

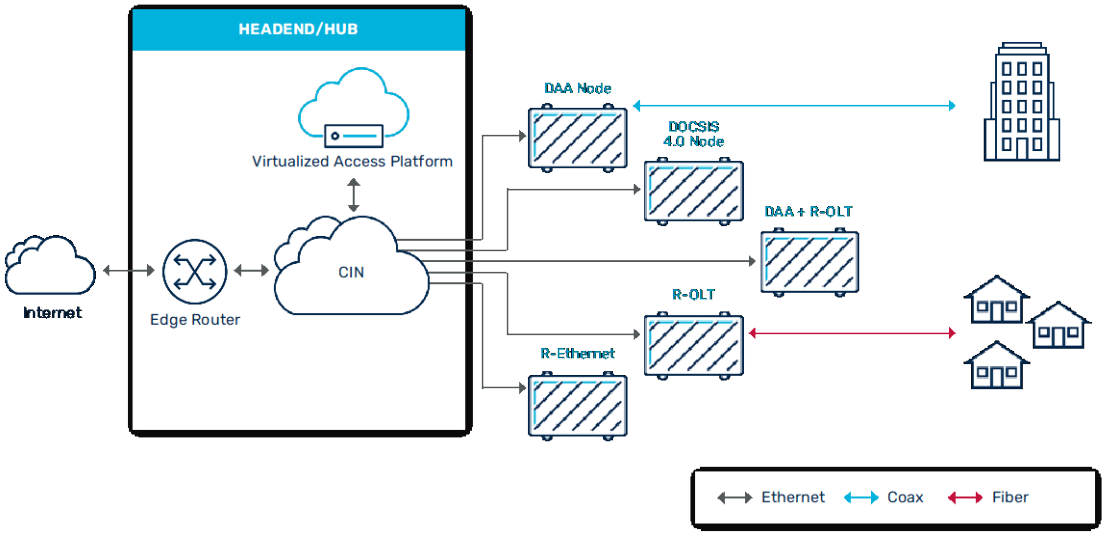


Figure 8 – Converged Multi-services Cable Access Network

The benefits of this converged multi-service cable access network include the following.

1. "Standardization" of the outdoor plant, with Ethernet links (via the converged interconnect network, CIN) connecting DAA nodes with headend/hub equipment that in turn may be connected to off-premises resources *in the cloud.*
2. *Stretching* the CIN closer to end users with outdoor digital fiber segments extending reach, aggregation, and overall network capacity.

It's worth mentioning that while current DAA deployments are based on 10G Ethernet connectivity between distributed access architecture switches (DAAS) and DAA nodes, the industry is looking into adopting higher rates (25G, 100G and higher speed Ethernet).

Using the same example of an enterprise business customer connected over a DOCSIS network looking for a higher speed tier, let's look at how a virtualized access platform enables instantiation of new services.

6.1. Use Case: Adding New Services On Virtualized Access Platform with DAA

6.1.1. Outside Plant Work

The scope of the outside plant work includes installation of a digital fiber from the DAA node closest to the customer's facility to enable 1G or 10G Ethernet or PON. Rather than the expense and complexity of

a headend/hub trunk connection, the scope of required construction to reach the customer's facility is far smaller.

6.1.2. CNF Instantiation

A vBNG CNF workload is deployed on the virtual cable access platform's compute resources that are closest to the customer's facilities. Generally speaking, the exact location of the compute resources serving a specific customer is flexible: location of the CNF can be anywhere in the network and can change over the time based on commercial and operational needs.

6.1.3. Topology Discovery

The logical connectivity (session) is established between CPE and the corresponding vBNG instance using automation tools or statically.

6.1.4. Services Provisioning and Activation

This step implies provisioning of the selected speed tier and SLA for the customer premises equipment (CPE) and applying the corresponding configuration to the vBNG application running on the virtual cable access platform. As noted earlier, using uniform MGMT and SPA CNFs permits common tools and workflows to be applied for services provisioning over different access networks.

6.1.5. Monitoring

At this point, services are up and running and an operator performs routine monitoring of the subscriber connectivity and availability using cloud-native telemetry and logging.

The advantage of this approach for adding a new subscriber or introducing a new type of access technology is that it does not require installation of any new equipment within the operator's facility or change in workflow and tools. Connecting a subscriber to the desired access technology (e.g., DOCSIS, Ethernet, PON) to the nearest DAA node becomes more plug-and-play. Such deployment flexibility and service agility are enabled by the introduction of the digital fiber all the way down to DAA nodes, on the one hand, and leveraging MEC for broadband access applications deployment, on the other.

7. Network Operations in a Converged World

The change in network operational practices for access networks convergence can be viewed through the prism of network virtualization. The move from hardware-based networking appliances towards CNFs running on a common virtualized access platform implies changes to the following network operations properties.

7.1. Separation of Platform Management from Applications Management

The virtualized access platform provides standard resource units, such as CPU cores, memory, storage, and network share, to applications (CNFs). The platform considers CNF resource needs, as well as other requirements, such as locality and high availability (HA) preferences. As a result, operators need only to operate a single Kubernetes-based platform that is responsible for multiple CNF workloads in a consistent way. In comparison, legacy access networks typically require at least one element management system (EMS) or network management system (NMS) dedicated to each class and vendor provided networking appliance. The result is oftentimes complex and expensive software systems that attempt to orchestrate this collection of appliances.

7.2. Ubiquitous Automation

CNF deployment automation provides a declarative configuration model and single configuration interface for managing the configuration of different CNFs. Automation in cable access networks leverages standard network management protocols, such as NETCONF, and data models, such as DOCSIS YANG [19].

1. Operator builds configuration template(s) using standard data model.
2. Operator applies configuration to a virtualized access platform via NETCONF.
3. CNFs are instantiated on a platform to accommodate services configuration provided by the operator.

Using services-oriented configuration data models allows an operator to focus on services requirements while corresponding CNFs instantiation and resources allocation are handled by the platform.

7.3. Deprecation of Legacy Monitoring Interfaces in Favor of Modern Telemetry-Based Protocols

In the current context, SNMP and IPDR fall into the category of legacy protocols providing data to external monitoring and billing systems.

In short term, the usage of service-based telemetry is not mutually exclusive to legacy protocols. In fact, the current state of access networks convergence is characterized by the simultaneous and compatible use of modern and legacy monitoring/billing interfaces and protocols. The long-term vision is that legacy protocols will be deprecated in favor of the scalability and deployment flexibility requirements available in virtualized converged access networks. Figure 9 provides an example of a telemetry streaming pipeline from CNFs running on top of a virtualized access platform. The pipeline is highly scalable, provides updates with short time intervals, and can be incorporated into operator's automated analytics and machine learning (ML) applications.

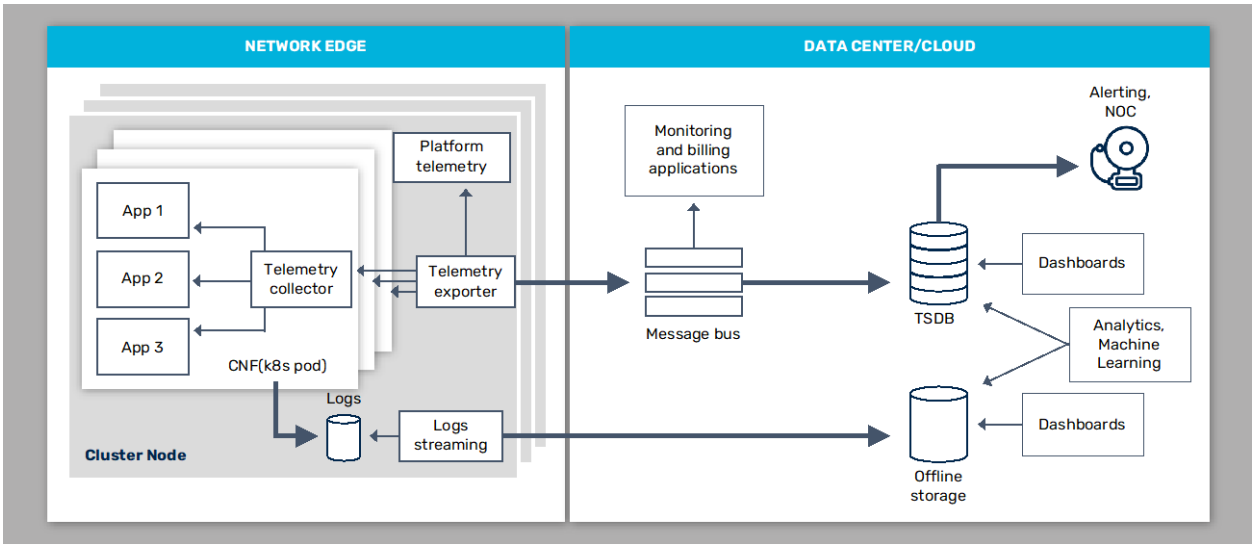


Figure 9 – Telemetry Streaming Pipeline

The result of telemetry data processing by automated analytics and ML applications can be then applied back to the access platform in a form of closed-loop automation.

7.4. Introduction of The New Software Distribution Models and Adoption of CI/CD Methodologies at The Organizational Level

In legacy networks, an upgrade of a networking device is typically executed in a form of transferring a software image file(s) to the target device and activating a new software version accompanied by the reboot of the device. In virtualized converged access networks, CNF software is released in the form of containerized images. This approach is characterized by the following.

a. The virtualized access platform is always connected to a container image registry, facilitating simultaneous software distribution and installation at *cloud scale*.
b. An upgrade of a CNF can be performed in a way seamless to the services provided to end users (in-service software upgrade, ISSU), with minimal service interruption. No host (server) reboot is required for a typical CNF upgrade cycle.
c. Individual CNFs running on a converged platform can be upgraded independent of each other.
d. The same virtualized access platform can host CNFs of the same type running different software versions. This enables *candidate or canary* upgrades and A/B testing on a per-service group basis for final field acceptance prior to networkwide updates.

8. Conclusion

Cable access network convergence and virtualization were developed as a general solution for the challenges and opportunities the cable industry has experienced over the years.

- Meeting and exceeding ever-growing network capacity and performance requirements.
- Network construction and operations cost reduction.
- Improving network reliability, availability and security.
- Managing the complexity of network operations to incorporate new services and access technologies.

The cable industry is taking advantage of concepts defined by MEC by applying hyperscale cloud methods to DAA with CNFs. These concepts help to accelerate the convergence of cable access networks. These innovations enable cable operators to take advantage of the following converged network benefits.

- Elasticity of cluster resources and infrastructure.
- Automation and service instantiation at high velocity.
- Horizontal scaling capabilities and ability to change capacity on demand.
- Visibility of the platform and service.
- Faster time to market for deployment of new applications and repair.

Standard Ethernet-based connectivity between MEC infrastructure and devices in the field becomes a key enabler for many access technologies.

Moving forward, future directions of MEC development in access networks may include the following.

1. Adding new types of broadband access technologies and services, with the end goal of deploying an access agnostic network.
2. Unifying service provisioning processes for multiple access network types.

3. Integration with hyperscalers, that may be executed in two directions:
    a. Moving certain types of CNFs from on-premises compute resources to hyperscaler infrastructure.
    b. Moving some of these workloads, traditionally operating in hyperscaler infrastructure to on-premises clusters or deep edge compute resources.
4. Adopting CI/CD practices and modern software distribution and management practices to further improve service agility.

Abbreviations

ASIC	application-specific integrated circuit
CAPEX	capital expenditures
CCAP	converged cable access platform
cCore	converged core
CD	continuous deployment
CI	continuous integration
CIN	converged interconnect network
CMTS	cable modem termination system
CNI	container network interface
CNF	cloud-native network function
COTS	commercial off-the-shelf
CPE	customer premises equipment
CPU	central processing unit
CSI	container storage interface
CSP	cloud service provider
CU	central unit
DAA	distributed access architecture
DNS	domain name server
DOCSIS	Data Over Cable Service Interface Specification
DPDK	data plane development kit
DPoE	DOCSIS provisioning of EPON
DU	distributed unit
EMS	element management system
EPON	ethernet passive optical network
ETSI	European Telecommunications Standards Institute
FCAPS	fault, configuration, accounting, performance, security
FMA	flexible MAC architecture
FPGA	field-programmable gate array
GPON	gigabit passive optical network
HA	high availability
HFC	hybrid fiber-coax
HW	hardware
IEEE	Institute of Electrical and Electronics Engineers
ISSU	in-service software upgrade
IT	internet technology
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
MAC	media access control

M-CMTS	modular CMTS
MEC	multi-access edge computing
MDU	multidwelling unit
MGMT	management
MHA	modular headend architecture
ML	machine learning
MMM	MAC management message
MSO	multi service operator
NETCONF	network configuration protocol
NMS	network management system
NOC	network operations center
OAM	operations, administration, and management
OLT	optical line terminal
OMCI	ONU management and control interface
ONU	optical network unit
OPEX	operating expenses
OS	operating system
OSI	open systems interconnection
OSS	operations support systems
PHY	physical layer
PON	passive optical network
RAM	random-access memory
RPD	remote PHY device
RPS	remote PHY shelf
SG	service group
SLA	service level agreement
SNMP	simple network management protocol
SMB	small medium businesses
SoC	system on a chip
SPA	service provisioning application
SSH	secure shell protocol
TSDB	time-series database
vBNG	virtual broadband network gateway
vCM	virtual cable modem
vCMTS	virtual CMTS
VNF	virtual network functions
VM	virtual machine
YANG	yet another next generation

Bibliography & References

[1] ETSI GS MEC 001 V2.1.1 (2019-01): Multi-access Edge Computing (MEC); Terminology. ETSI Industry Specification Group (ISG) Multi-access Edge Computing (MEC)

[2] CNCF Cloud Native Networking Preamble. Web. https://github.com/cloud-native-principles/cloud-native-principles/blob/master/cloud-native-networking-preamble%20(1).md
[3] CableLabs Network Convergence. Web. https://www.cablelabs.com/network-convergence
[4] Kubernetes. Web. https://kubernetes.io/
[5] ISO/IEC 7498-l Information technology – Open Systems Interconnection – Basic Reference Model: The Basic Model. International Organization for Standardization and International Electrotechnical Commission
[6] 10G: The Next Great Leap in Broadband. CableLabs; Web. https://www-res.cablelabs.com/wp-content/uploads/2019/08/26064924/10G-The-Next-Great-Leap-in-Broadband1.pdf
[7] ETSI GS MEC 003 V2.2.1 (2020-12): Multi-access Edge Computing (MEC); Framework and Reference Architecture. ETSI Industry Specification Group (ISG) Multi-access Edge Computing (MEC)
[8] Matatyaou, Asaf. 10G Technology Network Virtualization. February 19, 2020. Web. https://broadbandlibrary.com/10g-technology-network-virtualization/
[9] Matatyaou, Asaf. Real-World Deployment of a Virtual Cable Hub. Publication. San Jose: Harmonic, 2017. Web. https://www.nctatechnicalpapers.com/Paper/2017/2017-real-world-deployment-of-a-virtual-cable-hub
[10] Remote PHY Specification, CM-SP-R-PHY-I15-201207. December 7, 2020. Cable Television Laboratories, Inc.
[11] Matatyaou, Asaf. Transforming the HFC Access Network with a Software-Based CCAP. Publication. San Jose: Harmonic, 2015. Web
[12] DPDK. Web. https://www.dpdk.org/about/
[13] CableLabs Distributed Access Architecture. Web. https://www.cablelabs.com/technologies/distributed-access-architecture
[14] CableLabs DOCSIS® 4.0 Technology. Web. https://www.cablelabs.com/technologies/docsis-4-0-technology
[15] ITU-T Recommendation G.9807.1: 10-Gigabit-capable symmetric passive optical network (XGS-PON). June, 2016. International Telecommunication Union
[16] IEEE Standard for Information technology 802.3av-2009: Local and metropolitan area networks; Specific requirements; Part 3: CSMA/CD Access Method and Physical Layer Specifications Amendment 1: Physical Layer Specifications and Management Parameters for 10 Gb/s Passive Optical Networks. Institute of Electrical and Electronics Engineers
[17] DPoE Architecture Specification, DPoE-SP-ARCHv2.0-I07-190213. February 13, 2019. Cable Television Laboratories, Inc.
[18] Flexible MAC Architecture (FMA) System Specification, CM-SP-FMA-SYS-I02-210526. May 25, 2021. Cable Television Laboratories, Inc.
[19] DOCSIS YANG Model. Web. http://mibs.cablelabs.com/YANG/DOCSIS/. Cable Television Laboratories, Inc.

Extended-CIN

A Remote Head-End Solution for Space Re-Allocation in CIN Deployment

A Technical Paper prepared for SCTE by

Deepa Phanish, Ph.D.
Technical Analyst
Cox Communications
6305-B Peachtree Dunwoody Rd, Atlanta, GA 30328
+1 404-664-8816
deepa.phanish@cox.com

Alan Skinner
Principal Engineer
Cox Communications
6305-B Peachtree Dunwoody Rd, Atlanta, GA 30328
+1 404-210-3122
Alan.Skinner@cox.com

John Huang, IP Engineer, Cox Communications

Igor Tavrovsky, Reliability Engineer, Cox Communications

Ernest Fabre, Design Engineer, Cox Communications

Cox's Distributed Access Architecture (DAA) standard calls for deployment of a Converged Cable Access Platform (CCAP) chassis acting as Remote PHY core in every metro edge facility. To support the deployment of these chassis, we need substantial amounts of rack space, power, and HVAC at these critical facilities. In locations that have severe constraints we need an alternative solution to enable Remote PHY while avoiding highly expensive facility augmentations. In this paper, we explore our network design options to deploy a CCAP chassis non-locally in a host facility and to utilize our Converged Interconnect Network (CIN) to reduce the footprint at the remote edge facility. We discuss how to prioritize these solutions by their impact to service availability. The reliability analysis provides further insights into the design/decision thresholds for selecting a host site in a successful application of the remote head-end Extended (E)-CIN solution. We further discuss the implementation, limitations, and challenges of the E-CIN solution and assess impact to the business in terms of capacity planning and cost in comparison to the full-CIN solution. The outcome from this comprehensive analysis is very useful in deciding favorability of a given site as an E-CIN candidate.

1. Introduction

The access network has continued to evolve by leveraging new technologies – the recent one being Remote PHY– to meet the ever-increasing demands for bandwidth. Remote PHY is a distributed architecture that encompasses moving the physical (PHY) component from the traditional Cable Modem Termination System (CMTS) out to the edge, thereby extending Ethernet closer to the customer and providing the capability to support greater bandwidth. This architecture will ultimately enable cable operators to deliver Gigabit service tiers at a fraction of the cost of replacing the existing Hybrid Fiber Coax plant with fiber. In this context, CIN is a flexible, resilient, and extensible network that interconnects the CCAP core with Remote PHY Devices (RPDs). It is essentially the infrastructure that supports the distributed access and fiber-deep architectures of Remote PHY.

Implementation of the traditional CIN network requires deploying the CCAP core within the service provider's critical facilities. These CCAP chassis have significant space and power requirements that would need to be accommodated at those facilities. An example of the CCAP core is the Cisco cBR-8 shown in Figure 1. It has the below chassis specifications:

- Weight: 429 lb. (195 kg) maximum fully loaded
- Height: 13 RU (22.75 in / 57.78 cm)
- Width: 17.45 in (44.32 cm) with no rack mounts, 17.65 in (44.83 cm) with rack mounts
- Overall Depth: 28.075 in (71.3 cm)
- Operating temperature (nominal): 32 to 104ºF (0 to 40ºC) Sea Level

The Cisco cBR-8 router can be either mounted on the rack at the front or in the middle. Also, the router can be either mounted on a standard 19-inch wide four-post equipment rack unit or a two-post rack unit. It is also power intensive with the below requirements:

- Cisco cBR-8 Lifetime Facility Power Requirement: 9000 W
- Hardware Facility Power Requirement (D3.0): 7300 W
- Hardware Facility Power Requirement (D3.1): 7900 W
- Average fully loaded chassis between 4500 and 5200 W

Figure 1 – Cisco cBR-8 CCAP, 13RU Height

To deploy the CCAP chassis and digitalize a facility that is constrained on space, power, and HVAC, we would need cost intensive facility expansions/augments which can run into millions of dollars. In certain cases, an expansion may not even be a feasible option, requiring office move and re-design of the outside plant with new fiber builds. In this paper, we explore the concept of Extended (E)-CIN, which essentially means the digital CCAP is hosted in a separate facility than the site that houses the Remote PHY Aggregation switches (RPAs). In other words, E-CIN involves the de-coupling of R-PHY core resources (i.e., cBR-8 data & video cores) from the edge of the CIN. This can enable digitalization of facilities with physical space constraints and thereby recover rack space from decommissioning obsolete analog equipment, thus avoiding the regrettable spend on expansions. Although physical space constraint is the primary driver, E-CIN can also optimize network efficiency by consolidating host resources and serve as a conceptual proof towards the adaptation of future designs like "centralized CCAP" and/or "virtual CCAP".

In the rest of the paper, we refer to the facility where the cBR-8 is located as the "host" site, and the facility where the RPA resides as the "remote" site. Inherently, in most cases, E-CIN will increase the optical distance between RPDs and the CCAP core, which can add unique challenges, primarily regarding latency performance and network reliability. In section 2, we first consider all the possible network topologies and perform a reliability analysis to derive insights on the best choice of the host site for a given remote site. In section 3, we discuss the implementation of the designed solution. In section 4, we discuss the latency and throughput performance of E-CIN. In section 5, we discuss the financial impact to business for capacity planning and finally conclude in section 6.

2. Network Design

Important design questions on hand for E-CIN are:

1) What are the possible topological solutions for E-CIN?

2) How does one choose the most optimal solution and host site for a given remote site?

In this section, we first explore the topological possibilities by considering both L3 Internet Protocol (IP) and L1 optical Dense Wavelength Division Multiplexing (DWDM) layers. Secondly, we present an approach based on the analysis of network reliability to draw insights on the choice of host site.

2.1. Network Topology

The illustration in Figure 2 is representative of the CIN implementation within the Cox network. HUBC0x are the Metro hub routers controlling metro traffic, and DSRJ0x are the Distributed Service Routers connecting to the backbone network. The Remote PHY Aggregation switches (RPAs) aggregate the Remote PHY Devices (RPDs), and the Digital Physical Interface Card (DPIC) Aggregation switches

(DPAs) aggregate the connections from DPIC providing an ethernet network between the CCAP core and the RPDs. Note that the Host Site could also be the Regional Data Center (RDC).

For the problem on hand, we need to establish communication between the RPDs at the remote site and the CCAP within the host site. There are a couple options to consider:

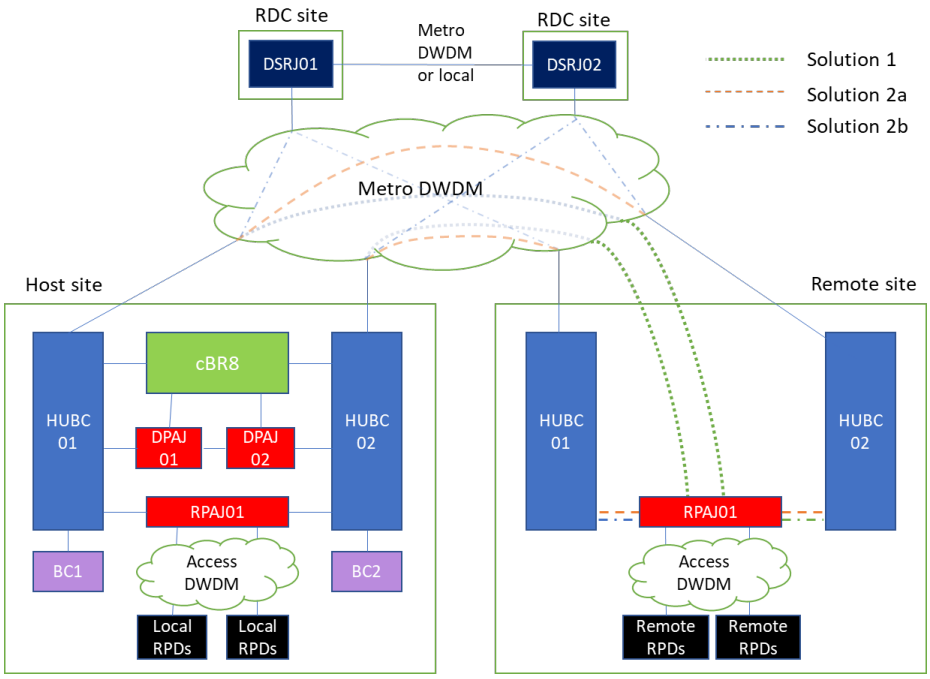


Figure 2 – Generalized Cox Metro Network Topology Showing Routing Options for E-CIN

Solution 1: One way to achieve this would be to leverage the DWDM metro optical network to transport traffic directly between the remote site RPA and the host site hub routers. In essence, the RPAs are re-homed manifesting as the expansion of the host site's access network.
Pros: This would be the least hop solution with low latency, and thus more reliable.
Cons: The solution is not scalable, as aggregation over the DWDM network is very expensive. Therefore, we do not discuss this solution any further in this paper.

Solution 2: The second possibility is to route the traffic via the hub routers. Traffic from the CCAP core hops via the hub routers in the host site before being transported over the metro optical DWDM network to the RPAs via the hub routers at the remote site.
Pros: Provides a scalable solution by aggregating traffic over the uplinks of the hub routers.
Cons: The solution has more hops, and consequently higher latency and lower reliability in comparison to the first solution.

- **Solution 2a**: The hub routers at the remote site can be directly connected to the hub routers at the host site over the metro DWDM network without passing traffic through the Distributed Service Routers (DSRs). This option creates a spur from the host site to the remote site, rather than utilizing the existing hub & spoke architecture. This is achieved by provisioning a primary wave and a secondary protecting wave between the hub routers at the two ends.
  Pros: Fewer hops in comparison to solution 2b.

Cons: Unless there is a pair of direct fiber between the two sites within the optical network topology, it results in a non-optimal aggregation over the DWDM network. Additionally, this alters the IP network topology from the standard (where uplinks connect back to the DSRs at RDC).

- **Solution 2b:** The hub routers at the remote site are not directly connected over to the hub routers at the host site, so the traffic instead hops via the DSRs over the metro DWDM network.
  Pros: Standardized metro topology in alignment with full CIN. Optimal aggregation of waves over the metro DWDM network.
  Cons: More hops and likely longer latency in comparison to solution 2a.

In summary, since scalability, DWDM aggregation and standardization are of top business priorities, our preferred solution is to reserve solution 2a for sites that have a direct fiber pair to another site. This implies we have a single node ring L1 optical network (L3 is secondary spur from primary set of hub routers), which would be the host in this case. For the more common multi-node ring topology of the L1 metro DWDM network (L3 topology is hub-and spoke from DSRs), solution 2b would be more preferential. Selection of host site in this case requires further analysis.

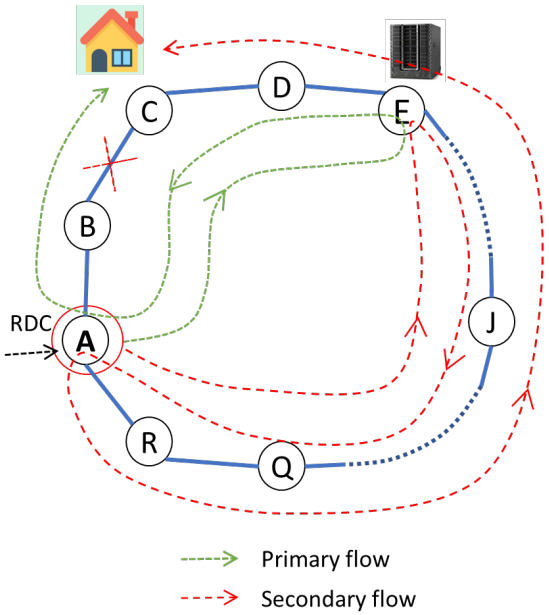
2.2. Reliability Analysis

Separating the CCAP core in one facility from the edge components in another facility increases the number of network elements, and specially fiber distance between the customer premise and the headend. This compels an evaluation of service availability to validate the Service Level Agreement (SLA) requirements, as well as draw insights into the impact of CCAP location on the network reliability.

2.2.1. Problem Definition

In a metro DWDM network, although mesh topologies exist, it is more common to have the ring topology to uplink the hub routers to the DSRs with a primary wave and a secondary protection wave. Considering the worst case - What is the impact on service availability as we vary the CCAP location?

Let's take a deeper look into the traffic path over the metro core in case of solution 2b. The traffic path over an optical ring is shown in Figure 3 for both the steady state and fail-over cases. Considered here is a ring with 18 sites (A to R) spanning over a total mileage of ≈1300km with a single RDC location, akin to one of the largest metro rings on the Cox network. In steady state, the traffic from backbone starting at the DSR in the RDC traverses the shorter side of the ring to the CCAP location, again on the shortest path, returns to the RDC before finally reaching the customer site where RPDs are located. The flow is similar during failover, except that it traverses the longer sides of the ring. In the example shown in the figure, A is the RDC site, E is the site where CCAP head end is located, and C is the customer site of interest. It shows the primary path of traffic flow, and the secondary path when a failover occurs, say on the event of a fiber cut between sites B and C. For this exercise we vary the CCAP and the customer site over the ring and perform a reliability analysis of the metro network.



Fiber between locations	Fiber (kms)	Metro Fiber (kms)	Longhall Fiber (kms)
A -> Rs	30.7	24.56	6.14
R -> Q	65.5	13.1	52.4
Q -> P	111.7	11.2	100.5
P -> O	91	9.1	81.9
O -> N	86	12.9	73.1
N -> M	67.9	13.6	54.3
M -> L	55.5	5.6	49.9
L -> K	93.5	14.1	79.4
K -> J	87.9	17.6	70.3
J -> I	111.2	83.4	27.8
I -> H	70.7	28.3	42.4
H -> G	124.4	24.9	99.5
G -> F	58.3	43.8	14.5
F -> E	31.5	4.8	26.7
E -> D	99.8	15	84.8
D -> C	66.3	10	56.3
C -> B	63.8	12.8	51
B -> A	53.5	21.4	32.1
Total	**1,369.20**	**366.2**	**1,003**

Figure 3 – Protected Traffic Paths on a Metro Ring with the Fiber Distances

2.2.2. Modeling and Simulation

For reliability analysis we used the ReliaSoft BlockSim package allowing the creation of reliability block diagram and discrete repairable system event simulation. The fiber distances between sites for the case study is shown in Figure 3. Outside plant fiber mileage is distinguished between metropolitan and long-haul by their urban density for analysis since optical impairments (disruption of signal in the fiber optical link) largely depends on human activities, and they have different failure rates. The failure rates are modeled with lognormal distribution computing the means and variance for long haul and each metro individually. The customer service availability is then modeled with the primary and secondary path elements in a parallel system with all intermediate routes and optical infrastructure at the module level with built in redundancies.

Disregarding the “last mile” OSP infrastructure within the access network, we have simulated the service availability for customers at different sites on the ring as we vary the CCAP locations over the ring as well. For this comparative analysis, the Mean Time to Repair (MTTR) is assumed constant and equal to 4 hours. The results are shown in Figure 4. It can be observed that for a given CCAP location, network reliability is highest when the headend is local and decreases as the CCAP is moved farther away. Secondly, for all sites, the network reliability when the CCAP is located at the RDC site A is almost as high as when it is local. This is apparent from the figure as the line plot for “CCAP in A” is an envelope for the rest. It is due to the nature of traffic flow - locating the CCAP at the RDC does not add any additional fiber distance to the traffic path. Hence, the RDC would be the second choice, next only to local, from the perspective of network reliability. Further, if the RDC cannot host the equipment, then the facility closest to it needs to be considered.

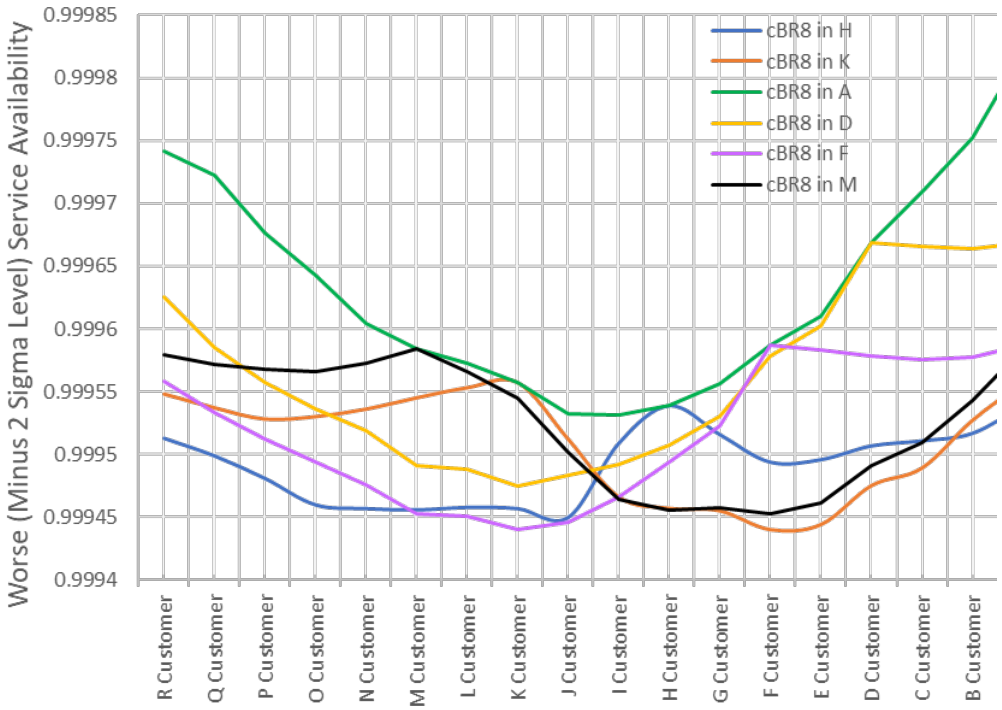


Figure 4 – Metro Core Service Availibity at Different Sites for Fixed CCAP Location

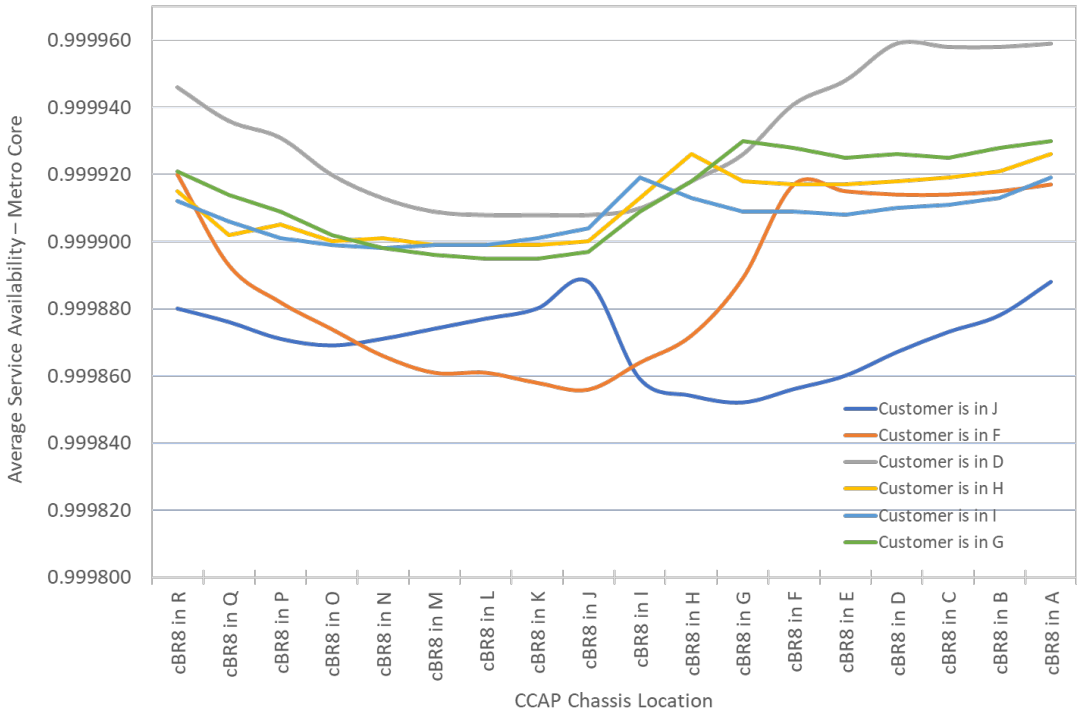


Figure 5 – Metro Core Service Availability at Specific Sites for Different CCAP Locations

In Figure 5, we can see that the inverse relation has a similar trend. It should be noted that the plots consider only the metro core network reliability and does not include the backbone and the Access OSP infrastructure. To qualify a host site, we need to verify that the absolute reliability for the entire network meets the SLA. Due to shape of the plot, it is possible for certain sites within the market to meet the requirements while certain others beyond a distance do not. In the above case study, we performed a detailed analysis including the "last-mile" R-PHY OSP infrastructure at Cox by considering the hardware, software, human factor, and commercial power outages. To be more accurate, we have modeled the MTTR with log-normal distributions with the log location $\mu = 3.3.4576$ and the log standard deviation $\sigma = 0.5287$. Similar distributions are used for RDC and hub site hardware with tailored parameters. This is a computationally very intensive model with over 1000 blocks simulating 5 years of operation and at least 5,000 iterations for acceptable convergence. For remote site F as example, the mean availability for CCAP location in RDC A resulted in 0.99964, and we can therefore qualify A as the host. We have observed a consistent drop in mean availability compared to Figure 5 by about 0.00028 for all host sites with the improved accuracy model.

2.3. Summary

In summary, directly linking the remote site RPA to the host hub routers (solution 1) is not scalable. To handle the RPD scale and to benefit from aggregation, the CIN traffic needs to hop via the remote site hub routers before passing through the DWDM network. Given the optical network topology, the remote site hub routers may have direct primary and secondary uplink to the hub routers at another site, if they have a pair of direct fibers between them. In this case, the second site would be the chosen host site. This is the "***Subtended hub-hosted***" solution. Generally, we have sites with optical nodes in the metro connected as a ring with one or two (split) RDC sites hosting the DSRs connecting over to backbone. In the standard metro IP topology with primary and secondary uplinks from the hub routers within the market to the corresponding DSRs, these uplinks are non-overlapping DWDM waves provisioned over the optical network on the primary and secondary fiber paths of the ring. Hence all traffic from a remote site hops over the RDCs before reaching the host site. Therefore, the RDC itself would be the primary choice for the host site in E-CIN. This not only maximizes reliability as shown in the above analysis, but also minimizes latency. This is referred to as the "***RDC hosted***" solution. In cases where the RDC facility does not meet the requirements for space, power, and HVAC, then the hub site closest to the RDC in fiber distance would be the next choice of preference for a host site. This is most generalized but least optimal, and it is called the "***Hub-hosted***" solution.

3. Implementation

In this section, we discuss the implementation of E-CIN and the necessary configurations specific to E-CIN deployment. Firstly, although general CIN routing guidelines apply, IP addressing needs to be tailored appropriately for successful operation. Secondly, supporting video service becomes more complex due to multiple channel lineups at the host core and requires additional configurations on the CCAP.

3.1. Networking

The host site will have the CCAP, Digital Physical Interface Card (DPIC) Aggregation switches (DPAs), and the Boundary Clocks (BCs). The remote site will have the "edge" CIN equipment: one or more RPAs, Remote PHY access DWDM equipment, and switches for local management/telemetry. However, remote sites will not have local boundary clocks for Precision Time Protocol (PTP); the boundary clocks at the host site serve the purpose instead.

In terms of routing, general CIN guidelines apply with appropriate updates to the IP space addressing and route advertisements to ensure reachability between the RPA, CCAP core and boundary clocks. Route advertisements specific to the Cox network is shown in Figure 6, and the same routing policy works for all solutions of E-CIN. In a typical Remote PHY network, a facility has two PTP boundary clocks and each is connected to a different router which in turn resides either on the "Hub1" or "Hub2" side of the network. BC1 is preferred by all RPDs on the CCAP, because the standard Remote DOCSIS Timing Interface (R-DTI) profile points to it. BC2 only comes into play if BC1 is unreachable. If the primary path between host and remote sites utilizes the Hub Router 1 path, then this standard CCAP configuration is used. If the optimal path between host site and remote sites uses the Hub Router 2 path, however, then BC2 is preferable because it is closer to the RPDs. In this case, a second R-DTI profile is configured on the CCAP and the RPDs in the remote site must use that profile instead of profile 1.

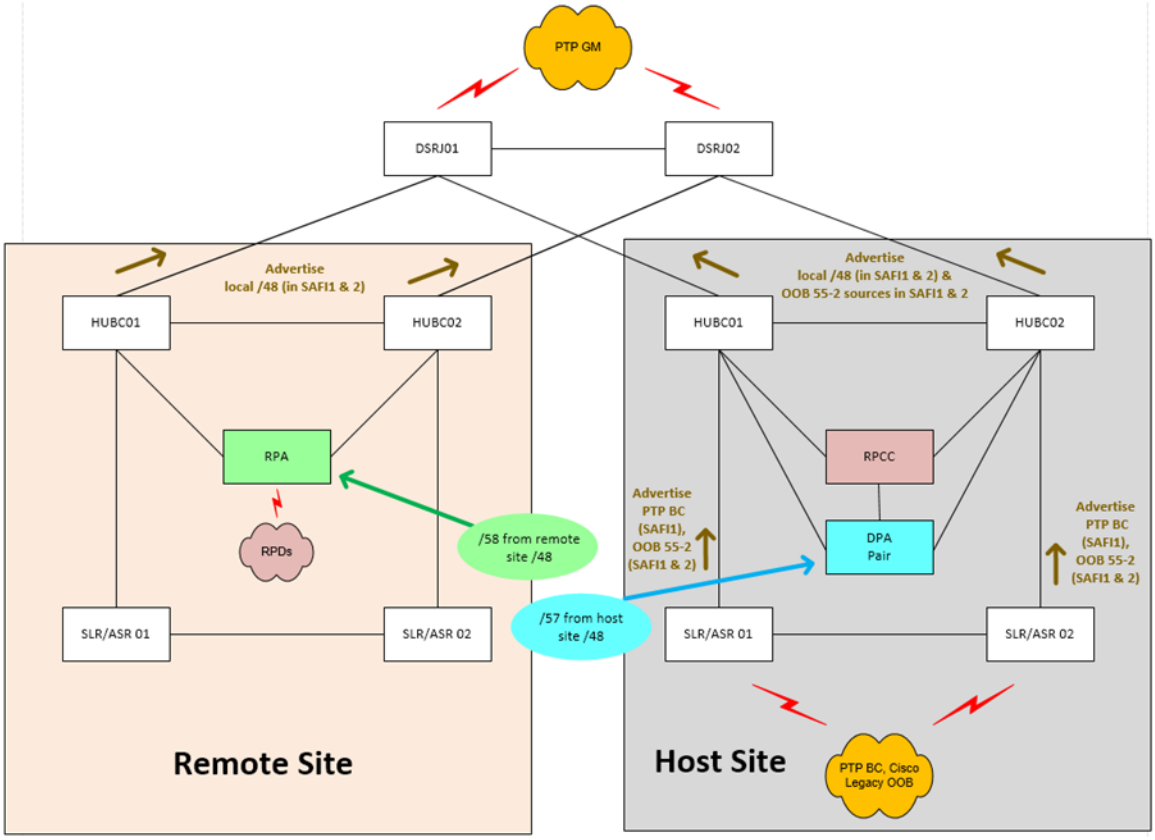


Figure 6 - Route Advertisements in the CIN Network

3.2. Video Support

One of the limitations to E-CIN is the increased operational complexity for supporting video services. In the case where the remote site has the same channel lineup, ad zones, DOCSIS Set-top Gateway (DSG) tunnels, legacy Out-Of-Band (OOB), and Public, Educational, and Government (PEG) channels as the host site, there may be no additional configurations required to support Quadrature Amplitude Modulation (QAM)-based video services including broadcast video and narrowcast video (Switched Digital Video (SDV) and Video on Demand (VOD)). However, if any of the above differ between the host and remote site, then care must be taken to configure the CCAP accordingly.

Consider a Synamedia PowerKey market as an example. Legacy OOB is a Narrowband Digital Forward (NDF) signal delivered via multicast from a Kronback NDX source to a set of destination RPDs. The legacy OOB is assigned to a Digital Hub on the Explorer Controller. It is imperative that both the legacy OOB NDF signal and the DSG multicast flows to the CCAP are assigned to the same Digital Hub on the controller. Failure to do so will render Basic DSG Tuning Adapters inoperable. A CCAP hosting an E-CIN will have to be evaluated to determine whether the remote site is serviced by a different Digital Hub. If so, the CCAP will require a matching NDF pseudo wire for each set of DSG multicast flows.

For any multicast video transport streams sourced at the extended facility, those services will need to be routed back to the host CCAP so they can be ingested and mapped to an output like any other video source. Additionally, care must be taken when supporting multiple Ad Zones in a single CCAP. Multiple Ad Zones means that not all narrowcast video service groups are equal. This calls for additional coordination and operational processes. Before a narrowcast service group is associated with its first RPD, it must be first determined to which Ad Zone the RPD belongs, and then the correct zone should be associated with the service group. Further, any other RPD that gets associated with that service group must also belong to that same Ad Zone.

Extended CIN is limited to environments contained within a single market. There are several reasons for this, but the primary factor is video support. Because all video in a R-PHY network is ingested by the CCAP and then re-generated toward the RPDs, it is not practical (or even possible, in many cases) to support channel lineups and video encryption from multiple markets on the same CCAP. Within a market, when the channel lineup in an extended site differs from the lineup in the host site, the CCAP must be configured to support both, which can add additional complexity. Sometimes, it may not be advantageous, or even possible, to group some sites together. In such a case, multiple CCAP chassis would be needed at the host site to support E-CIN.

Here are some of Cox's best practices on CCAP configuration. Some of these could be relaxed or eliminated if a standalone video core were used in place of a converged data/video CCAP:

- No more than 6 full broadcast service groups per CCAP.
- No more than 12 total broadcast service groups per CCAP including PEG service groups.
- Only 1 Conditional Access System per CCAP. Sites tied to different set-top box controllers cannot be supported on a single CCAP.
- Except for very specific circumstances, only 1 main SDV lineup should be supported on a CCAP. Within the CCAP video config, its linecards are pointed to a SDV session server. The session server is associated with a main lineup. While it is technically possible to point different linecards at different SDV session servers, this should only happen for very small sites whose total capacity needs can be satisfied with 1- or 2-linecards on a single CCAP.

When a host facility is planned to support one or more extended facilities, it is advisable to consider the following options:

1. Support for all extended sites on each CCAP in the host facility. This means that all CCAPs are essentially created equal, and RPDs from any extended site can be moved to any host CCAP. This option is the most flexible but comes at a cost of complexity on the CCAP, and potentially reduced Data Over Cable Service Interface Specification (DOCSIS) Service Group capacity.
2. Segregation of CCAPs by serving footprint. This option would involve setting aside one or more CCAPs as being "E-CIN" hosts and mapping the extended site(s) to certain CCAPs. Doing this would reduce the amount of waste involved in the universal configuration, but at the expense of having to keep track of RPD mappings.

3. Cox is also considering a third option to consolidate video onto a standalone dedicated video core. In this scenario, only the DSG tunnel configurations on the Data CCAPs would be needed to host multiple remote facilities. The CCAPs could therefore be fully utilized for DOCSIS.

4. Performance – Latency, Throughput, and Distances

There are many factors that contribute to a customer's perceived latency. Only the contribution of the access network (CCAP to cable modem) is considered here. Figure 7 illustrates the DOCSIS request/grant cycle and the impact of E-CIN on the same. It is critical to have symmetric latencies from core to RPD and vice versa for both PTP and DOCSIS traffic for Remote PHY operation. It can be observed that the optical distance between the RPD and the CCAP core is the primary factor influencing this latency performance.

The latency between core and RPD is higher in an E-CIN environment primarily because of the additional fiber distances from the metro DWDM network. In a lightly loaded system, where congestion is not a factor, customer ping times to the CCAP are generally 5-6 ms in the best case. Those times start to go up almost linearly as you add distance between the core and the RPD as shown in Figure 8. Typically, a non-Extended hub site can serve RPDs as far as 100km away with only minimal impact to ping times. With Extended CIN, that distance from hub to RPD must be added to the distance from host facility to remote site, which could be as high as 1200km. Because upstream scheduling is done in the core and not in the node, a customer's best-case ping time to the CCAP increases linearly by 4X the latency added by Extended CIN because of the DOCSIS request/grant cycle. For example, if 400km (2ms) is added to the distance between the host facility and the remote facility, the customer will see an increase of 8ms in their ping times (best case).

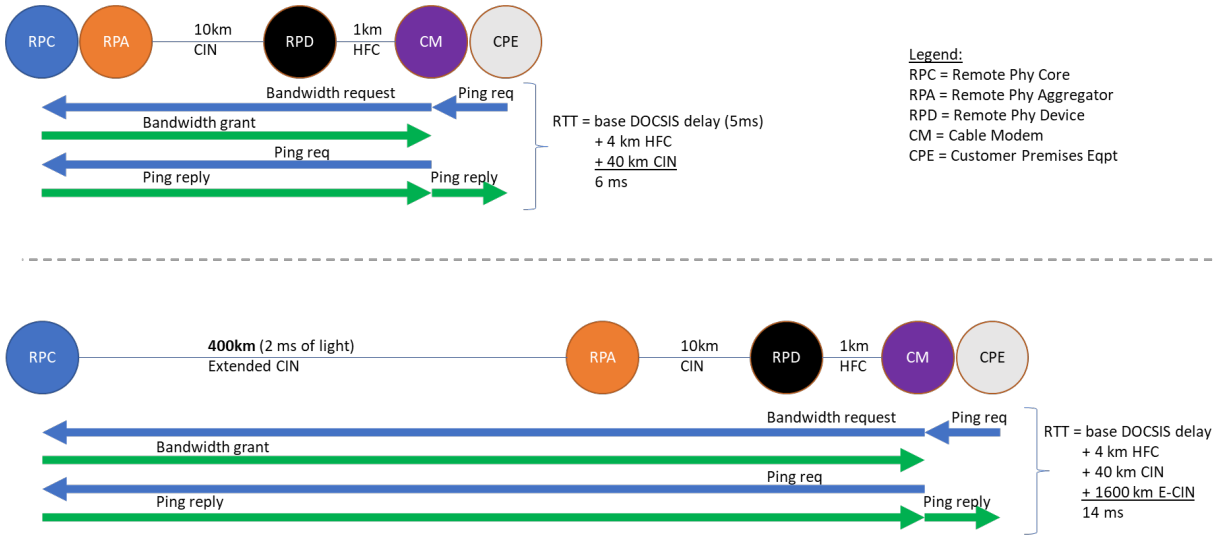


Figure 7 – DOCSIS Ping Time from CPE to CCAP in Regular vs Extended CIN

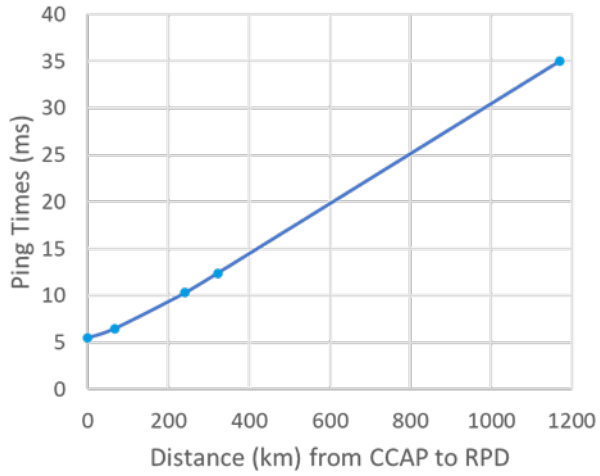


Figure 8 – Latency with Distance

Since DOCSIS is a timing-dependent technology, it is essential to maintain optimization, both in terms of preferring the shortest path (in steady state) and ensuring symmetrical (forward & return) traffic flow. The degradation in latency performance can be mitigated by ensuring that all Remote PHY traffic takes the optimal path from core to RPD, and from RPD back to core, without being load-balanced across multiple paths. This may require change in the Internet Gateway Protocol (IGP) metrics of the metro IP layer to default to the shortest path. Only during a failover should the suboptimal path come into play.

Field testing has shown that customers should be able to achieve full downstream and upstream throughput, even on the Gigabit tier, at distances up to 320km. Field experimentation is currently in progress to determine the exact distance limitation to preserve full Gigabit download. In the very worst-case scenario, which occurs during a path failover, at 1200km, downstream throughput was inconsistent and did not reach the usual gigabit-class downstream. The modems stayed online and continued providing service in a degraded state. However, use of secondary path needs to be minimized for achieving better throughput as well, in addition to latency. Limited field testing has shown that during a failover from short to long path, or from long to short, RPDs and cable modems generally stay online. Nevertheless, large scale testing would be needed to fully validate whether customers could see an interruption in service while their modems re-register.

5. Impact to Business

5.1. Capacity Planning

Depending on the E-CIN topology, the design may involve some “double-back” traffic across certain segments/links of the metro network. Therefore, bandwidth requirements need to be adjusted carefully. Let’s consider the generic hub hosted design for discussions in this section.

Traffic flow between Remote PHY core and the Extended CIN edge over metro core is unique in that it is deterministic and not load balanced. All the traffic – multicast, forward unicast, return unicast, and PTP – must always take the same path, for every RPD. This ensures traffic routes through the shortest path and there is symmetrical forward and return path latencies, which are necessary to meet the timing requirements for DOCSIS. Figure 9 shows how this can be achieved by designing the IGP metrics appropriately. The red path shows load balanced traffic, whereas the green path shows the L2TPv3 tunnel

with a preference on HUB1 side of the network. The optically shorter path, either via HUB1 or HUB2, needs to be engineered as the preferred default and the other as the failover.

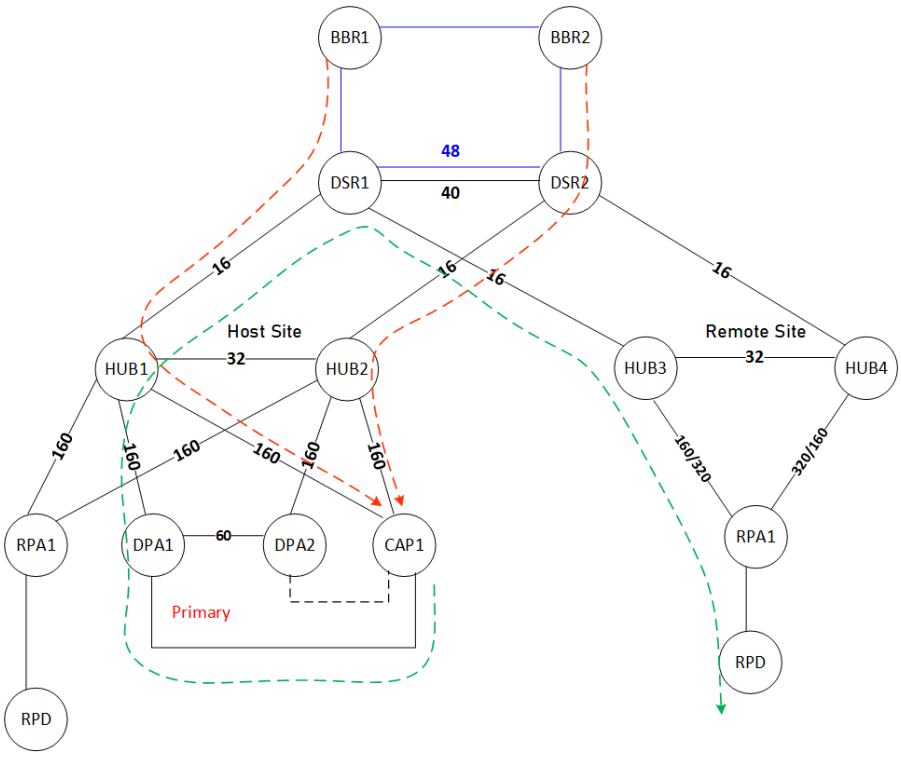


Figure 9 – Traffic Engineering with ISIS Metrics

In planning capacity for the metro network, the CIN component of the traffic needs to be monitored separately. For the metro-core links, we have the below capacity calculations (All traffic rates would be 95th percentiles in Gbps):

Host hub router to DSR uplink capacity = $\lceil max(Td_H + TCIN_R/2\ , M * (Tu_H/2 + TCIN_R/2)) \rceil$

Remote hub router cross bar capacity = $\lceil Td_H/2 + M * TCIN_R \rceil$

Remote hub router to DSR uplink capacity = $\lceil max(Td_R\ , \frac{M}{2} * (Td_R + TCIN_R)) \rceil$

Remote hub router cross bar capacity = $\lceil Td_R + TCIN_R \rceil$

Where Td_H is the total downstream traffic to the hub routers at the host site, Tu_H is the total upstream traffic from the hub routers at the host site, Td_R is the total downstream traffic to the hub routers at the remote site, $TCIN_R$ is the CIN component of the total downstream traffic to the hub routers at the remote site, and M is the bandwidth margin on the router uplinks. It is a good practice to set M=1.5, which means that the fill rate is below 66.66% in steady state operation. This margin is required to ensure healthy operation, since tunneled traffic - unlike load-balanced - would otherwise operate at 100% rather than 50% in steady sate unable to absorb temporary spikes.

Under special circumstances such as split-RDCs where DSRs in different locations, a single fiber cut could force traffic over the DSR crossbar, in which case, it needs to be augmented as well.

5.2. Cost Benefit

From a cost perspective, E-CIN has a wide range of possibilities. On one end, in cases where we could save expenses on facility expansion it could benefit the company, whereas, on the other end, where traffic rates are significantly large, E-CIN could be more expensive than regular CIN. Cost benefit is generally at best where facility augments can be avoided. Secondly, if resource sharing is maximized, remote site can share the CCAP chassis, its processor, and spare line cards; the boundary clocks; and the DPAs at the host. This could result in additional savings. However, depending on the video configuration, it may not be possible to fully leverage resource sharing, especially in cases where dedicated video core may be necessary.

E-CIN also comes with an additional expense – it needs more bandwidth capacity and hence growth expenditure on the metro core as discussed in the previous section. Bandwidth augments are not only necessary to address the double backing of CIN traffic, but also the tunneling effects viz., no load balancing, and operational margin. The main component here is the transport network cost, which depends on the traffic rates, optical network platform, transponder wave density, and its topology. Unless host and remote site traffic rates are very low, augments may be necessary on the uplinks of either one or both the hubs. We may need to typically add 2-4 wave augments on either end. Therefore, the financial impact for each remote and host site pair candidate needs to be assessed on a case-by-case basis to precisely estimate and compare costs.

6. Conclusion

Extended-CIN provides a novel technology to geographically decouple the Remote PHY core and the edge on already existing CIN infrastructure. It is useful reducing the footprint on a facility by consolidating core resources and therefore avoiding expensive augments. However, it can add unique challenges, primarily regarding network reliability, latency performance and operational complexity for video support. These risks should be minimized by optimally choosing the design and host sites with additional case specific analysis. It would be preferable to reserve E-CIN for cases where optical distances between the remote site's edge and the host site's core are relatively low and the cost benefit is high. As part of the future work, we will be evaluating the significant and applicability of E-CIN as newer technologies such as Remote MAC-PHY and virtual CCAP arise.

Acknowledgements

The authors would like to thank Deependra Malla, Jason Cole, Bill Wall, and James Stockdill at Cox Communications for useful discussions.

Abbreviations

BC	boundary clock
CCAP	converged cable access platform
CIN	converged interconnect network
DAA	distributed access architecture
DOCSIS	data over cable service interface specification
DPA	dpic aggregation switch
DPIC	digital physical interface card
DWDM	dense wavelength division multiplexing

DSG	docsis set-top gateway
DSR	distributed service router
E-CIN	extended-converged interconnect network
IGP	internet gateway protocol
IP	internet protocol
MTTR	mean time to repair
NDF	narrowband digital forward
OOB	out-of-band
OSP	Outside plant
PEG	Public, educational, and government
PHY	physical
PTP	precision time protocol
QAM	quadrature amplitude modulation
RDC	regional data center
R-DTI	remote docsis timing interface
RPA	remote phy aggregation switch
RPD	remote phy device
SDV	switched digital video
SLA	service level agreement
VOD	Video on demand

Bibliography & References

Cable Television Laboratories, Inc., 2015. *Distributed CCAP Architectures Overview Technical Report*. Cable Television Laboratories, Inc

Chapman, J.T., 2013. *DOCSIS Remote Phy*. White Paper, Society of Cable Telecommunications Engineers.

Malla, D., 2021. *Modernizing Cox Communication's Access and Aggregation Network for PHY Deployment.* Society of Cable Telecommunications Engineers.

Datasheet, cBR-8 Converged Broadband Router, Cisco

Fastest Path to Low Latency Services

How Can Cable Operators Deliver Consistent Latency by Following an Efficient and Future-Proof Design Path?

A Technical Paper prepared for SCTE by

Sebnem Ozer, Ph.D.
Senior Principal Architect
Comcast
1800 Arch St., Philadelphia, PA 19103
2152868890
Sebnem_Ozer@comcast.com

Aaron Tunstall, Engineer 3, Enterprise Data & Analytics, TPX NGAN Access Eng/Comcast

Carl Klatsky, Principal II Engineer, Prodt Dev Engineer, TPX CPT NCE/Comcast

Dan Rice, VP, HFC Architecture, TPX NGAN Access Eng/Comcast

Jason Livingood, VP - Technology Policy & Standards, TPX CPT NCE/Comcast

John Chrostowski, Executive Director, NGAN Access Eng, TPX NGAN Access Eng/Comcast

John Raezer, VP, XCS Strategy, Planning, Connectivity & Consumer Experience Eng/Comcast

Joshua Gerson Sr. Mgr., XCS Strategy & Planning, XCS Strategy & Planning/Comcast

Mulbah Dolley, Eng 2, Technl Research & Dev, TPX NGAN Access Eng/Comcast

Priyan Sarathy Sr Mgr, Enterprise Data & Analytics, TPX NGAN Access Eng/Comcast

Sarulatha Subbaraj Engineer 4, Enterprise Data & Analytics, TPX NGAN Access Eng/Comcast

Soomin Cho, Data Engineer, TPX CPT NCE/Comcast

Trevor Graffa, Engineer 4, Software Dev & Engineering, TPX RDK/Comcast

1. Introduction

The requirements of emerging interactive real-time services and changes in online usage patterns impose entirely different network challenges that Internet Service Providers need to overcome. Cable networks are going through a big transition to the next-generation 10G technologies with substantial speed increases, that can meet the online traffic volumes created by these services accumulatively. However, the new quality of experience judge will not praise or condemn the network operators only by their speed but also by their consistent support of low latency. Therefore, 10G technologies need to address a fundamental redesign of traffic classification and latency monitoring, prediction and optimization. Unprepared Internet Service Providers (ISPs) that design their architecture for mean and median values instead of peaks cannot support the interactivity of real-time services and cannot avoid the impact of these huge data volumes on other services.

In this paper, we will discuss the low latency services that are still evolving today, such as cloud gaming, video/voice conferencing and live video streaming, as well as emerging applications with progressively more interwoven human and machine interactions. We will first cover current network features and tools that can be used to measure, monitor and manage the latency of today's networks. We will then describe the next steps to support new Low Latency (LL) services by applying D3.1 features and a LL service differentiation framework. To support the LLD (Low Latency DOCSIS) features, traffic classification and monitoring must be redesigned and inherent rules may need to be replaced. Lastly, we will provide guidelines to deliver low latency services with the most efficient and future-proof investments.

2. Low Latency Services and Requirements

Not only do we experience continuous technological advancements and breakthroughs but we also observe faster democratization of technologies. Improved products and user experiences on interactive real-time services, IoT and sensor-based systems with big data learning, immersive applications, and autonomous systems altered the landscape of network traffic as consumers have easier access to these products and services. Mass production, digitization, software-defined, virtualized and cloud-based systems with open source software, platform models with partners and co-innovators have been key in the democratization of these technologies and building blocks of digital native companies. Legacy companies cannot survive if legacy chains are not broken for a digital transformation to meet the consumers' demands. Consumers are so immersed in the new technologies that they expect good quality, and nothing infuriates them more than if they don't work [1]. A technology that doesn't work for a consumer means bad Quality of Experience (QoE) [2].

Cable operators continue increasing connectivity speeds to improve the QoE of their subscribers through a series of new technologies and deployments, such as widening the upstream path via mid-split and high-split spectrum, and applying Full Duplex (FDX) architectures and distributed access networks. Speed as a performance metric has been regularly measured by MSOs, but speed is only one of the performance indicators. Different services and applications require different levels of Quality of Service (QoS) metrics, such as speed, latency, jitter, packet loss, reliability and security. Recently, many specifications and standardization documents from various networking technology organizations have been updated to include new traffic categories and QoS levels. Latency and latency variation (jitter) definitions and measurements are not as unified and standardized as speed (throughput) and packet loss. In the following sections, we will describe latency and jitter metrics that ISPs should monitor and improve for low latency services and thresholds, as defined in various standards and specifications.

2.1. Low Latency Services

For an efficient and future-proof design, cable operators must have a solid understanding of current and emerging services with reliable traffic forecasting. However, as we have seen during the pandemic, traffic forecasting and emerging services may not be always foreseeable. Therefore, it is crucial to design an agile system that can adapt to the changes faster and establish an accurate assessment of consumers' new QoE factors. Below we define the key points for current and emerging low latency services.

Real-time Gaming: Gaming lag means a delay between pressing a command button and the game responding on-screen. Network latency is one of the sources of gaming lag. Gamers are particularly sensitive to latency performance (a.k.a. jitter) when competing in multi-player online games, such as *League of Legends*, *Rocket League,* and *Fortnite*. A common complaint from gamers about latency is that the connection "lags out" during gameplay; even millisecond-level differences can make an impact for players competing in multiplayer online games. Latency variation (jitter) that lag-compensation algorithms cannot mitigate may reduce the gamer's QoE significantly. Additionally, the time difference of responses received by different multi-players causes unfairness [4]. Many game applications monitor and report measured latency and jitter, as shown in Figure 1, where players can compare the gaming performance to monitored latency.

Figure 1 – Ping Reports on the Gaming App

Cloud Gaming: Gamers send commands from a mobile device to cloud gaming platforms that execute those commands and then stream the results back to the gamer [6]. Network latency and jitter in the downstream affect the streaming and chat quality and while latency and jitter in the upstream affect the user input reception time and chat quality. Overall, high latency and jitter cause lag during play, choppy audio, poor video and distorted chat.

VideoConferencing: As more people have been working or studying from home, the quality of video and audio conferencing became essential in everyday life. High latency and jitter cause time lags, loss of lip synchronization and choppy or frozen video and audio [4]. Media and file sharing time can be also affected if the network quality is low.

Real-time interactive video streaming: While buffered streaming can be affected by high latency and jitter, the requirements for real-time interactive video streaming are stricter [3]. Services such as sports

betting, watch parties, shopping and synchronized second screen depend on an end-to-end platform with ultra-low latency and jitter. High latency and jitter can cause a subscriber to miss the hard cut-off times to place a bet before a game or a shopping window time. It can also cause spoiler issues during watch parties [8].

New services with web browsing are also sensitive to latencies on the order of hundreds of milliseconds [5]. Other emerging low latency services such as Holographic Type Communications, Multi-Sense Networks, Time Engineered Applications and Critical Infrastructure Services exposed several deficiencies in current network technologies that need to be addressed for future-proof deployments. [7]

2.1. Latency and Jitter Definition

Network latency (delay) is defined as the total time it takes for a data packet to travel between two networking points. One-way latency is the time required for a packet of data to travel from the sender to the receiver while Round trip time (RTT) is the time required for a packet of data to travel from the sender to the receiver and back again [5].

Network jitter or latency/delay variation refers to variation in the latency of arriving packets over time. Inter packet delay variation is the difference in latency of each received packet as compared to the previously received packet, while packet delay variation is the difference in latency of each received packet as compared to one reference value such as minimum or average latency [5].

The QoE depends on the end-to-end network latency and jitter while each network hop can be monitored and managed as discussed in [2].

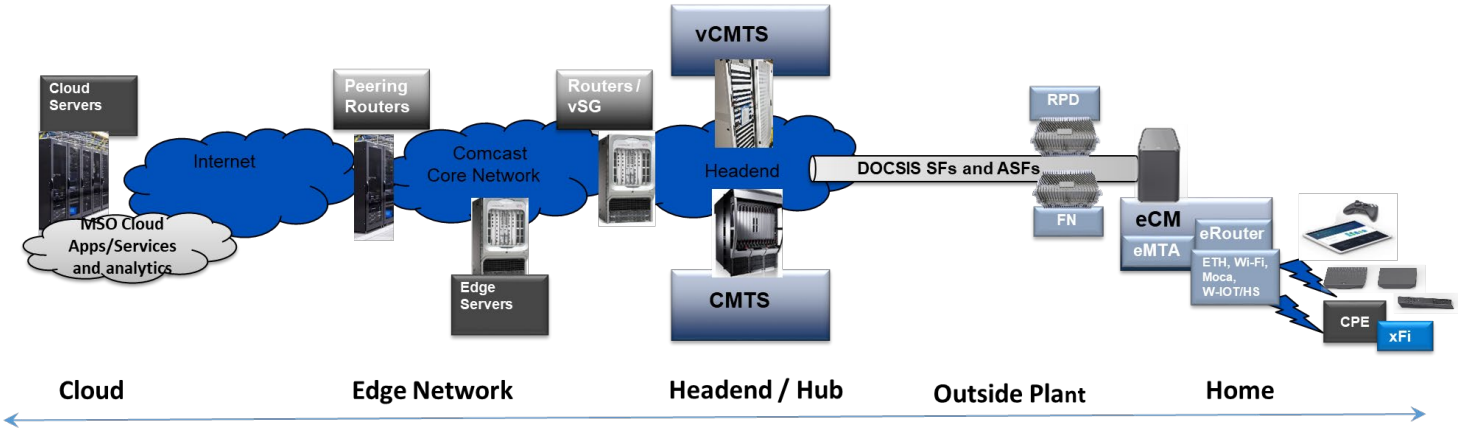


Figure 2 – Cable Network Segments

The low latency services in Section 2.1 require consistent latency, hence well-bounded jitter levels. It is important to assess the worst-case latency in the network while other latency and jitter measurements can help to analyze the latency sources and components. **Idle latency** measures responsiveness when a network connection is unused. It is mostly correlated to access network and distance of the path (round

trip time) e.g. fiber vs. Hybrid fiber-coaxial (HFC) vs. satellite. Many wireline network differences are insignificant. **Working latency (a.k.a. latency under load)** is the real-world measure of responsiveness when a network connection is actively used. Responsiveness of real-time applications during moderate usage of a network connection, whether upstream or downstream. When it gets really bad, it is often called “Buffer Bloat.”, e.g. when gaming or video conference is interrupted by large file download or many devices in homes. The worst-case latency can be measured by the maximum allowed load (e.g. speed tier rate).

A common property of low latency services is that packets are useless when they are received with latency higher than an acceptable level. Therefore, they benefit from fast transmissions over shallow queues. This traffic type is called **non-queue building traffic (NQB)**. They do not benefit from increasingly consuming resources beyond need. They perform well in idle network conditions and good link quality, but with larger queues or dynamic movement between Idle Latency and Working Latency QoE can be variable without the right technology [9].

On the other hand, large down/uploads, buffered video streaming, speed tests, email, etc. rely on protocols that fairly use as much of the network capacity as possible to transfer the data at a high rate. This traffic type is called **queue building traffic (QB)**. The applications often open many TCP sessions in parallel and they are not latency sensitive. When these applications traffic is present on the network, latency is Working Latency. Small network queues lower latency but can make high speed QB traffic not hit peak rates without the right technology.

An overview of QB and NQB traffic is displayed in Figure 3 [9]. Real-time gaming control data and some audio/videoconferencing data are low-data-rate NQB traffic while cloud gaming streaming, real-time streaming and some videoconferencing applications are being implemented by developing new scalable congestion control algorithms (e.g. defined in [9]) to conform to high-date-rate NQB traffic type.

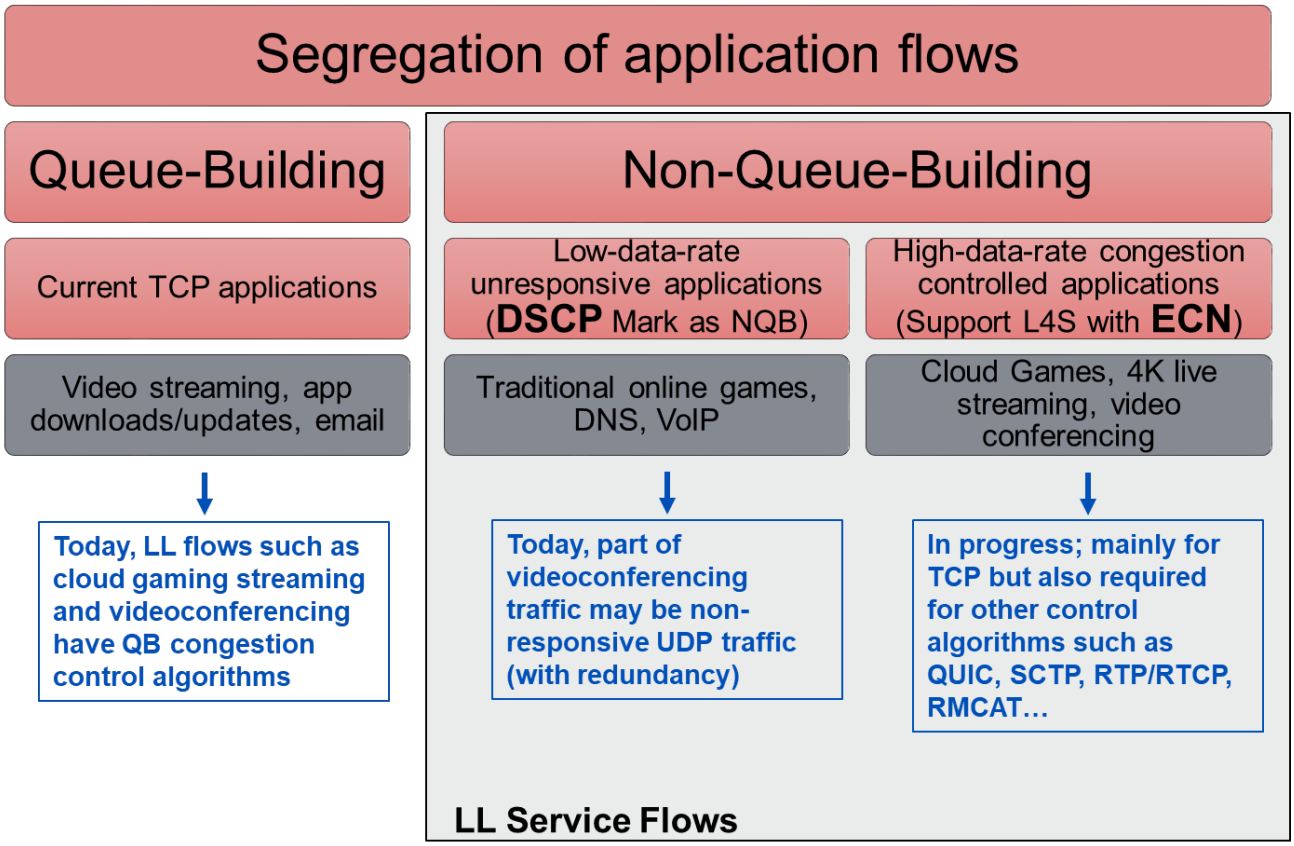


Figure 3 – QB vs NQB traffic

2.2. QoS requirements for Low Latency Services

DOCSIS 3.1 Specifications and Standards

DOCSIS 3.1 specifications include Low Latency Services support based on the coupled dual queue and proactive grant scheduling algorithms. The current target sets for applications such as real-time and cloud gaming and videoconferencing are <10ms RTT between the Cable Modem Termination System (CMTS) and Cable Modem (CM) for 99th percentile of packets. 1ms RTT can be achieved with proactive grant scheduling with the tradeoff of efficiency based on the currently available solutions. More details on the D3.1 low latency services support are provided in Section 4.1.

IEEE 802.11 Specifications and Standards

Time Sensitive Network support in 802.11 includes bounded 802.1Q traffic classification and stream reservation, low latency capabilities with 802.1Qbv over 802.11, scheduled operation with 802.11ax and 802.11be low latency channel access enhancements. Lower worst-case latency and jitter is a key feature for Wi-Fi 7 with multi-link and multi-AP operation and wider channels [10].

In [4], latency, jitter and packet loss requirements are proposed for several low latency services for the Wi-Fi networks (Table 1). Latency is defined as the RTT between the station (STA) and Access Point (AP), and jitter is defined as the standard deviation of latency. Since worst case latency is a key issue for these services, the definitions are based on the latency spikes that can also cause packet loss when certain

thresholds are exceeded, hence causing lagging and other issues. The document suggests new areas for further enhancement. Potential enhancements and new capabilities to address requirements of emerging real-time applications that can be grouped in the following categories:

Table 1 – Requirements Metrics of RTA Use Cases

Use cases		**Intra BSS latency/ms**	**Jitter variance/ ms**	**Packet loss**	**Data rate/ Mbps**
Real-time gaming		< 5	< 2	< 0.1 %	< 1
Cloud gaming		< 10	< 2	Near-lossless	< 0.1 (Reverse link) > 5 (Forward link)
Real-time video		< 3 ~ 10	< 1~ 2.5	Near-lossless	100 ~ 28,000
Robotics and industrial automation	Equipment control	< 1 ~ 10	< 0.2~2	Near-lossless	< 1
	Human safety	< 1~ 10	< 0.2 ~ 2	Near-lossless	< 1
	Haptic technology	<1~5	<0.2~2	Lossless	<1
	Drone control	<100	<10	Lossless	<1 >100 with video

3GPP Specifications

Technical specifications produced by the 3rd Generation Partnership Project (3GPP) and adopted by regional standards organizations use QoS class identifier (QCI) as a reference to a specific packet forwarding behavior (e.g. packet loss rate, packet delay budget) to be provided to a service data flow [3]. A subset of QCIs with a one-to-one mapping of standardized QCI values to standardized characteristics is shown in Table 2 for guaranteed and non-guaranteed bitrate resources (G/Non-GBR). A standardized QCI and corresponding characteristics are independent of the user's current access (3GPP or Non-3GPP). The characteristics describe the packet forwarding treatment that a service data flow aggregate receives edge-to-edge between the UE and the Policy and Charging Enforcement Function / Packet Data Network (PCEF/PDN) Gateway that is the interconnect point to the external network backbone.

Table 2 - A subset of 3GPP QoS Class Identifiers

QCI	Resource Type	Priority Level	Packet Delay Budget	Packet Error Loss Rate	Example Services
1	GBR	2	100 ms	10^{-2}	Conversational Voice
2		4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3		3	50 ms	10^{-3}	Real Time Gaming, V2X messages Electricity distribution - medium voltage Process automation - monitoring
4		5	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
67		1.5	100 ms	10^{-3}	Mission Critical Video user plane
5	Non-GBR	1	100 ms	10^{-6}	IMS Signalling
6		6	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
7		7	100 ms	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
8		8	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)

For instance, real-time gaming can be supported optimally if end-to-end latency is <60ms as working latency while <100ms latency can provide a good QoE [5, 6]. If Wi-Fi and DOCSIS networks can provide 15-20ms working latency as described above, 45-85ms latency can be allocated for the network segments outside of the ISP domain, encoding/decoding, rendering and/or cloud computing.

3. Current Latency Measurement, Monitoring and Management in the DOCSIS Networks

As discussed in a technical paper published at the 2020 SCTE Expo, and written by many of the same authors as this year's contribution [2], current DOCSIS features such as buffer control, Active Queue Management (AQM) along with better scheduling and efficiency implementations can help to reduce working latency and jitter for all High Speed Data services. Additionally, increasing speed tier rates through upcoming Mid-Split, High-Split and FDX technologies will help to improve the QoS support. However, these advancements alone are not adequate to provide bounded low jitter aimed for low latency services. QB applications can increase their traffic rates as speed tier rates increase and still cause high latency and jitter for NQBQ traffic. On the other hand, most NQB services do not require high speed, and network cost and operations can benefit from latency and jitter improvements to provide only the required

resources to each traffic type. To achieve better QoE for all services while creating an efficient and cost-effective network resource management, Cable Operators must have accurate and manageable latency measurement and monitoring tools.

We have been deploying with optimal configurations both D3.1 US AQM and DS AQM in our networks. The initial results have been captured in our 2020 latency paper [2]. Our internal measurement techniques for working latency (LUL) are similar to Samknows LUL data that is available as part of the FCC's Measuring Broadband (MBA) database [13]. We analyzed DS LUL results using FCC MBA data before and after our optimization. As displayed in Figure 4, DS working latency/LUL has been improved after optimal AQM deployments, which we were able to monitor and manage using our own platform.

As we extended our latency management platform, we optimized some of our methods and tools as we learned more with each deployment. In the following sections, we describe several parts of our platform by emphasizing key points that may be useful for other operators considering their latency strategies.

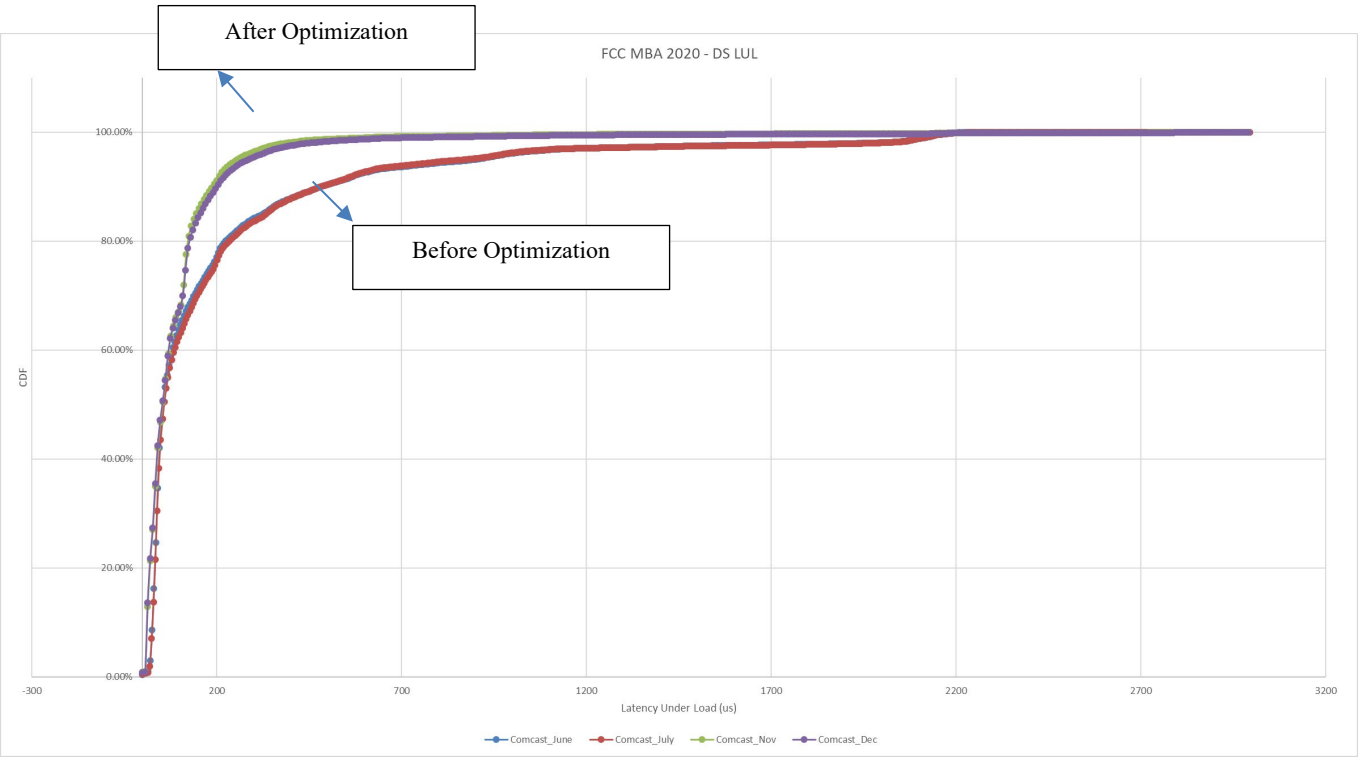


Figure 4 – FCC MBA 2020- Comcast LUL Results

3.1. Latency Measurement

As network operators focus more on the latency portion of the Quality of Experience, the challenge becomes how to measure the latency mitigation techniques being deployed on the network. Historically, the Internet Control Message Protocol (ICMP, RFC 792) has been used to perform a network layer latency check between two network endpoints. The limitation of ICMP is that it is a network layer check. Ideally, the latency check would be included as part of the application layer. Since that layer is predominantly outside the scope of the network operator, an alternate approach is to conduct the latency measurement at the transport layer. Further, most ICMP latency measurements are conducted in a

standalone instance, whereas the typical latency that impacts the customer QoE occurs while other users are accessing the network concurrently. Idle latency through ICMP pings can provide limited information.

Comcast has developed the Internet Measurement Platform (IMP) which provides a platform for measuring throughput and latency concurrently. Adapting the model used in other open source network measurement tools like Flent (https://flent.org), Comcast's IMP conducts both an "idle" latency measurement, meaning no other concurrent traffic from the test user, and a "latency under load (LUL)" measurement, meaning a latency measurement at the same time a throughput measurement is conducted, where the throughput measurement is trying to maximize its data consumption. Comparing the idle latency vs working latency (latency under load) measurements enables a clearer picture of the effectiveness of a deployed latency mitigation technique.

This new platform is implemented with an embedded agent on cable modems based on the Reference Design Kit/Broadband. The IMP agent interacts with the IMP control servers to process test requests using the specific IMP data plane test servers. The results are reported back from the client & server. By launching the tests from within the modem itself, the network operator is able to measure as closely as possible to the in-home devices that utilize the Internet service.

The idle latency portion of the measurement uses an HTTP CURL request / response, which uses TCP as its transport protocol. The latency under load portion of the measurement uses Netperf's request / response test, which uses UDP as its transport protocol. The throughput portion of the measurement uses the Iperf3 open source measurement tool, which uses TCP as its transport protocol. Both measurements are run concurrently. The TCP based data transfer will attempt to maximize its throughput up to the available provisioned capacity, potentially filling up node buffers along the network path while the UDP based request / response will attempt to complete its transaction competing against the load from the throughput measurement. In this fashion, the test is simulating the user's in-home experience where one user may be conducting a bulk data transfer (e.g. large file photo downloads) while another user is trying to complete quick request / responses (e.g. real-time gaming).

The latency reports can include min, mean, max, 50%, 75%, 95% and 99% and standard deviation values. Packet loss can be estimated by monitoring the successful transactions. Methods such as iRTT can be extended for more accurate packet loss and latency values, depending on an efficient implementation that can be supported by limited resources in the gateways.

The IMP platform has been audited by NetForecast, a 3rd party who independently reviewed the test results generated by the platform https://www.netforecast.com/netforecast-design-audit-report-of-comcasts-network-performance-measurement-system/

The IMP platform is currently in use on Comcast's production network, providing a comprehensive data set of latency measurements. Future enhancements to the platform include:

- Measuring the impact of different TCP congestion control algorithms
- As QUIC protocols are widely used, implementing UDP based data loading for speed test and working latency measurements
- Marking test data to measure latency for different services, such as low latency HSD flows.
- Exploring various control protocols to standardize test requests & results reporting

In addition to ping and working latency/LUL, there are other latency options that can be explored by the Cable Operators [5] such as actual customer traffic latency by monitoring TCP connections [2], monitoring queuing latency at DOCSIS with new D3.1 latency histogram recordings and Wi-Fi queuing metrics and two-way active measurement techniques.

The Two-Way Active Measurement Protocol (TWAMP), specified by IETF RFC 5357, provides a common protocol for measuring two-way or round-trip measurement between network devices. Today many routers used in Cable Operators' core network and CIN have TWAMP measurements capabilities. These capabilities may be extended to cover DOCSIS and other access networks and home networks as well [5]. The extensions can enable end-to-end latency profiling for QoE assessment of low latency services. A simplified version called Simple Two-way Active Measurement Protocol (STAMP) by IETF RFC 8762 can be used for one-way and round-trip latency, jitter and packet loss metrics.

As we measure the speed and latency for higher speeds, we see the shortcomings of certain measurement techniques. For example, when we measure symmetric 1Gbps DOCSIS network speed and latency for High-Split architectures, current TCP algorithms are not adequate. One key issue is that optimal test parameters that measure the speed tier vs working latency are not the same, depending on the protocol used for data load. Concurrent TCP flows may fill up the pipe but not the available buffer depending on the test parameters. When the goal is to measure the speed and working latency accurately and during a short time interval, other measurement options such as using UDP may be more efficient [11]. Cable operators can try these techniques by using Odroids connected to HS CMs as open source implementations (e.g. https://github.com/BroadbandForum/obudpst) are available before integrating them to their gateways for automated and stand-alone measurement capabilities.

3.1. Latency Visualization and Dashboarding

Visualization has always been a challenge for data. Of course, there are stated rules to follow also known as the graphic continuum. The challenge comes in telling a specific story that the data outlines in a readable form for the audience. Many can often make a scatter plot, bar chart, line chart etc., but readability becomes the challenge. Does your audience understand the story you portray?

For example, Figure 5 and Figure 6 can be used effectively to explain the difference between idle and working latency, to assess the best CM model and configuration for certain services. It is clear from the latency visualizations that, although idle latency performance does not vary significantly among the CM models, working latency is very different depending on the model and configuration. These visualizations can be used to compare the measured latency levels with the requirements given in Section 2.2.

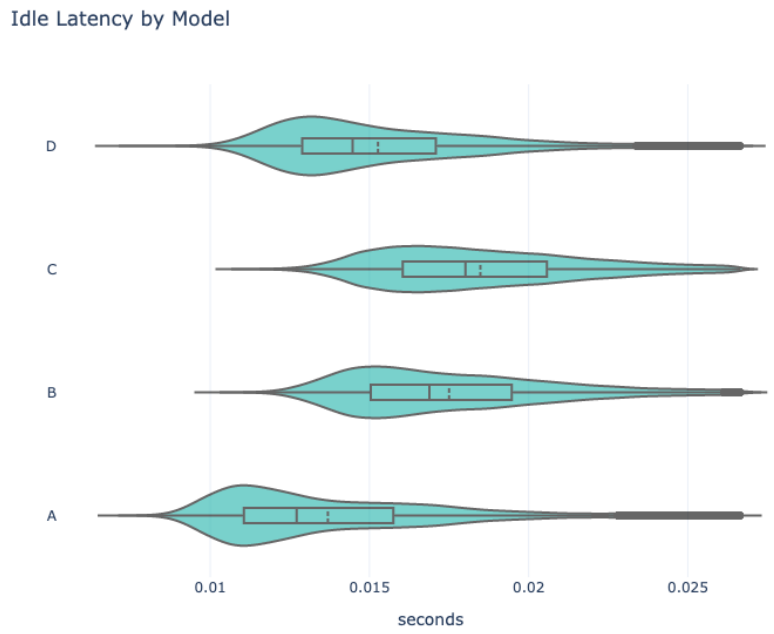


Figure 5 – US Idle Latency

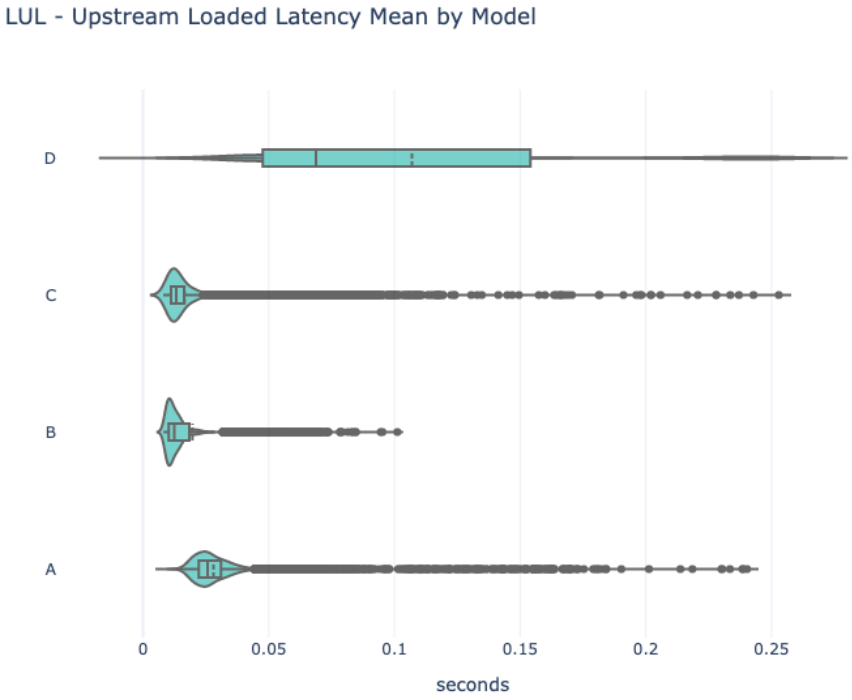


Figure 6 – US Working Latency (LUL)

Latency measurements over time as shown in Figure 7 can be used to monitor the latency improvements of a certain model that may have new FW or configurations, or of a new model that is deployed recently. Figure 8 is another example for DS working latency monitoring over time where new configurations are deployed to achieve a common improved latency range for all models.

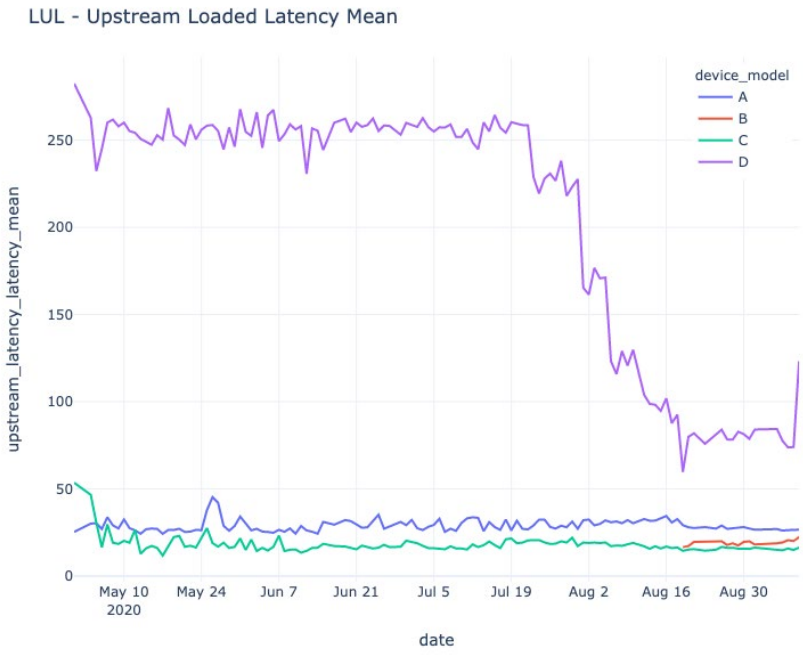


Figure 7 – US Working Latency (LUL) over time

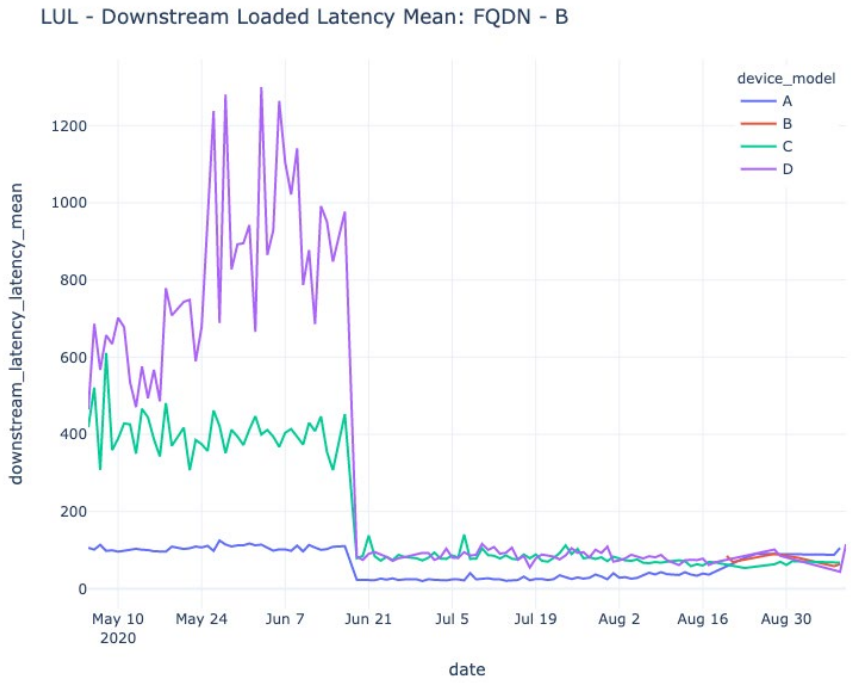


Figure 8 – DS Working Latency (LUL) over time

These initial visualizations are then used to create dashboards for continuous latency monitoring and management.

3.1.1. Current dashboards and data analysis

Figure 9 is an "executive" view of the latency data. Understanding your audience is what helped to shape the data into this view. The dashboard went through many iterations before becoming the high-level view shown above. The data shown is an aggregation of millions of rows of data that are processed daily. Breaking down the data with simplified thresholds allows an easy look into the performance. This also allows the end user to see spikes, anomalies, and trends from the data. This does not allow for a deep dive into the data and isolation of specific problems.

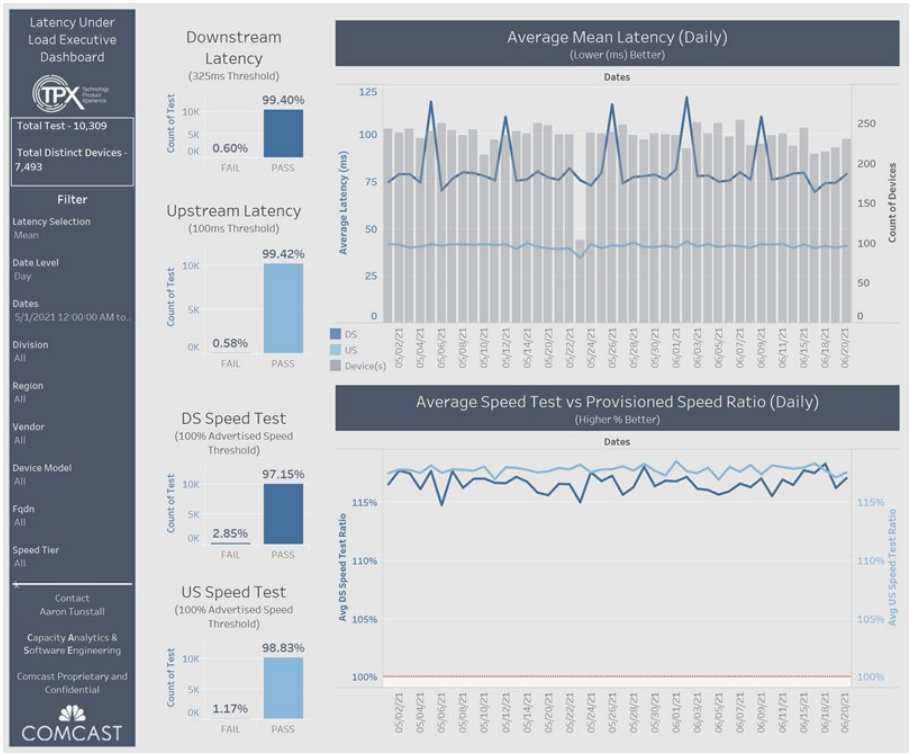


Figure 9 – Executive Dashboard for Latency & Speed test

To enable a more granular view of the performance we created a detailed version of the dashboard (shown in Figure 10). This version was the main source of truth for identifying latency performance and outliers. Each portion of the dashboard is interactive and allows the user a plethora of filters and customizations. Most users will never use all the functionalities included in this view, hence the "executive" view is the main view given to customers when they first load the dashboard. The main issue with data, especially latency, is that you have to balance what is readily available so all customers' needs are met and a proper analysis can be obtained.

Figure 10 – Latency Detailed Dashboard

3.2. Latency Management

3.2.1. Monitoring Latency

It is important to have a smart Business Intelligence (BI) tool when monitoring data. It must be dynamic enough to allow easy understanding and manipulation of the data. For latency we decided to use Tableau dashboards as they allow easy filtering and customization of views. The customer can easily click to change, aggregate or highlight specific data to find outliers and information.

3.2.2. Joint Analysis of latency and speed tests

Once latency data is processed and network configurations are set, the next analysis needed is to see if there are direct correlations between latency and network congestion (utilization). Correlation can be difficult when the measurements of latency can't be directly aligned with metrics that measure network congestion. This makes finding correlations difficult.

The easier method comes with comparing high latency to speed test results. When we compared "failed" speed tests (tests that don't achieve 100% of advertised speed), we found little correlation to high latency. The main reason is that latency is affected by small bursts of delay or packet loss that can be very challenging to detect in a granular way. When the utilization measurements are averaged between long measurement intervals compared to required granularity, the correlation cannot be seen. Therefore, we started a new trial where more granular utilization measurements are implemented to correlate their impact on the speed and latency tests.

3.2.3. Challenges and Guidance

When given the task to analyze data, the main challenge is balancing what is needed and what is not. It is an even harder challenge when you have data and no clear understanding on thresholds. What makes

latency good or bad? What is the expected customer experience? How can we isolate issues and outliers? Initially, when data was received, it was a small sample size and the dashboard was created with that in mind (Figure 11).

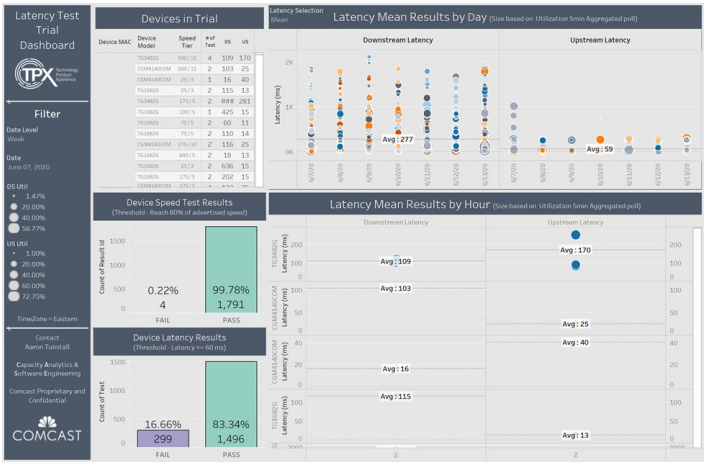


Figure 11 – Original dashboard for latency trial

At this point, there was no real threshold, just a small sample of test device data. We had an understanding that high latency is bad and lower latency is good. As the trial began and the scale of devices increased, the first version of the dashboard had to evolve. We created a detailed dashboard (Figure 10). This allowed a deep dive to analyze latency in multiple aspects. The dashboard showed how CMTS, device model, device firmware, vendor, and speed tier performed. This allowed us to isolate differences between all of the above, and how latency varied for each. The functionality made outlier identification a lot easier. This dashboard served its purpose, but when reporting findings it was not dynamic enough for presentation.

Newer temporary views were created, each to serve a different purpose for our Operations and Testing teams. These views were high-level comparisons of vendors, device types, and more. One of the lessons learned was to understand what your customer needs. This can be difficult when those needs are truly not yet understood. As time progressed, data became more readable, and thresholds started to be defined the needs of the customer become easier to define. What we learned allowed us to create the Executive dashboard (Figure 9) as well as easier views that tell the story the data shows. This allows a high level understanding of the overall customer experience.

In conclusion, data can be difficult to decipher when there is no clear understanding of how to read it. Understanding simple things, like high latency is bad and low latency is good, does not assist with telling the data's story. Sometimes it takes multiple views and analysis to find the needle in the haystack. As a data engineer, you have to understand that your initial idea or understanding may not be the final. Lastly, it is important to transform your data. As the analysis progressed the data sample grew, causing performance issues for customers viewing it. Summarizing data through aggregation is very important to giving the best customer experience when reviewing the data.

Another challenge is the storage, maintenance and query of the available data. We started getting LUL measurements from Elastic search production. The data was downloaded in Json format using Python script and stored it into the SQL server. The real-time data is visualized using Kibana - the data visualization dashboard software (open search dashboard) for Elasticsearch. The data download frequency was daily basis for the previous day. Due to ELK cluster workload maintenance, the LLD data is moved

to our internal streaming data platform, and Kinesis streaming for download [12]. The raw data download frequency is close to real-time. Scalability issues can arise later, as more data is collected and more queries are done by different teams. Therefore, the design must be flexible and extensible.

These large-scale measurement comparisons should provide additional data to justify the future deployment of AQM by ISPs and customer premise equipment manufacturers. This may be of interest to people working on the Low Latency, Low Loss Scalable Throughput (L4S) protocol or other TCP/UDP congestion controls. We believe, there is an opportunity to better standardize/define how working latency is measured. There is a need for open global internet measurement platforms to focus on working latency (or create new platforms & beyond access network segments) and/or sharing of such measurement data.

4. New Low Latency DOCSIS Features and Latency Management

As we discussed briefly in Section 1, broadband has historically differentiated based on download speed, however, customers are increasingly interested in additional characteristics when shopping for broadband. These emerging differentiators include faster upstream speeds, whole home WiFi coverage, and low latency.

Gamers are a large market and for serious gamers, latency influences ISP switching behavior and product selection. Gamers are also heavy streamers and in general, they tend to buy the best product that can support low latency services and high speed streaming. Beyond gamers, customers are generally interested in the concept of "no-lag broadband," suggesting that low latency DOCSIS could resonate with a larger audience. Specifically, broadband users that are employed and/or working from home have higher interest compared to other non-gamers, implying that the employed/work-from-home segment is another group to potentially target.

Enabling low latency DOCSIS would give cable providers an advantage over other ISPs that cannot compete with such consistently low levels of latency under load. Given the interest level in low latency among certain segments, go-to-market approaches could include incorporating low latency features across all internet households (where available), incorporating low latency features in premium speed tiers, or upselling low latency features within a separate premium data product.

To support these business cases, a new architecture that can differentiate low latency services and provide required QoS requirements must be defined and implemented. Although latency and jitter improvements described in Section 3 improve the overall QoE for several services, they are not adequate for current and emerging real-time interactive services.

Low latency will enable the creation of major new classes of applications where delay to local storage is equivalent to delay to a network-based resource. The cable industry is poised to take some leaps ahead of the competition as in the post-gigabit future, latency will be equally important to speed in marketing.

If we apply the same network resources with the same functionalities to each traffic with different QoS requirements (Section 2.2), then we provide equal resources but the outcome, which is the quality of experience, will be very different for each traffic type and unfair. AQM provides equitable performance because some traffic is managed (Figure 12 explains the difference between equality and equity nicely). Both large QB flows and smaller NQB flows perform better. LLD with L4s and dual queue promise even better. Early results show that we can hit the LLD target for working latency. As well, equity increases for all traffic because they are not having to share the same queue. This also helps to make sure that no app/service is harmed while, for example, providing different AQMs to low latency traffic groups because the equal quality of experience score is targeted.

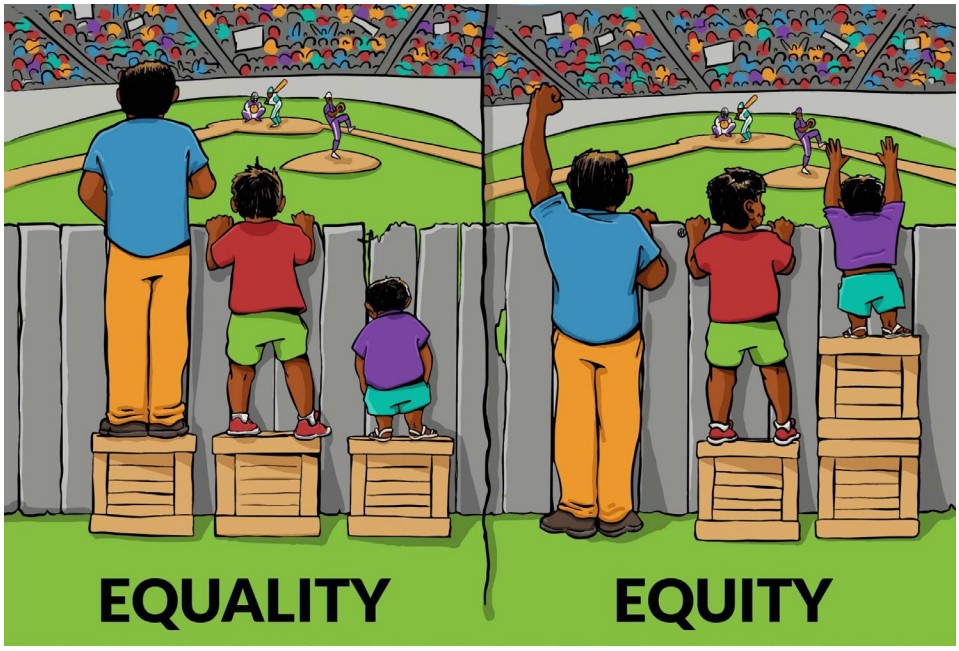


Figure 12 – Equality vs Equity

4.1. D3.1 LLD Features

New D3.1 LLD Dual Queue Features promise <10ms RTT between the CM and CMTS for 99th percentile of LL service packets. The main functionalities are described in Figure 13and Figure 14 [14].

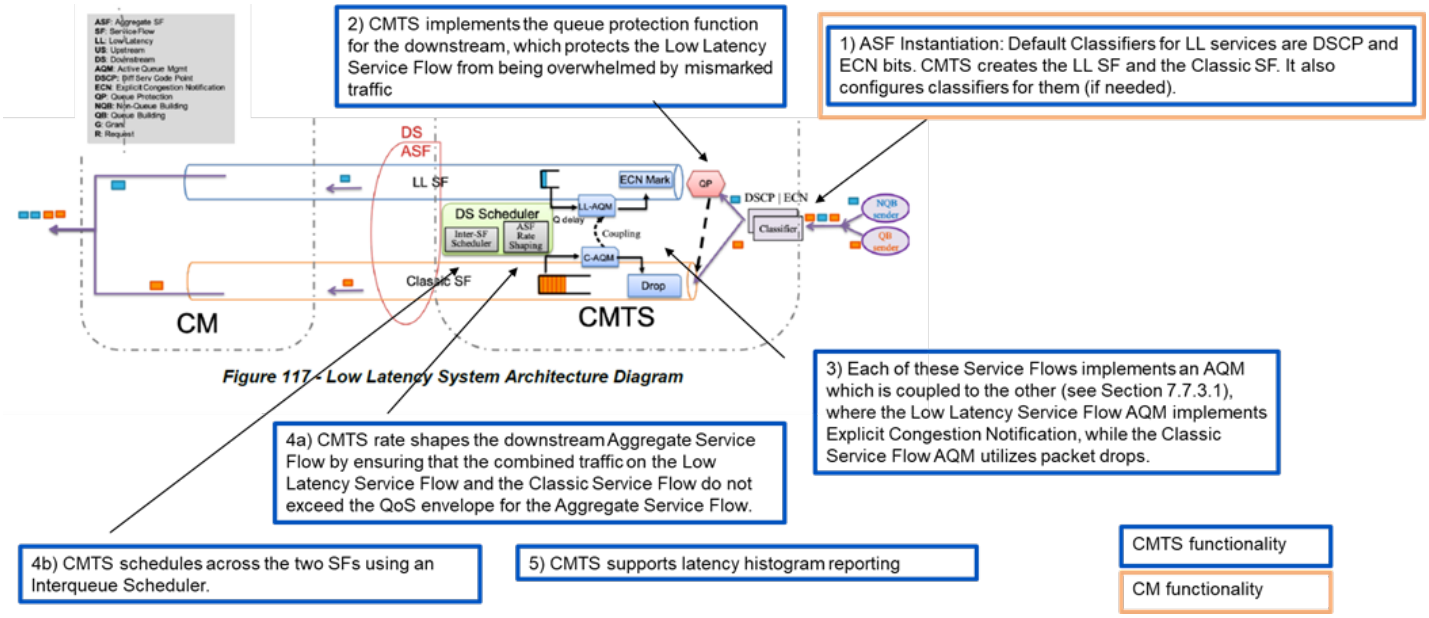


Figure 13 – D3.1 DS LLD Features

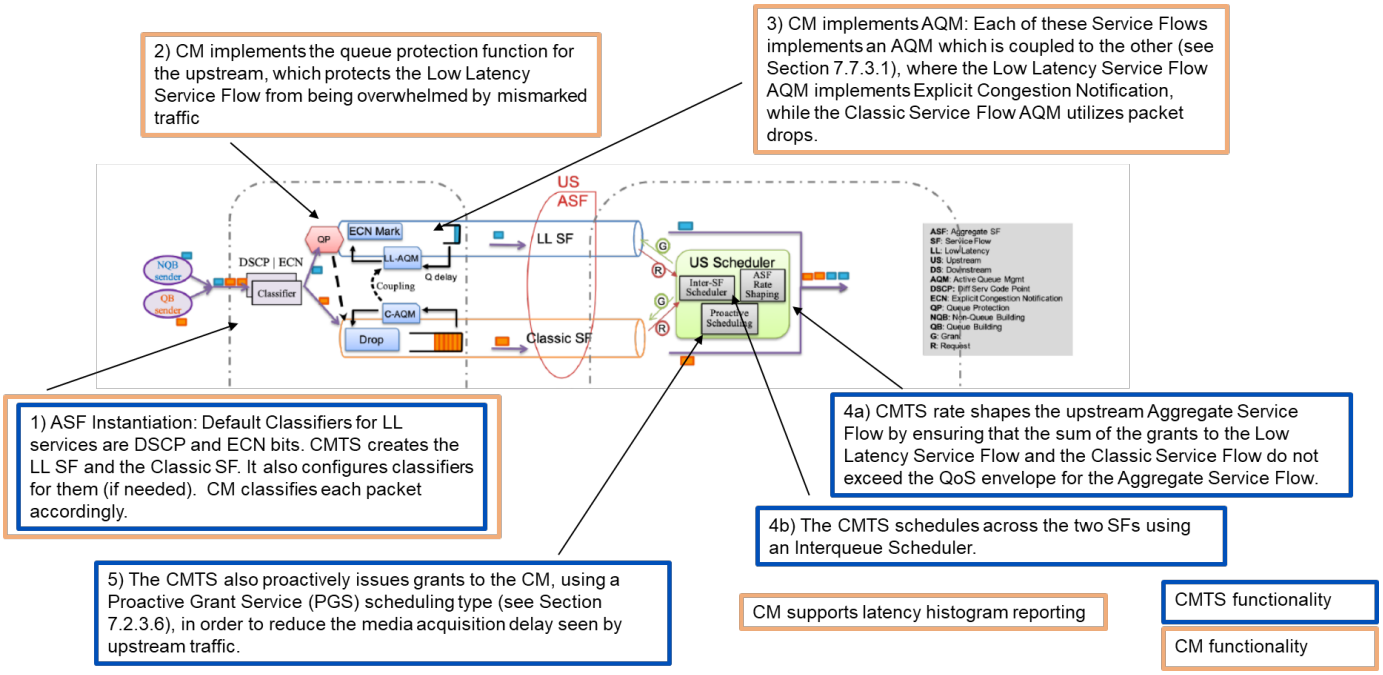


Figure 14 – D3.1 US LLD Features

We collaborate with Cablelabs closely for the new LLD features. The dual queue approach will enable us to support the latency and jitter requirements described in Section 2.2 for low latency services. In this section, we present an example scenario tested by Cablelabs[1] with dual queue approach and PGS.

Figure 15 shows the results for upstream LLD without PGS, with a mix of TCP upload file transfers as cross traffic. QB and NQB traffic packets are queued to classic and low latency queues respectively. In the latency distribution charts, both the raw latency distribution and the distribution of Inter-Packet Delay Variation are plotted. Latency distribution is plotted with solid lines, IPDV plotted with dot-dash lines. The 99.9th percentile of LL traffic latency is less than 10 ms.

Figure 16 shows upstream LLD with 2Mbps PGS, with the same cross traffic. The 99.9th percentile of LL traffic latency is less than 6 ms. The NQB traffic has different packet sizes while PGS grants for the selected settings are for 250 bytes. Therefore different latency values for different packet sizes are observed in this test.

We can conclude that the current DOCSIS network latency that may be in the order of 100 ms can be decreased to less than 10 ms for 99th percentile of LL service flow packets with the new LLD features.

[1] We would like to thank and acknowledge Greg White from Cablelabs for his LLD tests presented in this paper.

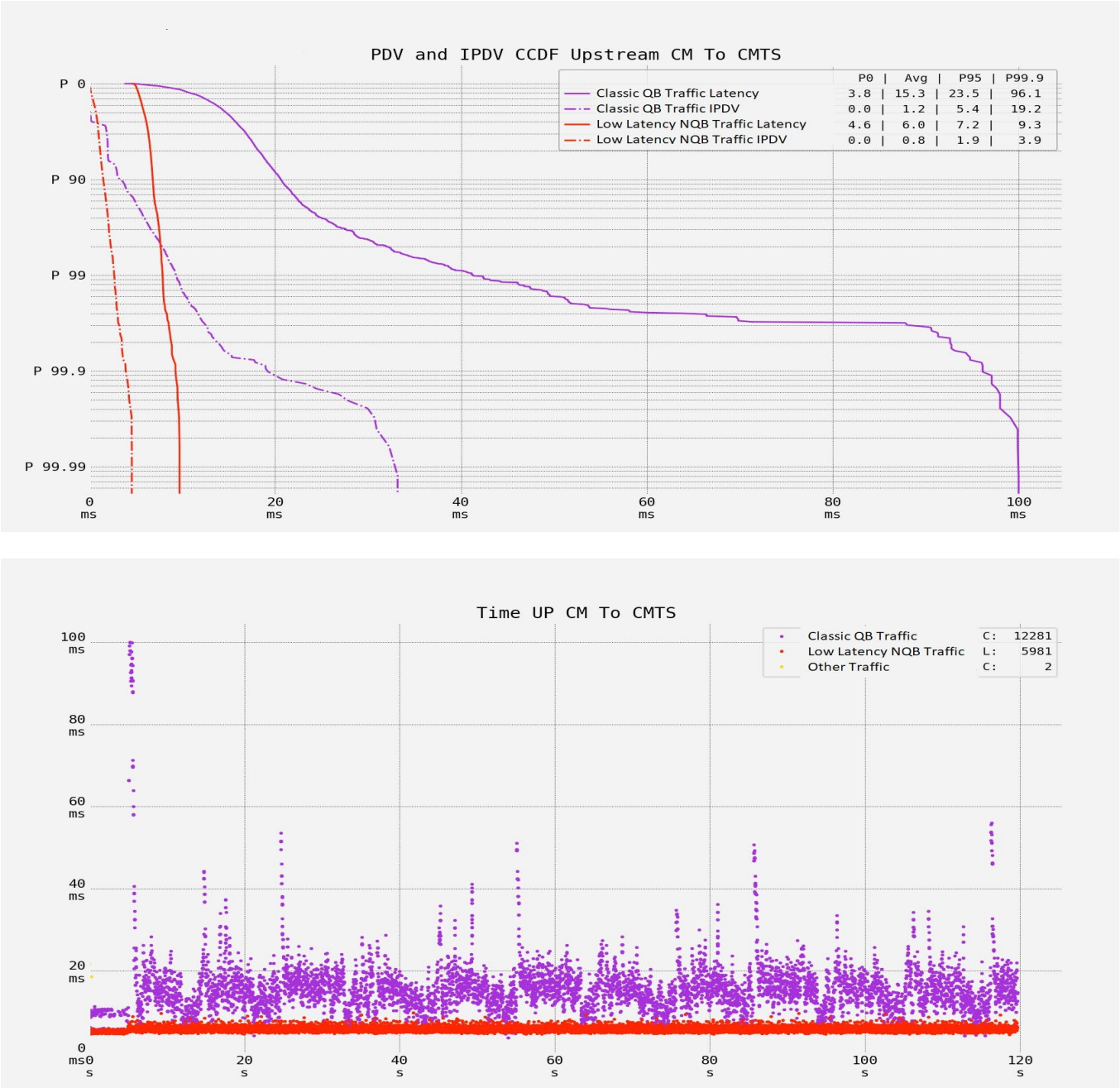
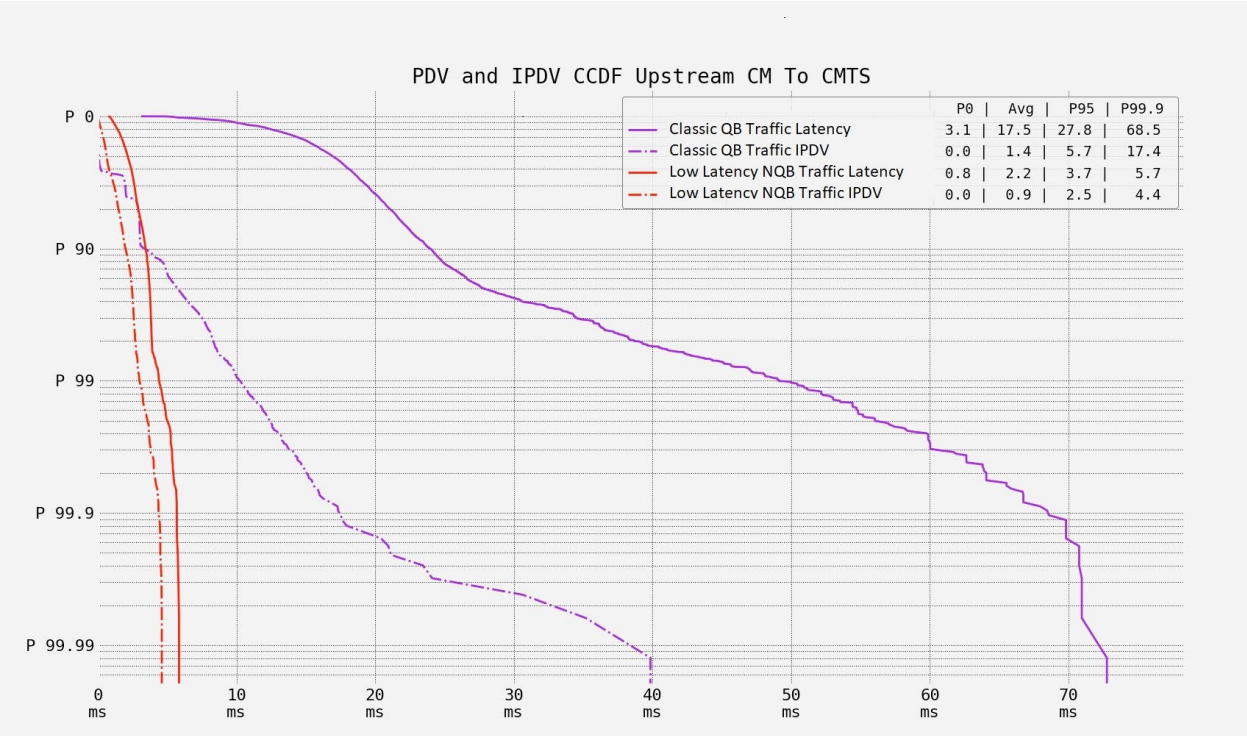


Figure 15 – US LLD With Dual Queue and no PGS



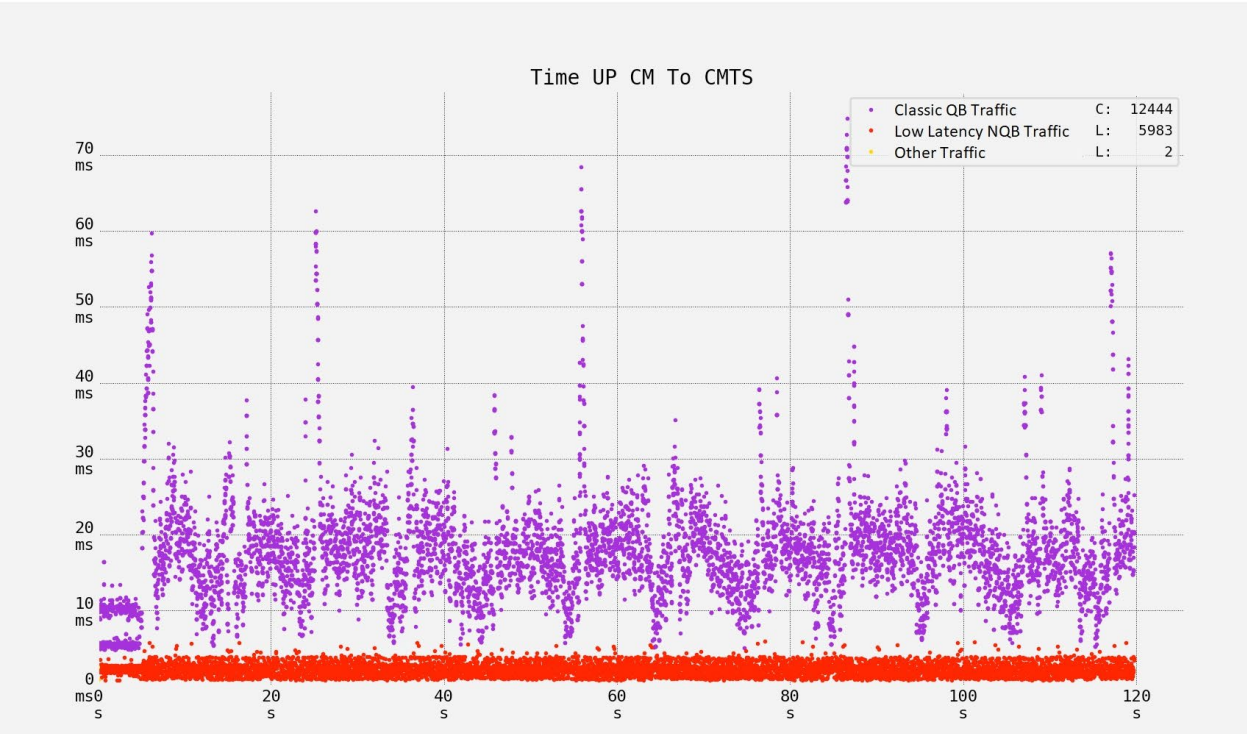


Figure 16 – US LLD With Dual Queue and PGS

These features can be enabled only if Low Latency services are differentiated. Traditionally, Cable Operators do not trust any marking from the user space due to security and operations challenges. The new framework suggests Differentiated Services Code Point (DSCP) marking for low-date-rate NQB traffic and ECN Capable Transport ECT(1) bit marking for high-data-rate NQB L4S traffic (Figure 3 and Figure 17). NQB-DSCP and L4S IETF documents address the challenges of this new framework.

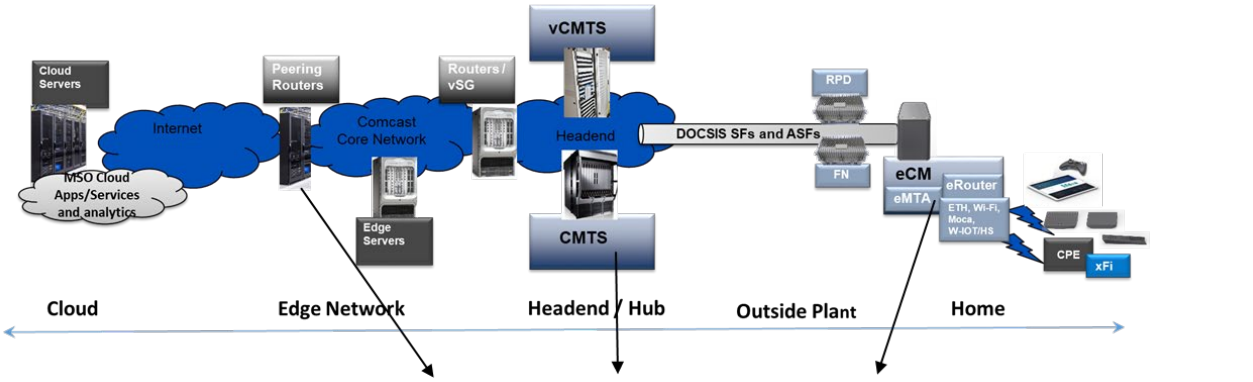


Figure 17 – Marking for LL Services

5. Conclusion

Current latency measurement, monitoring, and management frameworks are fundamental to support the next generation of Low Latency services. These platforms and network architectures must be extended to differentiate LL services to apply D3.1 LLD features, Wi-Fi and Core Network enhancements and to manage their performance. The faster democratization of technologies pushes ISPs to support new services at a faster pace. This can be enabled by achieving three main points within the industry:

1) Better standardize/define how latency, jitter, packet loss and other QoS metrics are measured and create open global internet measurement platforms to focus on end-to-end QoE assessment.
2) Start breaking legacy chains through digitization, software defined, virtualized and cloud based systems with open source software, platform models with partners and co-innovators to meet the consumers' demands in an agile way.
3) Apply an end-to-end approach for traffic differentiation and QoE management with new upcoming 10G technologies.

Abbreviations

AP	Access Point
AQM	Active Queue Management
CE	Congestion Encountered
CM	Cable Modem

CMTS	Cable Modem Termination System
DSCP	Differentiated Services Code Point
ECN	Explicit Congestion Notification
ECT	ECN Capable Transport
FDX	Full Duplex
GBR	guaranteed bitrate resources
HFC	Hybrid fiber-coaxial
IPDV	Inter-Packet Delay Variation
ISBE	International Society of Broadband Experts
ISP	Internet Service Provider
L4S	Low Latency, Low Loss Scalable
LL	Low Latency
LLD	Low Latency DOCSIS
NQB	Non-queue-building
PCEF	Policy and Charging Enforcement Function
PDN	Packet Data Network
QB	Queue-building
QCI	QoS class identifier
QoE	Quality of Experience
QoS	Quality of Service
RTT	Round Trip Time
SCTE	Society of Cable Telecommunications Engineers
STA	Station

Bibliography & References

1. *The Democratization of Technology*; Mihir Shukla, Forbes Technology Council Post, https://www.forbes.com/sites/forbestechcouncil/2019/11/07/the-democratization-of-technology/?sh=765aa8643796
2. *Approaches to Latency Management: Combining Hop-by-Hop and End-to-End Networking,* Sebnem Ozer, Carl, Klatsky, Daniel Rice, John Chrostowski, SCTE-ISBE Workshop 2020
3. *Policy and charging control architecture,* 3GPP TS 23.203 V17.1.0 , 3GPP, 2021
4. *IEEE 802.11 Real Time Applications TIG Report,* 2019.
5. *Latency Measurement: What is Latency and How Do We Measure It?,* Karthik Sundaresan, Greg White & Steve Glennon, SCTE-ISBE Workshop 2020
6. *Mobile cloud gaming: the real-world cloud gaming experience in Los Angeles*, RootMetrics, 2020
7. *A Blueprint of Technology, Applications and Market Drivers Towards the Year 2030 and Beyond*, ITU-T FG-NET-2030
8. *Four reasons why low latency streaming matters*, https://nscreenmedia.com/4-reasons-low-latency-streaming-matters/, 2021
9. *Low Latency DOCSIS: Overview And Performance Characteristics*, White, G., Sundaresan, K. and B. Briscoe, SCTE-ISBE Workshop 2019.
10. *Wi-Fi TSN Capabilities and Evolution Towards Deterministic Low Latency,* Dave Cavalvanti and Ganesh Venkatesan, 2020
11. *Maximum IP-Layer Capacity Metric, Related Metrics, and Measurements*, TR-471, BBF, 2020

12. *High Performance Data Streaming with Amazon Kinesis: Best Practices* (ANT322-R1) - AWS re:Invent 2018
13. *FCC MBA Raw Data Releases*, https://www.fcc.gov/oet/mba/raw-data-releases
14. Cablelabs DOCSIS 3.1 MULPI Specifications

Fixed-Wireless Convergence on a Multi-Access Edge

Technical Paper prepared for SCTE by

Juan Rodriguez
Sr. Director US MAJORS/MSOs
Nokia
Orlando FL, USA
juan.rodriguez@nokia.com

Arnold Jansen
Sr. Marketing Manager IP/optical Network Infrastructure
Nokia
Ottawa ON, Canada
arnold.jansen@nokia.com

1. Introduction

Broadband access is an essential utility to connect with each other and the cloud, from the convenience of our homes. Ultimately broadband evolution is about delivering a reliable and affordable user experience at gigabit speeds, but no single access technology can cover all use cases equally well. To connect everyone with the digital world we live in, requires a multi-access broadband strategy that combines an expanding range of wireline and wireless access technologies.

The paper discusses how cable operators may effectively leverage 5G fixed-wireless access (FWA) to complement and enhance their existing wireline broadband offer and deliver a converged multi-access broadband experience everywhere.

2. The case for fixed-wireless broadband access

Cable operators have done a remarkable job extending the life and utility of the existing coaxial access plant to bring the Internet to nearly every home. By pushing fiber ever closer to the home, they were able to speed up data transfer rates from kilobits to megabits per second and do so without breaking the bank. But demand for faster access continues unabated while the cost of extending fiber further to the home is increasingly steep. Fiber-to-the-home is typically the preferred and future safe way to go, but deployment can be costly and time consuming, even for new housing projects. Gradually and inevitably, a faster broadband experience may literally become out of reach for more and more consumers.

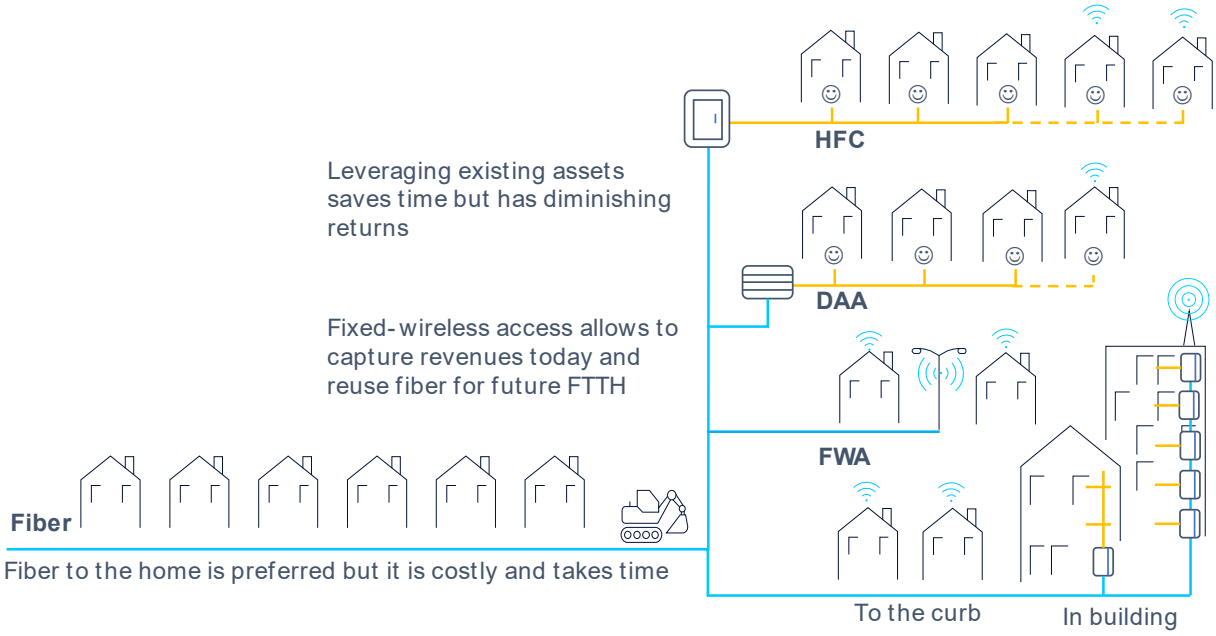


Figure 1. The case for fixed-wireless broadband access

Fixed-wireless access can address these last-mile broadband wireline coverage issues (

Figure 1). LTE/5G radio deployed in the sub 6 GHz frequency bands can rapidly cover large areas with access speeds that can sustain high definition video streaming. Small cells operating at mmWave

frequencies (24 to 39 GHz) can even deliver fiber-grade Gigabit broadband services. Besides licensed spectrum, many regulators have actively engaged to make shared and unlicensed spectrum available in various frequency bands. For example, the US Federal Communications Commission designated spectrum in the 3.5 GHz band for shared use (the "innovation band" of the Citizen's Broadband Radio Service or CBRS).

FWA perfectly complements wireline broadband deployments in the last mile:

- In brownfield it can be used in overlay to address capacity and coverage issues of underserved wireline broadband users.
- In greenfield, FWA can be an effective tactic to get quick service coverage and revenues and potentially support a phased fiber-to-the-home rollout.

The question is how to integrate FWA in a combined multi-access broadband infrastructure to create better cost synergies reduce operational complexity?

3. Multi-access broadband and edge convergence

Convergence in general aims to achieve operational simplification and better economies of scale by consolidating network requirements and pooling resources. The operational goal of fixed and wireless broadband convergence is a seamless and simplified multi-service user experience that is access agnostic. Although broadband consumers typically use only one wireline or wireless access method at a time, another objective is to leverage a common operational infrastructure to manage subscriber access and service policies. Converged operators that also offer mobile broadband services are keen to consolidate all their services on a more agile, cloud-native 5G Core (5GC), and reap the cost and performance synergies of fixed and mobile convergence (FMC).

The ultimate goal is to build an agile and cost optimized network that can deliver a ubiquitous and seamless service experience over any broadband access medium. Figure 2 summarizes the requirements in a multi-access broadband target infrastructure.

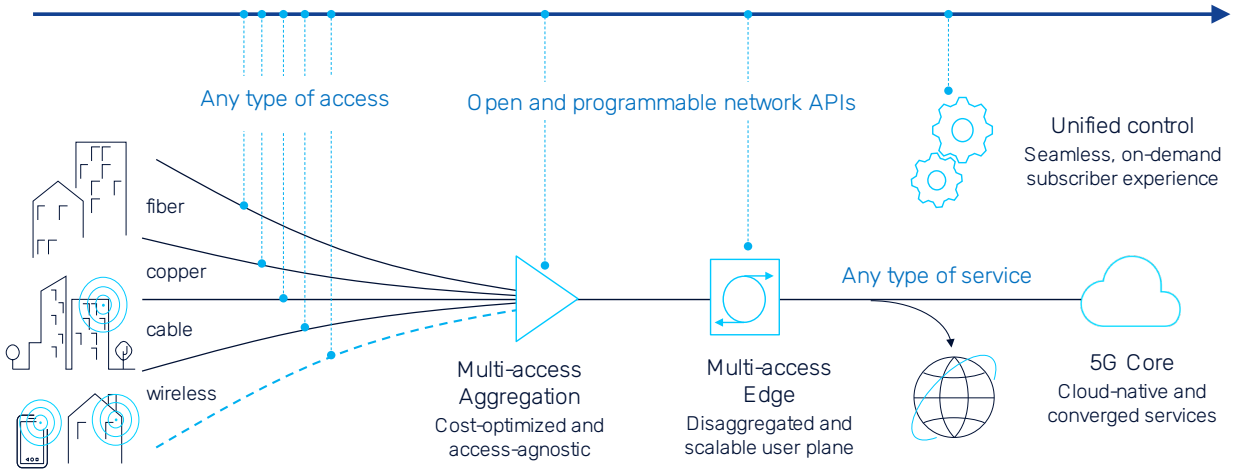


Figure 2. Multi-access broadband target architecture

Because all service traffic has converged on IP, all wireline and wireless broadband traffic can potentially be aggregated over a common multi-access backhaul network. However, current broadband access and aggregation networks may present an obstacle that hinders a smooth transition to this multi-access

broadband architecture. While most operators are moving towards access-agnostic IP/ethernet aggregation layer in the metro and regional network, they often maintain dedicated platforms that accommodate for different access network requirements. The inherent duplication of functions in a heterogenous aggregation network reduces becomes increasingly costly and complex as more access technologies are introduced and evolve in parallel.

Fortunately, the new generation of disaggregated wireline and 5G wireless access solutions enable access-agnostic IP/Ethernet transport aggregation on a converged interconnect network (CIN) by virtualizing technology-specific access network functions and moving them into the distributed edge cloud under a programmable software defined network (SDN) controller. Examples are virtualized converged cable access platform (vCCAP) for distributed access architectures, virtualized optical line terminal (vOLT) for XGS-PON, and virtualized baseband units (vBBU) for the 5G radio access networks.

The second obstacle is a broadband service edge that typically deploys dedicated edge gateways for each access network technology: A CCAP/CMTS for cable access, a Broadband Network Gateway (BNG) for fiber-to-the-home (FTTH) access and a Serving and Packet Data Network gateway (SPGW) for 4G/LTE mobile services. While such divide and conquer edge strategy separates concerns, it also reduces cost synergies while increasing complexity, both at the service edge and in operational back-end systems.

The broadband service edge is a critical network junction where subscribers and service policies are enforced. The inclusion of fixed-wireless access should at least add no further complications, and this is the main subject the rest of the paper will focus on.

4. Fixed-wireless access gateway requirements

A converged fixed-wireless edge can play an important role in delivering affordable, high-performance broadband services to every home. However, wireline and wireless access technologies come from very different worlds, so where does fixed-wireless access fit in? Wireline or wireless?
To answer this question, we must compare their service characteristics (Table 1).

Table 1. Comparing wireless and wireline service characteristics

Requirement	Mobile user	Wireline user
User devices	1 (typically)	>10 per home
Service type	Nomadic	Stationary
Session type	Dynamic	Always-on
Subscription type	Usage based	Unlimited (flat rate)
Monthly data usage	3-5 GB	100s of GB
Average speed	~10 Kilobit/s	~1 Megabit/s
Gateway location	Centralized	Distributed
Gateway functions	Virtualized (x86)	Physical (NPU)

To complement (or compete with!) wireline broadband, fixed-wireless access gateways must support the same residential broadband applications. There are no roaming requirements because all user devices are tethered to a single, stationary home gateway that serves up to a dozen user devices such as TVs, PCs, tablets and gaming consoles. Home broadband services are always on and must sustain high bandwidth usage for multi-cast IPTV and streaming ultra-high definition broadcast TV, binge-watching Netflix, Zoom video conferencing, and downloading software for PCs and game consoles. Sophisticated quality-of-service (QoS) techniques are needed for the delivery of multiple services over the same bearer.

Monthly data usage per home averages several hundreds of Gigabytes that are charged at a monthly flat rate, without usage caps for the highest service tiers.

In contrast, wireless broadband users typically connect a single device (smartphone) with short-lived dynamic user sessions. Although mobile devices can generate significantly more control plane traffic, mobile data volumes are orders of magnitude lower and charged by the Gigabyte. Mobile data rates are comparably high. One month of broadband usage charged at mobile data rates would easily cost a homeowner more than a one year wireline broadband subscription.

As a result of their different usage characteristics, mobile-wireless gateways are polar opposites of the wireline gateways used for residential broadband:

- Wireless gateways are optimized for dynamic, mobile user applications that yield a high revenue per bit. They are typically centralized and virtualized on x86 servers, which allows operators to cost-efficiently pool resources for roaming users and leverage cloud-native compute and storage to support dynamic user sessions with elastic scaling needs. They are not designed nor dimensioned to cost-efficiently support bandwidth-intensive internet or IPTV applications over extended periods of time.
- Wireline gateways are cost-optimized for delivering always-on Gigabit broadband services to homes and businesses. They are typically purpose-built network appliances that leverage custom routing silicon for granular bandwidth management and hierarchical QoS to ensure that available network resources are fairly and optimally shared among subscribers and user devices. Default residential broadband service features such as Internet access and IPTV multicast replication are far more economical to deploy and scale on wireline gateway platforms, compared to Internet offload (LIPA-SIPTO) and IP multicast (eMBMS) on mobile gateways.

In conclusion, the service requirements of fixed-wireless access are very similar to wireline broadband access. Because fixed-wireless access gateways and broadband wireline gateways share most requirements, there are potentially significant operational cost and performance synergies when leveraging the same type of edge platform for both.

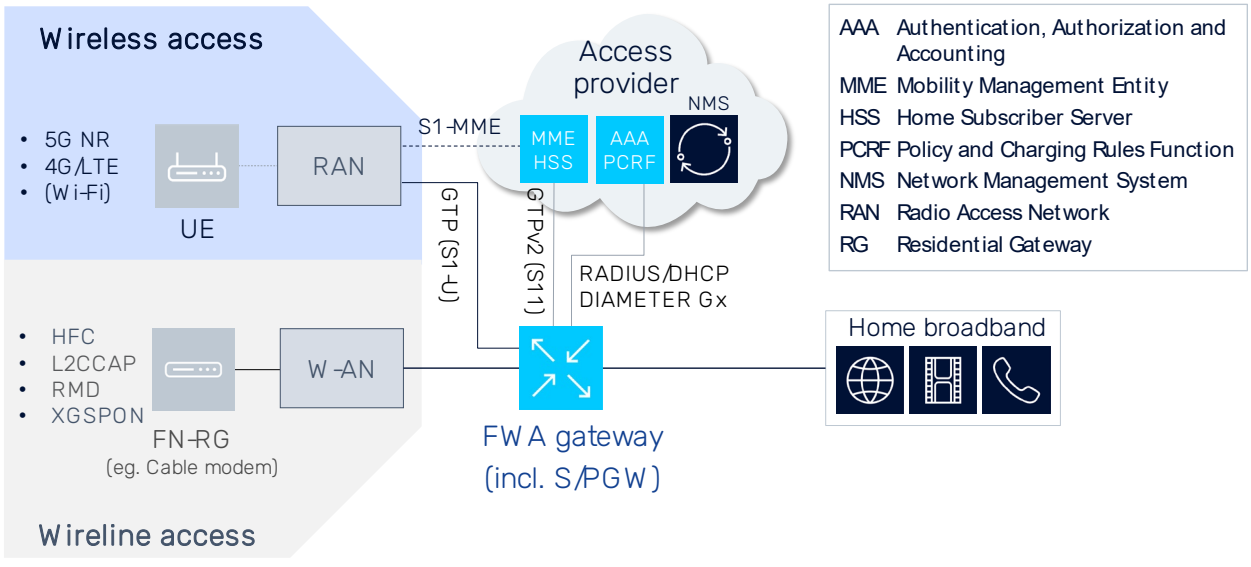


Figure 3. Fixed-wireless broadband access gateway

The addition of a Serving and PDN Gateway (S/PGW) function allows wireline operators to deploy FWA stand-alone or in conjunction with wireline access (see Figure 3). Broadband access providers deploying FWA will chiefly consider using shared, unlicensed or mmWave spectrum because licenses are more affordable. Because user devices are tethered to a stationary home gateway there is no need to provide blanket service coverage from multiple angles to support user roaming and only a subset of the MME and HSS mobility functions are needed to enable X2 handovers between adjacent base stations (eNB, gNB) within the tracking area. This significantly reduces the operational complexity of integrating FWA in a home broadband offer and enables access providers to essentially leverage the same operational backend as used for their existing wireline broadband service.

5. Disaggregated multi-access edge gateway

Converged operators may wish to interwork their wireline and fixed-wireless access services with their 5G packet core to enable fixed-mobile convergence. Most user traffic that is converging on the multi-access edge will originate from wireline broadband users (including 5G mobile users connected to Wi-Fi), followed by fixed-wireless access, and mobile roaming services. IPTV (live and on-demand) and Internet are the largest broadband consumers by far, and this traffic should be offloaded at the multi-access edge to distributed edge caches and peering points. The remaining (mobile) traffic is forwarded to an Evolved Packet Core (LTE and 5G non-standalone) or a 5G core and mainly consists of metered traffic from mobile voice, IMS and roaming data applications.

From an scaling perspective, a multi-access edge gateway needs to make a gymnastic split to combine the need for a distributed, high-performance user plane capable of processing massive data volumes, with the goal of operating a centralized, cloud-native management and control plane that can seamlessly integrate with a 5G Core. Control and User Plane Separation (CUPS) achieves this feat and is the key enabler for the wireline broadband evolution to fixed-wireless and fixed-mobile convergence.

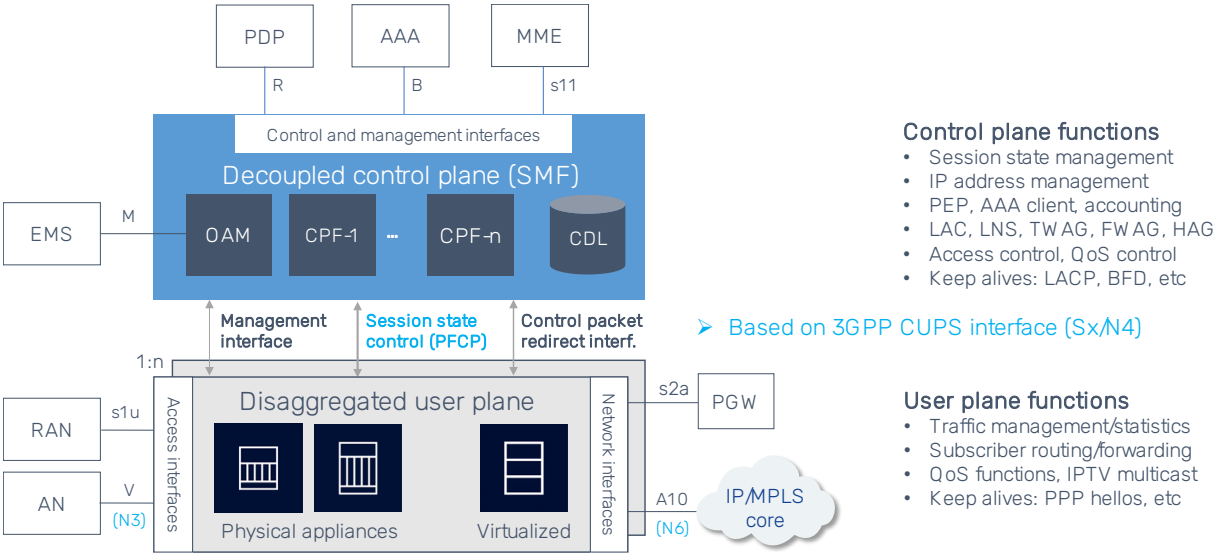


Figure 4. Disaggregated multi-access edge gateway with CUPS

A broad industry cooperation between the 3GPP, Broadband Forum and Cable Labs helped define the necessary standards (e.g., BBF TR-459) for a disaggregated multi-access edge gateway that can fully

integrate with the 5G ecosystem (Figure 4). The standards specify the functional separation of control and user plane functions and the interface protocols to be used for their subsequent interworking. The interworking protocols between the control and user plane are derived from the 3GPP CUPS standards with extensions for use in wireline access networks. They encompass the following:

- Model-driven management APIs allow for centralized management of multiple, distributed user plane functions as a single management entity.
- The Packet Forwarding Control Protocol (PFCP) is used by the control plane to manage subscriber session state. It leverages the 3GPP CUPS specification with wireline extensions.
- The Control Packet Redirect interface enables the User Plane to forward subscriber authentication messages from the CE to the centralized control plane over GTP tunnels.

The control plane (SMF) is typically cloud-native and deployed in a data center. Ideally, it uses a stateless compute model with a Common Data Layer to manage control plane state information. This model allows to easily scale out and load-balance control plane capacity, and quickly recover from failure situations that trigger a reboot of any virtualized control functions.

Control and User Plane Separation (CUPS) offers several operational benefits:
- Efficient operation, through a centralized control plane that can scale out in the cloud, while distributing physical user plane to optimize delivery cost and performance.
- Simplified maintenance, as decoupling control and user plane functions makes it easier to manage their different life cycles and minimizes the impact of hardware and software upgrades.
- Flexible scaling, by allowing control and user plane functions to be deployed on either physical or virtualized platforms and to scale their capacity independently.
- Fixed-mobile convergence and wireline broadband integration with a 5G Core by interfacing the UP via the Sx/N4 interface with a common 3GPP Session Management Function (SMF).

The last two bullets are essential for the evolution to a cost-optimized multi-access edge that seamlessly integrates wireline and wireless access with a 5G core.

6. Wireline and wireless convergence on a 5G Core

Let's now examine the evolution to a multi-access edge that interworks both wireline and fixed-wireless access networks with a 5G core. Although 5G is strongly oriented to new radio technology and mobile service evolution, wireline access networks can be integrated by supporting a set of well-defined 3GPP reference interfaces:

- **N1** between User Equipment (UE) and Access and Mobility Management Function (AMF) to exchange NAS (Non-Access Stratum) messages.
- **N2** between Wireline Access Nodes (W-AN) and AMF based on NGAP (Next Generation Application Part).
- **N3** between W-AN and multi-access User Plane Function (UPF) based on GTP-U (GPRS Tunneling Protocol User Plane).
- **N4** between Session Management Function (SMF) and multi-access UPF to manage data sessions at the user plane based on the Packet Forwarding Control Protocol (PFCP).
- **N6** between multi-access UPF and packet data networks to transport IP and Ethernet packets.

A 5G Access Gateway Function (W-5G-AGF) can be introduced to interwork legacy wireline access platforms that do not natively support the required N1, N2 and N3 reference interfaces (see Cable Labs

specification WR-TR-5WWC-ARCH). Through the W-5G-AGF, both the wireless and wireline access network can then be controlled by the 5G AMF.

The objective of delivering multi-access broadband access with 5G Core interworking is enabled by leveraging a common control plane (SMF) across both wireline and wireless user planes. The unified SMF operates on the same 3GPP N4/Sx CUPS interface and dynamically selects the proper UPF for fixed wireline/wireless or mobile access based on Access Point or Data Network Name, IP address range, subscriber profile, traffic load or configured resources and services. This allows efficient off-loading of broadband internet and video traffic from the 5G Core and allows independent scaling (and placing) of fixed/-wireless and mobile gateway functions (Figure 5).

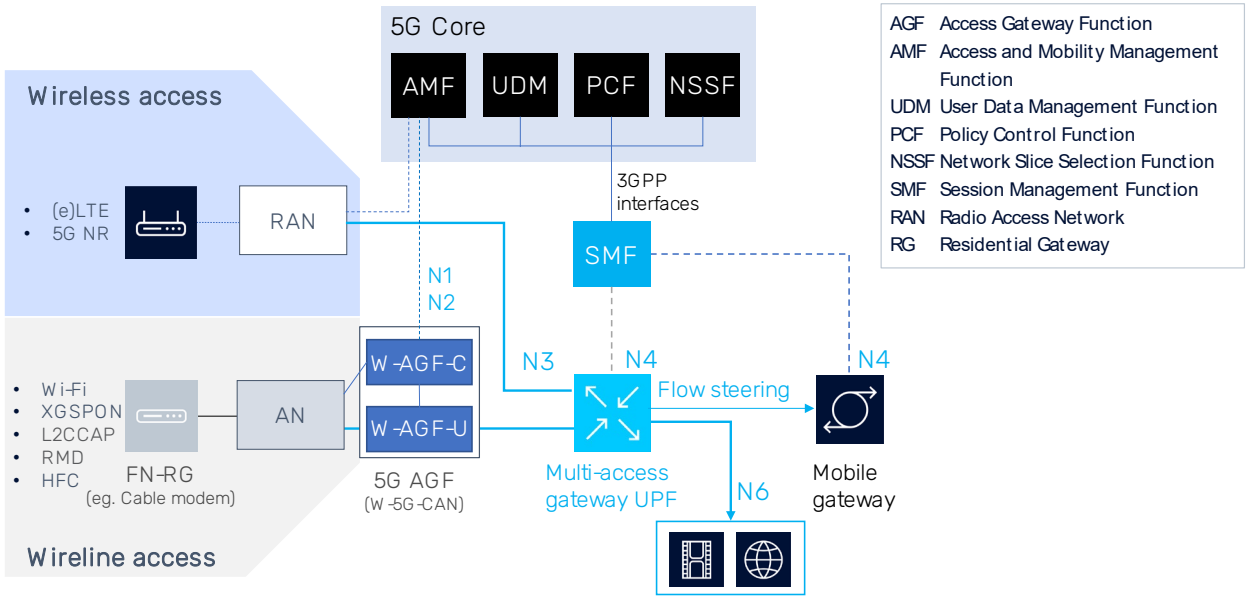


Figure 5. Wireline and wireless convergence on a 5G core

The multi-access edge gateway (UPF) handles all wireline and wireless broadband traffic. It offloads internet and video traffic to the IP/MPLS data network via the N6 interface and steers the remaining 5G user traffic to the more centralized cloud mobile gateway that control access to 5G Core services and applications.

To cost-efficiently manage high-volume/low-revenue broadband applications, the multi-access UPF is best performed by purpose-built edge router appliances that can be distributed in proximity of end users. The following requirements are typically supported by custom network processors and routing silicon:

- Highly scalable hierarchical QoS (HQoS) to enable thousands of subscribers on dozens of access nodes to fairly share the available broadband network capacity
- Granular shaping and policing of traffic flows to ensure that each subscriber receives the committed – and peak data bitrates according to their subscription policy
- Scalable and granular Access Control Lists (ACLs) to enforce security functions such as anti-spoofing to prevent service theft and denial of service (DDoS) attacks.
- IP multicast replication and IGMP snooping to facilitate premium broadcast TV streaming
- Streaming flow telemetry of subscriber flows to account for data usage and monitor quality of experience, and identify security threats,
- Flow mirroring to support lawful intercept and analyze potentially malicious traffic flows

The SMF provides a single access point for managing all distributed UPF instances and presents a common interface to the other 3GPP system functions (e.g., AMF, UDM, PCF, NSSF, etcetera).

7. Conclusion

5G fixed-wireless access is a new and promising technology that can be cost-effectively leveraged to expand the coverage and capacity for underserved broadband homes in wireline brownfield deployments, and to facilitate and accelerate the rollout of fiber-to-the-home in greenfield and out-of-region areas.

The introduction of a multi-access edge gateway with control user plane separation, enables converged service operators to combine wireline and wireless access in a seamless broadband experience, and to enable fixed-mobile convergence on a unified 5G Core.

Abbreviations

AGF	Access gateway function
AMF	Access and Mobility Management Function
CCAP	Converged cable access platform
CIN	Converged interconnect network
CPF	Control plane function
CUPS	Control user plane separation
DDA	Distributed access architecture
FWA	Fixed-wireless access
NSSF	Network slice selection function
PCF	Policy control function
SMF	Session management function
SPGW	Serving and Packet Data Network Gateway
UDM	User data management function
UPF	User plane function
XGS-PON	10 Gigabit/s symmetric passive optical network

Bibliography & References

Juan Rodriguez leads the IP Consulting Engineering team for US MSO, Majors, and Telco (Wireline and Wireless) accounts at Nokia. In this position, he combines his technical and customer service skills to contribute to the success of a dynamic, cutting-edge international organization through innovative customer solutions.

Juan is based in Orlando, Florida and holds a degree in Telecommunications Engineering from PUCMM in the Dominican Republic.

Arnold Jansen is a senior solution marketing manager in Nokia's Network Infrastructure business division and responsible for promoting IP routing products and solutions. Arnold has held a number of roles in research and innovation, sales, product management, and marketing during his 30 years in the telecommunications industry.

Arnold is based in Ottawa, Canada and holds a Bachelor degree in Computer Science from the Rotterdam University of Applied Sciences.

Flexible MAC Architecture in the Cloud: Architectures for a Virtual World

Douglas Johnson
Principal Software Architect
Vecima Networks, Inc.
Saskatoon, SK
douglas.johnson@vecima.com

Jeremy Thompson
Sr. Software Architect
Vecima Networks, Inc.
Saskatoon, SK
jeremy.thompson@vecima.com

1. Introduction

Over the last few years, and accelerated by COVID-19, the approach to corporate IT has fundamentally changed and companies are undergoing significant shifts in IT strategy and culture. The illusion of a "private, secure" network run by undersized IT teams has been shattered and companies are left grappling with the complexity and security of large, remote, organically grown networks.

Cloud as-a-service operators offer a reprieve: augment your team with our offerings. Their teams manage the day-to-day complexity and security of the services in a cost-effective way while your IT teams focus more on value-added services specific to the business. In some domains, that value-add is as simple as on-prem tech support. In other domains, the value-add can be a significant network unto itself.

The latter domain is our focus: MSOs operate significantly complex networks and domain-specific applications. MSO teams have specialized technical skills and experiences which allow operators to provide scalable and robust Internet connectivity to their end-users. Cloud offerings can augment existing investments by reducing time to deploy new services and enabling existing teams to focus on domain specific problems and solutions.

In this paper we look at a Flexible MAC Architecture (FMA) deployment following these principles. Some of the network is domain specific and managed by specialized in-house teams which are then augmented by resources and teams provided by 3rd party Cloud offerings. We examine the viability of a hybrid approach to FMA deployment through design, constraints, security, and costs using a prototype deployment.

2. FMA Overview

CableLabs®'s Flexible MAC Architecture defines an architecture for deploying a Remote MACPHY architecture, where the DOCSIS processing is done remotely in specialized hardware and the management is disaggregated into software components. The architecture is comprised of 3 primary components:

- MAC Manager (MM). A Management plane component that aggregates many RMDs into a single, unified controller. It provides a backwards compatible OSSI interface to legacy cable backoffice technologies.
- Remote MACPHY Device (RMD aka MAC-NE). A physical device containing a DOCSIS MAC and DOCSIS PHY expected, but not required, to be housed within an outside plant Node enclosure.
- PacketCable Aggregator (PAG). An aggregation component which bridges between existing PacketCable infrastructure and a population of deployed RMDs.

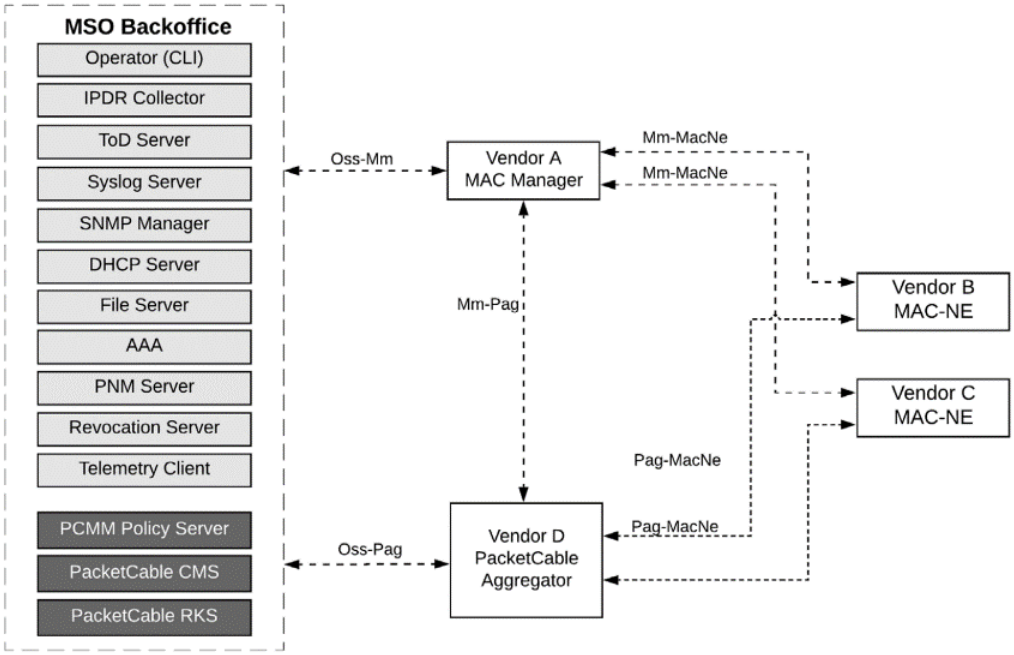


Figure 1 - CableLabs FMA

A key differentiation between FMA and Modular Headend Architecture v2 (MHAv2) Remote PHY Devices (RPDs) paired with (virtual) Cores is that the DOCSIS portion of the access network is terminated at the remote RMD and customer bearer traffic is readily available at the first-hop aggregation switches within an operators' network. FMA separates the data plane packet handling from the management components, placing data plane into the RMD and management plane concerns in the MAC Manager. This separation removes the need for the MAC Manager to handle high throughput packet processing and allows the MAC Manager to be more easily virtualized.

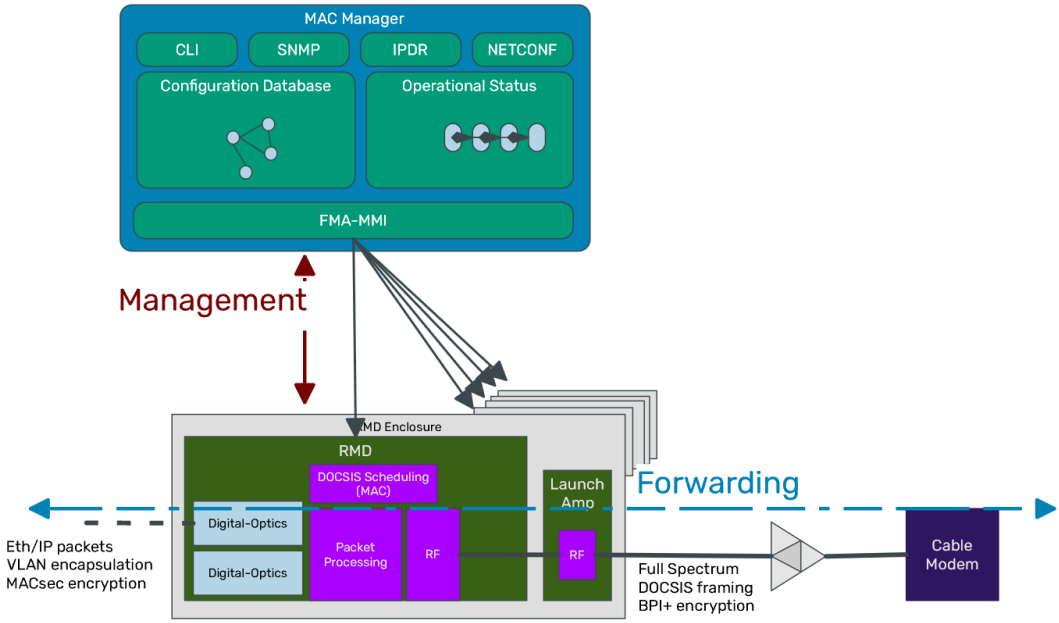


Figure 2 - FMA Management / Data Plane Separation

We take advantage of this property to build a best-of-breed hybrid network: domain specific networking concerns for bearer traffic are handled by in-house specialists while generic compute resources are augmented into the team by scaled cloud providers.

A cloud-based Flexible MAC Architecture can be readily designed in many ways, some of those options are explored within this paper as a thought-exercise. To explore a cloud-based solution more concretely, we deployed our solution for testing as follows:

- The MAC Manager was deployed into Amazon AWS, although any large cloud provider could be used.
- A VPN connection was established to a private lab network
- The MSO backoffice was on the private network
- The cable modem traffic was routed on the private network

This leads to a hybrid network where AWS was used as an extension or expansion of our private network. Customer traffic was not routed to or originated from an AWS address space.

3. Clouds

There are many options for cloud providers today, both large and small. The large cloud providers, such as Amazon, Google, and Microsoft, offer similar competitive portfolios and can be compelling partners when investigating cloud augmentation.

Cloud service offerings can be approached and purchased in different ways, and we categorize the offerings into the following from higher level to lower level. The service provided by the

cloud operator dictates the required software packaging, service model, and abstraction level for any application deployed into that service.

- **Software-as-a-Service (SaaS)**: Building an application on top of provider-specific applications services. SaaS applications can be augmented to other deployment technologies to address specific application requirements. SaaS offerings have the largest variance between cloud providers.
- **Containers**: Building an application as a set of containers deployed onto a cloud-managed Kubernetes or another container orchestration platform.
- **Virtual Machines**: Building an application bundled with an operating system and targeting an ideal hardware environment which would run on a hypervisor and be deployed as a unified whole.
- **Bare Metal**: Building an application bundled with an operating system targeting a specific hardware environment and running directly on the hardware resources.
- **Racking**: Renting rack-space, lab resources, and connectivity while purchasing and managing hardware life cycles and depreciation yourself.

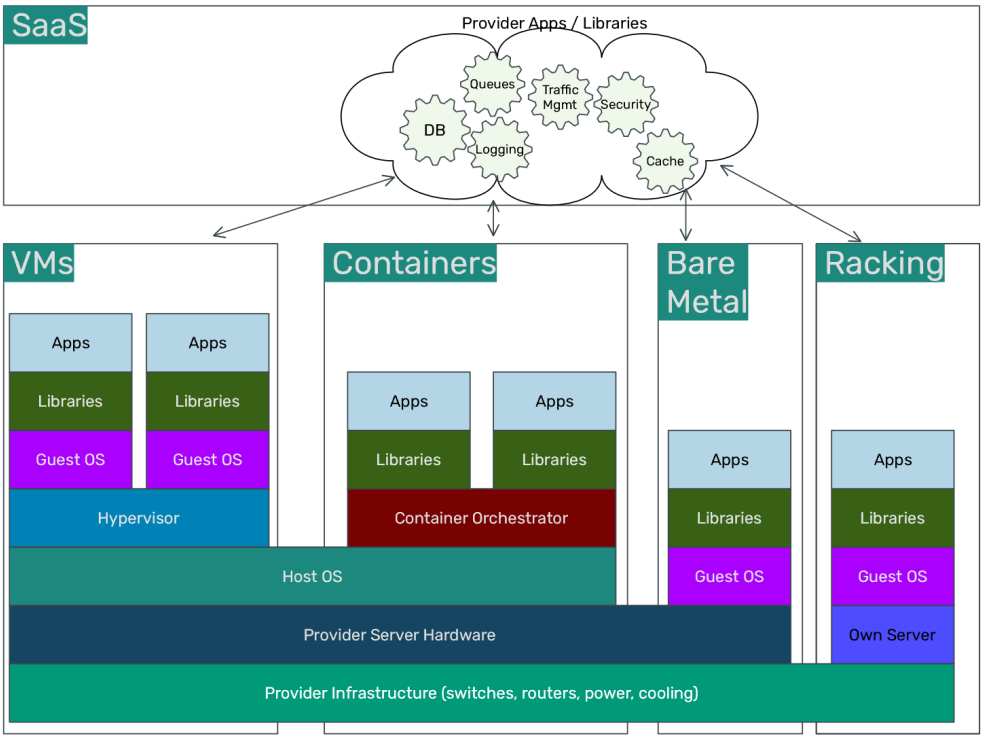


Figure 3 - Cloud offering types

The lowest level options, Bare Metal and Racking, are not as attractive to most operators because there is little value-added services added to an operator team; as such, they are not discussed in this paper.

The final 3 options of Virtual Machines, Containers, and SaaS each offer different advantages and disadvantages that need to be considered when investing in Cloud solution architecture.

4. Deployment Models

Virtual Machines, Containers, and SaaS augmentation are three ways to engage with cloud providers and each type of engagement has different strengths, weaknesses, and costs. In this section we provide an overview of these engagement models.

4.1. Virtual Machines

The most straightforward cloud-based deployment model is the placement of virtual machines (VMs) into a cloud provider network. A VM combines the application software with a bundled, often customized, operating system into a portable VM image which can be launched on a hypervisor.

Virtual machines are attractive, in part, due to their low coupling between the software application and the virtual infrastructure. This makes the VM easy to target as an application developer and easy to deploy into any cloud offering, reducing cloud vendor lock-in. Virtual machines can bridge gaps between defined hardware appliances and a fully virtualized world making it easy to work with internal teams and external vendors.

There are some inherent disadvantages to bundled virtual machine deployments. Given their agnostic attitude to their infrastructure and that they include their entire operating system, they can sometimes cost more than other options. In addition, all application dependencies are usually included directly in the virtual machine image, making it difficult to offload application features, such as database redundancy and resiliency, to the cloud operator.

A single MAC Manager is expected to manage many RMDs, so consideration must be given to redundancy, resiliency, and the impact of an outage (planned or unplanned) on customer services. In a similar way to the physical deployments that VMs emulate, high availability strategies come from the VM vendor implementation and are difficult to transparently offload to a cloud provider. When deploying redundant VMs, it's important to ensure cloud availability zones (AZs) are a part of the strategy which limits or restricts the use of layer 2 based high availability techniques. Most cloud providers offer some insight and monitoring into the health of VMs but do not have visibility into the health of the application(s) running in the VM.

The VM resources are analyzed in a similar fashion to physical deployments - in increments of CPU, memory, and storage. Evaluation and costing of compute resources in a VM model is straight-forward as the costs of a virtual machine are obvious from the cloud provider. Network traffic needs are more sensitive to the running application design and configuration, as certain parameters such as telemetry, logging frequency, and communication density may be adjustable in the application. However, between the two cost considerations of compute and traffic, network traffic costs will outpace those of the compute resources except in certain edge-cases, such as GPU compute.

There is significant parity among all the major cloud provider offerings regarding virtual machine deployments and compute resource offerings.

4.2. Containers

Containers offer a deployment option in which individual microservices can be launched and managed as discrete entities running on a virtual environment. Containers can be deployed manually but are more often paired with an orchestration service, such as Kubernetes, which performs the role of container lifecycle management, redundancy, storage virtualization, and load balancing functionality.

Containers decouple the applications into individual services, allowing each to be deployed independently. Where a VM commonly bundles "everything" into a single integrated deliverable, a container deployment provides a model to offload key functions to an operator or third parties. By discarding the Guest OS, containers reduce the virtualization overhead required to deploy applications and may reduce overall costs of a solution.

In a VM model, a hypervisor limits each VM to a specific set of resources and ensures two or more VMs operating on the same physical hardware do not interfere with each other. Containers run without a hypervisor and without a Guest OS, so resource constraints need to be considered on a per-container basis. Most container models allow for fine-grained tuning of resources and resource limits to give priority resources to the applications that need them most while ensuring that none overstep defined limits causing negative side effects to the broader system. Shared resources such as disk, memory, and network access can be defined on an individual basis that are best suited to the requirements of the application.

Orchestration systems, like Kubernetes, enable larger shared resource pools across many physical devices to be managed and container instances to be deployed automatically within the pools. Using load balancing or distribution strategies, individual containers managed by the orchestrator can have workloads distributed evenly (in, for example, a round-robin fashion) or as a redundancy strategy. Containers can exist as long-term entities for persistent, permanent operation or short-term entities for distributed workloads such as metrics processing or batch operations.

Container orchestration systems can be deployed in VMs or onto bare metal by an operator manually, however, most major cloud providers offer a Kubernetes container deployment target as a service. In these models, the operator doesn't think about or provision any virtual machines and can deploy containers directly into the cloud operator's container network. The underlying virtualization/HW is left to the cloud operator.

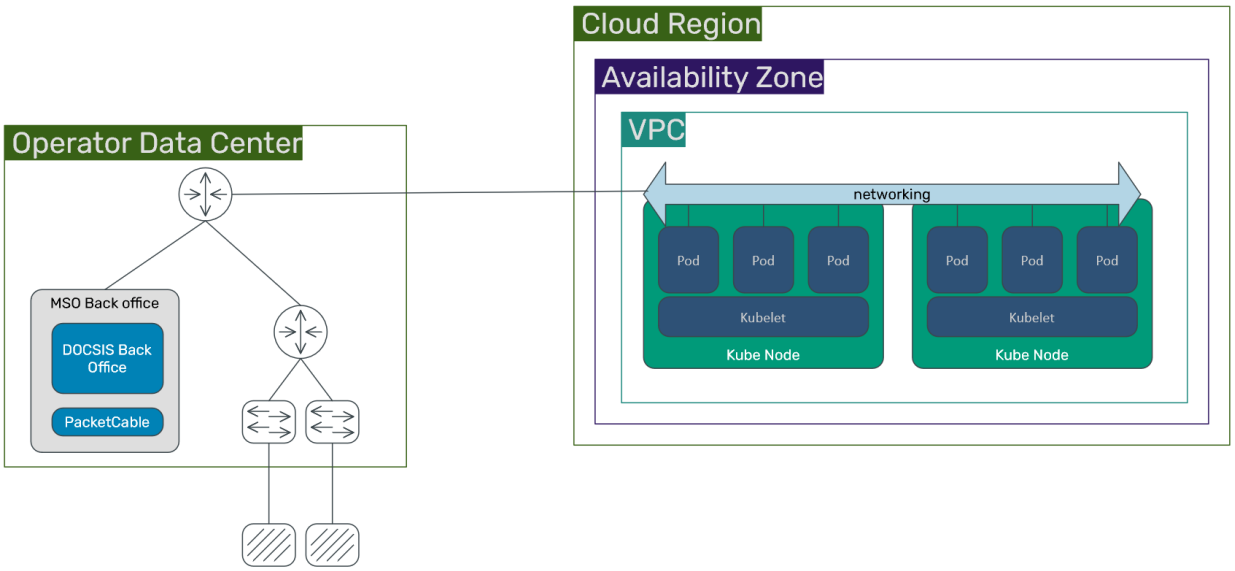


Figure 4 - Orchestrated Containers

A key principle of microservices and containers is the reduction of application-scope: A container will do one thing well and rely on other containers to provide any other services. For example, a container will often be stateless and rely on other containers or SaaS offerings to provide stateful storage, such as a database service. In some cases, the smaller container scope allows for best-of-breed application container choices and increases release velocity of individual container applications. The downside of this approach is the increase in the number of integration points an operator needs to manage. When a container needs to communicate with other containers to fulfill its responsibility, it does so with a protocol, protocol version, and specific API. These communication points need to be integration tested by the operator and vendor before confidence in the whole system can be established. These new integration tests can be empowering for an operator but also need to be understood and managed through life cycle and resource allocation.

4.3. Software-as-a-Service

Software as-a-Service (SaaS) can augment any other type of deployment, outsourcing critical generalized functionality to the cloud provider. SaaS functionality between cloud providers is the most specialized with different providers offering different SaaS products. In many cases it is also the most expensive but also the highest value functionality a cloud provider can offer. Databases, queues, traffic routers, and caches are all generic architectural components demanding high availability and reliability while also being complex to deploy and manage. SaaS offerings from cloud providers offload that complexity to their specialized teams to handle infrastructure, monitoring, security, and support infrastructure. This lowers the operator's burden and allows operator teams to focus on connectivity domain specific concerns.

Some SaaS offerings require little custom integration to gain the benefits. For example, most SaaS SQL databases are compatible with SQL client applications without any additional effort. However, in some cases, specialized high-value offerings, such as AWS Lambda, need specialized application logic and delivery mechanisms to integrate correctly.

While SaaS offerings can be cost-effective, reliable, and easy to use, they can cause an application to be locked into a specific cloud provider due to proprietary APIs or service offerings from a specific cloud provider.

Operators can compare SaaS offerings versus their own resources and areas of expertise and make individual evaluations for each of these categories of need. Ultimately, decisions will come down to costs and benefits. While maintenance of logging stacks, metrics stacks, or databases is possible independent of cloud providers, SaaS choices remain available, and these options are where distinctions between cloud providers are more evident. Furthermore, it becomes a matter of comfort, preference, or experience that may define more attractive choices to operators.

4.4. Cloud location & Connectivity

Operators will most often have, and want to take advantage of, the opportunity offered by major cloud service providers to geographically locate cloud-based deployments into areas that pair best with MSO datacenters. Latency tests indicate that even cross-continent, round trip times are low enough that RMD to MAC Manager communications are feasible without failures but reducing RMD to MAC Manager latency could improve the overall performance of interactive tasks.

A standard connection to a cloud provider would be over the public Internet with best-effort delivery. At a minimum, a VPN connection established between the cloud provider network and the operator network would be expected, but this is still delivered as best-effort over the Internet. The largest cloud operators offer additional services to reduce the latency of best-effort connectivity between an Operator core network and the cloud services. These additional services vary between the cloud operators, but we attempt to unify the concepts here.

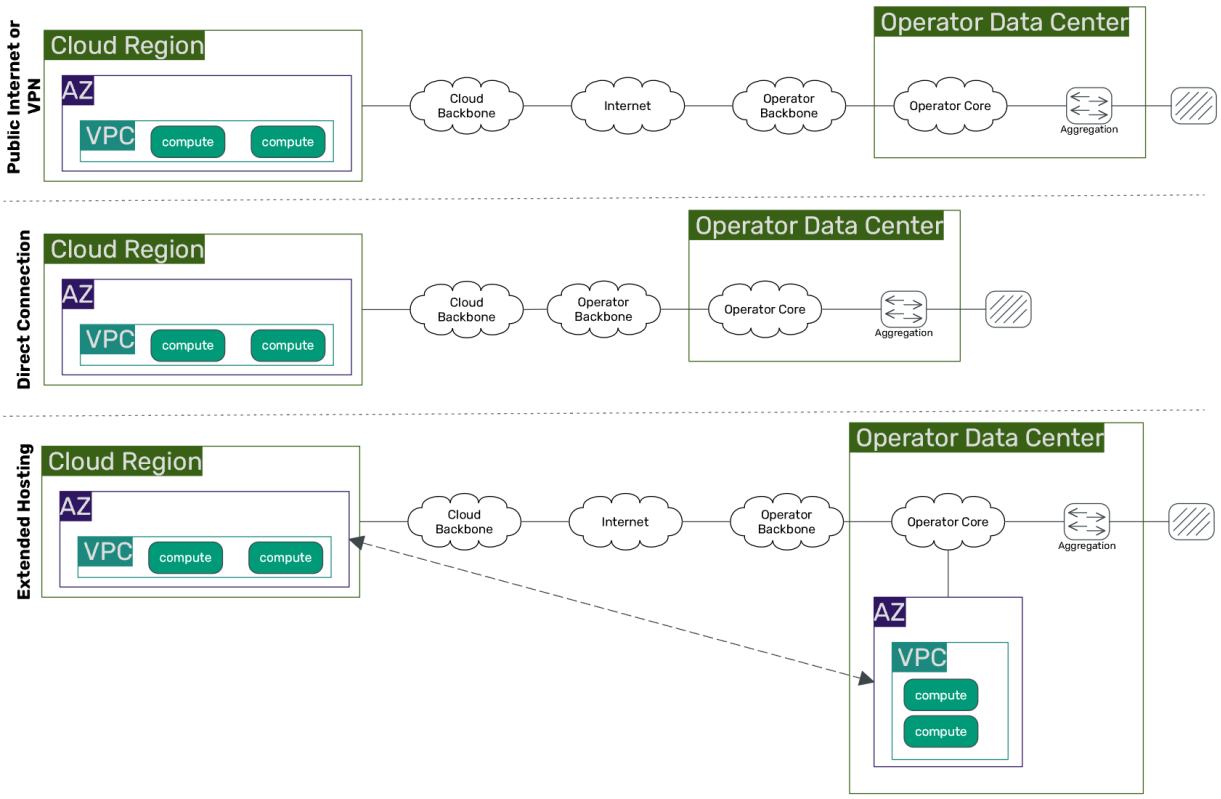


Figure 5 - Cloud Connection Types

One option is to use "accelerator" functions to attempt to route primarily on the cloud operator backbone and stay off the general Internet. This may be particularly attractive if the MSO already has peering established with the cloud operator. Another option is establishing a direct connection physically between two data centers; however, this requires presence in a common location and provisioning the connection. Another option for some cloud operators is to install their hardware directly into an MSO data center, which can then be provisioned through the cloud user interfaces.

With a diverse set of geographically located datacenters from cloud providers across North America, Europe, or Asia, operators should consider leveraging these locations for hosting their applications. It may also be required that cloud hosting be located in a specific region for legal, taxation, privacy, security, or other purposes. It's important to highlight that the MAC Manager will likely store and manipulate certain fields considered private identifying information in some jurisdictions and cloud data center geographic location will play a key role in complying with those regulations.

4.5. Backup and Retention

Depending on the nature of the cloud deployment, there are extensive options available for backup and retention of data. Virtual machines are often backed by block storage devices that can be snapshotted on a manual or scheduled interval. The same options are available for block storage devices attached to container images in an orchestrated containerized deployment. These

block backup techniques can be achieved transparently to the VM or containers, reducing integration cycles associated with backup and retention mechanisms.

The various SaaS options, such as SaaS databases, often manage and monetize their own retention models, where retention rules can be set by volume, time, or other configurable parameters.

Some cloud providers offer additional choices for backup operations, such as large-volume cold storage, where data can be retained at very low cost, but retrieval or restoration of the data often comes at a higher cost. In some offerings, physical export of the data is possible as well.

Beyond the options cloud providers offer, direct connection data links into the cloud would allow for more conventional, self-maintained backup and retention policies. In the case of certain SaaS offerings, manual export and retention of bulk data may not be fully compatible with the software offerings, or across cloud providers.

5. Cost Centers

Cost centers associated with cloud-based deployments of FMA architectures can vary greatly depending on methodology chosen and services deployed. There are some consistent elements across cloud providers that will affect cost independent of the chosen architecture:

- MAC Manager compute resource hosting/consumption (VM or container)
- MAC Manager runtime storage, with a focus on IOPS
- Bandwidth usage of FMA-MMI and FMA-OSSI traffic
- Fixed data links between MSO datacenter and cloud datacenter
- Backup and retention costs

In a resilient architecture, regardless of whether the MAC manager is situated as a monolithic software package or a microservices-based model, some measure of cost will exist. Virtual machines operated by cloud providers typically present two costing options: hourly or reserved (fixed cost by term). Hourly-costed instances offer more flexibility in terms of the actual runtime of the virtual machine. For example, operators could choose to have cold standby backup instances which may offer a cost savings approach. Similarly, in microservices models, instances could be launched or deprovisioned to support batching operations for processing data in bursts instead of instances running continuously.

Reserved instances are more cost effective outside of these types of operations. Reserved instances cost less than the equivalent hourly instance when run for the same amount of time but require a contractual lock-in over a monthly or yearly term. As a result, reserved instances require greater understanding of operational needs in advance of the actual deployment.

Storage access, particularly in IOPS, can represent a significant area of cost. Different implementations of MAC Managers may have very different IO profiles, Resilient data storage will play a key role in any MAC Manager implementation.

Bandwidth costs will vary based on several factors. Principal among them is the number of deployed RMD devices connected to MAC managers, as well as configuration on reporting thresholds and intervals of telemetry, IPDR, and logging data. As a general rule, when surveying major cloud providers, we found an asymmetric cost associated with data transfer: ingress data is cost-free and egress data is costed at total data transfer across tiered pricing intervals.

In FMA, outside of actual bearer traffic, there are two inherent modes of traffic, each bidirectional in nature but also asymmetric in the volume of data transfer. In a robust cloud-based deployment, recommended setup would have redundant or backup MAC managers distributed across availability zones to ensure impact of outage has a small footprint in terms of service impact, scope, and time of impact.

In general, bandwidth usage and anticipated cost is summarized by the following:

Table 1 - Bandwidth Usage Summary

	RMD to MAC Manager		**MAC Manager to RMD**	
FMA-MMI	High throughput	No cost	Low throughput	Costed
	MSO DC to MAC Manager		**MAC Manager to MSO DC**	
FMA-OSSI	Low throughput	No cost	Medium throughput	Costed
	MAC Manager inter-AZ communication			
MAC Manager HA	Medium throughput	Costed		

The final cost center is in bridging the MSO datacenter into the same network space as the cloud provider. The major cloud providers each offer their own variant on this high throughput, dedicated secure network link. Datacenters and links are regionally distributed, meaning that an MSO datacenter in the eastern region of the US could have a direct, non-public link into geographically matched cloud provider networks.

Similar to virtual machine costing options, these fixed and dedicated links can be found in hourly or fixed-term options, with the corresponding flexibility versus cost optimization consideration.

6. Experimental Cloud Architecture

By disaggregating the DOCSIS processing from management concerns, FMA has introduced flexibility and design choices in deployment strategies. Data plane concerns of throughput, latency, and cost-per-bit are left to be optimized by packet-processing silicon, while management plane concerns of flexibility, agility, and evolution are left to be optimized in the virtual, software-defined design space.

Shown below, we built a specific architecture to explore a Cloud FMA deployment, but other architectural choices are possible as well.

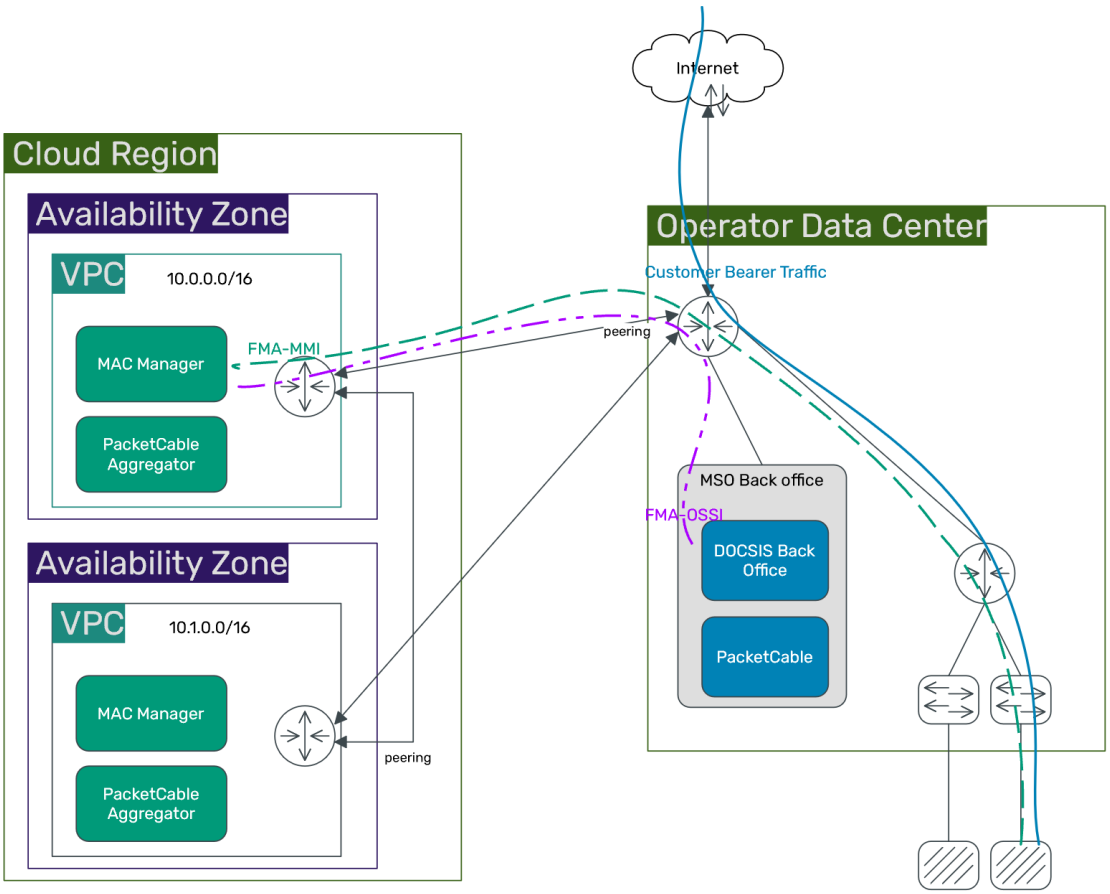


Figure 6 – Cloud-based FMA

A virtual private cloud (VPC) is an isolated network and set of cloud resources, such as compute servers, launched and configured in a cloud providers infrastructure. All resources within the VPC are private from other users, isolated from other networks including the Internet, and are assigned IP addresses from a private CIDR pool.

Resources launched within an Availability Zone (AZ) are on the same physical network, connectivity, and power infrastructure. Resources launched in two (or more) separate AZs are on unique physical infrastructure to provide redundancy and resilience, limiting the impact of an outage in one AZ. AZs are usually inter-connected with high-speed redundant links to allow for common data-replication techniques to be useable across AZs. To ensure high availability it's important to deploy applications across AZs.

Cloud providers offer different options when connecting an Operator data center to the cloud provider. Two common methods are, generically, a VPN and a direct connection. The VPN option is quick and easy to setup and provides an encrypted tunnel between the cloud and the operator network. While encrypted and secured, the VPN tunnel is routed over the open Internet and may have a variable performance profile and specific maximum throughput limitations. Another option is to directly connect from your own data center into the cloud providers data

center. This direct connection is not software controlled and involves people from both companies to install and provision physical connections between the sites.

A VPN connection can be 'upgraded' to a direct connection without impacting the logical architecture of the solution. This allows for initial deployment trials to be setup with a VPN and later, optimized into a direct connection. For our setup, we used a VPN based connection between our deployment and AWS.

Once a connection between a VPC and the MSO data center is established, routing rules need to be installed to allow communication into and out of the VPC. Routing to the cloud over either connectivity option can be done with static routes or dynamically using eBGP. Given the small network in our testing we used static routing, but larger deployments will likely want to make use of eBGP.

We placed the MAC Manager and PacketCable Aggregator into the VPC and setup a static route to our physical infrastructure. The VPC was isolated from the Internet. The RMD to MAC Manager control communication transited our aggregation network, through our traffic router, across the VPN peer connection, and to the MAC Manager in the VPC. The RMD customer bearer traffic transited our aggregation network, through our traffic router, and out to the Internet. This deployment did not make use of PacketCable. Our implementation hosted firmware files for software downloads for the RMDs in the MAC Manager component.

7. Bandwidth

Bandwidth consumption over the Cloud connection is a primary concern when moving to cloud deployments. In an FMA deployment, the management plane traffic is separated from the customer traffic and the MAC Manager is not doing per-packet data processing. This disaggregated architecture allows for hybrid network deployments by placing management components in virtualized Cloud networks and keeping customer bearer traffic within the operator core network.

Bandwidth consumption between the MAC Manager and RMD will vary between vendor implementations and services deployed. However, to attempt to understand possible real-world consumption of the cloud provider transit, we investigated bandwidth consumption of an implementation of a MAC Manager and RMD, deployed in a 2x2 configuration, with a modest number of Cable Modems.

Communication between a MAC Manager and RMD in FMA can fall into one of two traffic patterns: steady state and on-demand. Steady state traffic is a continuous exchange of data during normal operation and on-demand traffic is bursty and usually triggered by an external command, such as a software upgrade.

We then classified the different streams of communication into the following categories:

Table 2 - Communication Classifications

Category	Type	Description
Telemetry	**Steady State**	**Regular streaming of status, operational, and statistics which the MAC Manager uses to monitor RMD population.**
Configuration	**On-demand**	**MAC Manager actions to configure changes in the RMD.**
IPDR	**Steady State**	**Regular streaming of customer related statistics to fulfill IPDR interface north of the MAC Manager.**
Support Info	**On-demand**	**Extra support and trouble-shooting data gathered and stored for historical data during support cases.**
Logs	**Steady State**	**Streaming of system logs.**
Heartbeats	**Steady State**	**Regular heartbeat and RMD discovery processes.**
Firmware Upgrades	**On-Demand**	**Download of firmware to RMDs.**
SSH/CLI	**On-Demand**	**Direct SSH/CLI connections to RMDs if needed.**

We monitored the communication between the MAC Manager and the RMD over time and through regular use, classified all protocol connections into one of the above categories, and aggregated the consumed bandwidth into an average consumption rate. The values are specific to our implementation but can provide an "order of magnitude" value to allow us to understand cloud provider transit costs.

Table 3 – Bandwidth Consumption by Classification

Category	Size	Downstream Bandwidth	Upstream Bandwidth
Telemetry	**Constant**		**4.3 Mb/s**
IPDR	**Constant**		**1.0 Mb/s**
Support Info	**~ 30MB**		**76 Mb/s**
Configuration	**n/a**	**Negligible**	**Negligible**
Logs	**Constant**		**Negligible**
Heartbeats	**Constant**	**Negligible**	**Negligible**
Firmware Upgrade	**~ 131MB**	**116 Mb/s**	
SSH/CLI	**n/a**	**Negligible**	**Negligible**

The constant steady state traffic between each MAC Manager and RMD pair is about 6 Mb/s when communicating with a 2x2 RMD with a modest Cable Modem count. Telemetry and IPDR making up most of this traffic means that the consumption will vary with the number of services deployed within the RMD. To understand Cloud transit consumption, we need to take the steady state values and multiply them by the RMD population size served by the MAC Manager. So, with the experimental implementation, the MAC Manager deployed with 100 RMDs might constantly consume ~ 600 Mb/s (75 MB/s) of cloud transit. Some cloud providers also charge asymmetric rates, where "download" out of the Cloud is charged at a different rate than "upload" into the Cloud. Our investigation found most of the steady state traffic is "upload" from the RMDs to the MAC Manager.

The largest on-demand "download" operation was the software upgrade functionality. This function, managed by an operator, commands one to many RMDs to download their software upgrade file from the MAC Manager. The firmware file used in the test was about 131MB and the full download for a single RMD took about 10 seconds. While short lived, this consumption could be significant if an operator commanded an entire population of RMDs to download their firmware upgrade file concurrently. The FMA architecture does not require that the MAC Manager host SSD firmware files. If the cloud-based MAC Manager is aggregating many RMDs and an operator expects a high concurrent download demand, an attractive option is to host the SSD firmware files 'on premise' on a local HTTP server and simply issue the SSD command to download the firmware files from the locally hosted file server.

8. Latency

In an FMA deployment, the latency between the MAC Manager and the RMD can affect management plane traffic for configuration and status information but does not directly add to bearer traffic latency. This is due to FMA making the bearer traffic available at the first-hop aggregation switch rather than routing the bearer traffic through a core, such as the MAC Manager.

To better quantify the impact of latency between the MAC Manager and the RMD, we injected latency into our deployment and monitored the operational status of the deployment in the presence of latency. The latency was only injected between the MAC Manager and RMD connection and not in the data plane bearer traffic, which was separated at the first-hop aggregation switch and routed normally.

Typical Headend/Hub based MAC Manager to RMD one-way latencies we see in deployments are between 0.01ms and 8ms, inclusive of standard propagation delay and the overhead of switching elements.

The introduction of a routed network over a VPN into the cloud VPC adds latency to the MAC Manager to RMD connection. The amount of latency added by the cloud connection is highly variable and based on many factors, some of which are not within the operators control. We tested a connection to explore real-world latency values to ensure our latency tests would simulate a useful range of latency targets. The latency measurements were made from a single IP location in the US Southeast to load balancers in our cloud provider network (AWS in this setup). Values are round-trip time averages.

Table 4 - Cloud Latency Measurements

Zone (US Southeast to…)	Average RTT (ms)	Std Dev (ms)
US Northeast	30.37	9.16
US Central	64.79	13.52
US Northwest	85.99	17.43
US Southwest	79.26	14.82
Europe (Frankfurt)	132.08	17.31
Asia (Tokyo)	254.31	26.26

The overseas values were included as interesting data points but are not relevant to our testing and not discussed further due to their performance and the legal and regulatory implications of hosting a MAC Manager across international borders. Focusing only on the continental US results, we see significant result differences between the AWS regions from our RMD deployment in US southeast. It's expected that countrywide RMD deployments will likely want to be connected to MAC Managers deployed in cloud data centers with the best performance from the RMD location. In AWS, for example, we would want MAC Managers deployed in each of the 4 major regions and have RMDs connected within a single region.

For our testing, to ensure controlled injection of a range of latencies, we built a controlled network with as few switching elements as possible and used a network testing tool to inject the specific latency targets.

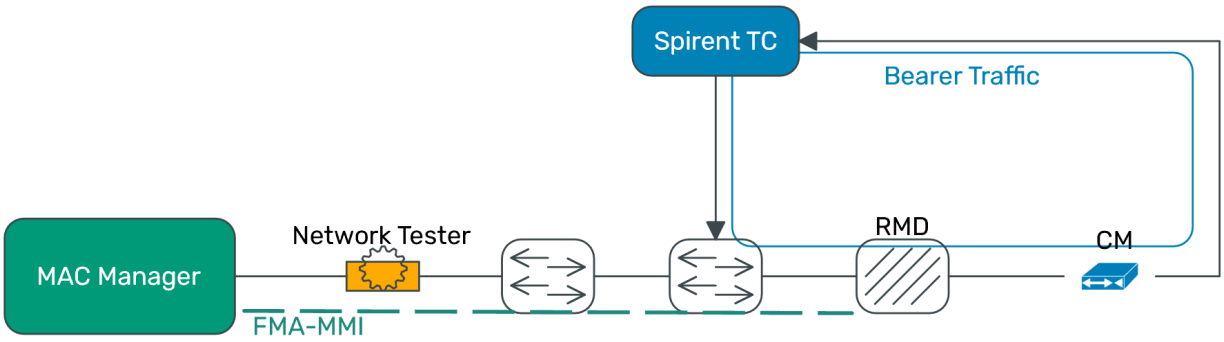


Figure 7 - Latency Test Setup

For our investigation, we choose a range of 1-way latencies between 10ms and 50ms, equivalent to 20ms - 100ms round-trip time (RTT). 10ms (20ms RTT) being worse than current on-premises, real-world deployment cases and 50ms (100ms RTT) being an upper range beyond latencies we saw with our real-world connections to our cloud provider.

When normalized to propagation distances, this results in a range as follows:

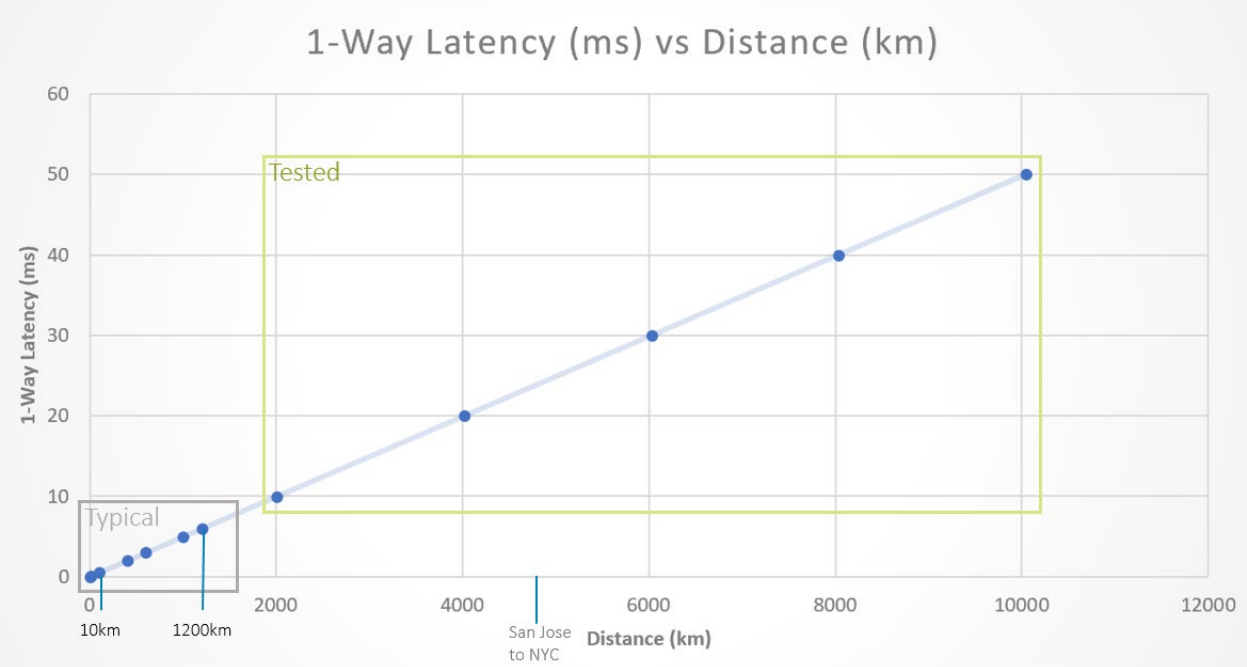


Figure 8 - Latency vs Distance (propagation)

We tested the same use cases as in the Bandwidth section, specifically:

- RMD Software Upgrade downloads
- RMD Support Info uploads
- MAC Manager to RMD Control connection
- MAC Manager IPDR (bulk data)

RMD Software Upgrade downloads represent the single largest bulk download operation from a MAC Manager to an RMD. These files are downloaded over a TCP connection to ensure reliable delivery of the upgrade file. Firmware image files sizes are highly variable between RMD implementations. In our test, the upgrade file was 131MB in size. Additionally, the specific RMD implementation we used in our test throttles firmware upgrade download speeds to 120Mb/s to ensure safe transport in the presence of other network traffic, which is visible in the baseline result:

Table 5 - Firmware Download

# of RMDs	Baseline	20ms	40ms	60ms	80ms	100ms
1	9s	10s	11s	13s	15s	19s
20	12s	32s	81s	101s	123s	144s

TCP connections used for bulk transfers are sensitive to latency due to TCP being a protocol that requires an acknowledgement from the receiver before more data is transferred. TCP utilizes a

window-size scaling algorithm that accommodates a progressive increase in the amount of data transferred per acknowledgement up to a threshold. When latency is present, a throttling effect can come into play ensuring that the data has been reliably transferred.

In the FMA model, firmware is downloaded and stored on the remote devices directly and is not downloaded during each reboot by a bootloader in the RMD. This means the firmware upgrades are only issued within the FMA system when new firmware is provided by the vendor and approved for distribution by the operator. We expect firmware upgrades to be somewhat infrequent and associated with a maintenance window. The additional impacts of latency to the firmware download within this context are minor.

The RMD Support Info files are on-demand uploads from an RMD to the MAC Manager used during support activities. The size and contents of these files are vendor-specific, but the transfer mechanisms are standardized in FMA. In our RMD implementation, these files are between 2-30MB, depending on RMD history data files, and we used a 17MB file size during the testing.

Table 6 - Support Info Upload

Measure	Baseline	20ms	40ms	60ms	80ms	100ms
Time	1.8s	1.8s	2.1s	2.6s	3.4s	4.2s
Bitrate	76 Mb/s	76 Mb/s	64 Mb/s	53 Mb/s	41 Mb/s	32 Mb/s

We also tested the operational behavior of the system under latency conditions. Latency plays a complex and not directly measurable role in the operational behavior of the other connection types, so we tested the impact of latency as to a user of the function.

Table 7 – Functionality Impacts

Function	20ms	40ms	60ms	80ms	100ms
MM to RMD Control Connection	No issues	No issues	No issues	No issues	No issues
IPDR	No issues	No issues	No issues	No issues	No issues
CM Remote Query	No noticeable delay	No noticeable delay	No noticeable delay	No noticeable delay	No noticeable delay

The MM to RMD Control connection is a TCP connection transporting YANG-based object models. The transported data is much smaller and more intermittent than the previous bulk transfer leading to negligible system impact, despite the latency introduced by the TCP Ack RTT. Another mitigating factor for the Control connection is that the DOCSIS MAC is housed

within the RMD itself, further reducing the systems impact of latency in the Control connection since there is no MAC signaling between the RMD and the MM.

The IPDR connection is between the IPDR collector and the MAC Manager and the IPDR protocol was not negatively affected by the latency injection. The MAC Manager has an internal cache for fulfilling IPDR data requests and a real-time TCP control connection round-trip for configuration and maintenance aspects of IPDR, which is where latency injection would impact IPDR operation. During each of the injected latency tests, our IPDR collector did not have any issues gathering required IPDR records from the MAC Manager.

The CM Remote Query function has the MAC Manager gather SNMP operational data from all subtended Cable Modems on regular intervals and cache the values within the MAC Manager. The SNMP protocol is "chatty" with many packet exchanges during SNMP operations which would be penalized by our injected latency. This test was to ensure that SNMP CM Remote Query would not have operational issues in the presence of high latency during the SNMP exchanges. We found from MAC Manager internal metrics and user tests that the additional latency did not impact the ability for the MAC Manager to provide the CM Remote Query functionality.

Our conclusion is that the FMA architecture is resilient and robust in the presence of latency between the MAC Manager and RMD components. The MAC Manager functions we found most impacted by the additional latency, bulk downloads/uploads, are still completed within acceptable ranges and, more importantly, are robust in the face of the additional latency.

9. Conclusion

The Flexible MAC Architecture decouples data plane and management plane concerns and provides a strong distributed access architecture in hybrid-network deployment models. With the MAC Manger control traffic decoupled from customer bearer traffic, the MAC Manager can be placed in a virtualized hybrid-cloud deployment without introducing latency on the bearer traffic. Furthermore, since customer bearer traffic is not routed through the MAC Manager, this traffic can be processed by high-throughput, low-latency specialized equipment designed specifically for Ethernet/IP packet processing.

We explored an implementation-specific MAC Manager to gather real-world tests to validate the FMA decoupling approach in variable-latency environments. We also gathered an implementation-specific baseline of bandwidth usage to validate relative or proportional utilization of the link bandwidth transit demand that could be seen in MAC Manager implementations.

For many operators, a hybrid-cloud approach utilizing AZs, VPCs, and compute engines to host FMA functionality may provide an agile framework to start small and pay-as-you grow into more advanced services provided by the major cloud operators. FMA is uniquely suited to agile approaches due to the separation of management and data plane traffic, allowing packet processing of data plane customer traffic to be managed separately and completely “on-premises”.

Abbreviations

API	Application Programmer Interface
AWS	Amazon Web Services
AZ	Availability Zone
BGP	Border Gateway Protocol
CIDR	Classless Inter-Domain Routing
DAA	Distributed Access Architecture
DOCSIS	Data Over Coax Service Interface Specification
eBGP	External/Exterior BGP
FMA	Flexible MAC Architecture (CableLabs standard)
HA	High-Availability
HTTP	Hyper-Text Transfer Protocol
IOPS	Input/Output Operations Per Second
IPDR	Internet Protocol Detail Record
MAC	Media Access Controller
MAC-NE	MAC-layer Network Element
MACPHY	MAC layer and PHY layer, apropos networking stack
MHAv2	Modular Headend Architecture v2 (CableLabs standard)
MM	MAC Manager
MMI	MAC Manager Interface
OSSI	Operations Support System Interface
PAG	PacketCable Aggregator
PHY	Physical layer, apropos networking stack
RMD	Remote MACPHY Device
RPD	Remote PHY Device
RTT	Round-Trip Time
SaaS	Software-as-a-Service
SSD	Secure Software Download
VPC	Virtual Private Cloud
YANG	Yet Another Next Generation – Data modeling language for NETCONF

Bibliography

CableLabs. (n.d.). *Flexible MAC Architecture System Specification I02.* Retrieved from https://www.cablelabs.com/specifications/CM-SP-FMA-SYS

Spirent. (n.d.). *Attero Ethernet Network Emulator*. Retrieved from https://assets.ctfassets.net/wcxs9ap8i19s/2ikhr9AEdDBsY7xJOWdtzg/8b56d702abb60c67baad8e3df0d6e473/DS-Spirent-Attero-and-Attero-X.pdf

Follow the Yellow Brick Road: From Integrated CCAP or CCAP + Remote PHY to FMA with Remote MACPHY

A Technical Paper prepared for SCTE by

Colin Howlett
Chief Technology Officer
Vecima
771 Vanalman Ave, Victoria, BC, Canada V8Z3B8
+1 (250) 881-6235
colin.howlett@vecima.com

Jeff Finkelstein, Cox Communications
jeff.finkelstein@cox.com

Rex Coldren, Vecima
rex.coldren@vecima.com

Douglas Johnson, Vecima
douglas.johnson@vecima.com

1. Introduction

From the early days of cable data services back in the early to mid-1980s until today there has been a continued improvement in supported capabilities by pushing the limits of Hybrid Fiber Coax (HFC) technologies, both in terms of modulation orders and frequencies utilized. Each step has had its fair share of corporate antigens introduced to elicit the desired outcome, and fortunately for the industry an adequate supply of antibodies has generally resulted. Monolithic architectures served their purpose for DOCSIS® head-end equipment by getting equipment deployed and reducing risk associated with rolling out these new technologies.

A monolithic architecture has limitations, primarily in terms of rack space, powering, and environmental requirements. The first part of overcoming these limitations was decomposing the Cable Modem Termination System (CMTS) by adding the use of Edge QAMs as the PHY layer for all downstream spectrum, which also allowed a single device to provide video, voice, and data services. The second piece of the puzzle was moving toward a distributed access architecture (DAA) to push these capabilities closer to the customer and remove analog optics from the deployment equation. The third strategy is moving both the MAC and PHY into the field to simplify the scheduling and timing requirements necessary for services of the future and to eventually the eliminate rack space, powering, and environmental problems of the legacy monolithic architecture.

By moving to a DAA based on a Remote PHY Device (RPD) or a Remote MACPHY Device (RMD) to provide services there are several important considerations. First, the head-end is greatly simplified as a significant portion of the combining network is no longer needed with many services being provided over Ethernet and with the transformation to radio frequency (RF) occurring in the field at the RPD and RMD. Second, the rack space, power, and cooling requirements in the head-end are reduced and potentially eliminated. Third, an all-digital network means that higher order modulations may be used as the Ethernet to RF transition is now within no more than a few thousand feet from the customer.

Transitioning to new technologies is challenging, but we must consider why we should do other than just extending the life of the monolithic architecture as a primary means of service delivery to customers.

- When cost of entry is low, cost to exit is high. The easy technologies to deploy are typically the hardest to move away from when needed.
- Moving to the next generation of a technology is cheaper than trying to keep the previous generation meeting the needs of customers. There comes a point where it is clearly no longer cost effective to keep an existing technology working. When you finally realize that you can be years behind in getting a new technology deployed.
- Changing technology is easy; changing people is hard. Never forget it is the field staff who keep things running and they need to be ready when new things are deployed.
- Everyone has a nice paint job. Most competing technologies look nice and shiny but are mostly similar for 80%-85% of necessary functions.

In this paper we address how we can progress towards a DAA and minimize the risk along the way. Major upgrades are never easy, but with careful planning and training they can be accomplished. One of the biggest challenges in the brave new world of DAA is how to manage new devices without forcing significant changes to OSS/BSS/EMS/NMS software that has been used for decades. To stand up to the challenge, we have been working for the past 8+ years on Remote PHY and the 3+ years on the Flexible MAC Architecture (FMA), which includes Remote MACPHY. Remote PHY and FMA provide standard interfaces both north and southbound of their respective Core network functions to manage DAA devices, removing OSS/BSS/EMS/NMS roadblocks and allowing the journey to DAA to begin.

2. Deployment Motivations for DAA and FMA

With DAA, we're not in Kansas anymore. Getting there may seem as scary as the idea of riding a tornado to land a house on top of a wicked witch. Where we came from was the legacy CMTS and an Edge QAM, a Converged Cable Access Platform (CCAP) and an Edge QAM, and a fully integrated CCAP. Familiar things that will always be a part of us. Where we are going is a land where DAA thrives and where CCAP with Remote PHY support live. And most certainly lots and lots of Munchkins.

In the DAA Land of Oz the Emerald City is where FMA and Remote MACPHY reside, and we must not only get to Oz but journey through it if we want to get to the Emerald City. But first, we need to consider why we might want to go there. Then we need to take a good look around at where we are now. We must consider what we expect to find at our destination and plan for what we need to take with us when we go there. Only then can we start our journey in earnest.

Literally speaking, an operator wanting to get to FMA and Remote MACPHY from where they are today has a lot to consider. What can be reused, what needs to change, and how? What service evolution factors influence how FMA is implemented with Remote MACPHY? How can the legacy approaches co-exist with FMA in a gradual roll-out?

But let's first start with the motivations…

2.1. Why DAA?

Why should we want to go to the DAA Land of Oz in the first place? Let's focus on five areas of interest as illustrated in Figure 1:

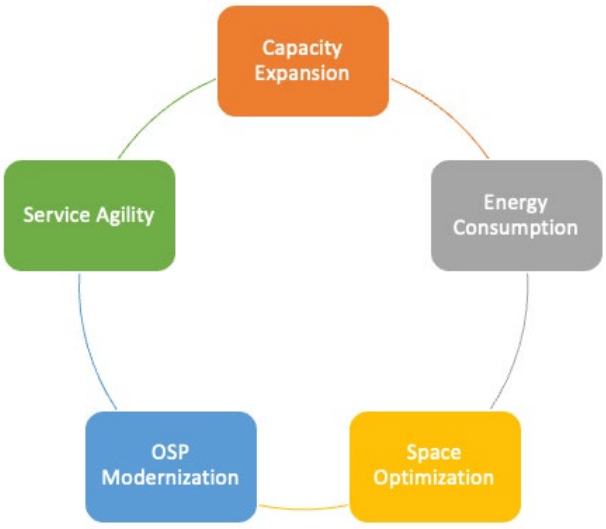


Figure 1 - Benefits of DAA

First, capacity expansion is a critical consideration. As shown in Figure 2, consumer speed tiers continue to scale on an exponential basis with consistent cumulative growth rates. Existing chassis-based CCAP devices that deliver the DOCSIS® and serve 50+ service groups cannot scale to the maximum throughputs needed while at the same time keeping all the RF circuitry within the platform. Hub space, power, and cooling requirements necessary to scale capacity of analog optical transmitters and receivers, the RF

combining networks to serve them, and the traditional CCAP devices themselves are too bulky if we continue to shrink service groups to deal with increased capacity needs.

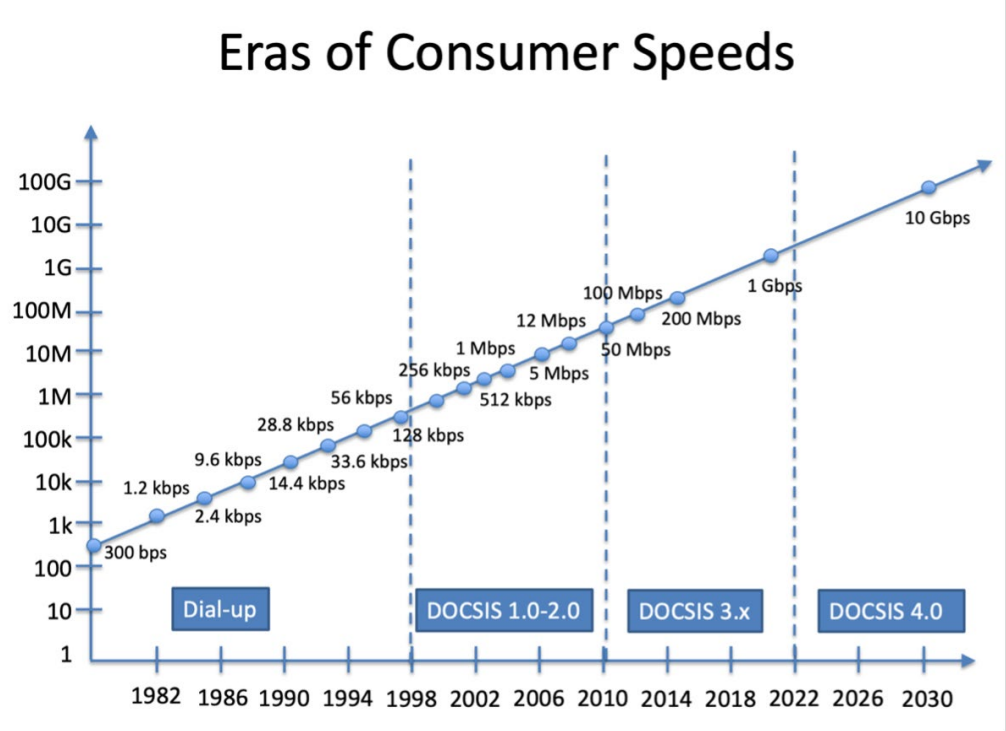


Figure 2 - Consumer Speeds over Time

To avoid regrettable spend on real estate rather than on network infrastructure to serve customer needs, we need to slow and eventually reverse hub space growth and associated increased power and cooling requirements. Remote PHY can help by removing the DOCSIS PHY and RF modulation/demodulation aspects from the CCAP and placing them in the outside plant (OSP). Remote MACPHY can go further by moving all DOCSIS functionality and RF modulation/demodulation to the OSP. Both alternatives can eliminate the need for a hub-based analog combining network and analog optical transmission equipment, replacing them with a Converged Interconnect Network (CIN) of compact Ethernet switches.

Modernizing the OSP by moving from analog optics to digital Ethernet optics increases optical reach, improves performance, and reduces the cost of fine-tuning in the OSP. Moving the RF modulation/ demodulation closer to the subscriber greatly improves RF performance. The combined effects of moving to digital optics and relocating the RF modulation/demodulation close to the subscriber also lead to a cleaner and easier to operate OSP. The movement to all-Ethernet in the OSP optics allows convergence of the network into a CIN with DOCSIS, passive optical networking (PON) for fiber-to-the-home, 4G and 5G wireless Xhaul, and business services all served from a common shared Layer 2/Layer 3 (L2/L3) network.

Finally, DAA is the only path to get to service levels beyond DOCSIS 3.1. DOCSIS 4.0 requires a DAA baseline, whether it be Remote PHY or FMA/Remote MACPHY. You can get to the DAA Land of Oz from Kansas, but you can’t get to DOCSIS 4.0 from Kansas without first taking a trip through Oz.

2.2. Why FMA? Why Remote MACPHY?

Once we're in the Land of Oz, why would we want to go all the way to the Emerald City? The answer is, that's where FMA and Remote MACPHY are.

FMA has standardized the Remote MACPHY architecture, which has been deployed globally in pre-standard implementations. Standardization allows RMDs from multiple vendors to work in the same operator network by interoperating with a single management entity and provides the operator with protection from the risks of being locked into single-vendor solutions.

The Remote MACPHY architecture is a step forward architecturally from Remote PHY. Eliminating the split DOCSIS architecture removes the tight coupling and synchronization requirements from the Core network element and the OSP devices. The "Core network element" in Remote MACPHY is a software management function that only needs to support device and subscriber management, not DOCSIS MAC and upper layer functions such as the DOCSIS upstream and downstream schedulers. It also does not need to deal with the DOCSIS data plane at all as FMA disaggregates the data plane and management planes into separate elements. With this disaggregation and decoupling comes several benefits:

- Simpler Interoperability – Interoperability for Remote MACPHY will be simpler relative to split DOCSIS Remote PHY interoperability, which has been and continues to be cumbersome due to the tight coupling of MAC and PHY between CCAP Core and RPD.
- Latency – DOCSIS latency of request/grant cycles between Cable Modem (CM) and DOCSIS MAC function is contained within the RF network between RMD and subscriber in Remote MACPHY. This provides optimal support by eliminating the delay between CCAP Core and RPD given the Remote MACPHY scheduler is co-located with the MAC and PHY layers in the same device.
- Simplified CIN – Remote MACPHY's CIN is simplified. Traffic engineering for DOCSIS data traffic can easily be provided in a native Ethernet/IP network, whereas it cannot in Remote PHY since data traffic is encapsulated in L2 tunnels. There is also no need to design the CIN to hairpin data traffic through a purpose-built Core element in Remote MACPHY and the need for precise phase and frequency synchronization between MAC and PHY from RPD using IEEE 1588 Precision Time Protocol (PTP) is not necessary for DOCSIS support as the MAC and PHY are co-located. Since long-term cable networks will likely run video over IP (i.e., DOCSIS) the Remote MACPHY approach can minimize or eliminate the need for PTP.
- Reduced hub space, cooling, power – Remote MACPHY enables the minimum requirements in hub space, cooling, and power of any of the DOCSIS alternatives. The "Core network elements" are COTS servers for management and control functions rather than chassis-based systems that pass hundreds of Gigabits of data plane traffic. These COTS servers can be located anywhere in the network and do not need to be near the OSP devices.
- Optimized virtualization of management/data plane – Remote MACPHY and FMA were designed to maximize virtualization opportunities for operators by cleanly separating the management and control functions from the data plane functions. The simplified CIN of Remote MACPHY plays nicely in a virtualized architecture by maximizing native transport virtualization opportunities as well.

There are plenty of reasons to want to go to the DAA Land of Oz, and there certainly are incentives enough to take a stroll toward the Emerald City once we are there.

3. Current Deployment Architectures

Before we can begin our trip to the Emerald City, it is important to understand where we are starting from so that we can plan our journey and know what to take with us. We first must document what is included and how the system functions are realized in existing centralized access architectures for DOCSIS and QAM video.

3.1. Pre-DAA

In Kansas we have a centralized access network with lots of splitting and combining and analog optics. We have legacy CMTS and an Edge QAM, a CCAP and an Edge QAM, and a fully integrated CCAP. The state-of-the-art in Kansas is the fully integrated CCAP.

The fully integrated CCAP includes a routing engine, cable control plane and subscriber management functionality. It also brings together DOCSIS and QAM video functionality and has the analog RF ports which connect to a splitting and combining network and then to analog optical transmission gear. In practice only narrowcast video is provided through the integrated CCAP. In many cases, broadcast video is processed external to the CCAP and needs to be combined in the combining network separately. Also not included in the integrated CCAP is video out-of-band processing, which is provided by special hub and head-end gear that connects over RF into the splitting and combining network. Figure 3 illustrates the fully integrated CCAP.

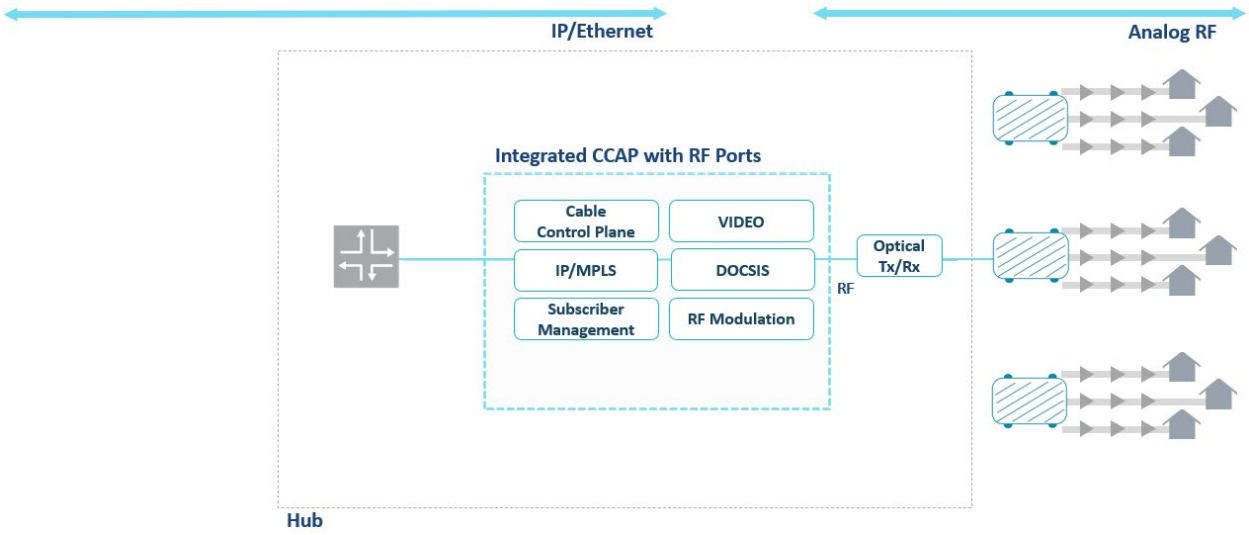


Figure 3 – Integrated CCAP

The RF access network in this environment starts in the hub. It includes the splitting and combining network and analog optical transmission gear. Analog fiber nodes are deployed in the OSP to perform RF optical/electrical conversion. The serious OSP operations maintenance activities begin back at the hub.

The difference between the integrated CCAP architecture and other RF-based hub equipment is typically in the video. A CMTS provides only DOCSIS functionality and is far less dense than a CCAP. A non-integrated CCAP is analogous to a dense CMTS. Narrowcast video is provided in both cases by a separate Edge QAM. Video out-of-band in all cases uses special gear which must be combined in the RF domain.

3.2. DAA with Remote PHY

In Oz we have a distributed access network, DAA, which provides us both Remote PHY and Remote MACPHY. Due to existence of legacy CCAP deployments a first DAA step for many operators was to convert their “big iron” chassis-based CCAPs to support Remote PHY.

In the chassis-based Remote PHY environment are most of the same things provided in the integrated CCAP. This includes a routing engine, cable control plane and subscriber management functionality. It also brings DOCSIS functionality but does not have analog RF ports, a splitting and combining network, or analog optical transmission gear. Those are replaced with Ethernet-based digital optics and a stack of Ethernet switches.

The RF line cards of the integrated CCAP are replaced with Ethernet-based Remote PHY line cards. These line cards perform upper layer DOCSIS and DOCSIS MAC functions. The DOCSIS PHY and RF modulation/demodulation are moved to the OSP in the RPD, typically in a fiber node form factor. This split DOCSIS is achieved via a complex set of Remote PHY requirement specifications provided by CableLabs®. Precision timing is required at both the Core network element and RPD to support DOCSIS.

Narrowcast video that was provided through the integrated CCAP cannot be supported as it was before. RF video of all types in DAA needs to be packetized and sent over the CIN. For this purpose, specialty elements called Video Cores and Traffic Engines are added to the architecture to provide packetized RF video which is then modulated at the RPD and delivered to the traditional video end devices. Legacy video out-of-band needs to be treated in a similar manner by what is dubbed an Out-of-band Core.

Figure 4 illustrates the integrated CCAP Remote PHY architecture.

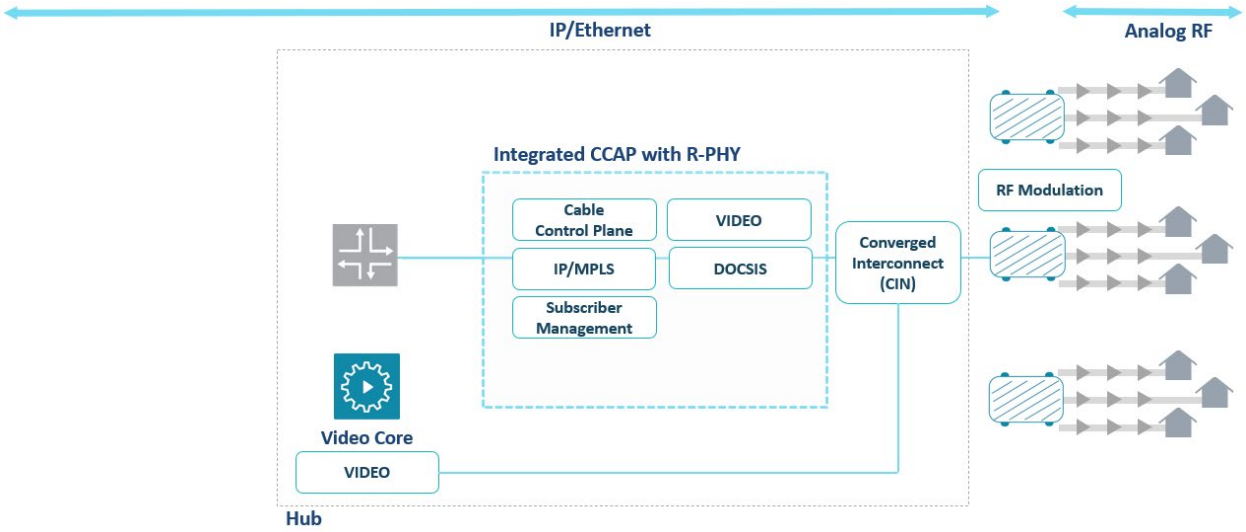


Figure 4 – Remote PHY Architecture

The RF access network in this environment starts in the Remote PHY node in the OSP. The serious OSP maintenance activities begin at the Remote PHY node, which is a significant DAA improvement versus the centralized access architecture.

4. Future Deployment Architecture – DAA with FMA

4.1. Overview

As much fun as the trip might be, getting to FMA does not literally require a tornado flight to the DAA Land of Oz nor a lengthy journey on foot to the Emerald City. Simply put, it is an evolutionary step that builds upon the successful foundation of DOCSIS, CMTS, integrated CCAP, and Remote PHY.

The key benefit of a DAA deployment is to converge the data plane packet interfaces across a range of access technologies on common and well-understood Ethernet/IP technologies. This allows operators to reap the benefits of DOCSIS and deep RF signal modulation/demodulation in their coax network, while also leveraging common high-volume packet-switching products on the fiber part of the access network.

FMA is a set of standards that further improves on DAA concepts by decoupling the management plane and data plane components of a traditional integrated CCAP, allowing each concern to be treated and deployed separately. The data plane can be optimized for throughput, latency, and uptime, while the management plane can be optimized for scale, agility, and velocity.

FMA defines three primary new components:

- MAC Manager (MM) – A management plane functional component that aggregates many MAC Network Elements into a single, unified controller. It provides backwards compatible OSSI interfaces to legacy Multiple System Operator (MSO) backoffice technologies.
- MAC Network Element (MAC-NE) – A physical device containing a DOCSIS MAC and DOCSIS PHY. It is expected, but not required, to be housed within an OSP Node enclosure. The MAC-NE embodied currently in FMA is the RMD. We refer to the FMA MAC-NE as an RMD throughout the paper.
- PacketCable Aggregator (PAG) – An aggregation functional component which bridges between existing PacketCable infrastructure and a population of deployed MAC-NE/RMDs.

FMA optionally reuses some Remote PHY components:

- Video Cores and Traffic Engines – FMA is compatible with the Remote PHY legacy video infrastructure.
- Out-of-Band – FMA is compatible with Remote PHY SCTE 55-1, 55-2, and Narrowband Digital Forward/Narrowband Digital Return (NDF/NDR) Remote Out-of-band (R-OOB) solutions.
- PTP – FMA is compatible with Remote PHY PTP profiles.

The FMA architecture is illustrated in Figure 5.

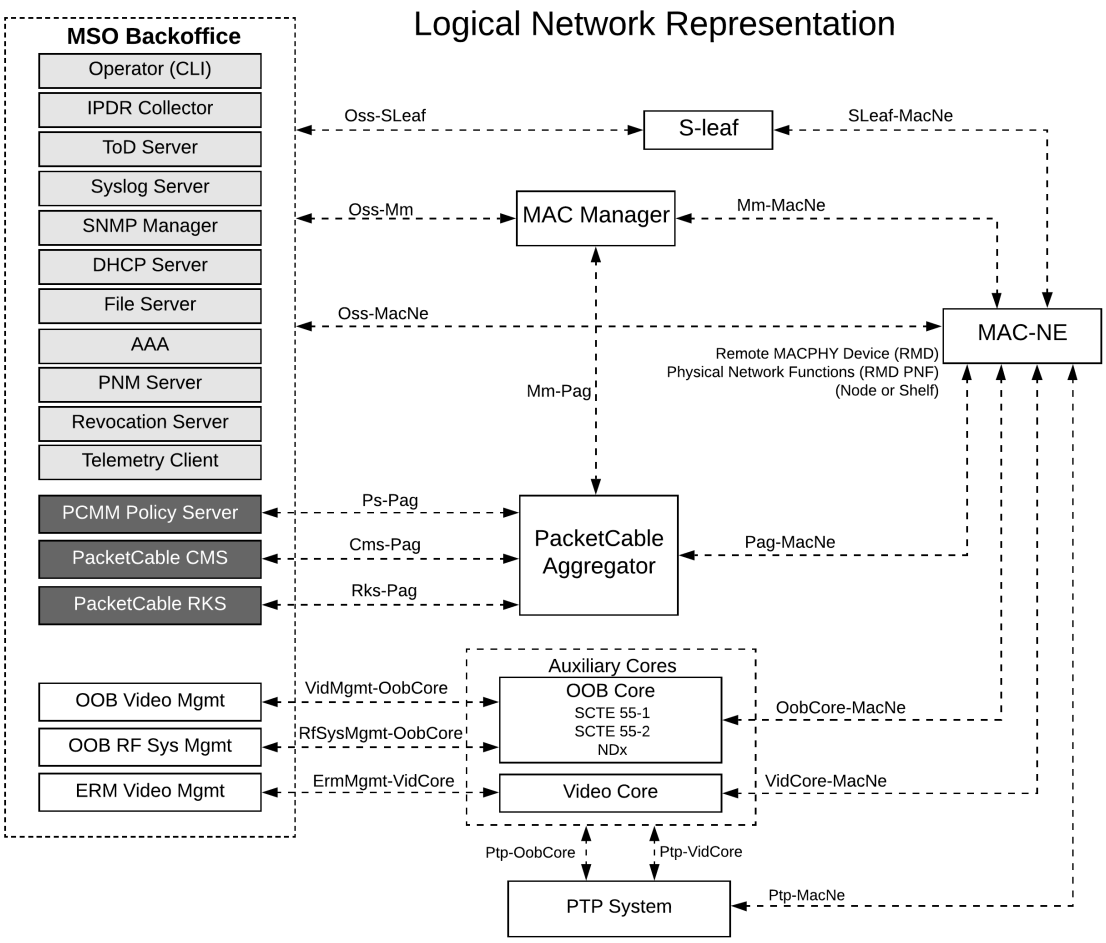


Figure 5 – CableLabs® Flexible MAC Architecture

In FMA, each interface is named by specifying both ends of the interface. A few key interfaces within the specification are:

- Oss-Mm – OSS to MAC Manager. A backwards compatible interface with existing MSO backoffice interfaces (e.g., SNMP), realized on the MAC Manager.
- Mm-MacNe – MAC Manager to MAC-NE. A new API, defined in YANG modules, between the MAC Manager and the MAC-NE to command, control, and monitor MAC-NEs. Mm-MacNe is commonly referred to as the MAC Management Interface (MMI).
- Pag-MacNe – PAG to MAC-NE. A new API, defined in Google Protocol Buffers, to command-and-control PacketCable functionality on the MAC-NE. Pag-MacNe is commonly referred to as the PacketCable Aggregator Interface (PAI).
- SLeaf-MacNe – CIN Secure Leaf to MAC-NE. A data plane interface between the first-hop aggregation points in the CIN and the MAC-NE.

Like Remote PHY, FMA places the full-spectrum RF generation and reception in a remote location, such as a node enclosure deployed in the OSP. This allows FMA to realize the same RF signal quality benefits as Remote PHY in improving Modulation Error Rate (MER) and allowing for higher modulation rates in channel plans. Unlike Remote PHY, which tunnels DOCSIS packets over IP to/from a packet-processing

CCAP/DOCSIS Core, FMA terminates DOCSIS in the remote device and provides raw customer bearer traffic directly in Ethernet/IP packets to the first-hop aggregation device. This removes the need for a CCAP/DOCSIS Core in the architecture and enables per-packet-processing and routing to occur in standard switch and router products. This allows operators to access the larger switching market and bundle switch and router cores into higher volume products. Further, because FMA provides standard Ethernet/IP packets at the SLeaf-MacNe interface, FMA is converging with other access technologies on common and proven packet interfaces.

4.2. Management Considerations

In FMA, the MAC Manager acts as the management plane entity which aggregates configuration and operational status of many individual RMDs into a single functional entity. The MAC Manager does not participate in data plane packet handling, a function that is housed completely within the RMD. However, the MAC Manager has larger-scope visibility and intelligence, and the ability to dynamically adjust the behavioral parameters of the distributed RMD packet-processing entities.

The MAC Manager provides an OSSI compatible interface to existing MSO backoffice systems. This enables a backwards compatible transition from CMTS/CCAP deployments into FMA deployments. SNMP, IPDR, and Command Line Interface (CLI) for the RMD population is all realized in the MAC Manager for all subtended RMDs. DHCP and TFTP snooping/intercept is implemented in the RMDs and learned services are provided to the MAC Manager for operational management over the MMI.

The FMA-OSSI specification extends the existing CCAP-OSSI with new object models to manage the subtended RMDs. Streaming telemetry is expected to be added in future phases to allow metrics to be streamed directly from the RMDs or from the MAC Manager to telemetry collectors in the cable operator backoffice.

From a management perspective, transitioning to FMA from an existing CMTS/CCAP with or without RPDs is straightforward. MAC Managers will likely be available as software-only deliverables (virtual machines or container clusters) or may be available in ready-to-deploy server appliances as well. After launching/installing the MAC Manager and giving it an IP address in a routable network, the MAC Manager would be ready to be used.

Like the approach used in the Remote PHY architecture, in FMA the MAC Manager requires L3 connectivity to any RMDs that it will be managing. Also like the approach used in Remote PHY, the RMD learns the network address of the MAC Manager via DHCP options. This discovery mechanism has the RMDs announce their presence to the MAC Manager using addresses provided over DHCP, either via IPv4 or IPv6, to the RMDs. RMDs support discovery of MAC Managers via DNS, IPv4, or IPv6 address. The DNS mechanism is new in FMA and is not supported in Remote PHY. It provides a clear path to virtualized MAC Manager solutions. When deploying container-based solutions into an orchestrator, such as Kuberenetes, the favored approach is through dynamic IPv6 addressing of container instances. Using MAC Manager discovery through DNS eases the deployment of container-based MAC Managers and helps bridge the dynamically orchestrated container environment to the statically deployed RMD environment.

Existing MSO backoffice implementations continue to operate when moving to a MAC Manager and RMD based deployment, while new functionality such as new FMA-OSSI object models, can be layered into existing operational tooling to gain better visibility to the RMD status directly.

4.3. Data Plane Considerations

4.3.1. DAA Advantages

In FMA full spectrum RF is generated at the RMD, just as is done in Remote PHY at the RPD. In CMTS and integrated CCAP, full spectrum RF is generated in the head-end/hub. This relocation of full spectrum RF generation has multiple benefits, including moving from analog modulation (AM) to digital-optics in the HFC fiber plant, improved fiber redundancy, and improved modulation orders in the HFC plant.

Converting the legacy HFC AM fiber to instead carry digital-optical signals represents one of the single most important aspects of moving to a DAA in both Remote PHY and Remote MACPHY deployments. Removing optical AM modulation of RF signals in HFC fiber removes the errors introduced by optical losses such as clipping and optical noise. Digital optics in SFP module form-factors are more robust, enable higher throughput, and make better use of optical wavelengths. Digital optics carry Ethernet/IP packets in a point-to-point connection between the CIN and the OSP Node enclosure housing the RPD and RMD modules. When deployed with DWDM the fiber cable is shared with other point-to-point SFP pairs at unique wavelengths.

In DAA, on one end of the SFP pair is an aggregation switch or router in a Hub or Headend aggregation point. On the other end is an individual RMD or RPD device. Redundancy of this point-to-point link can be handled passively at the optical/Layer 1 or actively at L2 by deploying two SFP pairs with unique fibers or wavelengths between the CIN and a single RPD/RMD, and then configuring active redundancy in the RPD/RMD software. In FMA, active redundancy is standardized using 802.3ad Link Aggregation Group (LAG) and Link Aggregation Control Protocol (LACP).

DAA moves full-spectrum RF generation into the OSP. This supports improved spectral density with higher modulation orders and better MER on the RF portion of the network, which no longer includes the AM optics. All that remains to impact modulation order and MER is the coaxial distance between the DAA node and the CM. This creates an opportunity to greatly improve modulation profile configurations in DAA versus legacy CMTS and integrated CCAP architectures.

The DAA data plane is clearly superior to legacy architectures for the reasons previously mentioned. Within DAA there is however a data plane divergence between Remote PHY and Remote MACPHY to consider. In Remote PHY the MAC continues to be based in a centralized component through integrated CCAP chassis line-card upgrades or transitioning to a virtual Core architecture. DOCSIS packets are opaquely tunneled across the CIN. In Remote MACPHY the MAC is deployed along with the PHY in the RMD. This makes subscriber bearer traffic more readily available at the first-hop Ethernet switch or router. The Remote PHY choice supports integrated CCAP chassis reuse through line-card upgrades. The Remote MACPHY choice enables early routing and maximizes CIN convergence with other applications such as PON, 4G and 5G wireless Xhaul, and business services.

4.3.2. Remote MACPHY Data Plane

Figure 6 illustrates the Remote MACPHY data plane logical architecture. Upstream, the RMD uses digital-optical modules to provide connectivity back to the aggregation network in a hub/headend location. Commonly, this link would be 10G SFP+ DWDM LR modules, though FMA isn't prescriptive on the digital backhaul. The optical interface of the RMD carries subscriber bearer traffic over Ethernet/IP which in the upstream direction has already been BPI+ decrypted and reassembled. FMA RMD uses MACsec L2 encryption technology between the first-hop aggregation device and the RMD to ensure subscriber traffic is encrypted while transiting the OSP fiber links.

Downstream, the RMD outputs RF which could be provided to a launch amplifier and further extended with line amplifiers in N+X or N+0 plant designs before reaching the taps and customer premise. FMA supports full spectrum up to 1.2 GHz DOCSIS 3.1 and 1.8 GHz DOCSIS 4.0, but plant upgrades from existing upper frequencies are not explicitly required by RMD deployments. The cable-plant coax/RF return is also terminated in the RMD with link processing and packet-reassembly occurring in the RMD.

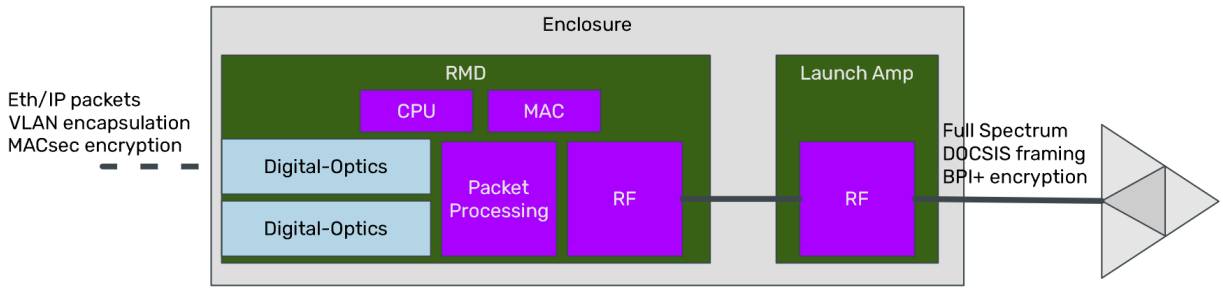


Figure 6 – Logical RMD Architecture

The RMD implements the DOCSIS MULPI and PHY specifications as a L2 CMTS. Remote MACPHY terminates the DOCSIS network in the RMD, a key difference compared to Remote PHY. The DOCSIS MAC is contained within the RMD and both downstream and upstream packet scheduling is managed within the RMD itself.

In Remote MACPHY no additional DOCSIS-specific packet-processing components, physical or virtual, exist in the network. All bearer traffic transiting into or out of the RMD are transparent native Ethernet/IP packets handled by standard switching and routing solutions. Hub and headend space are consequently significantly reduced since no additional compute equipment specific to DOCSIS packet processing is required. The amount of switching equipment needed in the CIN may also be reduced in FMA compared to Remote PHY because the DOCSIS traffic is no longer opaquely tunneled to a packet-processing entity for DOCSIS-specific processing, which can cause the bearer traffic to transit twice across the CIN due to hair pinning to reach the packet-processing entity.

When compared to an integrated CCAP deployment the FMA architecture decomposes the monolithic chassis into composite parts and places functionality into best-fit locations. Access-technology specific functionality moves to the edge closer to the customer and is contained in its optimal location. Packet aggregation at L2 is placed within first-hop aggregation switches and readily capable of horizontal scaling. L3 routing is decoupled from L2 aggregation to allow it to be more centrally deployed and managed.

This architecture can be seen to be very similar to a traditional Broadband Network Gateway (BNG) used with L2 PON devices or in telco copper access with digital subscriber line access multiplexers (DSLAM). Shifting routing to a BNG-like implementation with L2 in the RMD allows for easier convergence for operators deploying a mix of HFC and fiber-to-the-home.

The Remote MACPHY architecture functional decomposition between CIN functions and RMD functions is illustrated in Figure 7.

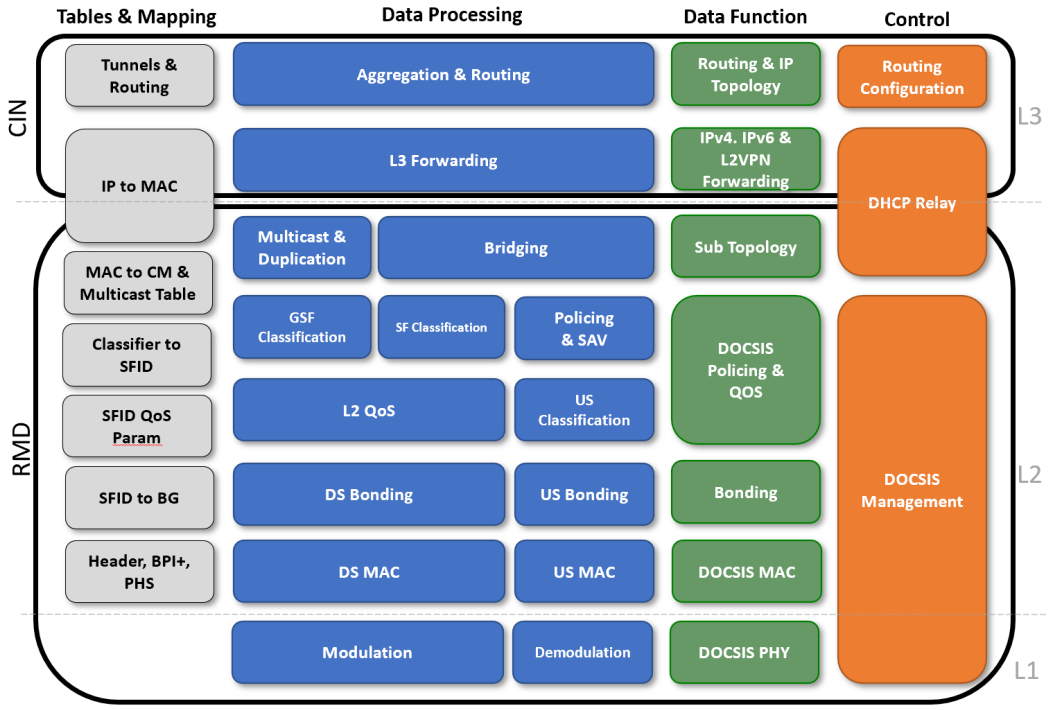


Figure 7 – Remote MACPHY Functional Decomposition

4.3.3. CIN Considerations

FMA outlines one example of CIN design, that of a Spine-and-Leaf illustrated in Figure 8, but does not mandate that type of network layout.

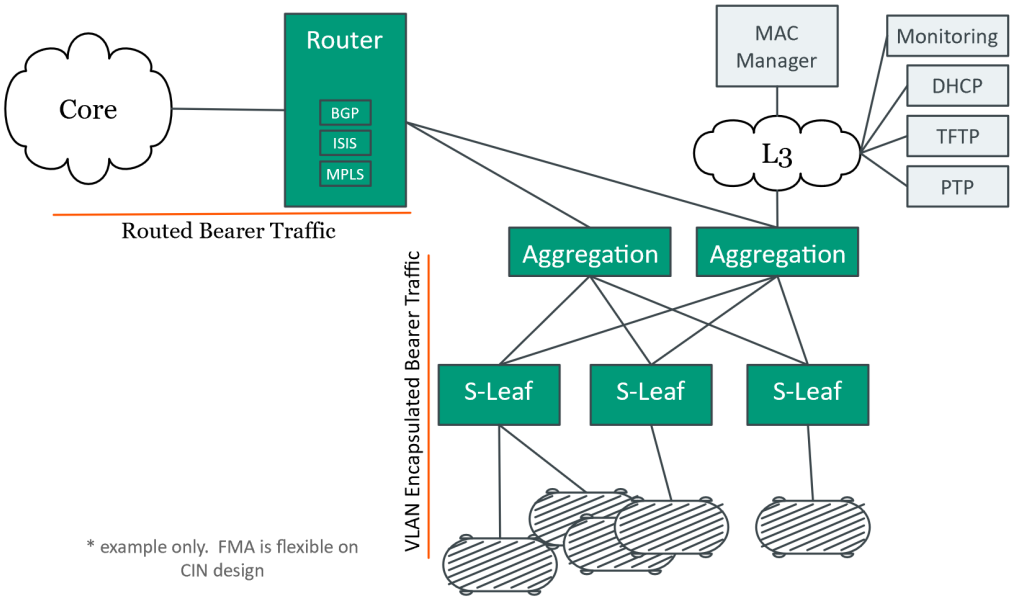


Figure 8 - Spine-and-Leaf CIN

An FMA based DAA does presuppose certain functions are handled within a CIN of any design though. MACsec, DHCP relay, and routing of both are required, as are separation of management and subscriber traffic are supported. The RMD handles certain DOCSIS-specific network functions: DHCP intercept (add/remove DHCP options), TFTP intercept, Source Address Verification (SAV), MAC learning, and cable bundles via VLANs. As such, routing at the first hop networking device is an equally valid topology for FMA.

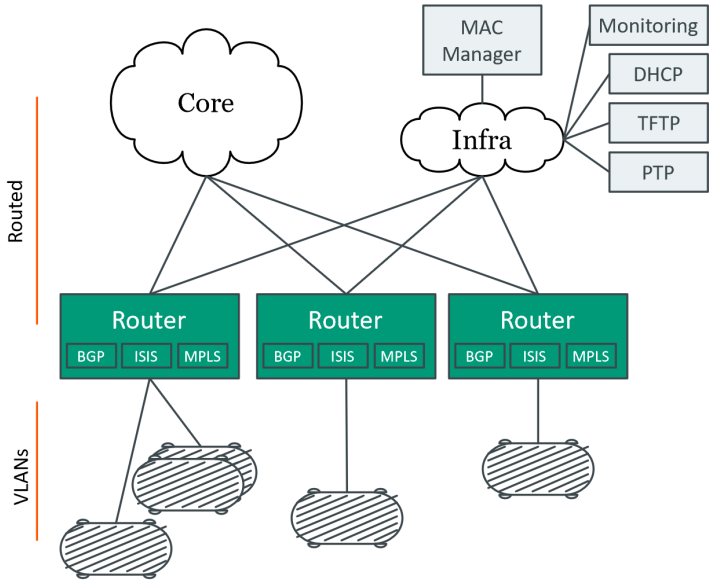


Figure 9 - Flat Network CIN

Spine-and-Leaf can be attractive in designs were high aggregation and oversubscription is expected, though FMA doesn't expect to see large east-west traffic patterns which mutes some of the benefits of this type of CIN. Flat designs are attractive to reduce VLAN stretching and in deployments with high north-south traffic patterns but may result in more routing entities to manage as the network horizontally scales. Operators may want to explore hybrid or other CIN topologies as well.

4.4. PacketCable

4.4.1. PacketCable Overview

PacketCable 1.5 Dynamic Quality of Service (DQoS), PacketCable Multimedia (PCMM) and PacketCable Event Messages (PCEM) are all supported transparently in FMA by the combination of the MAC Manager and PAG applications and the collection of RMDs within their scope. FMA makes the distributed access system appear to the legacy PacketCable backoffice systems as a CMTS or integrated CCAP, greatly simplifying the transition to FMA from legacy systems.

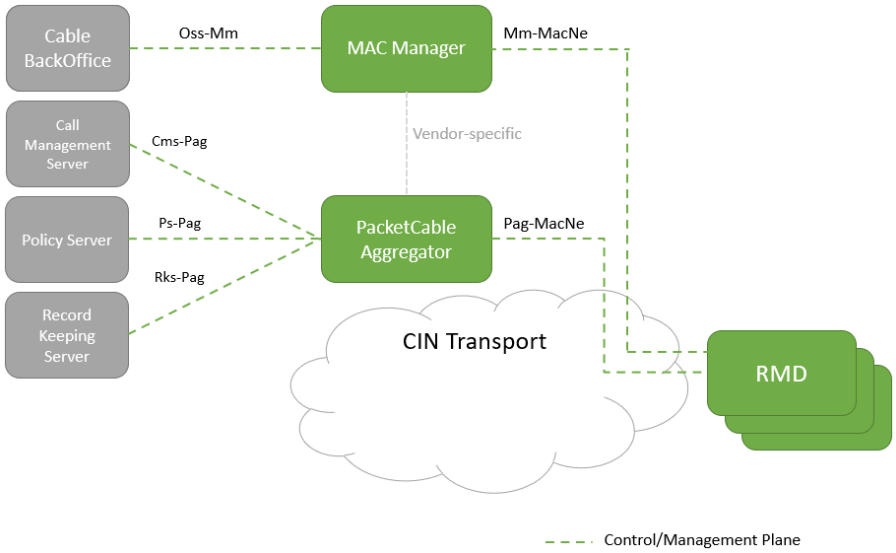


Figure 10 – FMA PacketCable Architecture

The MAC Manager supports MSO backoffice provisioning of PacketCable on the FMA System via the Oss-Mm interface. The PAG enables DQoS via Call Management Servers (CMS) for IP packet-based voice telephony and PCMM via Policy Servers for IP packet-based multimedia applications by providing the necessary COPS-based Cms-Pag control interfaces to these legacy functional entities. The PAG also supports PCEM delivery to Record Keeping Servers (RKS) for accounting purposes via the associated RADIUS-based Rks-Pag event reporting interfaces. The MAC Manager and PAG are software-based applications in the FMA System. They interact in a vendor implementation-specific manner.

The RMDs provide finite state machine operations required of the CMTS in PacketCable specifications. However, RMDs need only communicate with their PAG rather than connecting with one or more CMS, Policy Server and RKS functional entities. This simplifies RMD operations by moving most connectivity management to a central location, the PAG. It also efficiently scales the FMA System by keeping the number of legacy functional entity connections on par with a CMTS or CCAP sized system.

4.4.2. PacketCable Provisioning

Existing CCAP PacketCable provisioning parameters are available in FMA, as specified in CCAP-OSSI and FMA-OSSI. In FMA the Oss-Mm is implemented on the MAC Manager.

The MAC Manager is responsible for provisioning the PAG with its PacketCable parameters. After the PAG receives its PacketCable provisioning it connects to the legacy PacketCable backoffice functional entities. PacketCable DQoS specifies requirements for connecting the PAG with a CMS. PCMM specifies requirements for connecting the PAG with a Policy Server. Both PacketCable DQoS and PCMM specify COPS-based protocols which run over TCP connections. When it is provisioned for event message reporting, the PAG reports PCEM to the RKS via a RADIUS-based protocol which runs over UDP. The RKS destination is not provisioned by the MAC Manager but is signaled dynamically to the PAG in PacketCable DQoS and PCMM control signaling.

When new RMDs come online, they are configured and provisioned by the MAC Manager over the MMI (Mm-MacNe). The MAC Manager is responsible for provisioning the RMD with its PAG address and its PacketCable parameters. The RMD connects to its provisioned PAG over the PAI (Pag-MacNe), which

runs on top of TLS. The PAI supports the complete message set that is specified in PacketCable. It is based on messages defined using Google Protocol Buffers (GPB), which is supported by tooling and automatic source code generation that greatly enhances interoperability.

4.4.3. PacketCable 1.5 Operation

In PacketCable 1.5, the CMS is the call control entity responsible for enabling packet-based voice over IP telephony over a DOCSIS access network. A CMS signals end-to-end to CMs with embedded multimedia terminal adapters (E-MTAs) using the Network-based Call Signaling (NCS) protocol for call processing control. Functionality such as relaying on-hook and off-hook conditions, applying dial tone, collecting digits, and ringing phones are all communicated to and from the CMS via NCS signaling. The PAG and RMD do not participate in end-to-end NCS signaling.

The CMS also controls DOCSIS access network quality of service (QoS) via DQoS signaling. In legacy environments, DQoS runs between the CMS and CMTS or CCAP, opening and closing QoS "Gates" on the access network. In FMA, DQoS operates between the CMS and PAG, which translates COPS-based messages to and from the corresponding PAI messages sent and received over the PAI between the PAG and RMD. The Gate Control operations signaled by a CMS are implemented in the finite state machines of the RMD.

When PCEM is supported, the RMD will generate QoS-related events over the PAI to the PAG. The PAG relays the events to the RKS over the PCEM Rks-Pag RADIUS interface. The PAG does real time event message relay and controls batch processing of events as well. It also manages RKS redundancy with the Primary and Secondary RKS. Event message controls are provisioned on the PAG by the MAC Manager and provided by the CMS in its DQoS messages to the PAG.

4.4.4. PacketCable Multimedia Operation

In PCMM, the Policy Server is the QoS control entity responsible for enabling QoS for PacketCable 2.0 voice telephony and various multimedia applications over a DOCSIS access network. The Policy Server or Application Manager associated with the Policy Server signals end-to-end with end user CPE and applications in an application-specific manner. The PAG and RMD do not participate in end-to-end application-specific signaling.

The Policy Server also controls DOCSIS access network QoS for these applications using PCMM signaling. In legacy environments, PCMM runs between the Policy Server and CMTS or CCAP, opening and closing QoS "Gates" on the access network. In FMA, PCMM operates between the Policy Sever and PAG, which translates COPS-based messages to and from the corresponding PAI messages sent and received over the PAI between the PAG and RMD. The Gate Control operations signaled by a Policy Server are implemented in the finite state machines of the RMD.

When PCEM is supported, the RMD will generate QoS-related events over the PAI to the PAG. The PAG relays the events to the RKS over the PCEM Rks-Pag RADIUS interface. The PAG does real time event message relay and controls batch processing of events as well. It also manages RKS redundancy with the Primary and Secondary RKS. Event message controls are provisioned on the PAG by the MAC Manager and provided by the Policy Server in its PCMM messages to the PAG.

4.5. Lawful Intercept

4.5.1. Lawful Intercept Overview

FMA provides full support for Lawful Intercept. The MSO interface point to Law Enforcement Agencies (LEA) requiring intercept of traffic within the cable operator network is referred to as Mediation Device, or alternatively the Delivery Function. The Mediation Device touches several network entities in the cable operator network to provision Taps on those devices and to collect mirrored traffic and certain traffic-related metadata. The Mediation Device is not specific to FMA. It is inherited from the legacy CMTS and CCAP Lawful Intercept architectures.

The FMA architecture provides alternatives which operators can leverage based upon the functionality present in their networks and the desired method of delivering intercepted traffic to the Mediation Device.

- For PacketCable 1.5 Electronic Surveillance, which supports intercept of PacketCable 1.5 voice traffic, the combination of PAG and RMD enable the functionality that is provided by the CMTS and CCAP in legacy systems.
- For broadband High-Speed Data (HSD) intercepts, the combination of MAC Manager and RMD enable the functionality that is provided by the CMTS and CCAP in legacy systems.
- Alternatively, for broadband HSD intercepts the operator may choose a convergence strategy where routers located above the access networks serve as the focal point for intercepting traffic over disparate access technologies such as HFC, PON and wireless.

These mechanisms are all supported in FMA. The details of each mechanism are provided in the sections which follow.

4.5.2. PacketCable 1.5 Voice Lawful Intercept

In PacketCable 1.5, the CMS is the call control entity responsible for enabling packet-based voice over IP telephony over a DOCSIS access network. The CMS also controls DOCSIS access network QoS using PacketCable 1.5 DQoS signaling. In this environment, requirements for supporting Lawful Intercept have been specified in the PacketCable 1.5 Electronic Surveillance specification.

Intercept Taps are provisioned by the Mediation Device on the CMS. The CMS signals that intercepts are required on a call-by-call basis via electronic surveillance parameters in PacketCable 1.5 DQoS messages, which are sent to the RMD by way of the PAG. If PCEM event messages are to be reported as part of the intercept, the PAG makes note of this and serves as a mirroring point for these events when they are reported to it by the RMD. In this case the PAG mirrors the events toward the Mediation Device.

The RMD implements the voice traffic mirroring point. It mirrors intercepted voice traffic to the Mediation Device as indicated in the DQoS electronic surveillance parameters on a call-by-call basis.

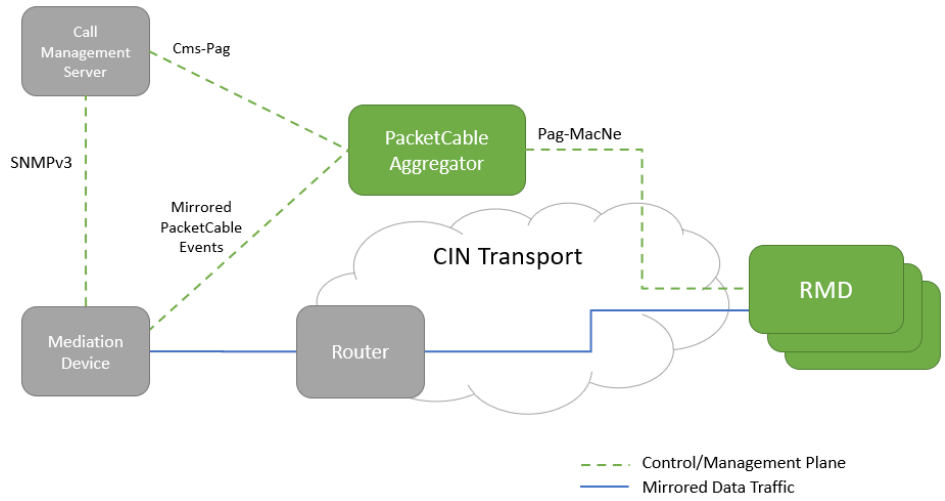


Figure 11 – FMA PacketCable Lawful Intercept Architecture

4.5.3. FMA-specific High Speed Data Lawful Intercept

Broadband HSD traffic on the cable network is also subject to Lawful Intercept. When an HSD Tap is provisioned, the traffic to be intercepted is specified in an FMA Tap MIB that unifies various Tap MIBS which exist in the industry to provision Taps on CMTS and CCAP devices. The FMA Tap MIB provides commonly used parameters and eliminates those parameters that the industry has not found to be relevant.

Intercept Taps are provisioned by the Mediation Device on the MAC Manager via SNMPv3 and the new FMA Tap MIB. The MAC Manager in turn provisions the Tap on the RMD over the MMI. These taps are persistent, unlike the call-by-call Taps used in PacketCable 1.5 Electronic Surveillance.

The RMD implements the HSD traffic mirroring point. It mirrors intercepted data traffic to the Mediation Device as configured by the MAC Manager over the MMI.

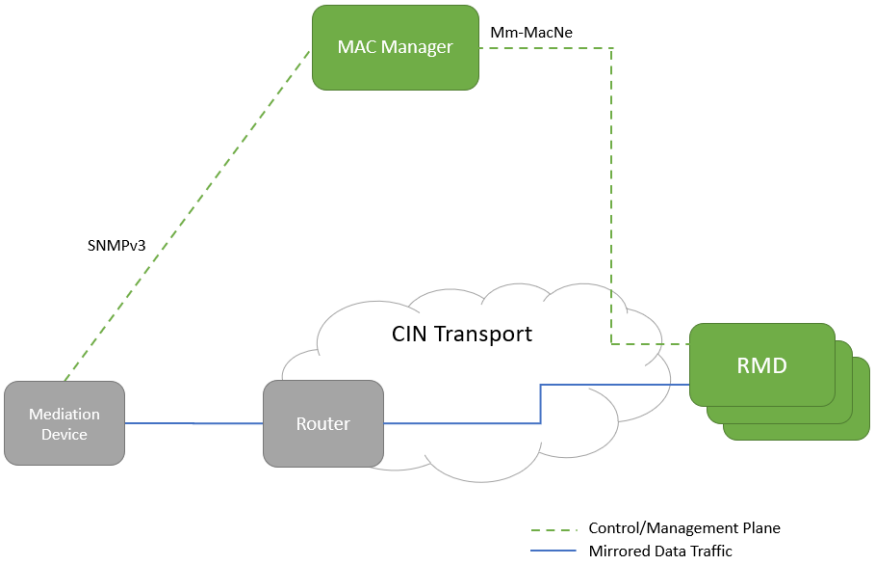


Figure 12 – FMA High Speed Data Lawful Intercept Architecture

4.5.4. Converged Access High Speed Data Lawful Intercept

Broadband HSD traffic is subject to Lawful Intercept regardless of the access technology. Operators with more than one access technology in their network may be wise to consider an architecture where Taps can be done at a common point in the network, such as in a router at a level above the access network. This is an architecture that is independent of access technology. Therefore, while such an architecture would also support FMA HSD traffic intercept, it does not impact the FMA System.

Intercept Taps are provisioned by the Mediation Device on the router via one of the industry-supported Tap MIBs. These taps are persistent.

The router implements the HSD traffic mirroring point. It mirrors data traffic to the Mediation Device as configured by the Mediation Device.

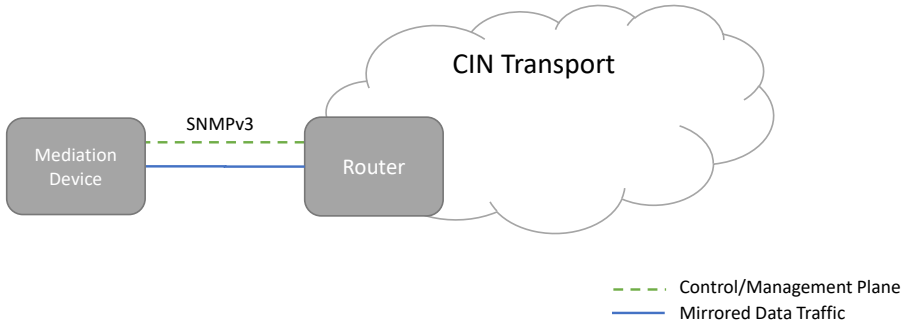


Figure 13 – Converged Access High Speed Data Lawful Intercept Architecture

4.6. Remote PHY Reuse – QAM Video, OOB, and Plant Maintenance

FMA specifications include the idea of eventually incorporating a mixture of flexible MAC (the "FM" in FMA) devices or MAC-NEs under a single umbrella architecture. This could possibly include a CCAP Core configured and monitored by a MAC Manager, or even a Remote MAC Core (RMC) that contains the DOCSIS MAC but feeds subtended RPDs.

There has also been focus on reusing components from the original Remote PHY specifications where possible so that these can be shared between different MAC-NE implementations, potentially in the same market area. This includes three main functions: QAM video operation, set-top box (STB) and other out-of-band (OOB) signals, and signals needed for plant maintenance.

In all cases the same protocols, Generic Control Plane (GCP), Remote Downstream External PHY Interface (R-DEPI), Remote Upstream External PHY Interface (R-UEPI), and Remote Out-of-Band (R-OOB) are used by FMA so the associated centralized components, the Auxiliary Cores, supporting these functions do not need modification to serve all forms of MAC-NE in an FMA System.

Auxiliary Core operation in FMA relies on the same mechanisms as Remote PHY, with the exception that the MAC Manager configures all the Auxiliary Core information on the RMD before the RMD can start GCP operation. The MAC Manager does not need to interact with Auxiliary Core except to be able to configure the core details in the RMD.

4.6.1. QAM Video

DAA QAM video implementations have evolved as Remote PHY systems have moved into scaled deployment and the unique needs of operators are being addressed. [QAMVideoDAA] provides an excellent summary of the options available and considerations in those Video Cores and Traffic Engines. Figure 14 summarizes options generally available and being deployed by operators for QAM video in any DAA environment, whether it be Remote PHY or Remote MACPHY.

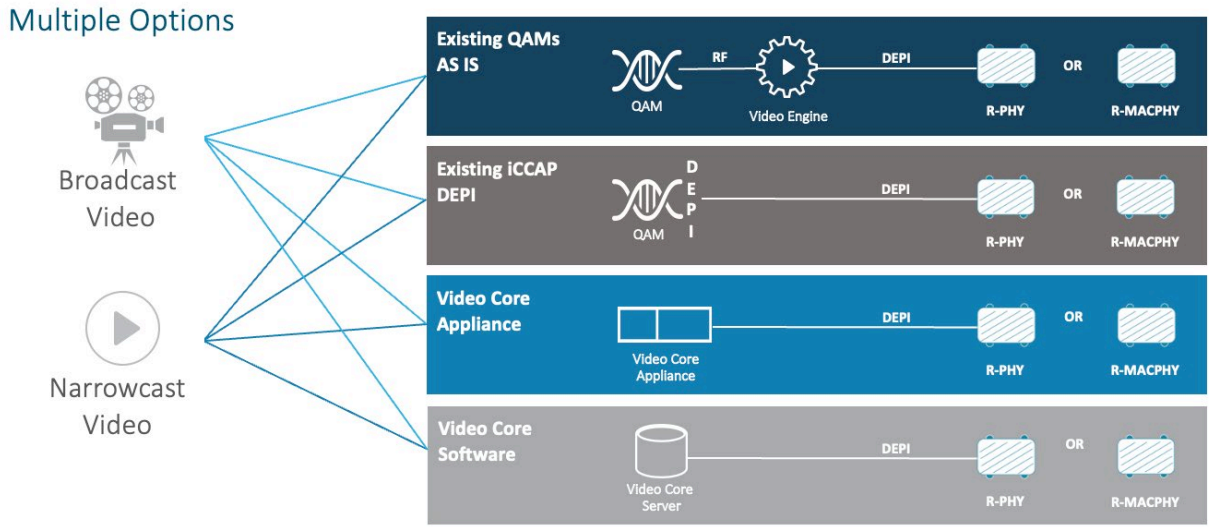


Figure 14 - QAM Video Implementation Options

4.6.1.1. Edge QAM and Traffic/Video Engine

Operators can leave existing Edge QAMs in place and deploy large scale QAM demodulators to convert pre-encrypted, full rate QAM, multi-program transport streams (MPTS) with NULL packets into R-DEPI format for transport to the RPD or RMD. This solution is also referred to as a Traffic Engine, which means a device that supports the QAM video data plane only, leaving Auxiliary Core operation to configure the video channels to other components. The MAC Manager can support video configuration of Traffic Engines on the RMD through the MMI as a static pseudowire.

4.6.1.2. Fully Integrated CCAP

Many operators have existing integrated CCAP solutions that support all their QAM video needs for analog node deployments. Some integrated CCAP solutions may also support R-DEPI out for Remote PHY operation and Auxiliary Core operation. These solutions can provide DAA QAM video for RPDs and RMDs alike.

4.6.1.3. Standalone Video Core

A third alternative is a Standalone Video Core appliance or software solution that contains both traditional Edge QAM functions (narrowcast video control plane, single program transport stream (SPTS) to MPTS multiplexing, encryption) and provides an R-DEPI output compatible with any RPD or RMD solution. These Video Cores may support both Traffic Engine for video data plane and Auxiliary Core for

GCP control or may just act as the Traffic Engine and support configuration using the MAC Manager with static pseudowire objects specified over the MMI.

4.6.2. Out-of-Band

QAM video deployments in HFC with traditional STBs require a way to provide authorization, encryption messaging, and software update capability for those STBs. Options for the transport of this traffic includes:

- DOCSIS 1.0-3.1 – No specific functionality is required in the FMA system to accommodate.
- DOCSIS Set-Top Gateway (DSG) – The RMD operates as a DSG Agent. An additional accommodation of adding NAT64 to the DSG Server in the network and a DSG Agent in the RMD is required if the operator chooses to deploy in an IPv6-only environment. This functionality is unique to FMA and not reused from Remote PHY.
- SCTE 55-1 – ALOHA protocol, typically used with Motorola conditional access.
- SCTE 55-2 – DAVIC protocol, typically used with Scientific Atlanta conditional access.

Similar to QAM video, SCTE 55-1 and SCTE 55-2 in FMA reuses implementations already supported in Remote PHY. Figure 15 illustrates these options – NDF/NDR or direct OOB operation using an OOB Core.

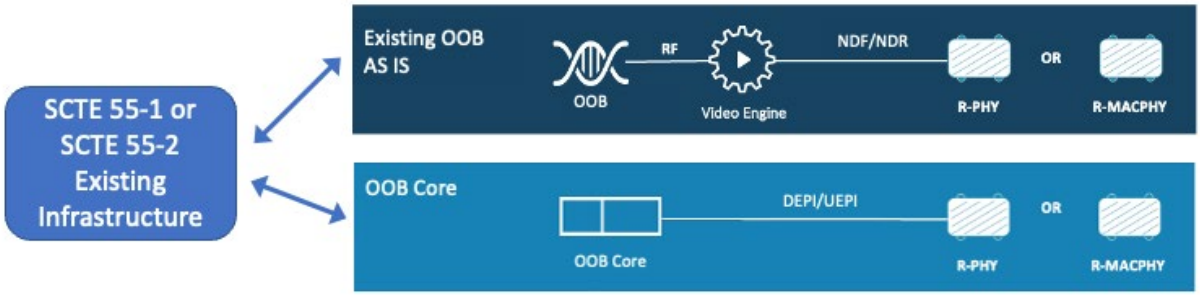


Figure 15 – Out-of-Band Implementation Options

4.6.2.1. NDF/NDR

NDF digitizes narrowband signals from the existing RF-based SCTE 55-1 and 55-2 systems with a centrally located Video Engine component which then transports the digitized signal using R-DEPI to the RPD/RMD where it is regenerated and transmitted to the STB.

NDR digitizes narrowband return signals from the STB in the RPD/RMD and transports them back to the Video Engine component using R-UEPI. The Video Engine regenerates those signals and transmits them back to the existing 55-1 or 55-2 RF systems.

SCTE 55-1 is tolerant to jitter and latency, so NDF/NDR operation is robust in many different deployment architectures.

The SCTE 55-2 protocol requires fast acknowledgements (under 3ms) of received STB transmissions. As a result, SCTE 55-2 operation using NDF/NDR requires shorter distances between the Video Engine and

RPD/RMD, as well as tight control of round-trip network latency. This general DAA constraint applies to both RPD and RMD systems.

4.6.2.2. Direct SCTE 55 Operation

In direct operation mode, centralized components are added which replace or augment the existing RF-based systems in place for SCTE 55-1 and SCTE 55-2.

For SCTE 55-1, downstream Moving Picture Experts Group (MPEG) transport packets from existing control components are R-DEPI encapsulated by a centralized 55-1 OOB Core function and subsequently decapsulated and modulated onto SCTE 55-1 OOB RF channel(s) at the RPD/RMD. Upstream ATM-like cells are R-UEPI encapsulated at the RPD/RMD and subsequently decapsulated at the 55-1 OOB Core, which groups together multiple RPD/RMDs to look like a single legacy RF component to minimize issues with scale of the existing control components.

The centralized 55-1 OOB Core function can be implemented as:

- Part of the CCAP Principal Core in Remote PHY – Extra functions are integrated within an Integrated CCAP Core already serving DOCSIS and QAM video.
- An Auxiliary Core in Remote PHY and FMA – A single element contains both data plane transport over R-DEPI and configuration of RMD/RPD using GCP.
- A Traffic Engine in Remote PHY and FMA – A data plane only element based on static pseudowires, with configuration coming from the Principal Core using GCP in Remote PHY or the MAC Manager using MMI in FMA.

For SCTE 55-2, existing centralized RF-based control components are replaced by functions in the RPD/RMD and a centralized 55-2 OOB Controller. The 55-2 OOB Controller acts as an Auxiliary Core, containing the full 55-2 MAC and connecting to the video backoffice. The 55-2 OOB Controller encapsulates downstream packets in R-DEPI format and transmits them to the RPD/RMD, which performs R-DEPI decapsulation and RF modulation. In the upstream the RPD/RMD demodulates RF, encapsulates as packets in R-UEPI format and transmits to the 55-2 OOB Controller. The RPD/RMD must also perform some lower layer MAC functions such as upstream acknowledgement insertion. Lower layer MAC functions for 55-2 are included in the RPD/RMD to ensure tight ACK turnaround times can be met with long distances and higher jitter between the OOB Controller and RPD/RMD.

4.6.3. Plant Maintenance

Reliable operation of the HFC system requires support of the operational tools used by field technicians to setup and monitor critical RF OSP conditions. Each of these tools for FMA RMD operation leverages the systems already deployed widely for Remote PHY:

- Downstream and upstream sweep – Sweep provides ongoing validation of proper signal levels across the entire downstream and upstream bands to detect excessive cable loss and impairments between the node (RPD/RMD), amplifier cascade, passives and CMs in the plant. Available options for sweep operation in Remote PHY are reused for FMA and are described in R-OOB Appendix III.
- Upstream spectrum capture – Capture of the upstream spectrum as part of overall proactive network maintenance (PNM) provides significant benefit to operators to understand sources of upstream packet errors due to persistent or intermittent noise and ingress. Upstream spectrum

capture operation for Remote PHY is reused for FMA and is described in R-UEPI sections 8.2.8 and 8.2.9.

- Leakage detection – Regulators mandate that cable operators ensure the cable plant does not "leak" RF energy into frequency bands that may be susceptible to interference (e.g., aeronautical communication frequencies or spectrum used by licensed mobile wireless systems). Identifying and repairing faults as part of PNM leakage detection activity also reduces errors due to ingress interference from those other uses. R-OOB Appendix IV and Section 9 describe operation of downstream leakage detection. In the future, High Split systems where upstream OFDMA signals will operate in the aeronautical frequency band on potentially licensed mobile wireless frequencies will require additional support, which has recently been added into CableLabs® DOCSIS specifications. See [LeakageHighSplit] for details on the specifications and High Split leakage detection solutions. It should be noted that all High Split leakage detection solutions in the specifications are applicable to both Remote PHY and FMA RMD deployments.

5. FMA Migration Considerations

As operators begin their journey toward the Emerald City to enable the deployment of FMA with RMD, it will be important to know what they should take with them. If their journey started with legacy systems in Kansas, a transition to DAA is required. If their journey began with Remote PHY, a DAA starting point is already in place. In either case it is important to understand the similarities between Remote PHY and FMA standards and components as well as where they diverge. Knowing this can help operators identify common operational benefits where Remote PHY and FMA may coexist and understand the changes required if transitioning deployments from Remote PHY to FMA with RMD.

Table 1 summarizes the reuse opportunities for each of the steps.

Table 1 - FMA Migration and Remote PHY Reuse

FMA Migration Step	Remote PHY Architecture and Component Reuse
Fiber node change	Yes
OSP analog to digital optics change	Yes
CIN - L2 RMD aggregation layer	Yes
CIN - 802.1X RMD port authentication	Yes
CIN – MACsec encryption	No – MACsec not used in Remote PHY with BPI+ applied at CCAP Core
CIN – DHCP for RMD	Yes
CIN - Synchronization (PTP/SyncE)	Yes, but PTP only needed in FMA for delivering precision time services such as Mobile Xhaul
CIN – L2/L3 termination	No – largest area of divergence
QAM video and STB OOB support	Yes
Plant maintenance tool updates	Yes
MAC Manager addition	No – Remote PHY pCore/vCore management possible with MAC Manager in future FMA phases

5.1. Migration Steps with Remote PHY Reuse

The FMA network migration steps listed below all have significant reuse of Remote PHY architecture and components.

5.1.1. Fiber Node Change

An analog fiber node is swapped for a Remote MACPHY node or a new Remote MACPHY node is deployed instead of splitting an existing analog node.

Remote MACPHY node installs in the OSP are effectively the same as Remote PHY conversions by migrating to nodes with digital Ethernet optics capability.

5.1.2. OSP Analog to Digital Optics Change

The traditional OSP analog fiber distribution and associated wavelength division multiplexing (WDM) equipment is converted to digital optics with appropriate pluggable Ethernet transceivers. Outdoor mux/demux are installed as needed to match the wavelength and fiber planning.

Remote MACPHY OSP digital-optical fiber conversion is the same as Remote PHY and, in fact, investments in digital conversion for Remote PHY can be reused for Remote MACPHY deployments.

5.1.3. CIN - L2 Remote MACPHY Aggregation Layer

The new digital Ethernet optics and associated Remote MACPHY node are connected into a L2 aggregation network as the node-facing part of the CIN.

The L2 aggregation CIN network is very similar between Remote PHY and Remote MACPHY with the exception of some considerations for segmenting offered services into VLANs that is available in Remote MACPHY.

5.1.4. CIN - 802.1X Remote MACPHY Node Port Authentication

The L2 CIN is provisioned to support 802.1X mutual authentication for network access control.

802.1X operation is very similar between Remote PHY and Remote MACPHY as a way to ensure there is no unauthorized access into the CIN L2 aggregation layer.

5.1.5. CIN – DHCP for Remote MACPHY

A DHCP server, IPv4 or IPv6, is configured to serve RMDs and provide the MAC Manager network location via IP address or DNS name.

RMD boot and DHCP mechanisms are based on the initialization and operation of RPDs.

5.1.6. CIN – Synchronization (PTP/SyncE)

As an optional step, IEEE 1588 PTP Grandmasters (GMs) are added and connected to the CIN if needed for precision time services such as Mobile xHaul. Network elements in the CIN between GM and RMD are provisioned as Boundary Clocks or Transparent Clocks as needed to meet precision timing performance requirements. SyncE network clock synchronization may also be added as needed to meet precision timing performance requirements.

In contrast with Remote MACPHY, Remote PHY networks cannot operate without PTP phase and frequency synchronization, which is required between the CCAP Core and RPD for DOCSIS scheduling. While Remote MACPHY does support R-DTI requirements from Remote PHY, PTP is only needed in

FMA for delivering precision time services such as synchronization for small cells or mobile Xhaul. Depending on implementation in any deployed Remote PHY equipment, coexistence of Remote MACPHY and Remote PHY in the same network may also require the use of PTP for frequency synchronization for synchronous QAM video or NDF/NDR solutions.

5.1.7. QAM Video and STB OOB Support

QAM video and STB OOB support are added in the network using Video Engine(s), Traffic Engines and Auxiliary Cores using whichever best fits the operator video backend and STB infrastructure. The MAC Manager directly configures static pseudowires on the RMD or directs the RMD towards the Auxiliary Core(s) for further configuration.

QAM video and Remote PHY R-OOB functions (55-1, 55-2, NDF/NDR) are all directly reused in FMA.

5.1.8. Plant Maintenance Tool Updates

Updated versions of plant maintenance operational tools for sweep, spectrum capture, and leakage detection are installed and configured, with the RMD provisioned for working with these tools through the MAC Manager as needed.

Plant maintenance tools used in the network are the same for Remote PHY or FMA.

5.2. Migration Steps Divergent from Remote PHY

The FMA network migration steps listed below all diverge from the Remote PHY architecture and do not reuse components from Remote PHY.

5.2.1. CIN – MACsec Encryption

FMA/Remote MACPHY adds configurable MACsec encryption in the L2 aggregation layer for operators who choose to encrypt all traffic in the OSP network segment(s).

Remote PHY encrypts user traffic with BPI+ from the CM through the RPD and CIN to the CCAP Core, so MACsec encryption is not required for bearer traffic. Remote PHY does not mandate encryption for non-bearer traffic such as management and control signaling between the Core and RPD, or for certain types of signaling between back-office and CM, such as DHCP and TFTP.

A CIN which converges PON access along with HFC will generally require MACsec for the same reasons as FMA with Remote MACPHY since the PON layer terminates at the OLT or Remote OLT.

5.2.2. CIN – L2/L3 Termination

A router is configured to route bearer traffic from the RMD to the Core network through the CIN.

The largest divergence between Remote PHY and FMA is the decoupling of L3 routing from DOCSIS packet processing. In FMA there is no centralized DOCSIS packet processing engine terminating an L2TPv3 tunnel containing DOCSIS frames and providing a L3 CMTS northbound router for the customer traffic. Rather, the RMD terminates the DOCSIS frames and acts as a L2 CMTS northbound toward the aggregation network. L2 packet flows are aggregated at a standard L3 router to provide routing into the operator core network. Customer bearer traffic is directly accessible for aggregation, switching, and routing at the first-hop switch or router which can be selected from a range of readily available switch and

router vendors. Traffic engineering of natively transported traffic in an FMA CIN is a considerable upside of the architecture.

5.2.3. MAC Manager Addition

A MAC Manager is deployed and configured with IP routable connectivity between the RMD network location and the MAC Manager, with provisioning for the associated RMD(s) configured on the MAC Manager.

No MAC Manager is present in current Remote PHY systems but the "Flexible MAC" part of FMA can be utilized to manage pCore/vCore MAC-NEs in future FMA phases once that operation is standardized. In this way, the Remote PHY portion of the network can eventually be brought under one FMA umbrella.

6. Other Architecture Considerations

There are several other important architecture considerations between traditional integrated CCAP, CCAP + Remote PHY, and FMA with Remote MACPHY deployments.

6.1. CIN Latency

Scheduling upstream traffic on a DOCSIS access network relies on a request and grant cycle between the CM requesting bandwidth and the entity scheduling minislots against those requests.

Both Remote MACPHY and traditional integrated CCAP co-locate the MAC and PHY components so there is no additional delay added to each request and grant in the protocol. No CIN impact on DOCSIS request/grant cycle latency is present in FMA with RMD. In addition, as shown in [FMACloud], latency between the MAC Manager and RMD is not a critical consideration for system operation since no data plane functions exist between MAC Manager and RMD.

In Remote PHY the MAC and PHY are in separate devices separated by a CIN of one or more hops. CIN latency and jitter between the CCAP Core MAC and RPD PHY add to request and grant cycle, thereby lengthening the time for CM bandwidth requests to be served and increasing upstream latency. The impact is dependent on the distance and delay within the CIN as well as the DOCSIS MAP interval. Centralization of the CCAP Core exacerbates the problem by extending the distance, delay and probability of jitter, so careful consideration of where CCAP Cores are placed is needed in Remote PHY systems.

6.2. Deployment Granularity

In traditional integrated CCAP and CCAP + Remote PHY deployments, space and power targets for hub-based equipment have resulted in high density chassis which serve tens to hundreds of service groups in a single element. This tends to result in a natural bundling and a deployment architecture where only the largest geographic areas get the latest high density chassis features and capabilities. CCAP + Remote PHY deployments can help with this by centralizing the CCAP Core to share over a large geographic area but then the CIN latency impacts of section 6.1 can result in degraded upstream latency.

In contrast, the latency tolerance of the disaggregated MAC Manager and RMD in FMA allows for service groups to be added one RMD at a time, with operational efficiency by using a centralized MAC Manager which can be deployed 1000s of kilometers away. This makes RMD with FMA well suited to rural and other low-density locations, or as a way to boost capacity in targeted locations within a larger urban center while continuing to use integrated CCAP or CCAP + Remote PHY.

7. Conclusion

We know why we want to go to the DAA Land of Oz and, once there, why we want to go to the Emerald City. There's a clear path to the Emerald City once in Oz, and that path as not as scary as it would seem since we can take along plenty of familiar friends for support.

This paper has presented a comprehensive discussion of the considerations for migration of HFC delivered services from traditional integrated CCAP or CCAP + Remote PHY architectures to FMA with Remote MACPHY. Significant reuse of architecture and network components between Remote PHY and FMA allows operators deploying FMA with Remote MACPHY to:

- Obtain the well understood and proven benefits of DAA
    - Converged digital Ethernet optics
    - Improved end of line RF performance for greater capacity
    - Improved field operations through intelligent nodes
- Apply the lessons learned by the industry during Remote PHY deployment in an FMA deployment
- Support coexistence of Remote PHY and FMA with Remote MACPHY within the operator network
- Support a graceful migration from Remote PHY to FMA with Remote MACPHY

Some areas will be different between Remote PHY and FMA for operators – this isn't Kansas anymore – but these offer a pathway to the future of cable operator access networks:

- Disaggregation of the management and data planes through the FMA MAC Manager
- Access network abstraction through the FMA MAC Manager
- A common L2 architecture similar to PON for convergence on the overall access network L2/L3 topology

The near future is sure to be very active in FMA with standards activities moving from specification writing to product implementation, interoperability events happening to demonstrate the multi-vendor capability of FMA, and the upcoming industry movement towards DOCSIS 4.0 which requires DAA nodes for deployment.

Abbreviations

AAA	Authentication, Authorization, Accounting
AM	Analog Modulation
BG	Bonding Group
BGP	Border Gateway Protocol
BNG	Broadband Network Gateway
BPI	Baseline Privacy Interface
BSS	Business Support System
CCAP	Converged Cable Access Platform
CIN	Converged Interconnect Network
CLI	Command Line Interface
CM	Cable Modem
CMS	Call Management Server
Cms-Pag	CMS to PAG interface
CMTS	Cable Modem Termination System
COPS	Common Open Policy Server
COTS	Commercial of the Shelf
CPU	Central Processing Unit
DAA	Distributed Access Architecture
DAVIC	Digital Audio Video Council
DEPI	DOCSIS External PHY Interface
DHCP	Dynamic Host Control Protocol
DNS	Domain Name Server
DOCSIS	Data Over Cable System Interface Specification
DQoS	Dynamic Quality of Service
DSG	DOCSIS Set-Top Gateway
DSLAM	Digital Subscriber Line Access Multiplexer
DWDM	Dense Wavelength Division Multiplex
EMS	Element Management System
E-MTA	Embedded Multimedia Terminal Adapter
ERM	Edge Resource Manager
FMA	Flexible MAC Architecture
GCP	Generic Control Protocol
GHz	Giga Hertz
GM	Grandmaster
GPB	Google Protocol Buffers
HFC	Hybrid Fiber Coax
HSD	High Speed Data
IPDR	IP Detail Record
IP	Internet Protocol
IPv4	IP version 4
IPv6	IP version 6
ISIS	Intermediate System to Intermediate System
LACP	Link Aggregation Control Protocol
LAG	Link Aggregation Group
LEA	Law Enforcement Agency
LR	Long Range

L2	Layer 2
L2VPN	L2 Virtual Private Network
L3	Layer 3
MAC	Media Access Control
MAC-NE	MAC Network Element
MACsec	MAC Security
MER	Modulation Error Rate
MIB	Management Information Base
MM	MAC Manager
MMI	MAC Management Interface
Mm-MacNe	MM to MAC-NE interface
MPEG	Moving Picture Experts Group
MPLS	Multi-Protocol Label Switching
MPTS	Multi-Program Transport Stream
MSO	Multiple Systems Operator
MULPI	MAC and Upper Layer Protocol Interface
NCS	Network-based Call Signaling
NDF	Narrowband Digital Forward
NDR	Narrowband Digital Return
NMS	Network Management System
N+x	Node plus x
N+0	Node plus 0 (zero)
OOB	Out of Band
OSP	Outside Plant
OSS	Operations Support System
OSSI	OSS Interface
Oss-Mm	OSS to MM interface
PAG	PacketCable Aggregator
Pag-MacNe	PAG to MAC-NE interface
PAI	PacketCable Aggregator Interface
PCEM	PacketCable Event Messages
PCMM	PacketCable Multimedia
PHS	Payload Header Suppression
PHY	Physical Layer
PNM	Proactive Network Maintenance
PON	Passive Optical Network
PS	Policy Server
Ps-Pag	PS to PAG Interface
PTP	Precision Time Protocol
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RADIUS	Remote Access Dial-Up Server
R-DEPI	Remote DEPI
RF	Radio Frequency
RKS	Record Keeping Server
Rks-Pag	RKS to PAG interface
RMC	Remote MAC Core
RMD	Remote MAC Device

R-OOB	Remote OOB
RPD	Remote PHY Device
SAV	Source Address Verification
SCTE	Society of Cable Telecommunications Engineers
SF	Service Flow
SFID	SF Identifier
SFP	Single Formfactor Pluggable
SLeaf	Secure Leaf
SLeaf-MacNe	SLeaf to MAC-NE interface
SNMP	Simple Network Management Protocol
SNMPv3	SNMP version 3
SPTS	Single Program Transport Stream
TCP	Transmission Control Protocol
TFTP	Trivial File Transfer Protocol
TLS	Transport Layer Security
ToD	Time of Day
UDP	User Datagram Protocol
VLAN	Virtual Local Access Network
WDM	Wavelength Division Multiplex
Xhaul	Cross-haul
YANG	Yet Another Next Generation

Bibliography & References

[FMACloud] *FMA in the Cloud*, D. Johnson, J. Thompson; 2021 Fall Technical Forum.

[LeakageHighSplit] *Leakage Detection in a High Split World,* R. Coldren, M. Cooper, G. Tresness; 2021 Fall Technical Forum.

[QAMVideoDAA] *Delivering QAM Video in Distributed Access Architectures,* C. Howlett, D. Johnson, K. Meisen; 2019 Fall Technical Forum; https://www.nctatechnicalpapers.com/Paper/2019/2019-delivering-qam-video-in-distributed-access-architectures

DOCSIS® Flexible MAC Architecture System Specification, CM-SP-FMA-SYS-I02-210526, May 26, 2021, Cable Television Laboratories, Inc.

DOCSIS® Flexible MAC Architecture MAC Manager Interface Specification, CM-SP-FMA-MMI-I02-210526, May 26, 2021, Cable Television Laboratories, Inc.

Flexible MAC Architecture, YANG Modules, http://mibs.cablelabs.com/YANG/DOCSIS/FMA

DOCSIS® Flexible MAC Architecture PacketCable Aggregator Interface Specification, CM-SP-FMA-PAI-I01-200930, September 30, 2020, Cable Television Laboratories, Inc.

Flexible MAC Architecture, Google Protocol Buffers, http://mibs.cablelabs.com/GPB/DOCSIS/FMA/

DOCSIS® Flexible MAC Architecture OSS Interface Specification, CM-SP-FMA-OSSI-D03-210630, June 30, 2021, Cable Television Laboratories, Inc.

DOCSIS® MHAv2 Remote PHY Specification, CM-SP-R-PHY-I16-210804, August 4, 2021, Cable Television Laboratories, Inc.

DOCSIS® MHAv2 Remote PHY OSS Interface Specification, CM-SP-R-OSSI-I16-210903, September 3, 2021, Cable Television Laboratories, Inc.

DOCSIS® MHAv2 Remote PHY Downstream External PHY Interface Specification, CM-SP-R-DEPI-I16-210804, August 4, 2021, Cable Television Laboratories, Inc.

DOCSIS® MHAv2 Remote PHY Upstream External PHY Interface Specification, CM-SP-R-UEPI-I13-201207, December 7, 2020, Cable Television Laboratories, Inc.

DOCSIS® MHAv2 Generic Control Plane Specification, CM-SP-GCP-I05-200323, March 23, 2020, Cable Television Laboratories, Inc

DOCSIS® MHAv2 Remote DOCSIS Timing Interface Specification, CM-SP-R-DTI-I08-200323, March 23, 2020, Cable Television Laboratories, Inc

DOCSIS® MHAv2 Remote Out-of-Band Specification, CM-SP-R-OOB-I12-200323, March 23, 2020, Cable Television Laboratories, Inc

DOCSIS® 4.0 Physical Layer Specification, CM-SP-PHYv4.0-I04-210826, August 26, 2021, Cable Television Laboratories, Inc.

DOCSIS® 4.0 MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv4.0-I04-210826, August 26, 2021, Cable Television Laboratories, Inc.

DOCSIS® 4.0 CCAP Operations Support System Interface Specification, CM-SP-CCAP-OSSIv4.0-I04-210521, May 21, 2021, Cable Television Laboratories, Inc.

DOCSIS® 3.1 Physical Layer Specification, CM-SP-PHYv3.1-I18-210125, January 25, 2021, Cable Television Laboratories, Inc.

DOCSIS® 3.1 MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.1-I21-201020, October 20, 2020, Cable Television Laboratories, Inc.

DOCSIS® 3.1 CCAP Operations Support System Interface Specification, CM-SP-CCAP-OSSIv3.1-I21-210716, July 16, 2021, Cable Television Laboratories, Inc.

PacketCable™ 1.5 Specifications Dynamic Quality-of-Service, PKT-SP-DQOS1.5-C01-191120, November 20, 2019, Cable Television Laboratories, Inc.

PacketCable™ 1.5 Specifications Event Messages, PKT-SP-EM1.5-C01-191120, November 20, 2019, Cable Television Laboratories, Inc.

PacketCable™ 1.5 Specifications Electronic Surveillance, PKT-SP-ESP1.5-C01-191120, November 20, 2019, Cable Television Laboratories, Inc.

PacketCable™ 2.0 Specifications Electronic Surveillance Intra-Network Specification, PKT-SP-ES-INF-C01-140314, March 14, 2014, Cable Television Laboratories, Inc.

PacketCable™ Specification Multimedia Specification, PKT-SP-MM-C01-191120, November 20, 2019, Cable Television Laboratories, Inc.

Fostering of Patent Pools Covering Cable Technology

Lessons from VVC Pool Fostering

A Technical Paper prepared for SCTE by

Carter Eltzroth
Managing Director
Helikon.net
PO Box 3, Newburyport, MA 01950
+1 202 302 2466
celtzroth@helikon.net

Judson Cary
General Counsel
SCTE
j.cary@cablelabs.com

1. Introduction

Patent pools are increasingly well known as a mechanism to license patents essential to a technical standard. Pools have a variety of benefits, including relative ease in licensing for both licensors and implementers ("one-stop shop") and a lower aggregate royalty level. CableLabs had early experience when it launched in the mid-1990s the first modern day pool covering MPEG2-essential video codec patents. Since then, pools cover a variety of technologies, including video codecs that succeeded MPEG2. Other pools today encompass a wide range of standardized technologies implemented by cable operators or used in cable households. Over time, the fractured licensing environment for the video codec HEVC brought many industry players to look for greater clarity for licensing of the most recent ISO/IEC MPEG video codec, Versatile Video Coding (VVC). Earlier this year, the Media Coding Industry Forum (MC-IF) completed its fostering of pool formation covering VVC-essential patents.

When implementing a video codec in a cable network, SCTE and its members bear in mind not only the technical merits of the standard but also the overall costs, including royalties paid to patent owners by their vendors (e.g., set-top box manufacturers). Patent pool formation and licensing have a cost but provide certainty and other benefits that are attractive to the cable operator and to their vendors. Some of these benefits can be passed on to subscribers and other consumer households.

As undertaken by MC-IF, fostering is a pre-commercial activity that is intended to result in the selection of a single pool administrator that takes forward the work of pool facilitation. Once facilitation is completed (the licensors have agreed royalties and other terms, royalty split, the role of the administrator), the administrator manages the pool, including licensing, royalty collection and distribution to patent owners.

Pool formation takes time. For this reason, MC-IF recognised that it was important to launch the pooling effort soon after the adoption of the new VVC standard. This effort was based in part on the experience of the DVB Project in fostering pools essential to DVB standards. MC-IF scaled DVB's process to meet the challenges of the VVC patent environment: dozens of holders with diverse business models drawn from the audio-visual industry, but also widely divergent industries. Some based their revenues wholly on collection of royalties; some largely on sales of devices and services (while owning one or more VVC-essential patents). In this paper, Section 2 describes patent pooling and its growing use as a tool for licensing patents essential to the implementation of a standardised technology. Section 3 sets out the experience of DVB (and other standards bodies) in fostering the formation of pools covering their standard essential patents. Then Section 4 shows how MC-IF applied DVB's approach to pool fostering in its effort to foster a pool for VVC-essential patents, notably with the goal of the selection of a single pool administrator. Section 5 concludes, first by presenting some "lessons learned" from the VVC pool fostering activity. These lessons could inform future efforts at patent pooling. That section also offers a comparison of VVC Pool Fostering with the licensing frameworks for other next generation video codecs.

2. Patent pooling

Patent pools are increasingly well known as a mechanism to license patents essential to a technical standard. A patent pool is licensing program under which patents essential to a standardised technology are jointly administered by a licensing administrator. This model was introduced by a MPEG LA in the mid-1990s with its pool covering MPEG2-essential video codec patents. It has been adopted by other licensing administrators such as Sisvel and Via Licensing and the recent new entrant Access Advance. Very broadly, the formation of a patent pool is facilitated by a licensing administrator who calls for participation by holders of patents essential to the target technology; it manages the process for review, often by an independent expert, of patents claimed to be essential; and the administrator convenes meetings of holders to negotiate terms, for example, on the joint patent license and on distribution of royalty income among the pool participants. This first step, pool facilitation, can take well over a year to complete. When the participants have agreed on the terms, the licensing administrator launches the licensing program, encourages implementers to take up licenses, and collects and distributes royalties.

Patent pooling is an attractive alternative to other licensing models like bilateral licensing (licensing after negotiation between individual patent owner and individual implementer). Pooling offers a variety of benefits, including relative ease in licensing for both licensors and implementers ("one-stop shop"); a lower aggregate royalty level; and through its essentiality review process, greater certainty of patent quality. CableLabs had early experience when it launched in the mid-1990s the first modern day pool covering MPEG2-essential video codec patents. When CableLabs formed MPEG LA, it worked with the US Department of Justice in developing a number of safeguards to reduce the risk of anticompetitive practices. These included: only standard essential patents may be offered through the pool; the patents holders remain free to license bilaterally; the licensing administrator is independent; and sensitive market information is not communicated to the royalty recipients.[1] This well-settled regulatory framework that reduces exposure to claims of antitrust violation is another attractive feature of today's pooling environment.

Since the launch of the MPEG2 pool, pools cover a variety of technologies. In addition to the video codecs that succeeded MPEG2, these technologies include DVB and ATSC transmission standards; mobile telephony standards; differing audio formats; Wi-Fi (802.11) and more. The success of the MPEG2 pool, and of MPEG LA as pool administrator, has encouraged other commercial entities to offer pooling services, notably Sisvel (whose first pool indeed predates MPEG LA), Via Licensing, and more recently Access Advance and Avanci.[2]

[1] The regulatory framework is set out in a series of Business Review Letters issued by the US Department of Justice. MPEG LA's 1997 letter can be found at on the Department of Justice website here. The framework has evolved as reflected, for example, in a recent letter, *University Technology Licensing Program* (13 Jan 2021).

[2] A list of pools currently operational and under development can be found at the websites of the licensing administrators. See Further Resources below. Until recently Access Advance was known as HEVC Advance.

As a result of over two decades of commercial activity, it's possible to identify elements of the technology marketplace that are well-suited for an intermediary to provide patent pool license administrator services and to derive profits for such services.. A licensing administrator can be useful when

(a) ownership of standards-essential patents is widely diffused (ownership is not concentrated in few holders);

(b) manufacturers (implementers of the technology) are numerous,

(c) for the production of numerous implementations, notably interoperable devices (and corresponding services) for large consumer markets (for example, where the market can be measured in billions of implementing devices);

(d) the standardised technology will not be overtaken quickly by other technological developments; and, relatedly,

(e) the standardised technology will be a market success, ideally a blockbuster, government-mandated or otherwise compelling.[3]

Broadcast transmission standards have fit within the framework. For example, the ATSC pool had (at inception) some eight licensors, a number of manufacturers of ATSC-implementing devices and enjoyed a market primed by analog switch-off (replacement of analog broadcast services with digital broadcasting), mandated by government, and a government subsidy for household purchase of ATSC consumer equipment. Similar factors have influenced territories adopting DVB terrestrial transmission standards. And broadcasting has had in the past long technology development cycles, for example to reduce the risk of obsolescence of television receivers targeted by broadcasters under a public service mandate but constrained by terrestrial frequency scarcity.

Pooling can fail. Pools can fail during the facilitation phase: for example, the call for participation fails to attract the holders of a critical mass of patents, or significant holders representing the key commercial innovation for a standard remain outside the pool. Or holders may not agree on royalty split or other terms. After formal launch, the pool may not be commercially successful because implementers find the terms unattractive and they prefer to wait to see if anyone else takes a license (a form of holdout). The technology underlying the pooled patent may not be a commercial success if it's displaced by later innovation. Pooling can also fall short if there are multiple, competing pools for the same standardised

[3] These factors are drawn from an unpublished study prepared by one of the authors for a standards body then considering offering commercial licensing administration services.

technology. A pool administrator can also abandon a pool if it does not enjoy the anticipated commercial success and it finds that it's more lucrative to commit its resources elsewhere.

In the case of HEVC, competing pools, offering different licensing models and royalty rates, confused the landscape for obtaining licenses of HEVC-essential patents. The confusion arguably slowed market adoption of this video codec technology. As a result of this recent, indeed ongoing, experience, many industry players looked for greater clarity for licensing of the most recent ISO/IEC MPEG video codec, Versatile Video Coding (VVC). Earlier this year, the Media Coding Industry Forum (MC-IF) completed its fostering of pool formation covering VVC-essential patents. Its effort was based in part on the experience of the DVB Project in fostering poo ls essential to DVB standards. Section 3 of this paper sets out DVB's experience in pool fostering; section 4 discusses the application by MC-IF of this experience to its effort in VVC pool formation.

3. Pool Fostering

Pool fostering is generally the effort undertaken by a standards body to encourage the formation of a patent pool covering one of its standards. It is an extension of the work of the standards body after it completes development of the standard. As such it is a precommercial activity, preceding pool facilitation and administration. As a precommercial activity, it generally adheres to the antitrust rules governing standards bodies, including no exchange of market sensitive information, no anticompetitive collusion among participants, etc.[4]

Pool fostering is a response to the perceived risk of market failure of a recently adopted standard due to concern over onerous aggregate royalties or other difficulty in licensing the underlying patents. By and large, there's little need for fostering if the assessment of commercial pool administrators is that a pool is viable, and they independently take preliminary steps leading to pool facilitation. On the other hand, fostering is valuable when participants of a standards body consider that an early start in clarity in licensing for a new standard will be an advantage for market acceptance of its standard. For example, the contributors to the SDO's standard (often the holders of patents essential to these contributions) may have knowledge of the capabilities of the technology and its market potential superior to that of licensing administrators. The SDO participants are aware that a pool may enhance the attractiveness of the SDO's innovation, overcoming market "hesitation." [5]

[4] Pool fostering is discussed extensively in Eltzroth, *Fostering by Standards Bodies of the Formation of Patent Pools* (2018) available at SSRN.

[5] Other instances of "market hesitation" include: when modern pooling was relatively unknown (CableLabs fostering, leading to the MPEG LA pool); and when the commercial returns for otherwise "orphaned" standards are not evident. Fostering may be suitable as a first step to a pool to be sponsored by the standards body (AVS1 pool, sponsored by the Audio Visual coding Standard Workgroup of China; IEEE's similar activities in the 2010s).

Pool fostering can also be attractive to licensors to get an early start on pooling and to reaffirm early licensor interest in "one-stop shop" and the other benefits of pooling. This was the case for VVC pool fostering (discussed in Section 4) when many patent holders shared the market perception that confusion in HEVC licensing caused in part by multiple pools had slowed take-up of HEVC technology. Another feature of pool fostering is that it offers the opportunity for all licensing administrators, incumbents, and new entrants, to present their capabilities and expertise to licensors on a "level playing field." This would reduce the advantages of incumbency. In addition, those new to pool licensing (and indeed new to exploitation of their patents by licensing) can engage with colleagues to discuss pooling, within a framework virtually without enduring legal commitments.

DVB's experience with pool fostering started shortly after its inception in the mid-1990s, prompted by the provisions of the IPR policy in its Memorandum of Understanding.[6] The initial approaches to pool fostering were not the most efficient and through trial and error, across a number of fostering efforts, DVB has developed a "toolbox" to advance a pool fostering effort. Today these include early identification of probable holders of patents essential to a recently adopted DVB standard; after completion of standards development, prompt convening of these holders to an initial meeting of holders; as guided by these participants, exchanges of information between the participants and candidate pool facilitators; formal presentations by the candidates; selection by the participants of a facilitator to take forward the work of pool completion and ultimately pool administration.

Based also on close to three decades of fostering, DVB has a set of well-settled rules for the conduct of fostering, including the basis of participation by a holder of DVB-essential patents; confidentiality; decision by consensus; and equality of treatment. Each of these is geared to be "light-touch," that is not imposing undue burdens and costs on participants. These ground rules were largely adopted in VVC Pool Fostering (discussed in Section 4) with modifications to account for the larger number of participants and the need to hold meetings entirely remotely as a result of the COVID pandemic. In both DVB and VVC Pool Fostering, there was careful attention to adherence to antitrust rules: as with conduct within standards bodies, pool fostering is treated as a precommercial activity. In DVB's pool fostering, commercially sensitive information was not exchanged, and participants were reminded of the constraints on their jointly decided actions.[7]

[6] A discussion of how DVB was brought to include pool fostering in its IPR policy is included in Eltzroth, *IPR Policy of the DVB Project*, Int'l J IT Standards & Standardization Res (2008, 2009) available at dvb.org.

[7] DVB's toolbox is summarised in a DVB document promoting pool fostering, DVB Project, *DVB's Fostering of early Formation of Patent Pools: Note to DVB's Liaison Partners and to Standards Bodies that author Materials normatively referenced by DVB Standards.* See Further Resources.. DVB has other tools relating to pools, involving the use of its IPR Module as a sounding board for exchanges of views of possible licensing terms. (VVC Pool Fostering did not exclude a later meeting to discuss pool developments if circumstances warrant.) DVB also intervenes when a pool is failing. Eltzroth, *Fostering by Standards Bodies of the Formation of Patent Pools* (2018), available at SSRN.

4. Pool fostering for VVC

Media Coding Industry Forum was founded in 2018 with a goal to facilitate cross-industry discussions around the non-technical aspects of deployment of media coding standards, including patent licensing. MC-IF has become the focal point for discussions around deployment and licensing of the next-generation video coding standard, Versatile Video Coding (VVC).[8] At the time of MC-IF's formation. VVC was on a course to complete formal standardisation within ISO/IEC by summer 2020.

In common with other industry participants, MC-IF members noted the lack of clarity in licensing of patents essential to HEVC, the predecessor standard of VVC. In a nutshell, it was claimed that market adoption of HEVC was slowed by the multiplicity of pools covering essential HEVC patents; the competitive tension between the pools; other licensing structures formed by patent holders; and still other holderss that had made clear that they would license their HEVC-essential patents only bilaterally. In the face of this apparent confusion, some contributors to VVC standards development looked to MC-IF for a solution that would, for VVC, resolve the issues that have persisted for HEVC. The general view was that confusion could be reduced if a single licensing administrator for a VVC pool could be named early by the patent holders.

Several MC-IF members were familiar with DVB's model of pool fostering. In addition, DVB had recently begun "evangelising" its fostering model through contacts with sister standards bodies, including ISO/IEC.[9] Senior leadership of both DVB and MC-IF met at IBC in Amsterdam in September 2019 for a further discussion of pool fostering, its benefits and DVB's experience. By spring 2020, MC-IF had decided to launch a pool fostering effort and engaged a specialist in pool fostering to lead the effort.[10]

Several weeks after the VVC standard was adopted by ISO/IEC,[11] MC-IF's pool fostering issued a call for participants and set an initial meeting to occur on 1 September 2020. The activity was designated

[8] The website of MC-IF has further information on its activities and materials on its VVC pool fostering activity. www.mc-if.org Other next generation video coding standards include Low Complexity Enhancement Video Codec (LC-EVC) and Essential Video Codec. Together with VVC, both are standardised through ISO/IEC. Other recent video codecs, not developed through ISO/IEC, include AV1 (developed by the Alliance for Open Media) and AVS3 (Audio Visual coding Standard Workgroup of China). The licensing policies related to LC-EVC, EVC, AV1 and AVS3 are discussed in Section 5.

[9] See Further Resources for a link to DVB's Note addressed to other standards development organisation on DVB's Fostering of early Formation of Patent Pools.

[10] Both convenors were experienced in pool fostering. Carter Eltzroth, as DVB's Legal Director, had led its pool fostering across a range of DVB standards. He was named convenor. Co-convenor Judson Cary played an active role in pool formation for a technology for which standards were adopted by DVB and CableLabs. He was also President of MC-IF.

[11] The VVC standard was consented by ITU-T Study Group 16 on 3 July 2020, to be published as ITU-T Recommendation H.266. Concurrently, MPEG submitted the VVC standard for Final Draft International Standard ballot, to be published as ISO/IEC 23090-3.

"VVC Pool Fostering." While the activity was sponsored by MC-IF, the designation was intended to indicate that fostering was pursued independently of MC-IF. It would be for the participants in VVC Pool Fostering – the VVC-patent holders – and not the board and members of MC-IF, to take decisions on its direction and final decision. The initial meeting and later meetings were all held virtually because of the COVID pandemic. This presented disadvantages but at least one benefit: because there was no travel, meeting cycles could be significantly shortened.

In addition to the call for participants and other press releases, the co-convenors actively solicited for other participants. Ultimately 49 companies joined VVC Pool Fostering. By its participation each affirmed that it had a well-founded belief that it held one or more VVC-essential patents. Among these 49 companies were nine of the top 10 companies whose contributions were accepted during the course of VVC standards development. The 49 companies represented a mix of R&D companies and non-practising entities, together with implementers (with essential patents). There was broad geographic diversity: 15 companies were US-based; eight came from the EU and UK; and 26 from East Asia. The number of Chinese companies (10) represents their increased contribution to standardisation, including the VVC standard, and the recent emphasis on patent filing in China.

As in DVB's practice, the initial meeting of VVC Pool Fostering set out the operating rules governing participation, conduct and decision-making. The participants agreed that:

> Participation by each company was based on its well-founded belief that it holds one or more patents potentially essential to VVC; consistent with the DVB model, the company's "well-founded belief" was sufficient; there was no call for declarations of essential patents (or third-party determination of essentiality) that could have imposed unnecessary (and ultimately duplicative) costs on participants;
>
> Each participant agreed, by its presence in VVC Pool Fostering, to treat as confidential the contents of meetings and documents (notably presentations by candidate facilitators); no formal non-disclosure agreement was proposed; the sole exception to confidentiality covered public statements agreed by participants in VVC Pool Fostering, for example press releases reporting on progress or calling for additional participants;
>
> Any decision (including notably on candidate facilitators) was to be undertaken by consensus or, in the absence of consensus, a two-thirds supermajority; on consensus, VVC Pool Fostering followed the well-understood notion of this practice within standards bodies, including ISO/IEC;
>
> All participants were to be treated equally; there was no favorable treatment for MC-IF members and no separate "voting block" formed by MC-IF;
>
> There was to be strict adherence to antitrust rules and antitrust counsel was present at all meetings; and

In view of the varied business practices of participants and the need for virtual meetings, the participants adhered to a basic set of unexceptional netiquette guidelines.

The convenors also solicited the participation of licensing facilitators based on the convenors' knowledge of pooling and the suggestions of VVC Pool Fostering participants. Some eight were contacted; of these four agreed to participate as candidate facilitators. After the initial meeting, the participants drew on tools from the DVB toolbox: they solicited questions from potential candidate facilitators; once received the participants agreed on answers that were then, together with the underlying questions, circulated to all candidate facilitators. Conversely, the participants addressed questions to potential candidates. After these exchanges, VVC Pool Fostering held successive meetings to receive presentations from the candidate facilitators. As the meetings progressed, the field of candidate facilitators was narrowed progressively from four to two.

Over time, as VVC Pool Fostering continued its work, some additional process elements were added. One concern was that some participants may not fully express their views because of the difficulties in communicating in a remote setting and because of reticence (whether cultural or due to lack of familiarity with patent licensing). For this reason, the convenors redoubled their efforts at participant outreach. In addition, VVC Pool Fostering facilitated participant engagement by the use of confidential non-binding surveys, notably to solicit views on individual candidate facilitators. Anonymous comments were welcome (that is, delivered to, and anonymized by, the convenors). Moreover, an important process development was the use during meetings of the roll call, calling upon each company to express its views. This prompted colleagues to prepare a statement (but some would merely offer "not ready to express a view").

VVC Pool Fostering progressively narrowed the field of candidate facilitators from four to two. In each case of elimination, a secret ballot was taken. The result of the vote guided the decisions taken at the next meeting; a consensus of participants was found that a candidate was to be removed from contention. MPEG LA and Access Advance remained. By the seventh meeting, it was clear that the participants would not find consensus (as "consensus" was defined by participants at their initial meeting) around a single facilitator to take forward the work of pool formation. In addition, neither would achieve the two-thirds supermajority vote to make a definitive decision. Instead VVC Pool Fostering "identified two strong pool administrators."[12]

VVC Pool Fostering had set as its goal the selection of a single licensing administrator to form a pool of VVC-essential patents. A single facilitator was thought to be best suited to reduce the type of market confusion, recently encountered in HEVC licensing, caused by multiple licensing administrators offering patents in multiple pools. But the same contest between leading administrators of HEVC pools spilled over to the deliberations with VVC Pool Fostering and shaped the ultimate outcome. Each final candidate facilitator had a core group of participant supporters to vote and to advocate on its behalf. In

[12] The press release of VVC Pool Fostering announcing this result, together with other public statements during the course of pool fostering, are available on the MC-IF website. See Further Resources.

addition one candidate explicitly linked its existing HEVC program to its VVC proposal so that, for example, an implementer could obtain a license covering both video codecs.

VVC Pool Fostering was not writing on a clean slate; it was arguably inevitable that the turmoil of HEVC licensing would frustrate it from achieving its goal of one-stop shop for VVC licensing. On the other hand, some participants reflected that the outcome of VVC Pool Fostering had value because a "two-stop shop" was better than none at all. And the competing efforts in pool facilitation by the "two strong pool administrators" may (at the time of this writing) yet result in a single administrator.[13]

5. Conclusions

While its efforts did not result in the selection of a single licensing administrator to take forward the work on a VVC patent pool. VVC Pool Fostering was successful in bringing together more than a critical mass of VVC patent holders, drawn from a variety of industries and using diverse business models, to discuss VCC licensing arrangements in a precommercial setting. It provided a platform for presentations by several licensing administrators, including a new entrant. This was achieved in the midst of the pandemic when remote communications alone were available. Overall, the DVB model, grounded on a toolbox of arrangements for convening and advancing pool fostering, successfully scaled up to cover a standard with far more contributors.

At the same time, some lessons may be drawn from the experience of VVC Pool Fostering.

> *(Slightly) tightened participation test?* The basis for participation in a pool fostering effort is today established as "well-founded belief that the participant holds one or more patents essential to the standard". In VVC Pool Fostering this imposed a low threshold for those new to pooling (and indeed to patent licensing). This was welcome. At the same time, the cost for participation was low for interlopers or "peepers," those companies with essential patents that had no intention ultimately to join a pool.
>
> The test for participation could be changed by requiring the affirmation from the participant that it is "actively exploring joining a patent pool".

[13] By mid-August 2021, Access Advance had issued a press release announcing the launch of its VVC licensing program, *Access Advance Launches VVC/H.266 Video Patent Pool: includes innovative Multi-Codec Bridging Agreement that Provides Substantial Royalty Savings to Licensees in both the VVC and HEVC Advance Pools* (1 Jul 2021) and (earlier) a Royalty Rates Summary (1 Jun 2021). MPEG LA has announced development of a VVC pool license and called for submission of VVC-essential patents: *MPEG LA Announces Development of VVC (Versatile Video Coding) Pool License: VVC expected to improve video compression efficiency and functionality* (27 Jan 2021).

Confidentiality. Confidentiality was based on a "gentleperson's agreement" not to disclose. It worked well (at least the convenors were not aware of significant leaks of confidential information). At the same time it may be useful to consider, in a future pool fostering a "plain-vanilla" non-disclosure agreement, that could be adopted without significant negotiation (and without delay).

Confidentiality protected the deliberations and materials of VVC Pool Fostering. It may have negatively impacted its activities when some participants, when solicited by a candidate facilitator, entered into an NDA with that facilitator. It's possible that these participants felt constrained in their freedom to discuss the merits of the candidate facilitators. A further pooling effort could discourage participants from agreeing to arrangements with potential candidate facilitators that undercut the fostering process

Finding consensus. During its initial meeting, VVC Pool Fostering adopted a rule that, in the absence of consensus, decisions were to be taken by a two-thirds supermajority. It was essential to have adopted this rule at initial meeting! But the value of this rule could be reconsidered when a large number of participants refuses to vote, declaring themselves "not ready," with "no comment," choosing "to abstain." Other voting or tally mechanisms could be considered.

Affiliates in pooling. Participants and candidate facilitators in VVC Pool Fostering were not required to disclose any cross-ownership or other material affiliations. For example, several participants were shareholders in candidate facilitators. These links are well-known among licensing professionals, but participants new to licensing may be unaware. During meetings, they could have a sense that proceedings were gamed. For this reason, it may be suitable to require, in pool fosterings, disclosure by a participant if it has an affiliation with candidate facilitator and its pledge that it will not pass confidential information on to its affiliate.

RfP? In DVB's experience, during the presentation phase, each candidate facilitator largely shapes its own materials and chooses what to disclose. In VVC Pool Fostering, some participants already had extensive knowledge with the candidate facilitators and drew on this background when developing questions addressed to them. Their answers were worked into the presentations (or delivered separately). In pool fostering where there are several candidate facilitators (and many participants) it may be well to specify the structure of the presentations, and the items to be covered. This could be formalized in a request for proposals issued by the pool fostering participants. Setting a more formal framework for presentations could permit the participants more easily to compare competing proposals and the respective strengths (and weaknesses) of the licensing administrators.

As a result of VVC Pool Fostering, is VVC in an advantageous position compared to other next generation video codecs? It is hard to make a definitive assessment because of the range of licensing models offered by competing codecs. In addition to VVC, two other video codec standards have been recently approved by ISO/IEC. Essential Video Coding (EVC) is an open-source codec completed by MPEG in April 2020 where the development effort was led by Samsung, Huawei and Qualcomm. The

development process has been defensive against patent threats establishing a baseline with codec tools made public more than 20 years ago. There are a further 21 tools for the main profile to be available under separate royalty-bearing negotiated licenses. In a May 2020 press release, the three companies announced that they would be announcing their respective licensing terms for these further tools within two years.[14] In respect of Low Complexity Enhancement Video Codec (LC-EVC), the company responsible for its core technology foundation, V-Nova, announced licensing terms in May 2021: use of its essential patents are royalty-free for integration by device or chipset manufacturers, browsers, encoder / player vendors; but a fee is payable for usage by service operators based on service size (from $0.01/per user per year to a cap at $3.7 million/year).[15]

In addition to the codecs adopted by ISO/IEC, two other standards bodies have completed development work on competing technology solutions. The codecs previously adopted by the Audio Visual coding Standard Workgroup of China were each the subject of a patent pool characterised by a low royalty. For its AVS3, AVS has launched pool facilitation, calling for holders of AVS3-essential patents to submit declarations for an initial meeting of holders during autumn 2021. AVS has made clear that decisions on royalties and other licensing terms will be determined by the essential patent holders participating in AVS3 pool formation.[16] A further codec, AV1 has been developed by the Alliance for Open Media, which, for licensing has adopted the W3C IPR policy. Its form of patent license provides for a "no-charge, royalty free" license for implementation. But this license is not free of controversy because it also provides that the implementer / licensee must make its own essential patents available under the same royalty-free terms. While this creates a genuine "eco-system" for royalty-free licensing, it arguably places some implementers, looking for royalty revenues, at a disadvantage.[17]

The licensing terms for each of these alternatives are associated then with special factors: EVC, with a baseline of technology subject, if at all, to expired patents; LC-EVC, a single principal owner of the core technology; AVS3, a pool structure now seeking to extend beyond East Asia; and AV1, available on a royalty-free basis, provided the licensee agrees to reciprocal RF licensing. VVC can be based on a more classic model of licensing of standards-based technology, coupled with pooling, a well-established mechanism for easier licensing administration. It can follow the structure already adopted for a long line of video codecs, MPEG2, AVC and HEVC. Fostering is the means to encourage VVC holders to take on that structure again.

As cable operators define next-generation set-top boxes and streaming services, they will need to choose video codecs that meet their technical needs for compression, efficiency, and resolution. They will also

[14] Samsung et al, *MPEG-5 EVC is the next generation video codec for the media industry* (8 May 2020)

[15] V-Nova, *V-Nova LCEVC Licensing Terms announced for Entertainment Video Services* (21 May 2021)

[16] Communication made by AVS to one of the authors (among others).

[17] Information on the Alliance for Open Media and its licensing policy can be found at www.aomedia.org. Sisvel has formed a pool around AV1-essential patents not bound by the AOM licensing policy.

need to take into consideration the patent license royalties and licensing models associated with their codec of choice.

Abbreviations

AOM	Alliance for Open Media, www.aomedia.org
AV1	video codec of AOM
AVS	Audio Visual coding Standard Workgroup of China
AVS3	video codec of AVS
DVB	the DVB Project, a standards development organisation
EVC	Essential Video Coding, an ISO/IEC standard
HEVC	International Society of Broadband Experts
LC-EVC	Low Complexity Enhancement Video Codec, an ISO/IEC standard
MC-IF	Media Coding Industry Forum
VVC	Versatile Video Coding, an ISO/IEC standard

Further resources

Press release, VVC Pool Fostering identifies Access Advance and MPEG LA as possible administrators to take forward pool formation covering VVC-essential patents (MC-IF, 27 Jan 2021)

DVB Project, DVB's Fostering of early Formation of Patent Pools: Note to DVB's Liaison Partners and to Standards Bodies that author Materials normatively referenced by DVB Standards (2018)

Frequently asked questions on pool fostering, and other materials, available at VVC Pool Fostering | MC-IF (mc-if.org)

Information on commercial pools can be found at the websites of licensing administrators, for example, Access Advance, Avanci, MPEG LA, One Blue, Sisvel, Velos Media and Via Licensing.

For licensing of next generation video codecs, information on the progress on the pooling efforts undertaken by Access Advance and MPEG LA, as the next step following VVC Pool Fostering, can be found on their respective websites. For LC-EVC, V-Nova, V-Nova LCEVC Licensing Terms announced for Entertainment Video Services (21 May 2021); ; for EVC, Samsung et al, MPEG-5 EVC is the next generation video codec for the media industry (8 May 2020); for AV1on the License webpage of the Alliance for Open Media ; for AVS3, through the website of AVS.

Authors

Carter Eltzroth is Managing Director, Helikon.net, and Legal Director, DVB Project, Geneva. celtzroth@helikon.net **Judson Cary** is President, MC-IF and General Counsel, SCTE. j.cary@CableLabs.com In 2020/2021, each served as co-convenor of MC-IF's activity to foster pool formation covering VVC. They earlier worked together to foster the formation of a DVB pool. While

they are grateful for the input of the many participants in the VVC Pool Fostering, the views the authors express in this paper (and the remaining errors) are their own.

From Bolted-on to Built-In: The Journey of Cybersecurity

A Technical Paper prepared for SCTE by

Casandra Bowes
Principal Security Architect
Comcast
650 Centerton Road, Moorestown, NJ 08057
(609) 313-5636
casandra_bowes@comcast.com

Harwant Mahal
Director, Security Authentication and Access Management
Comcast
650 Centerton Road, Moorestown, NJ, 08057
(215) 756-4387
harwant_mahal@comcast.com

1. Introduction

Over the past two decades, the transformation of service delivery and use of social media and digital platforms have opened the flood gates for the threat landscape. When it comes to protecting ourselves from bad actors, are tips, tricks, and risk indicators enough?

In short, the answer is no. It is not enough in some cases. We see evidence of this every day in the news. Even though we see signs that bad actors are sometimes successful, security professionals often prevail. With a conscious effort to move from "information" and "education" to meaningful behavioral change, we are turning the tides.

Laws are being written and updated with new mandates on data protection and privacy. Corporations are taking strides, reacting to government mandates, and implementing initiatives such as shifting security left. Security professionals are stepping up their game on both user and device identity, reducing the blast radius of potential attacks, and making it harder for bad actors to breach networks and steal data. Protection is provided to consumers by building security into products and services, as well as account protection measures like strong authentication and alerts of new or suspicious activity.

But there is more to do. Security professionals must go beyond tips, hacks, and risk indicators and use all the tools in their arsenal to create plans that incorporate a blend of policy, process, and technology as a catalyst to change end-user and consumer behavior.

This paper demonstrates how security threats have shifted marketplace adaptation. Modern DevSecOps and security practices are discussed, along with how establishing cyber risk ratings to provide accurate and transparent cyber health can uplift security efforts. Adoption of these transformations can bring security out of the shadows and benefit corporations and consumers alike, making a siloed approach to security tenable.

2. Security Goes Mainstream

The 2010s can be labeled as the decade cybersecurity went mainstream. Data breaches started to make big headlines. Personal information was being stolen from big name retailers, popular gaming sites, and even the US government. Millions of people were affected, and corporations faced extensive downtime, costly damage, expensive fines, lawsuits, and a loss of consumer trust. Government secrets were exposed, threatening the safety and security of nations. The lives of everyday people were being affected by cybercriminals and it was being felt across the globe.

2.1. How Hackers Pushed Security Forward

Cybercriminals with no budget constraints or change control processes developed a level of sophistication that surpassed the outdated "point solution" security controls that were protecting our most valuable data. They were taking a multifaceted polymorphic approach that easily bypassed the mechanisms that were supposed to keep them at bay.

Amidst all the damage, society was awakened to just how much disruption and destruction could be done with surprisingly little effort. This changed the perception that security can be left as an afterthought. Cybersecurity budgets grew. Corporate board members started paying attention. Governments started paying attention. Consumers started paying attention. Security was getting its long overdue 15 minutes of fame.

While there is no industry that is safe from the threat of attack, the cable industry is one of the top 10 targets for hackers [1]. “Always-on” internet access and video streaming have changed consumer behavior. Consumers are moving away from watching linear channels to adopting streaming channels which require high-speed broadband. With this shift in demand, broadband security risks are being compounded. As the details and motivations of attacks continue to evolve, there are several popular methods attackers use, which are described next.

2.1.1. Compromised Accounts

A favorite weapon of choice when looking to gain unauthorized access to networks and wreak havoc on unsuspecting victims is compromised accounts. Companies can see from hundreds to thousands of attempts per day to penetrate both enterprise and consumer accounts on their networks via compromised credentials. Due to increased sophistication of password-related attacks, coupled with the slow adoption of methods to combat them, an astounding 61% of data breaches can be traced to compromised credentials. [2] There are several ways in which accounts are compromised. Some of the more common methods include phishing attacks, brute force password attacks, and credential stuffing.

Once an attacker retrieves a set of credentials, they are validated, and after a successful test, made available on the dark web for sale. Fraudsters then purchase these credentials and use them not only to access consumer broadband accounts and steal services, but also as a pivot point to identity theft. Additionally, the fraudsters can sometimes use the credentials to access consumer email. Once access to an email account is gained, they use scanning tools to find third party targets like bank information, financial transactions, credit card information or crypto currency accounts associated with the victim. The email credentials are used to reset passwords, and possibly as 2nd factor authorization to access the victim’s financial accounts. Sometimes fraudsters can even get as far as transferring money out of a victim’s account. The use of stolen credentials by fraudsters to gain access to consumer banking services creates a ripple effect to the financial sector. Though cable operators can implement more secure authentication mechanisms to protect their own subscriber accounts, a chain is only as strong as its weakest link. If authentication is weak on one account, all accounts that are linked are exposed to risk.

The use of video streaming services by consumers during the pandemic has increased many folds. Some consumers share their credentials with their family, but in some cases, these credentials are sold for commercial use. The identification of legal sharing of passwords within family or illegal sharing has become another threat to video streaming and account security.

2.1.2. Brand Phishing

“Lower your cable bill” is an example of a scam that has been around for many years and has hurt the brands of many cable companies. The scammers pose as ‘companies’ officially doing business, call customers, and trick them with better deals, getting monthly payments out of their pockets and impacting the company brand. Often, these scammers will request the account credentials to offer discounts, further exposing the subscriber to identify theft risks.

2.1.3. Ransomware

Ransomware attacks go beyond just stealing and exploiting data, with recent attacks impacting the daily lives of many in ways we have not experienced before, causing gas shortages, interruptions in public transit, and disrupting the food supply chain. Unfortunately, experts do not expect this onslaught of attacks to slow down anytime soon. It is projected that during 2021, ransomware attacks against business will occur every 11 seconds, with costs expected to exceed $20 billion globally. [3]

The concept of a ransomware attack is as follows: The hacker finds a way to install malicious software onto an unsuspecting network. Usually this occurs via malicious emails with malware, an unpatched vulnerability, or exposed ports or services with weak authentication for remote access. Once the ransomware is deployed, the data is encrypted, and the victim receives a ransom note demanding payment, usually with a cryptocurrency such as Bitcoin. The victim is then faced with the choice to either pay the fee or incur the untold damages of downtime, rebuilding, or declaring a total loss on the infected resource(s). When faced with these difficult options, more than half of the victims of ransomware attacks pay the ransom. This is not a favorable outcome, given that it fuels further ransomware attacks.

2.1.4. Voice Fraud

Voice technology has taken a strong position in customer interactions for businesses in recent years. It is the future of every service offering and still being adopted in the industry. The voice interactions are using artificial intelligence (AI) for chat bots to respond to customers. There are device apps which provide faster calling services. These new innovations in voice technology have invited scammers for voice phishing [4] [5]

Additionally, scammers make their way into unsuspecting victim's voice service accounts and place high volumes of expensive calls on the victim's dime. This practice is lucrative for criminals and expensive for voice service providers with an estimated cost of $12 billion dollars in damages annually.

2.1.5. Internet Facing Application Attacks

Business-centric interactive websites have become the new norm to do business and provide quick responses to customers. Although these on-line sites have tremendously

improved the user experience, they have also brought new cyber security threats. The most common threats are on web servers and databases for cross-site scripting and SQL injection. According to industry data, 90% of the applications have security flaws and it takes an average of 38 days to patch the application-related components. During the development process of these web applications, code analysis is done for less than 50% of applications. [6] A small rogue application, which is not on a company's radar, poses an even larger risk due to lesser controls and assessments. A vulnerability on a rogue application's web server with internet connectivity and unknown change control practices, can invite a threat actor with this ingress and allow lateral movement.

2.2. Turning the Tides

If there is one guiding principal security professionals should follow, it is this one: Assume you will be breached.

Most companies have started initiatives related to the overall cybersecurity uplift, to strengthen the ingress points, decentralize applications, implement, and govern the security frameworks with Zero Trust to reduce the blast radius -- to name a few. The use of Machine Learning (ML) for cybersecurity threat intelligence, multi-factor authentication to inform risk engines, for both users and devices, has contributed greatly to the understanding and reduction of risk within the threat landscape. The enterprise-side user education on threats, via ransomware readiness exercises, phishing exercises, and Zero Trust programs will help prepare the organizations for growing attacks. The implementation of MFA brought a positive change, and over last couple of years it reduced the number of threats for identities.

2.2.1. Zero Trust Security

If one were to make a list of most popular security buzzwords, Zero Trust would likely be close to the top. But while buzz words usually get a bad rap with security professionals for not living up to the hype, Zero Trust would be an exception. Maybe because Zero Trust is more than a "latest and greatest" product suite that you enable to solve all your security woes. In fact, it's not a product at all, but rather a philosophy or mindset of "never trust, always verify." [7] [8]

Traditional security models worked like a "castle and moat" system for accessing corporate resources. A perimeter was created around the company network and security efforts were focused on ensuring the perimeter was not breached. Everything within the perimeter was considered a safe and trusted resource. Everything external to the perimeter was not trusted. This model was more effective when corporate resources lived within the perimeter and were accessed from the confines of the corporate network using corporate owned devices. Today this is not the case. Corporate resources exist both within and external to the corporate network and can often be accessed from any device and/or location. The network perimeter, where security efforts were traditionally focused, has disappeared. [9]

In contrast, a zero-trust security model assumes every device, user, and application is a threat until verified. Every resource is treated as though it is an attack vector, and measures are taken to contain the damage of an attack. IP based access controls are replaced with identity-based controls and security is taken out of the shadows by shifting the focus to people, devices, and applications as the network perimeter.

While each organization will implement a zero-trust security model in its own way, there are common elements that one would expect to find including strong identity management for users, devices and applications, network segmentation, and advanced monitoring and logging. [10]

2.2.2. Improved Incident Detection and Response

To stay ahead of the constantly shifting threat landscape, corporations are building their own threat intelligence programs. These programs, focused solely on threat prevention and detection, use a combination of threat intel data collected from trusted sources and advanced SEIM technologies to alert organizations of potential and actual threats. They also focus on incident preparedness, using tools such as the MITRE ATT&CK framework.

The MITRE ATT&CT framework was designed to help organizations in assessing risk against common tactics used by attackers. The framework is defined as “a comprehensive matrix of tactics and techniques used by threat hunters, red teamers, and defenders to better classify attacks and assess an organization's risk” [11]. Using tools such as the MITRE ATT&CK framework can assist with mapping threats detected in a SEIM system to potential use cases for security automation.

Another instrumental tool in preparing for attacks is an exercise known as a purple team event. In these events, both a red team (made up of “attackers”), and a blue team (made up of “defenders”) get together in a controlled environment and use an “offensive defense” strategy to identify security gaps within an environment. These exercises, sometimes partnering with specialized third-party organizations, help to improve skills and techniques of the participants so they are ready to respond when a real incident occurs. They also bring visibility to security gaps that do exist, which helps prioritize the efforts of resources so they can focus on the biggest wins.

2.2.3. Multi Factor Authentication

We already know that one of the most common ways data breaches occur is compromised credentials. We have also looked at common ways accounts are compromised. The best defense against these kinds of attacks is the use of multi factor authentication (MFA).

The premise of MFA is that you prove you are who you say you are – aka authenticate yourself - by providing something you know – your password – with something you have

(token, one time passcode [OTP], approved sign-in request) and/or something you ARE (fingerprint, facial recognition).

Although some forms of MFA have been around for the better part of two decades, adoption has been slow, mostly due to cost, additional hardware requirements, and the cumbersome user experience. Prior to the abundance of data breaches that took place over the past decade, companies and consumers alike were not particularly interested in investing money or losing convenience by adding steps to access online resources. As companies grew better at detecting compromised accounts and realized how often and easily passwords were being compromised regardless of "best practice" password policies, it became widely recognized that the benefits of MFA far outweighed the drawbacks of cost and inconvenience.

Technical difficulties, conflicting priorities, and an endless list of ever-changing use cases are just a few of the very real challenges companies may face when deploying an MFA solution, especially when considering the vast differences between protecting enterprise and consumer accounts. However, there is no disputing its effectiveness at stopping password related attacks almost entirely, with a success rate of more than 99% [12].

When implementing MFA, there are a variety of different solutions from which to choose. Depending on the complexity of the environment, there may not be a single solution that meets all use cases. In situations such as this, multiple solutions may need to be deployed to ensure all resources and users are MFA protected. For many companies it is likely that MFA will be an ongoing journey that will continually be improved upon as new technologies emerge to combat the ever-constant threat of account takeovers.

2.2.4. Ransomware Readiness

With ransomware attacks on the rise, it is more important than ever to be prepared on how to respond to a ransomware attack. There is no one single strategy to combat ransomware attacks. To best prepare for a successful outcome, consider the below recommendations:

a) Address your known vulnerabilities and keep up to date on patches, especially on parameter assets.
b) Disable unused services and processes, specifically remote desktop protocol (RDP) and secure shell (SSH), on externally facing systems. If these services must be exposed, use access control lists (ACLs) and multi-factor authentication.
c) Use least privilege access models.
d) Reduce the blast radius of attacks with network micro segmentation
e) Use advanced security tooling for logging, monitoring, and alerting to bring visibility to what's happening within your environments
f) Back-up your systems regularly and encrypt backups
g) Have a response plan ready and practice recovery efforts, especially for critical resources.

2.2.5. Shifting Security Left

The concept of "shifting security left" is all about closing security gaps further upstream. To accomplish this, security feedback must be incorporated as part of the feedback loop throughout the entire software development lifecycle (SDLC). Adopting the mindset of addressing security concerns before any code is written saves teams considerable time and money and elevates the level of trust customers have in products and services. When a multifaced approach is taken that includes training, coaching, and automation to help prioritize the sometimes-overwhelming number of security tasks, application development teams can improve the chance of successful delivery for secure products without adding any additional time to their development cycles.

The software development cycle has many skip hops when trying to reduce cycle time and get products delivered faster into customer hands. To support shortened cycle times that incorporate security into the development process, many organizations are on a journey to transition from DevOps to DevSecOps deployments. This shift brings the demand for new tools, and specialized knowledge. This creates the need for organizations to determine how to support these new requirements. Some organizations have incorporated coaching programs, appointing experts in both security and the development process to guide their development teams through this transition.

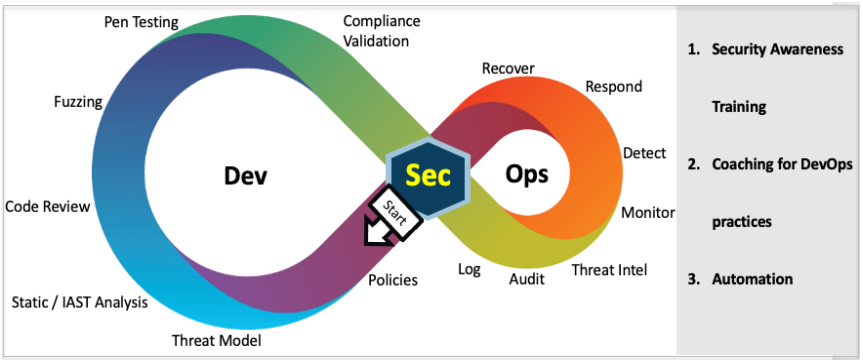


Figure 1 - Shifting Security Left in each phase of SDLC

2.2.5.1. Security awareness training

One aspect of creating a security focused development process is ensuring that all resources contributing to any part of the SDLC are well trained on security best practices and policies, including those in management roles. Ideally this training would be specifically tailored to the role(s) of each participant. Training material could be catered in a tiered level ranging from a base level of understanding all the way to expert. The ideal would be that, after training, participants are able to

demonstrate their knowledge of the security concepts covered before advancing to the next level.

2.2.5.2. Coaching for DevOps practices

With things evolving so fast, it may be hard for development teams to stay on top of emerging security threats and changing security priorities as well as stay focused on improving dev/ops practices to deliver quality products. This need creates an opportunity for security coaching. Security coaches can work with teams to align security practices to the phases in the development lifecycle. These individuals, with an expertise in both security and the software development lifecycle, guide development teams through onboarding new security requirements or shifting processes to incorporate better security practices by assimilating them into their team norms. Coaches can work with teams to understand their workload and constraints, and assist with prioritizing the security work, breaking it down to manageable deliverables.

2.2.5.3. Automation

Another helpful tactic is adding automation around security scans and other security required processes that have made their way into the continuous integration and continuous delivery (CI/CD) pipeline. By introducing automation around these new security requirements, teams are not bogged down with new requirements and things are able to run smoothly. Training the development teams and appointing developers as "security champions" can ensure that as developers built up their skills to handle the fallout of these new processes, the teams could still operate efficiently. [13]

3. Securing Your Customers

The defense strategy against attackers seeking to steal service, customer data, and video content is still evolving. It benefits broadband operators to collaborate both within the industry as well as with external partners to set guiding principles of basic security, communication to customers, and guidance on how to consume their services in secure ways. There are many efforts to secure web interfaces, user logins, and devices by broadband providers. Additionally, the following actions will help to improve the protection for consumers from email threats, account abuses, and device security.

3.1. Identity Protection Capabilities

Implement identity protection capabilities to combat the threats at scale by using AI to support consumers. Comprehensive detection of darknet market ads for account credentials provides the capability to measure the availability of account credentials. This is just an example of one measurement for identifying credentials available for sale in the darknet market. Some additional measures are listed below:

a. External credential spill monitoring and remediation can be built as a basic feature for standard security operations
b. Deep and dark web credential advertisement detection and remediation
c. Robot software agent (BOT) attack prevention at web, application programing interface (API), and mobile authentication interfaces
d. Use internet protocol (IP) information for "geo velocity" and to determine "geo location" to reduce credential theft
e. Implement 2FA or MFA support for consumers and disable less secure authentication methods
f. Detection capabilities using Machine Learning (ML) based tools for credential sharing and compromised accounts
g. Work with law enforcement and other enforcement bodies to identify and disrupt the distribution of unlicensed content

3.2. IoT Device Infection Prevention

In the connected world of customers, internet of things (IoT) devices pose a large threat [14] and have not received full attention. Customers usually do not focus on security measures for devices and forgo more secure devices for their less costly counterparts. On top of this, there is not enough standardization of security control requirements for device vendors. Also, vulnerabilities for these devices do not get addressed regularly. Bringing privacy awareness to device security features as well as how to keep devices up to date with the latest patches is the key to protect individuals from being attacked via their connected devices. Internet providers are offering internet security solutions which block access to compromised or malicious domains and protect customer devices from these threats.

IoT device infections have grown 100% during Covid. [15] This surge in the rate of device infections, directly matches the trajectory of the visibility of the devices on the internet. The cybercriminals probe these devices for security vulnerabilities and exploit them to control the device with unauthorized access, damage it, or penetrate to other devices on the same network. Antivirus software and intrusion detection capabilities are the first layer of security to protect these devices from infections. [16]

IoT is here to stay, and it will continue to grow tremendously, as will the attack surface. The industry demands security leaders to develop a comprehensive IoT lifecycle approach and an IoT security posture to protect and enable IoT devices from both existing and unknown threats. Also, consumer awareness for various precautions to protect the IoT devices is of utmost importance.

3.3. DDoS Attack Mitigation

The exposure of IoT vulnerabilities has dramatically changed the landscape of Distributed Denial of service (DDoS) attacks in recent years. These devices have enough compute power to launch any processes using BOTs and create an attack on other devices. Along with detection measures, effective preventive measures are important to combat the DDoS attacks. Many Threat Intelligence subscription services "Always-on DDoS Protection" are evolving

to help combat these threats and help the customers. [17] [18]. These capabilities include preventive, detective, and mitigation controls which automatically or on-demand mitigate the DDoS attacks.

3.4. Protect Customer Data

In today's data driven customer interactions, businesses rely on large amounts of data being available every second of the day without a blip. The Network Data Loss Prevention (DLP) solutions are relied upon by many organizations, but data breaches happen all the time. Hence there is more to do to protect the data. Privacy initiatives and recent Executive orders provide guidance on some controls around securing data for user identities. They also highlight a need for data-centric tools to build personally identifiable (PI) data inventory, creating data classification policies, performing data discovery, and other initiatives that address data privacy concerns around customer data. To safeguard the data, the use of data encryption and data de-identification practices provide protection but are still evolving. Another key factor in securing data is ensuring that there are strict policies around granting access to and removing access from sensitive data as needed.

3.5. Customer Education

A recent, sophisticated phishing scheme was uncovered in which consumers of unlicensed content were sent an email indicating their free trial was ending and their credit cards were about to be charged. They were directed to a fake website and informed that to cancel their subscription, they must download an Excel file, which contained a ransomware program commonly used by the REvil ransomware group. This is one of the many dangers consumers encounter in their daily lives. It is important to educate consumers of such risks and provide guidance on how to stay safe online.

The Cable and Telecommunications Association for Marketing (CTAM) has seen the need for educating consumers on various security topics and have created a website for sharing information on how to stay safe when streaming content on the internet (streamsafely.com). [19] The dangers of accessing television/streaming content through illegal pirate services and the risks involved in unauthorized password sharing are a few examples of the topics covered on customer education website. As the customer pattern changes and new services are offered, educating customers is essential in keeping the eco-system secure. Developing and sharing tools with industry partners, programmers, studio, and law enforcement, foster partnerships that create better protection for all connected consumers, not just service subscribers.

Figure 2 - Streamsafely.com educational articles

3.6. Industry Alliances required to combat the email Fraud

Customer communications have shifted significantly over the last decade to boost customer service, customer retention and engagement. Email communication is still heavily utilized for notifications or verification of identity during certain events for service changes. Email accounts are one the most sought-after types of credentials because they can be used as a pivot to identity theft. Some email clients on customer devices may be old and vulnerable to attacks due to flaws in email client protocols. An alliance is required between third party email providers to deprecate older email clients and encourage customers to deploy secure new clients.

Below are some actions which can be considered to secure email:

a. Disable unused 3rd party email clients
b. Modernize authentication for 3rd party email clients using oauth
c. Implement email platform anti-abuse capabilities (anti-spam, anti-malware, anti-phishing, anti-viral)

As an abundance of user interfaces are consumed for various user activities, account misuse (knowingly or unknowingly) is extremely high. This demands a defense in depth approach to all attack surfaces. The layering of intelligent defenses has shown a demonstratable impact and is recommended by various organizations. Advanced email authentication attack detection and mitigation development, partnering with vendors and dedicated teams of security professionals managing these platforms, help mature the security of email accounts.

4. Monitor Cyberhealth across your eco-system with CyberScores

The cable industry has been working hard over the last decade on security hygiene, regulatory compliance, and protecting themselves and their consumers from bad actors. Many companies have adopted one or multiple frameworks to evaluate, manage, and ultimately reduce risk over time. Some examples of frameworks that provide guidance on best practices are NIST Cybersecurity Framework, ISO 27000 series, SOC2, FISMA, and the Essential 8 Framework. Using frameworks such as these attest to mitigate 85% of cyber threats [20]. A few common best practices from these frameworks are to implement security controls around the following:

a. Patch Management
b. Application Controls for Workstations and Servers
c. Restrictive Admin Privileges
d. Backups
e. 3rd Party Risk Assessments
f. MFA

Businesses implement a variety of tools for assessing security controls based on framework adoption to better understand, manage, and reduce cybersecurity risk and help protect from various threats. These tools are used to collect metrics from possible sources of risk and present this information in dashboards. As security tools evolve, and new tools are added to the ecosystem, there are challenges around how to continuously measure the effectiveness of the security controls from these various tools and identify aggregated overall company risk.

To combine and analyze data from various security tools and frameworks, a robust risk assessment platform can be utilized to quantify and aggregate the risk and drive meaningful actions for the business.

A few products have emerged with the capability to collect data from various sources and quantify the overall performance of security controls and produce a meaningful cyber risk score dashboard or a Cyber Risk report [21] [22]. Some products also provide compliance assessment based on industry standards and even go beyond to simplify the scores by mapping them to risk ratings [23]. While the maturity of these products is in the early stages, businesses are looking for ways to develop their own centralized security metrics dashboards. According to Gartner, in the future, cybersecurity risk ratings will likely be utilized the same way credit ratings are used when assessing partners for business relationships [24]. The US department of Defense (DOD) now requires supply chain companies to report a cybersecurity maturity model compliance (CMMC) to the DOD if they provide services to them [25].

The enterprise eco-system is usually complex and the risks for each system are evaluated based on the rate of the issues and the severity. A cybersecurity rating platform enables continuous assessment of the tens of security criteria along with thousands of security checks [26]. Companies can either leverage cyber score tools or develop their own. The risks can be evaluated, assigned weightage, and aggregated for contributions to overall cyber risk score based on the company's risk appetite.

Starting small with just a few security controls and combining to one aggregated score may be a good first step toward a single platform. For example, combining the data from vulnerability assessment tools and 3rd party risk assessments into a single dashboard. Once this step is taken, team can be trained to understand the relevance of the combined have a clearer picture on the actions needed to secure their applications all in one place. As the cybersecurity rating program matures, the CyberScores can help drive the various decisions about cybersecurity risk posture from the application level all the way through to the organizational level.

5. Conclusion

The past decade has thrown security into the spotlight. It is no longer overlooked or addressed only after a damaging and expensive incident. Security is now a part of the conversation at every level of business and government. It is no longer bolted on but built into products and services with an expectation from consumers that those we trust are taking every precaution to protect us from the adversaries. Security professionals strive for the right level of protection to keep attackers out of systems and protect customers. Taking a multi-layered and proactive approach has created many successful outcomes. The industry alliances are also key in each sector to uplift the standards for common services and consumer education. The work cannot stop here. As long as there is value to be gained from their efforts, hackers will continue to create challenges. It is up to all of us to keep security part of the conversation, to learn from every attack, and grow our defenses to prevail.

Abbreviations

2FA	Two Factor Authentication
ACL	Access Control List
AI	Artificial Intelligence
AP	Access Point
API	Application Programming Interface
BOT	Robot Software Agent
bps	Bits per Second
CCPA	California passed the California Consumer Privacy Act
CI/CD	Continuous integration/continuous delivery
CMMC	Cybersecurity maturity model certification
CPNI	Customer Proprietary Network Information
CTAM	Cable and Telecommunications Association for Marketing
DOD	Department of Defense
DLP	Data Loss Prevention
FEC	Forward Error Correction
GDPR	General Data Protection Regulation
HD	High Definition
Hz	Hertz
IP	Internet Protocol
ISBE	International Society of Broadband Experts

K	Kelvin
MFA	Multi-Factor Authentication
ML	Machine Learning
NCTA	The Internet & Television Association
OTP	One Time Password
PI	Personally Identifiable
RDP	Remote Desktop Protocol
SCTE	Society of Cable Telecommunications Engineers
SDLC	Software Development Life Cycle
SEIM	Security Event Information Management
SSH	Secure Shell

Bibliography & References

[1] https://www.securelink.com/blog/81-hacking-related-breaches-leverage-compromised-credentials/
[2] "A Taxonomy of Fraud Experienced by Network Service Providers", online: https://www.nctatechnicalpapers.com/Paper/2020/2020-a-taxonomy-of-fraud-experienced-by-network-service-providers
[3] (https://www.investisdigital.com/blog/technology/why-ransomware-attacks-are-rise)
[4] *https://www.appypie.com/how-voice-technology-is-disrupting-different-industries*

[5] *https://www.idtheftcenter.org/voice-fraud-is-on-the-rise/*

[6] *https://www.darkreading.com/cloud/it-takes-an-average-38-days-to-patch-a-vulnerability*

[7] https://www.microsoft.com/en-us/security/business/zero-trust

[8] https://www.nist.gov/blogs/taking-measure/zero-trust-cybersecurity-never-trust-always-verify
[9] *https://whatis.techtarget.com/feature/History-and-evolution-of-zero-trust-security*
[10] *https://www.drivelock.com/blog/what-ingredients-does-a-zero-trust-model-consist-of*

[11] https://attack.mitre.org/
[12] *https://cheatsheetseries.owasp.org/cheatsheets/Multifactor_Authentication_Cheat_Sheet.html*
[13] *https://securityintelligence.com/posts/how-to-transform-from-devops-to-devsecops/*
[14] *https://www.iotsecurityfoundation.org/the-iot-ransomware-threat-is-more-serious-than-you-think/*
[15] *https://www.securitymagazine.com/articles/93731-infected-iot-device-numbers-grow-100-in-a-year*
[16] *https://www.justice.gov/criminal-ccips/page/file/984001/download*
[17] *https://www.radware.com/solutions/ddos-protection/*
[18] *https://business.comcast.com/enterprise/products-services/cybersecurity-services/ddos-threat-mitigation?CMP=KNC-GOOGLE&utm_source=google&utm_medium=cpc&utm_campaign=ENT_Ethernet_DDoS_BR_E_National&utm_term=comcast%20ddos%20protection-43700049663074668-VQ16-c-VQ6-398589687487-VQ15-&kw=comcast%20ddos%20protection&ad=398589687487&c=ENT_Ethernet_DDoS_BR_E_National&VQ16-c-VQ6-398589687487-e&ds_kid=43700049663074668&gclid=Cj0KCQjw7MGJBhD-ARIsAMZ0eevkouqgwHGG5OCIYzmygAnf2Eb5RORp5bTDcasJKzgoGGFL12TLDsAaApsKEALw_wcB&gclsrc=aw.ds*
[19] *https://streamsafely.com/*
[20] *https://www.cyber.gov.au/acsc/view-all-content/publications/essential-eight-maturity-model*

[21] *https://support.securityscorecard.com/hc/en-us/articles/360059534531-How-does-SecurityScorecard-implement-the-fair-and-accurate-ratings-principles-*
[22] https://www.bitsight.com/glossary/cyber-risk-report
[23] https://www.upguard.com/blog/what-are-security-ratings
[24] https://www.gartner.com/en/documents/3884271/innovation-insight-for-security-rating-services
[25] *https://www.rjo.com/wp-content/uploads/2020/11/What-DODs-Use-Of-Cyber-Scores-May-Mean-For-Contractors.pdf*
[26] *https://csrc.nist.gov/CSRC/media/Presentations/Creating-a-Cybersecurity-Scorecard/images-media/Developing%20a%20Cybersecurity%20Scorecard.pdf*

FTTx PON Architecture Considerations

Distributed Optical Taps

A Technical Paper prepared for SCTE by

Brian Yarbough
Outside Plant Engineering at
Cox Communications, Inc
6305-B Peachtree Dunwoody Rd, Atlanta, GA 30328

1. Introduction

Passive Optical Networks (PON) have come a long way in the Cox network since our initial Gigabit PON (GPON) deployments over 12 years ago. A key milestone for Cox Communications was the launch of IP Video and Telephony products via GPON in mid-2020, which enabled the elimination of Radio Frequencies over Glass (RFoG) technology and presented an opportunity to re-consider the architecture. We took this opportunity to relook at our FTTx deployments through a fresh lens and explore opportunities to improve operational efficiencies, including 10G PON evolutions, optical transport, and distribution network architectures.

In particular, the Optical Distribution Network (ODN) approaches were tailored to "right size" the cost and deployment options for the individual application, which led to different approaches for Single Family Units (SFU), Multi-Dwelling Units (MDU) and Commercial Business customers. Each had their own unique ODN architectures, which drove variation and complexity for field teams to support. There was an opportunity to harmonize those approaches with the additional benefit of making them easier to deploy and maintain.

This paper will explore a variety of architectural considerations and logic which drove decision-making around Cox Communication's next generation Fiber-to-the-X (FTTx) deployments, with a focus on a new approach to optical distribution and splitting methodologies.

1.1. History of PON at Cox Communications, Inc

Our first deployments of PON were Broadband PON (BPON) in 2004, which were deployed to a very limited extent for commercial applications. As the PON technology matured, Cox began deploying GPON in 2008, again exclusively for commercial applications. Fast forward to 2014, we repurposed the GPON platform and offered a gigabit symmetrical product to our residential customers, deploying GPON in both brownfield and greenfield applications. The key difference being the type of network carrying legacy data, video, and telephony products. Like many other cable operators at the time, we chose to leverage an RFoG technology to support these legacy products in greenfield fiber only deployments. Supporting RFoG and PON drove some unique considerations to Fiber-to-the-Home (FTTH) deployments which do not apply to PON alone, including but not limited to, a smaller optical budget and Optical Beat Interference (OBI). RFoG was intended to bridge the gap to All IP over PON, it turns out that gap lasted about 6 years and went through many evolutions until we launched all of our products over PON for residential services in 2020.

The PON portion of the FTTH network also went through a series of evolutions in the 2014 to 2020 timeframe. Initial deployments of GPON Optical Line Terminals (OLT) were rack-mounted in large environmentally controlled cabinets feeding a 1:32 split ratio. While we still deploy large OLT cabinets to a limited extent, today we primarily deploy hardened passively cooled Remote-OLTs for smaller targeted areas. The transport architecture used up until recently for the GPON Remote-OLT was a routed (layer 3) multi-hop ring solution, allowing up to 8 Remote-OLTs per ring. Recently it was decided to leverage synergies from our Distributed Access Architecture (DAA) solution used for Remote-Phy node deployments and migrate OLT transport across a homegrown Dense Wave Division Multiplexing (DWDM) solution called the Optical Communication Module Link extender (OCML). Furthermore, in an effort to position ourselves to support the ever-growing bandwidth demands, we're in the process of launching 10 gigabit symmetrical PON (XGS-PON) OLT's, capable of supporting 10G symmetrical speeds.

The GPON distribution network architecture started at a 1:32 split ratio and increased to a fixed 1:64 split ratio a couple years in to optimize OLT port consumption efficiencies. In an effort to optimize fiber and labor efficiencies, the optical splitter array varied based on the application. SFU applications used a centralized splitting architecture, while MDU and commercial applications each used different distributed splitter approaches.

With the elimination of RFoG on our roadmap, it gave us an opportunity to relook improving operational efficiencies in the FTTx ODN. In 2019, we began investigating a distributed optical tap concept, using a combination of unbalanced and balanced couplers to control optical insertion loss in a more efficient manner (see Figure 1).

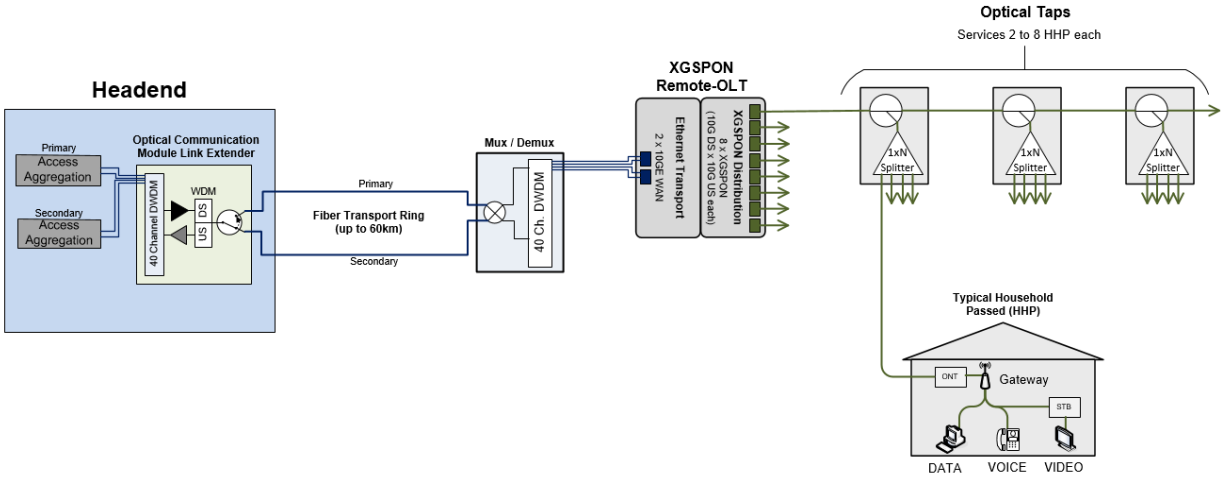


Figure 1 – Cox Communications FTTx Architecture

2. Optical Transport

Cox Communications current Optical FTTx Transport Network includes a pair of components called the Optical Communications Module Link extender (OCML) and Mux DeMux (MDM) which make up our standard DAA solution. The OCML is used to transport up to 40 DWDM wavelengths (20 channel pairs), redundantly up to 60km. The OCML is located in the headend and the MDM in the field. This same DAA solution is used by other deployments and is commonly shared with other network elements such as Remote-Phy nodes and Metro Ethernet links.

2.1. OLT Strategy

Up until this point, the vast majority of our FTTx deployments have been with GPON in alignment with ITU-T G.984.1, offering nearly 600,000 homes passed (HP) with gigabit symmetrical speed tiers today. In preparation to support 10G speeds in the future, we've recently started the process of transitioning all new PON deployments to XGS-PON exclusively in alignment with ITU-T G.9807.1.

While XGS-PON enables a path to 10G symmetrical speed tiers, long-term Next Generation PON 2 (NG-PON2) and/or Coherent PON (CPON) appear to be potential evolutions. Both NG-PON2 and CPON are still just being discussed and the picture isn't fully clear exactly how it will be operationalized, but future considerations are being made to support either option. It will be important for each PON iteration to operate at different wavelengths to allow co-existence through migration periods. NG-PON2 will operate

at 1530nm Downstream and 1600nm Upstream, but CPON wavelengths have yet to be determined. CPON is capable of speeds greater than 100G, and the calculated approach of the distributed optical tap system aligns well with the increased optical budget of the coherent optics allowing for extended optical reach and the potential for a truly passive plant without the need of remote OLT devices.

3. Optical Distribution Network

Up until recently, our PON distribution network was a conventional fixed 1:64 split ratio, which used a bank of optical splitters in an ODN cabinet. The two primary architecture types used were centralized splitters and distributed splitters. In a centralized splitter architecture, the entirety of the static split ratio is contained within the ODN cabinet. In this configuration each customer may get their own dedicated fiber spliced in parallel from the cabinet to customer premise. A distributed splitter architecture is also based on a pre-determined static split ratio, but a portion of that split ratio is distributed to a drop terminal (aka cross connect) within 400' of the customer premise (see Figure 2). For example, it may be common for an operator to distribute a 1x4 splitter near the customer and assume the first 1x16 of the total 1:64 split ratio is in the cabinet. The advantage of distributing splitters over centralized splitters is the reduction in fiber and splices required to build the network, which may result in cost savings. However, it can be wasteful because with any static split ratio it is uncommon to have exactly 64 customers to feed, so those additional ports may get stranded. Furthermore, in a static split ratio architecture, the more of the split ratio that is distributed, even more ports may be stranded. Considering MDU's typically have a higher density per demarcation Cox chose to distribute a 1x8 splitter. For commercial applications a distributed 1x4 was the right balance between splitter capacity and splicing labor considering the varying densities in commercial zones.

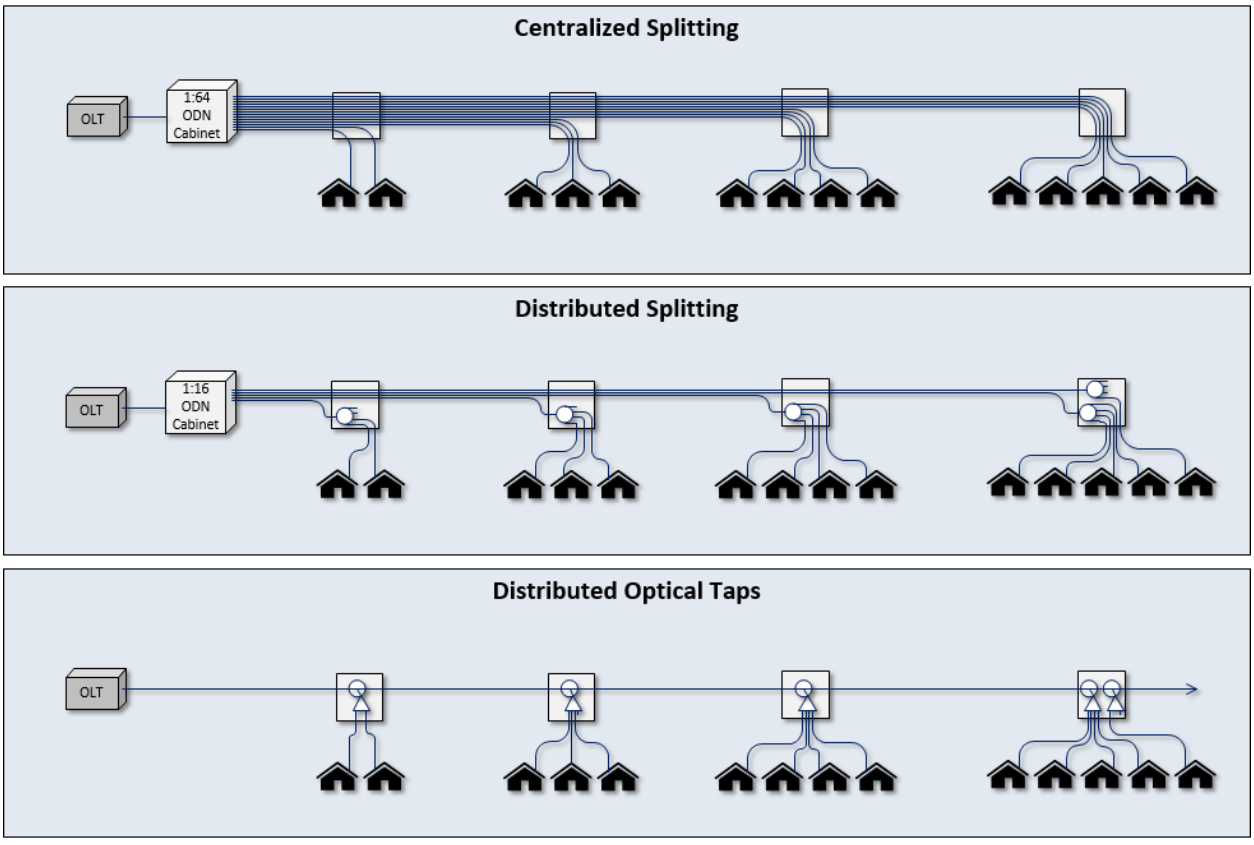


Figure 2 – FTTx ODN Architecture Concept Comparison

The insertion loss of the selected split ratio, fiber attenuation, and optical budget of each technology must be considered. Additional variables such as Co-Existence (CEx) WDM's, fusion splice and connector loss, may vary by operator, but also must be factored into optical reach calculations. Figure 3 below shows the relationship between split ratio and physical reach based on typical splitter loss characteristics. GPON assumes Class B+ optics, XGS-PON and NG-PON2 assume Class N1 optics, while the CPON specs are still being defined, but assumes worst case 35 dB optical budget operating in the C-Band. Both GPON and XGS-PON ITU standards assume a 20km physical reach limit; however, at Cox with standard 1:32 and 1:64 static split ratios, realistic operating ranges of GPON and XGS-PON are roughly 20km and 10km respectively, which include insertion loss characteristics of aforementioned variables.

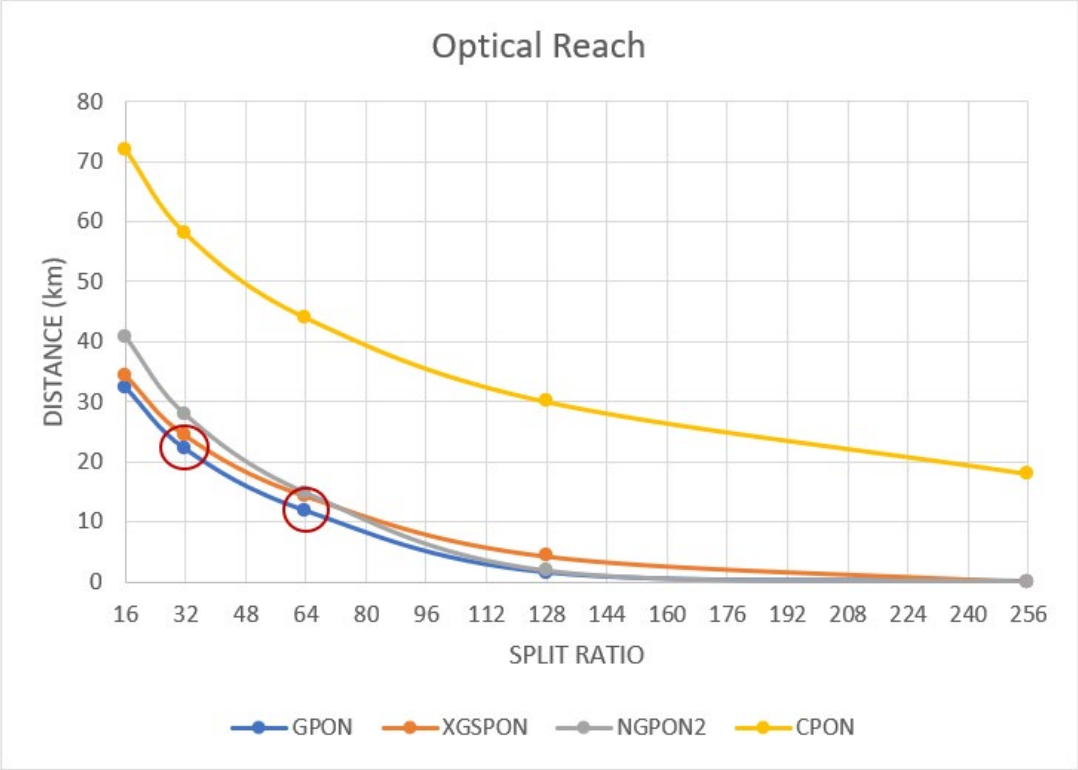


Figure 3 – Optical Reach with Static Split Ratios

3.1. Distributed Optical Tap Concept

Through the course of exploring all of these various splitting approaches it led us to a familiar concept, a tap (see Figure 4). The tap system is a controlled approach to managing signal levels to each customer throughout the network while optimizing fiber usage efficiencies and maximizing reach. A tap is characterized by a split-ratio, which is indicative of a percentage of signal received by the tap that continues through the tap to downstream devices versus a percentage of signal that is split off for creating network terminations at the customer premise.

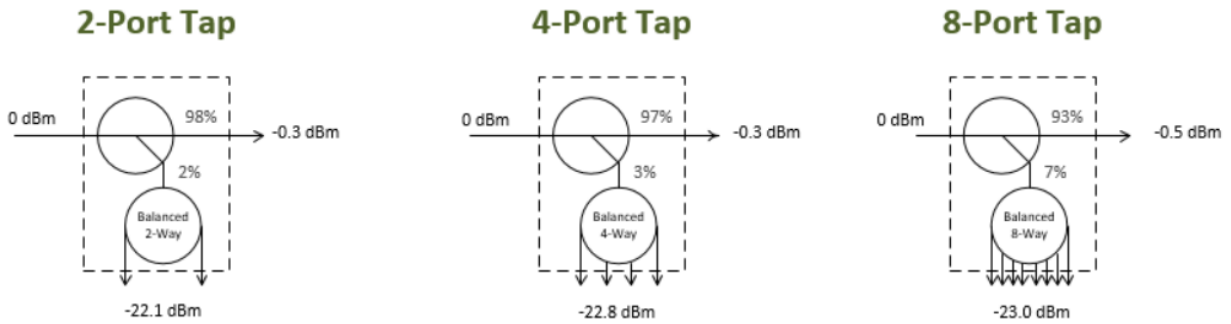


Figure 4 – Distributed Optical Tap Internal Scematics

The distributed optical tap solution takes the fiber efficiency concept a step further than a distributed splitter because much less fiber and fewer fusion splices are required than either conventional method described above. It also enables the user to control the 'size' (number of tap ports or legs) of the splitter included with the tap based on the number of legs needed and control the signal loss based on how much signal is received at a given location, and therefore is much less wasteful than the other approaches. The optical tap solution simplifies the application by pre-engineering combinations of a first stage coupler and a second stage splitter into a structured system of pre-integrated modular tap devices. A tap being "modular" refers to the fact that multiples of any tap type can be installed within any drop terminal. This contrasts with other solutions which integrate distributed splitters into a fiber enclosure, which does not allow for as much flexibility to add ports or control signal levels.

For example, if a user needs to add additional ports to a tap for new customers, they could replace a 2-port tap module with a 4-port tap module. As another example, if a technician needs more or less signal at a given location, they can simply replace the tap module with a different incremental value module. Both examples may trigger a network design change action in accordance with operational policies to maintain good record keeping, but the flexibility is feasible. In addition to the significant cost savings from minimizing fiber materials and fusion splice labor, this solution does not require a centralized cabinet for housing banks of splitters which may be expensive and challenging to get permission from municipalities to build in the right-of-way. The distributed optical tap solution also improves damage restoration time because a customer's service is dependent on fewer fibers and can be respliced quicker in the event of a damaged cable.

The distributed optical tap system intentionally mimics an HFC architecture, sharing many common principles (see Figure 5).

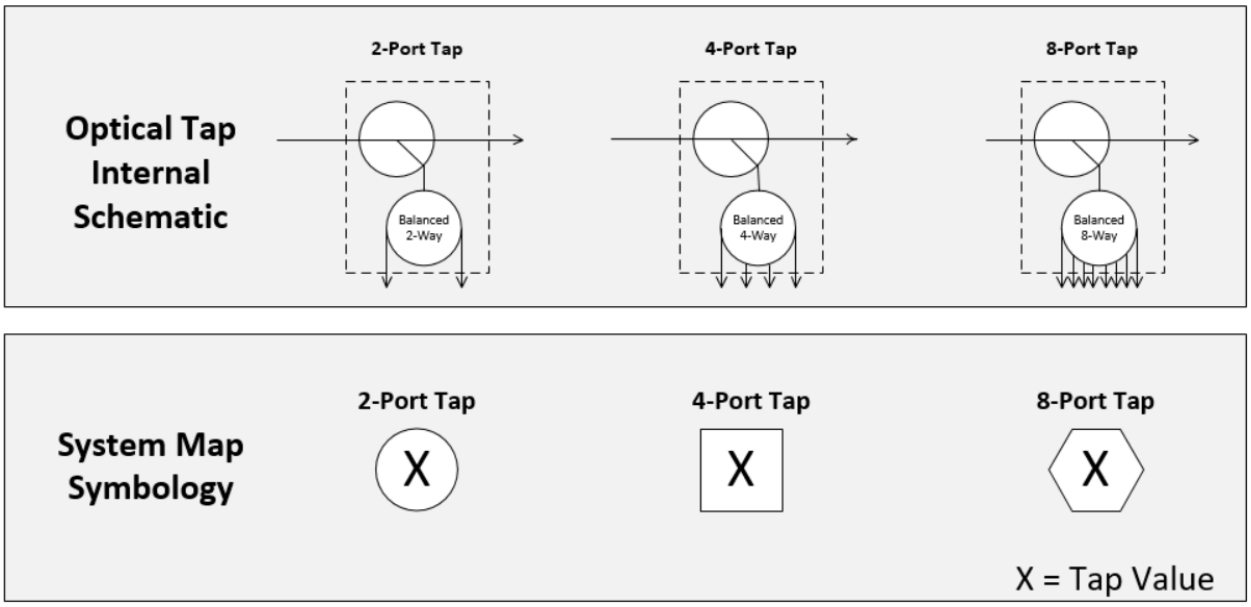


Figure 5 – Distributed Optical Tap Schematic and System Map Symbology

Optical tap devices will be installed into drop terminals (aka cross connects), which includes weather-tight fiber enclosures for pedestal, vault or aerial deployment, or PON wall-boxes for higher density MDU or commercial applications.

Additionally, more traditional optical splitters, such as a balanced splitters (2-Way & 4-Way) and asymmetrical directional couplers (DC) will be deployed into traditional splice enclosures to create additional branches as needed (see Figure 6). Similar to HFC, directional couplers will come in varying split ratios, like taps, the 'value' is representative of the down leg loss.

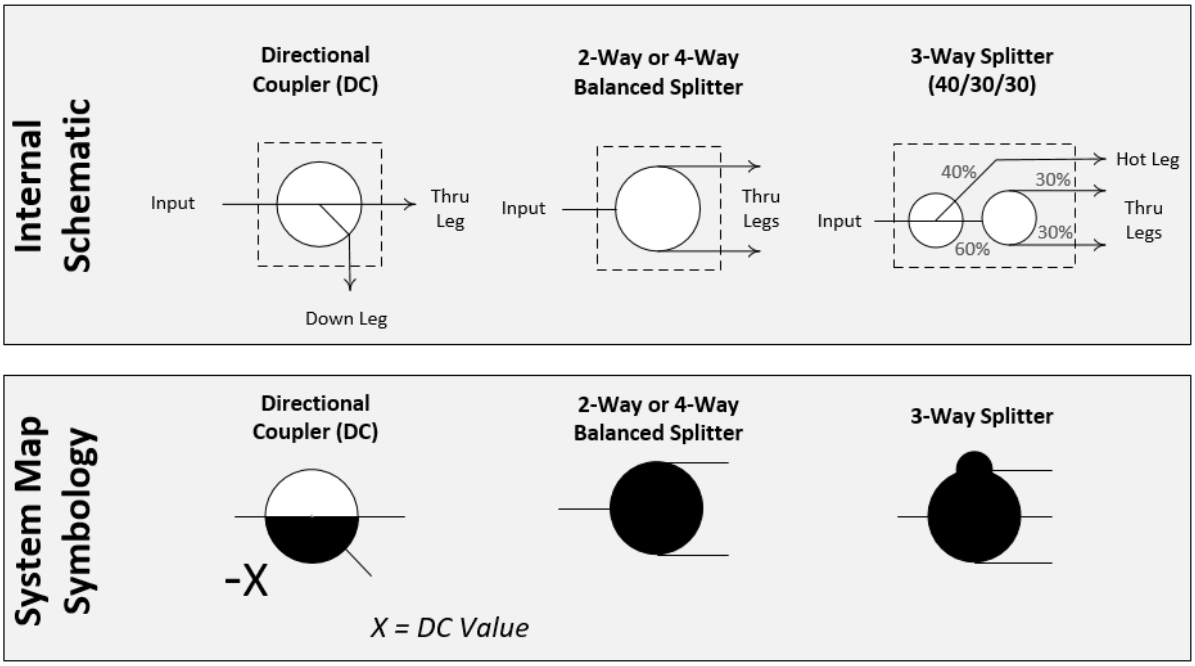


Figure 6 – Coupler/Splitter Internal Schematics and System Map Symbology

3.1.1. Optical Performance Characterization

Downstream, fiber splitter/couplers divide optical power from one common port to two or more split ports and upstream combine all optical power from the split ports to a common port (see Figure 7). 2-way optical splitter/couplers are often expressed in split ratio (SR), (e.g., 75:25).

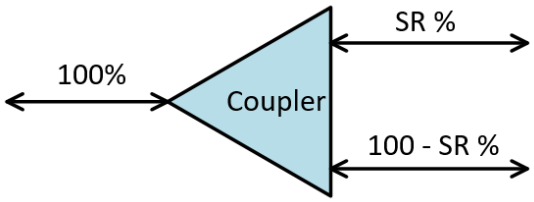


Figure 7 – Coupler Split Ratio Example

SR can be determined based on total optical power relative to the power passing through:

$$SR = \left(\frac{P_i}{P_T}\right) x\ 100$$

An industry standard formula was used to calculate insertion loss (IL):

$$IL\ = -10\ log\left(\frac{P_i}{P_T}\right)$$

Where:

IL = Splitter/coupler insertion loss for the split port, dB

P_i = Optical output power for a single port, mW or dBm

P_T = Total optical power output for all split ports, mW or dBm

Additionally, splitter/coupler excess-loss is a critical assumption that must be applied to factor in lost signal power due to imperfections in the manufacturing process and can range anywhere from 0.1 to 2.0 dB. Published insertion loss specifications from optical passive manufacturers often are expressed as "Typical" and/or "Maximum", but always factor in additional excess-loss. Based on lab testing and IL specifications from major optical passive manufacturers, we observed the higher calculated IL, the higher excess-loss variability, while the lower loss legs were more consistent. Lab testing data in Figure 8 below includes the average measured insertion loss of at least 3 samples of each type from each vendor.

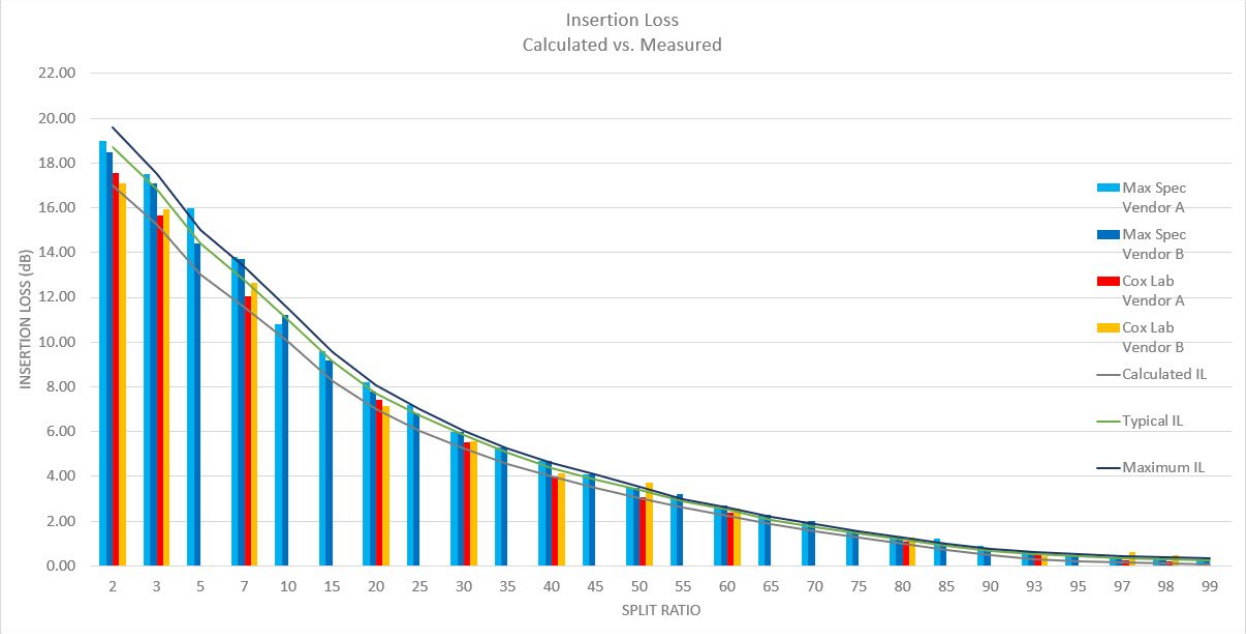


Figure 8 – Calculated Loss vs. Reported Loss

Based on these observations it was decided 'Typical' IL would include an additional 0.2 dB or 10% IL for typical excess loss. 'Maximum' IL would include an additional 0.3 dB or 15% of IL for maximum excess loss.

Both typical and maximum IL specifications are important for daily operations. It's important for field technicians to be mindful of maximum insertion loss while looking at individual events. However, if maximum loss were used for network design calculations of multiple taps spliced in series it would result in excessive amounts of cumulative loss margin. For this reason, it was decided to use typical insertion loss for network design calculations, which aligns closer to average measured loss.

3.2. Design Study

A design study was conducted to validate the optical tap concept in real-world network design scenarios against some of the aforementioned conventional methodologies. One of the many goals of the design study was to validate that we could repurpose some of the tools that we've been using for many years to design HFC networks. Additionally, we were seeking a larger sample size of network design data to help compare construction costs and network efficiencies against traditional splitting methods.

3.2.1. Design Parameters

Cox has been designing HFC networks via Lode Data Design Assistant to run RF calculations for many years. It turns out it is relatively easy to convert the core RF spec files to calculate optical levels instead. However, a significant amount of consideration was given to the core parameters that drive the design calculations to ensure a fair amount of margin for field variations without being overly conservative.

These are the key parameters and the assumptions we used (see Table 1 & Figure 9):

Table 1 – Key Network Design Parameters

Description	UOM	Parameter
Downstream OLT Transmit Power	dBm	+3.0
Downstream ONT Receive Power Minimum	dBm	-26.0
Downstream Tap Port Minimum	dBm	-23.0
Fiber Drop Loss (< 500') Maximum	dB	2.0
1490nm Attenuation (DS GPON)	dB/km	0.28
1310nm Attenuation (US GPON)	dB/km	0.34
1577nm Attenuation (DS XGSPON)	dB/km	0.25
1270nm Attenuation (US XGSPON)	dB/km	0.35
Maximum HP per OLT Port	HP	64

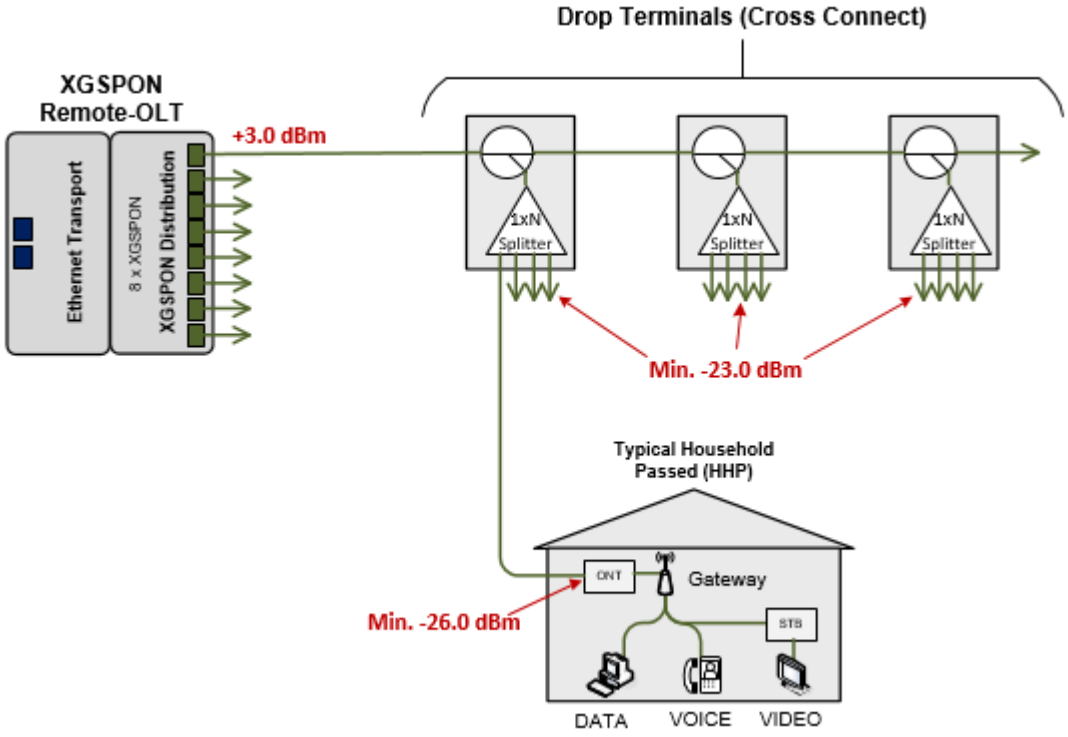


Figure 9 – ODN Diagram with Key Design Parameters

Considerations should be given to maximum versus typical, particularly when deciding what to design to, to balance performance and cost. In the case of the OLT launch power, per ITU-T G.984.1 the OLT transmit level can range from 1.5 to 5.0 dB, leading us to analyze the optical transmit power of over 17,000 OLT interfaces in our network to generate a histogram of OLT transmit power. Results found the OLT ports actually transmit at +3.63 dBm on average, with less than 1% of the population transmitting below +3.0 dBm (see Figure 10).

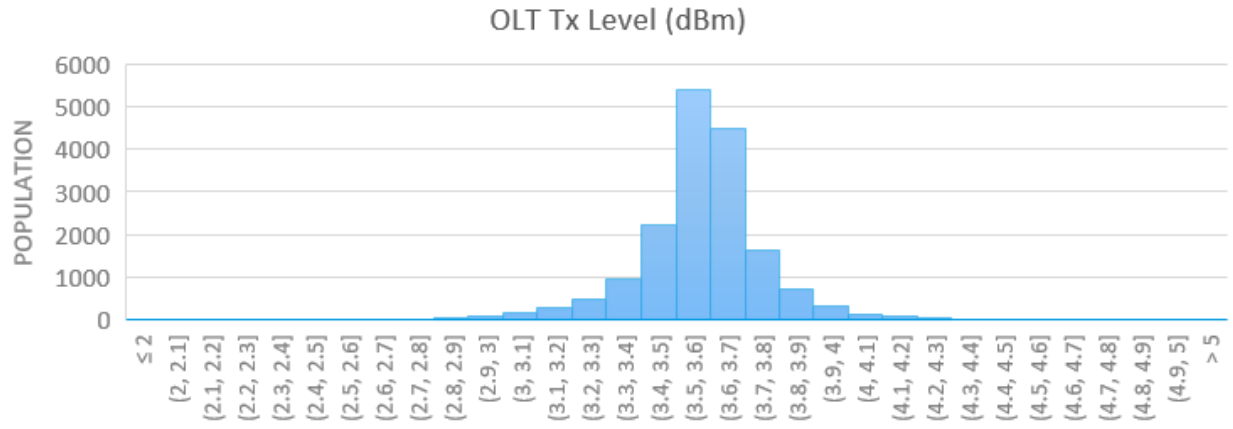


Figure 10 – OLT GPON Interface Population - Downstream Launch Power

Downstream tap port minimum (GPON & XGS-PON), for the sake of design calculations, was set at -23 dBm, but we allow -24 dBm acceptance criteria for field activations. This combined with OLT launch power margin provides 1 – 2 dB of margin for additional field variables, such as maximum insertion loss conditions, fusion splice loss, connector loss, etc. For the last 100' – 500', the drop network is given 2 dB of loss for the connectorized fiber drop from the tap port to the optical network terminal (ONT).

Also, considering this network type doesn't naturally control the number of HP per OLT port like a traditional fixed split ratio, we opted to implement a 64 HP limit policy which will be enforced at the individual design level. While there certainly is a healthy amount of optical budget in our current GPON and XGS-PON deployments to support many more than 64 HP, the spirit behind this policy was to mitigate potential network contention in the future. As technology and bandwidth demand changes this policy may be revisited.

3.2.2. Design Analysis

The actual design study included 5 SFU properties totaling 920 HP across 4 different markets, as well as 4 MDU properties totaling 799 HP across 4 different markets. Many data points were collected, but the two primary metrics being evaluated were fusion splicing requirements and OLT port optimization, because they have the most bearing on cost efficiency.

The SFU portion of the study considered standard centralized splits and distributed 1x4 or 1x8 splitters at a fixed 1:64 split ratio compared to the distributed optical tap concept (See Table 2).

Table 2 – SFU Design Study Results

SFU Properties		Centralized Splitters (Current)			Distributed Splitters			Distributed Optical Taps		
Market	HP	OLT Ports Used	HP per OLT port	Fusion Splices	OLT Ports Used	HP per OLT port	Fusion Splices	OLT Ports Used	HP per OLT port	Fusion Splices
Site 1	405	7	58	892	9	45	178	7	58	142
Site 2	85	2	43	178	2	43	39	2	43	29
Site 3	55	1	55	120	1	55	20	1	55	25
Site 4	240	4	60	526	5	48	86	4	60	91
Site 5	135	3	45	280	4	34	58	3	45	58
Averages:	***184***	***3.4***	***52***	***399***	***4.2***	***45***	***76***	***3.4***	***52***	***69***

When compared to centralized splitting in an SFU application, there was an 83% reduction in fusion splices required which is a significant labor cost driver. Furthermore, a centralized split is often considered the most efficient in regard to HP per OLT ports, but with distributed optical taps there was no change in OLT port efficiency. The distributed splitter model maintained the traditional fixed 1:64 split ratio but allowed the distributed splitter in the field to be either a 1x4 or 1x8 based on the number of passings at a given location. When looking at the distributed splitter solution versus centralized splitters it resulted in an 81% reduction in fusion splices, but costly OLT port usage requirements increased due to stranded capacity in the cross connect device at the curb.

For the MDU portion of the study, distributed 1x8 splitters at a fixed 1:64 split ratio were compared to the distributed optical tap concept. Centralized splits were not considered because previous cost modeling exercises in years past had already considered it and led to standardization of a distributed 1x8 for MDUs (see Table 3).

Table 3 – MDU Design Study Results

MDU Properties		Distributed 1x8 Splitter (Current)			Distributed Optical Taps		
Market	HP	OLT Ports Used	HP per OLT port	Fusion Splices	OLT Ports Used	HP per OLT port	Fusion Splices
Site 6	204	4	51	62	4	51	45
Site 7	287	6	48	106	5	57	75
Site 8	190	5	38	139	3	63	114
Site 9	118	3	39	48	2	59	31
Averages:	***200***	***4.5***	***44***	***89***	***3.5***	***58***	***66***

Distributed optical taps averaged 32% more efficient OLT port usage than our current distributed 1x8 splitter solution, which can be attributed to the flexibility allowed by the taps to choose the appropriate splitter size at each demarcation box. Furthermore, there was a 26% reduction in fusion splices required per property.

3.2.3. Cost Modeling

In addition to splicing and port efficiency improvements shown above, the ODN Cabinet was another important factor considered in the cost model. The distributed optical tap architecture solution assumes the ODN cabinet may be eliminated considering splitting will be distributed to the tap locations instead of a centralized bank of splitters in a cabinet or enclosure.

With the splicing and port efficiency improvements shown above, coupled with additional material and labor savings not directly tied to these drivers, we estimate approximately $65 and $40 per HP savings for SFU and MDU respectively, relative to our current architecture solution.

3.3. Product Development

As a cable operator with large groups of technical resources who are trained and educated on how to operationalize HFC, the goal of product development was to make it look and feel as much like HFC as possible. Small things like product naming and labeling were intentionally created to blend HFC terminology with fiber products for ease of knowledge transfer.

The physical form-factor of the devices was also carefully considered; it was important that the optical tap device was decoupled from the enclosure that it lives in. This provides the flexibility to deploy taps into any generic fiber terminal and deploy multiples of any combination of tap devices within the terminal based on the need of the given location. Generally speaking, our preference is fusion splicing whenever

possible, because hard splices are inherently more reliable than mated connectors. For this reason, we decided the standard tap modules would have standard 250µm bare fiber leads on the input and through-legs to allow fusion splicing the cascades of taps in series, while the tap legs would be connectorized for drop connections. Additionally, we specified a connectorized cassette with bulkhead connectors on all ports. Cassettes will primarily be used for wall-mount applications, where long strings of multiple taps are not spliced together in series.

3.3.1. Product Specifications

Tap modules and cassettes contain two integrated balanced and/or asymmetrical optical couplers designed to a series of fixed insertion loss values (See Figure 11) just as RF taps are designed today.

Definitions:

1st Stage Coupler: Either balanced or asymmetrical two-way optical coupler, with one of the output legs fusion spliced to the input of a second stage optical splitter.

2nd Stage Splitter: Balanced 2-way, 4-way, or 8-way optical splitter.

Input Leg: The input fiber of the first stage coupler will feed out of the tap module.

Thru Leg: The second output fiber of the first stage coupler will feed out of the tap module.

Tap Legs: Output fibers of the second stage splitter will feed out of the tap module.

Thru Loss: The insertion loss from input to thru leg.

Tap Loss: The combined insertion loss from input through drop legs

Tap Value: Numeric identifier indicating the loss structure of the tap, which is closely related to the tap loss.

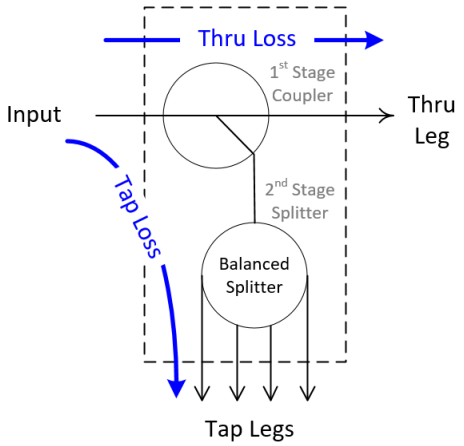


Figure 11 – Distributed Optical Tap Schematic

Insertion loss characteristics are expressed in both typical and maximum insertion loss characteristics of each tap variation across entirety of the passband (1260 – 1650 nm), not including connectors. Tap Loss Uniformity must be ≤ 1 dB.

Taps are available in 24 different size/value combinations and will be labeled with the following symbology (see Figure 12). Taps will be available in two form-factors: Outside Plant (OSP) Modules (see Figure 13) and Inside Plant (ISP) Cassettes (see Figure 14):

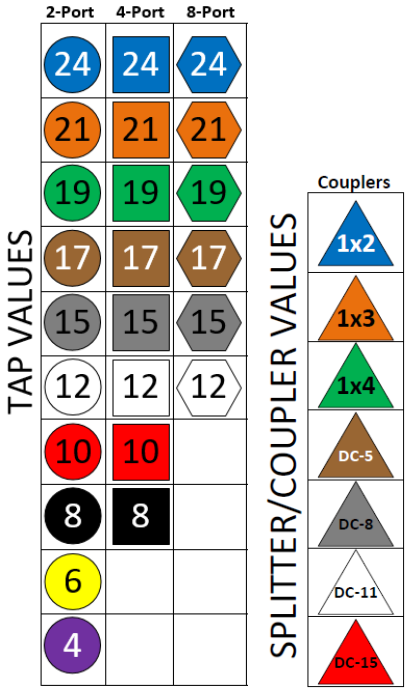


Figure 12 – Device Naming and Labeling

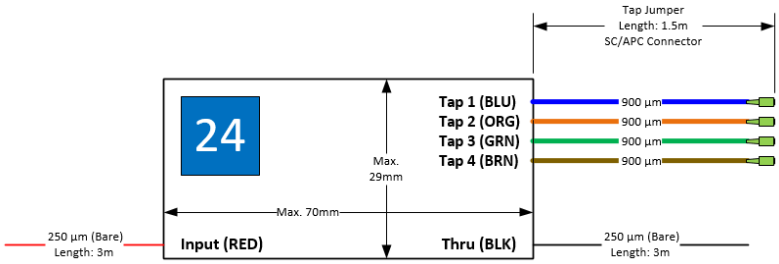


Figure 13 – Example of Tap Module Form-Factor

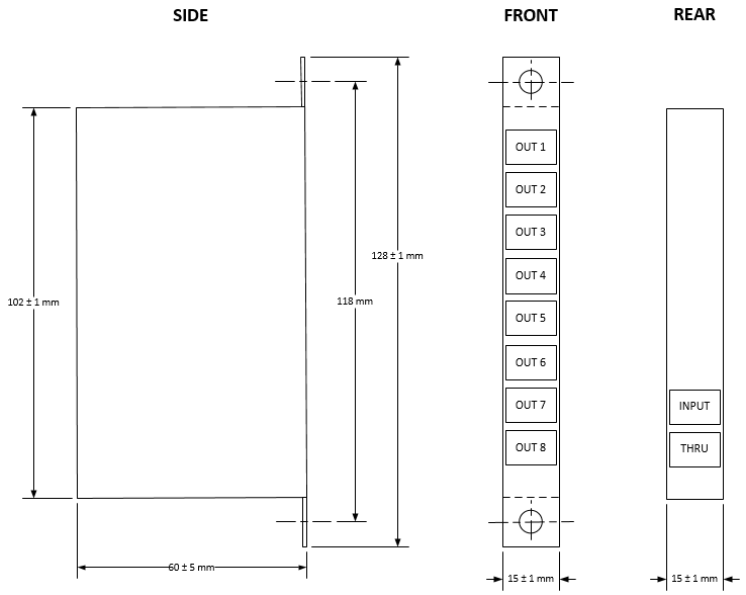


Figure 14 – Example of Tap Cassette Form-Factor

3.4. Field Trials

Prior to any customer facing deployments, full proof-of-concept networks were built in the lab with strings of couplers/splitters spliced together operating both GPON and XGS-PON networks successfully. As well as final product testing from multiple vendors, and mechanical form/fit/function testing within our standard fiber enclosures. Design tools such as our GIS database also needed further development to prepare for modeling of this type of network in digestible manner for field technicians. New object types were created, and fiber splicing documents were modified to support this type of network.

A reliability study was also conducted, which resulted favorably for a fusion-spliced distributed optical tap network with improvements in mean availability time, Mean Time Between Failure (MTBF), Mean Time to Restore (MTTR), and a reduction in annual maintenance truck rolls. While more customers are dependent on a single strand fiber in this type of network, there are fewer points of failure and in the event of a fiber impairment, services can be restored quicker.

The final step before full scale deployments was to gain some production experience to ensure processes are aligned and to identify any potential operational gaps. We conducted 5 field trials in 5 different Cox markets to demonstrate performance for the selected applications: greenfield, brownfield, SFU, MDU and commercial applications.

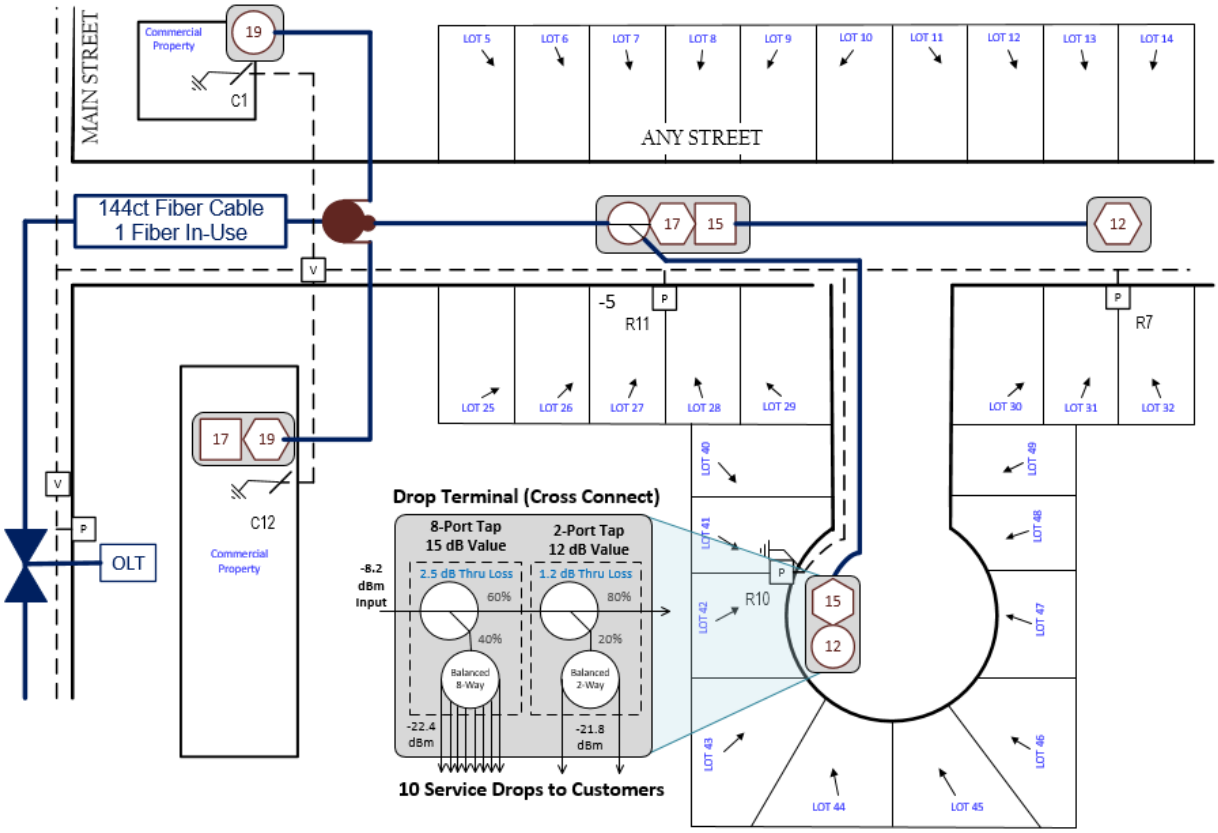


Figure 15 – Example of Field Trial Network Design

3.4.1. Lessons Learned

We successfully built and activated 5 different FTTx networks with the distributed optical tap solution and compared calculated design level versus actual measured levels. At the time the data was analyzed, levels had been collected from only 72 optical tap locations (see Figure 16).

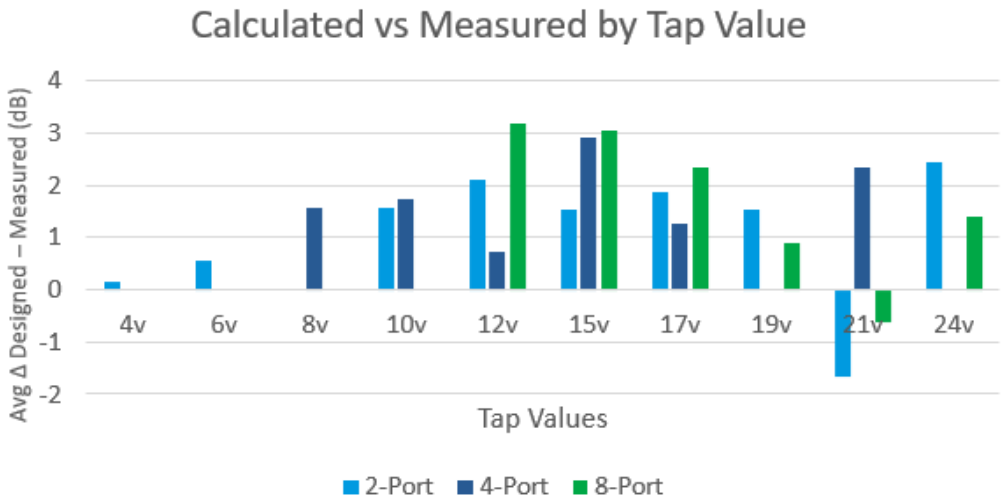


Figure 16 – Example of Field Trial Network Design

In most cases, measured levels were 1 - 3 dB better than calculated with few exceptions, which aligns with the built-in margin. In some cases, measured level was within the field acceptance criteria (-24 dBm), but lower than calculated, upon further investigation these were attributed to higher than typical splices but deemed acceptable. Considering all the variables involved and built-in margin we were comfortable with the results.

Be mindful of Optical Time Domain Reflectometer (OTDR) capabilities and limitations when designing turn up and troubleshooting processes. Characteristics of this type of network make obtaining OTDR shots particularly challenging due to relatively short distance links (< 2 km) with high insertion loss. We have observed good results when shooting 'in-line' into Thru Legs of taps but had very limited success shooting an OTDR into a tap leg. Tap legs, in particular of tap values greater than 17 dB, require a pulse width setting of at least 500ns, which decreases your resolution and creates lengthy dead-zones on the opposite side of the devices under test. This may be good enough to establish continuity back between two points, but events in between may not be visible. In contrast, in-line shots allow you to reduce pulse-width settings resulting in shortened dead-zones and with better resolution for the entirety of the link. We learned the addition of a connector on the last tap of the end-of-line circuit as an 'in-line' access point specifically for troubleshooting is valuable.

In the examples below (see Figure 17 & 18), OTDR shots are taken from the same 2-Port 12 dB value (2p12v) tap location into the tap leg versus the Thru leg; notice there is an additional ~10 dB of loss 400' from the test location which becomes much more difficult to discern with higher pulse width settings.

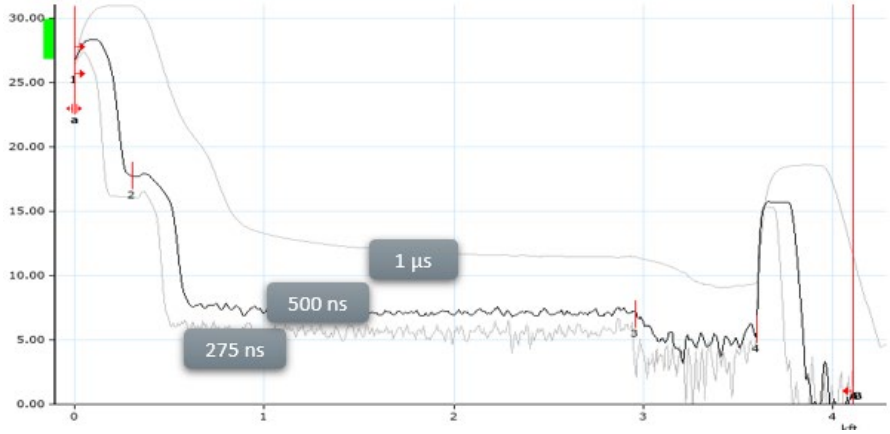


Figure 17 – Effects of Pulse Width Variation into Tap Leg of 12 dB Value Tap

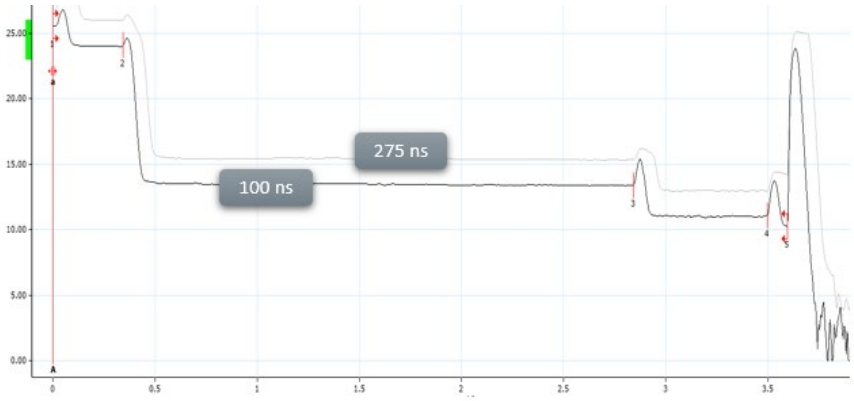


Figure 18 – Effects of Pulse Width Variation into Thru Leg of 12 dB Value Tap

Test equipment vendors are actively working to improve software profiles to make them more intuitive at identifying asymmetrical couplers and faults. In the interim, it takes a little more skill to read a traditional OTDR trace result and distinguish acceptable events from problematic events. We've found that the measured event loss of an OTDR can be very accurate (± 0.05 dB), the challenge is understanding the difference between a 'good' and 'bad' event. Historically any event greater than 0.5 dB was considered bad, in this type of network 5 dB of loss at one location could be 'bad' while the next is acceptable.

Figure 19 below is of a network with low levels (-27 dBm) at two locations, an OTDR was shot from end-of-line back toward the OLT. The technician must have a comprehensive view of the network design to compare anticipated loss to actual loss; in this case, the issue found was a bad splice on the input of the DC-8, which the anticipated loss was 1.5 dB, but the actual loss was ~7 dB.

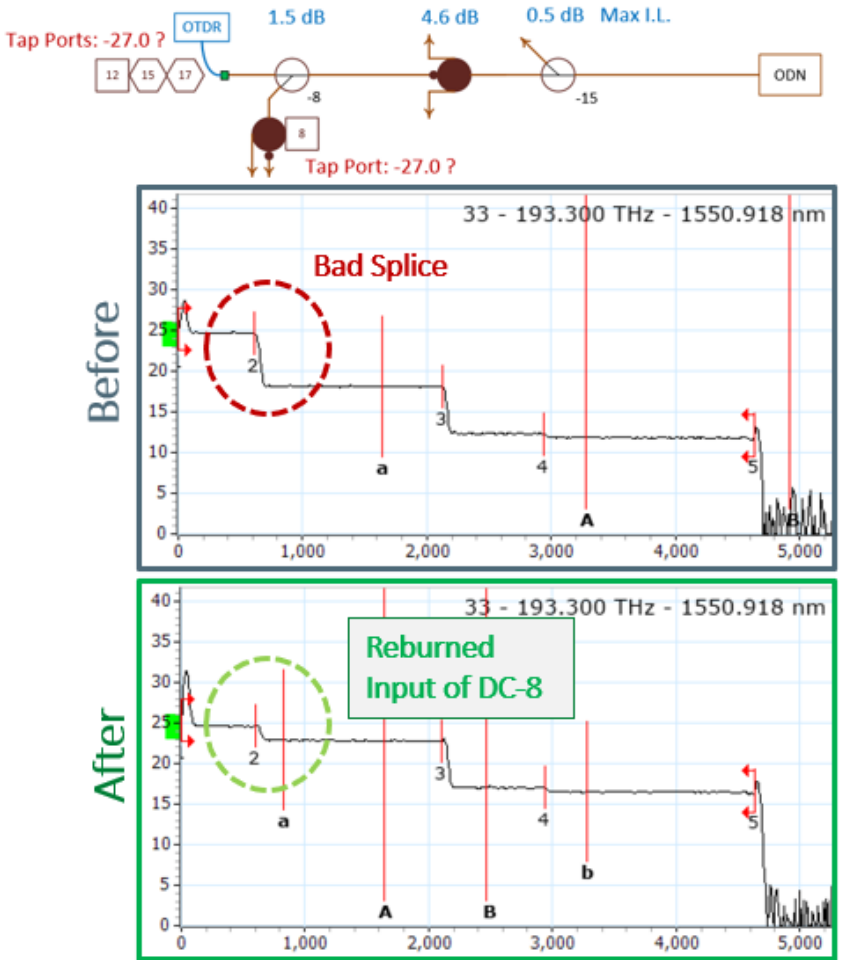


Figure 19 – OTDR Troubleshooting Example

4. Conclusion

Having completed lab testing, field trials and the early phases of production rollout, we feel confident in the distributed optical tap system serving as the standard optical distribution network for our PON deployments now and into the foreseeable future. Operationally, it fits well into many of our technicians existing skill sets and legacy RF design tools can be repurposed to model optical networks instead. This network is cost efficient, and more closely aligns network capacity with demand.

For network design parameters excess-loss must be factored in, but typical (as opposed to maximum) optical splitter insertion loss specifications from vendors is sufficient. Additional margin may be strategically applied to OLT launch power and/or drop loss instead.

An OTDR is a valuable tool for troubleshooting but must be performed in-line as opposed to a tap port. Consider fusion splicing versus optical connectors; well-placed connectors are valuable access points for troubleshooting purposes but can reduce optical reach and become future points of failure.

Abbreviations

BPON	Broadband Passive Optical Network
CPON	Coherent Passive Optical Network
DAA	Distributed Access Architecture
DC	Directional Couplers
DWDM	Dense Wave Division Multiplexing
FTTH	Fiber-to-the-Home
FTTx	Fiber-to-the-X
Gbps	Gigabit per second
GPON	Gigabit Passive Optical Network
HFC	Hybrid Fiber Coax
HP	Homes Passed
IL	Insertion Loss
ISP	Inside Plant
MDM	Mux DeMux
MDU	Multi-Dwelling Units
NGPON2	Next Generation Passive Optical Network 2
OBI	Optical Beat Interference
OCML	Optical Communication Module Link extender
ODN	Optical Distribution Network
OLT	Optical Line Terminal
ONT	Optical Network Terminal
ONU	Optical Network Unit
OSP	Outside Plant
OTDR	Optical Time Domain Reflectometer
PON	Passive Optical Network
RFOG	Radio Frequencies over Glass
SCTE	Society of Cable Telecommunications Engineers
SFU	Single Family Units
SR	Split Ratio
WDM	Wave Division Multiplexing
XGSPON	10 Gigabit Symmetrical Passive Optical Network

Bibliography & References

ITU-T G.984.1-200803: *Series G: Transmission Systems and Media, Digital Systems and Networks; Gigabit Passive Optical Networks (GPON): General characteristics; Telecommunication Standardization Sector of ITU*

ITU-T G.9807.1-201606: *Series G: Transmission Systems and Media, Digital Systems and Networks; 10-Gigabit-capable symmetric passive optical networks (XGS-PON); Telecommunication Standardization Sector of ITU*

ITU-T G.989.2-201902: *Series G: Transmission Systems and Media, Digital Systems and Networks; 40-Gigabit-capable symmetric passive optical networks 2 (NGPON2): Physical media dependent (PMD) layer specification; Telecommunication Standardization Sector of ITU*

Greenfield Mobile Network Considerations

Converged Networks and Mobility

A Technical Paper prepared for SCTE by

Ravi Guntupalli
Director of Technology, Mass-Scale Infrastructure Group
Cisco Systems, Inc.
raguntup@cisco.com

Ibrahim Ayad
Engineering Architect, Mass-Scale Infrastructure Group
Cisco Systems, Inc.
iayad@cisco.com

Irfan Ali
Principal Engineer, Mass-Scale Infrastructure Group
Cisco Systems, Inc.
irfaali@cisco.com

1. Introduction

Cable operators continue to expand the service offerings towards the end-users with wireless and cellular capabilities. Some of the operators already leverage Mobile Virtual Network Operators (MVNO) relationships with existing nationwide Mobile Network Operators (MNO) partners to offer mobile services on the macro cellular networks. This enables the Cable Operator to be able to offer a unified experience for the end users and able to tap into additional revenue by offering differentiated services.

However, depending on the markets and needs, the availability of spectrum or regulatory requirements Cable operators are considering deployment of new 4G and 5G networks from the ground up, which are a Greenfield network. Depending on the timing of the market in the region and the device ecosystem availability within the region, it may be possible that the Cable Operator would have to deploy a 4G radio but given the timing of the industry, we anticipate more 5G network deployments to ensure investments are directed towards the network of the future. In some cases, the operator may choose to leverage a 4G+5G Radio access network and terminating on a 4G Core. This architecture – widely known as 5G NSA would allow for the cable operators to leverage existing device ecosystem but also partially build towards the network of the future. However, more and more MSOs seem to be evaluating a 5G Standalone deployments and this paper tries to focus primarily on the 5G SA deployment models but also addresses ability to handle specific scenarios of 4G+5G deployments. In this paper, we consider any 3GPP standards based wireless network that is being built up for the first time by a provider without having the dependencies of legacy 3G network interworking or device support related to 2G/3G.

Additionally, we use the terminology Multiple Service /System Operator (MSO) to refer to cable operators who have decided to deploy Greenfield mobile networks. By contrast, Mobile Network Operators are the traditional mobile operators like AT&T, Verizon, T-Mobile, Vodafone, etc.

While architectures for deployment based on 3rd Generation Partnership Project (3GPP), Cable Labs and other industry standard development organizations (SDOs) are widely available and considered before finalizing deployment, it is also critical to consider some of the lessons learned by wireless network providers over the years ramping up the network capabilities.

Based on some lessons learnt in deploying 4G networks globally as well proven best practices in the industry for macro and microcellular networks, this document addresses and captures multiple key challenges to anticipate and plan for both the architecture and operational models. Given that 5G SA architecture leverages new capabilities like Service Based Architecture and some of the challenges related to 4G networks based on point-to-point interfaces using Diameter and GTP-C interfaces may not always be applicable, the authors believe that initial set of deployments will have some form of point-to-point capabilities – though leveraging new protocols defined in 3GPP. Hence, some of the lessons learnt in 4G networks would absolutely be applicable.

2. 5G – Architecture and Capabilties

There is significant amount of collateral and information related to 5G and the new capabilities that 5G network architectures bring to the table. Rather than reiterating the capabilities of 5G and cover the architecture options in detail, this section provides a very high-level view of 5G, and the nomenclatures used in the rest of the document.

Beyond just being an advancement in mobile generation – 5G was expected to lay a path for new capabilities and services that are yet to come. The underlying premise for the architecture enhancements was to enable capabilities that would allow Service Providers leveraging 5G to break through some of the constraints of legacy internet connectivity paradigms and enable a new set of differentiated services.

Within the scope of a 5G Network, we believe that the network must be flexible enough to deliver a wide range of new services but also be able to address existing legacy applications. Given the complexity of owning a brand new 5G network, the network needs to account for ease of deployment but also for ease of day-to-day operations and management.

The figure below takes a high-level representation of the various domains of an end-to-end 5G network.

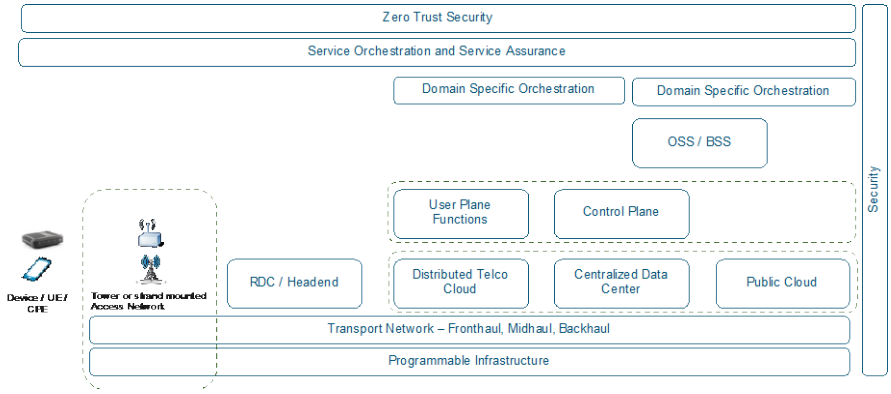


Figure 1: Different Domains of a typical 5G network

At a minimum, the following domains must be individually understood and designed for a successful deployment of 5G network.

- Radio Access Network
- Transport Network
- Packet Core
- Automation and Assurance
- Compute and Infrastructure
- Security

While this document does not go into each of these individual domains, care must be taken to ensure each area is tuned and enhanced to account for 5G network deployments for the specific MSO environment. For subsequent sections, capabilities and unique learnings and architecture options around packet core network are detailed.

A simplified representation of a 5G network architecture is shown below with the key capabilities described.

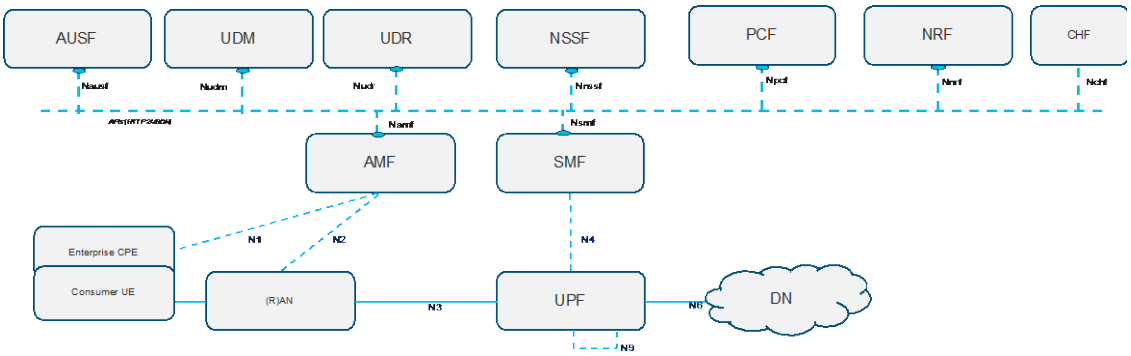


Figure 2: Simplified 5G network architecture with packet core focus

- **Consumer UE / Enterprise CPE**: The end user device that is connecting via the radio access network or in case of MSO environment, via the existing cable network back into the 5G Core network.
- **Radio Access Network (RAN)**: In the case of 5G wireless access technology, the Access Network typically consists of a 5G gNodeB which performs Radio resource management, radio bearer control, scheduling of packets over the air interface among other access management functionality. In MSO deployment scenario, based on 3GPP Release 16 or later specifications, Access Network could be existing access capabilities like Cable plant or WiFi technologies terminating via appropriate trusted or untrusted access.
- **Access and mobility management function (AMF)**: AMF is responsible for Registration management, access control and mobility management function for all accesses (incl. WLAN), Terminates NAS signaling for all accesses (single AMF per UE). AMF receives mobility related policies from PCF (e.g., mobility restrictions) and forwards mobility related policies to the UE (via N1)
- **Session management function (SMF)**: With 5G SA architecture, SMF provides a mechanism to deploy a common session management for all accesses (incl. WLAN) and wireline capabilities. SMF handles all session management signaling with UE (relayed by AMF), controls the user plane functions. SMF interfaces directly with UDM to receive subscription information (no need to go via AMF) along with policy information from PCF.
- **User plane function (UPF)**: UPF supports set of operations (forwarding to other functions, encapsulation/decapsulation, bitrate enforcement, application detection etc.) and acts as a primary anchor point of IP packets in the network. Within 5G SA architecture, SMF dynamically configures UPFs (activates/configures subset of the operations defined above) to provide the traffic handling functionality needed for a session. In addition, one or

multiple chained UP functions can be activated and configured by SMF per session as needed for a scenario. Within the MSO network architecture, UPF can also enable convergence capabilities by unifying access gateway capabilities when leveraging 3GPP Release 16 and additional SDO specifications.

- Policy control function (PCF): The PCF Provides QoS and charging rules to Session Management Function and interfaces with external Application function (e.g., IMS) when applicable. PCF also provides mobility related policies directly to AMF (e.g., mobility restrictions for stationary devices (FWA) and optionally provide policies to the UE e.g., on network discovery/selection. Additional support for network discovery/selection policies, while defined in 3GPP require UE/CPE capabilities to be deployed in the network.
- Unified Data Management (UDM): UDM uses subscription data (including authentication data) that may be stored in UDR and is responsible for generation of 3GPP AKA Authentication Credentials, identification Handling and Access authorization based on subscription data.
- Authentication Server function (AUSF): AUSF Supports authentication for 3GPP access and untrusted non-3GPP access and supports recover from synchronization failure in certain cases.
- Unified Data Repository (UDR): UDR carries the subscription data and offers services for UDM, PCF etc., to allow for retrieval of appropriate user data to be used for allowing the device onto the network and applying the corresponding policy. In MSO deployments, UDR could be a mechanism to unify the wireless and wireline subscription data in the future.
- Network Slice Selection Function (NSSF): Network Slice allows for a self-contained, logical portion of an E2E network resources within a service provider but to ensure the UE/CPE is allowed access to the appropriate slices defined, NSSF was introduced as a new capability within the 5G SA Architecture and helps selecting the set of Network Slice instances for the specific devices.
- Network Repository Function (NRF): With the introduction of Service Based Architecture, Service Providers have the ability of the steering away from somewhat of a point-to-point network architecture in 4G networks. NRF enables service registration from NFs and acts as a service discovery function – either for enable direct or indirect communication between various functions defined above.
- Charging Function (CHF): The charging function is responsible for generating charging data records (CDRs), based on usage information obtained from the UPF and SMF. The CHF interfaces to the operator's billing system.

Note that while this section captured some of the high-level capabilities of the critical network functions in the 5G architecture, 3GPP TS 23.501, Section 6.1 carries a comprehensive list of capabilities for the complete set of the network functions. Full set of functions as defined by 3GPP are listed in the Appendix A.

3. What could 5G be for MSO's

Cable operators get presented with more opportunities every day in different countries to expand the footprint and offerings to the end users by enhancing the service capabilities and enable new set of customer use cases and experiences. With more spectrum options being made available by regulatory bodies in different countries – example CBRS and C-band in US, Canada's 3.5 GHz auction and upcoming auctions in India as well as other countries - the ability for Cable operators to enhance their network to offer mobility solutions to their customers is more viable than ever.

Especially with 5G architecture capabilities defined in 3GPP specially to address the wireline and wireless convergence (5WWC Work Item in Release 16 of 3GPP specifications[1])and additional capabilities being addressed by Cablelabs[2] Cable operators are certainly closer to achieving the goal of becoming MSOs. Cable operators' interest for foraying into the mobility domain is being driven by a multitude of factors – some country specific and some globally applicable.

Offering a mobile service to the users could:

- Significantly increase revenues – new wireless service to existing customers or new standalone wireless customers
- Enhance customer experience by offering seamless service experience – extending the home experience to wider area
- Offer differentiated services that could not be offered on wireline alone – broadband services including voice

Depending on the country, regulatory requirements, and competitive positioning – Cable operators are typically presented with multiple options.

1. Tight partnership with an existing Mobile network provider
2. MVNO network agreements with one or more existing Mobile network provider
3. Build a standalone network with nationwide coverage
4. Build pockets of wireless network with MVNO or partnership with Mobile network provider for enhanced coverage by leveraging small cells or micro coverage

When it comes to a mobility network built from the ground up, some Cable operators have traditionally relied on unlicensed spectrum to offer mobility for the end users, leveraging capabilities of WiFi which is being enhanced every day with new capabilities as well. WiFi6 will continue to be a critical part of the MSO architecture it is critical to understand the capabilities of 3GPP to deploy 5G networks in the licensed bands as well as shared access and unlicensed access being leverage would be available. In addition to network capabilities, enhanced device capabilities also offer new potential options for MSOs that have traditionally not been present. As an example, being able to leverage eSIM capabilities will enable MSOs to migrate subscribers faster and more seamlessly onto the new MSO networks.

[1] https://portal.3gpp.org/desktopmodules/WorkItem/WorkItemDetails.aspx?workitemId=830050
[2] https://www.cablelabs.com/specifications/WR-TR-5WWC-ARCH

While some MSOs could have common requirements and the end use cases being similar, as captured in Cable and Mobile Convergence, A Vision from the Cable Communities Around the World[3] and Section 4, "A Survey of Mobile Deployment Plans by MSOs Around the World", the deployment models and the architecture options being considered vary per country and per MSO.

3GPP network architecture and the spectrum options could be classified into the following broad set of deployment architectures:

- Consumer Fixed Wireless Access
- Consumer Enhanced Mobile Broadband Access
- Enterprise focused wireless access – Private Networks or a Private access via public access

Deployment of RAN assets could vary as well – including but not limited to strand mount access nodes, small cells, dedicated base stations or shared resources access nodes. We believe there are 3 main architecture deployment models as listed below when the MSO chooses to deploy a wireless network based on 3GPP standards.

3.1. MNO Partnership

This is a scenario where the nationwide coverage is provided by the MSO with a tight partnership with an existing MNO. The MSO can build and deploy a selected set of markets or regions in the country to offload data via MSO network – which reduces cost of delivering the service and / or offers a more differentiated service compared to the MNO network. This can be achieved via roaming relationship with a single identity or DSDS capabilities[4].

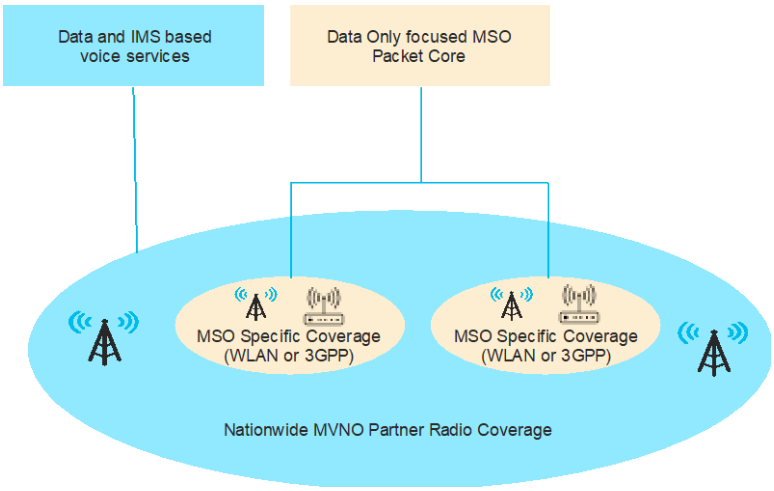


Figure 3: MNO Partnership for nationwide coverage with MSO pockets of coverage

[3] https://www.nctatechnicalpapers.com/Paper/2020/2020-cable-and-mobile-convergence
[4] DSDS details depicted in "Cable and Mobile Convergence, A Vision from the Cable Communities Around the World" section 4.1.5

3.2. Nationwide MSO

Ability to offer a nationwide coverage provided by the MSO on Day 1 leveraging existing WLAN assets and deploying new 3GPP RAN where applicable and available based on regulatory requirements. These scenarios would require the MSO offer full voice capabilities in the network and deliver regulatory compliance like Emergency calling from Day 1. In lieu of a tight MNO partnership, the MSO can leverage roaming partnerships but may be challenging to achieve roaming within the same country without a tight relationship with the MNO. While the network is being expanded and built out with 3GPP, the MSO may have some coverage gaps and impacting the subscriber experience.

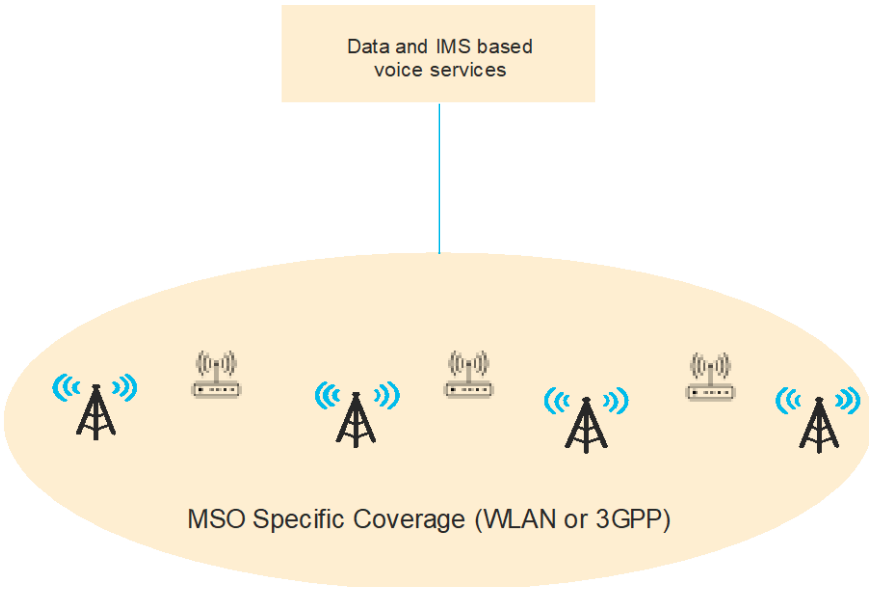


Figure 4: Nationwide MSO network without MNO partnership (except roaming)

3.3. Private Networks

A scenario where the MSO could choose to focus primarily on new opportunities only using the licensed spectrum and target the Enterprise Private Network markets. The MSO may choose to offer a consumer service only via a pure MVNO partnership as an option – however, the spectrum assets are leveraged only for the enterprise deployments to begin with. With this model, the MSO could always expand to nationwide MSO or MNO partnership models in the later phases. MSO may choose to deploy a dedicated or a multi-tenant core for enterprises

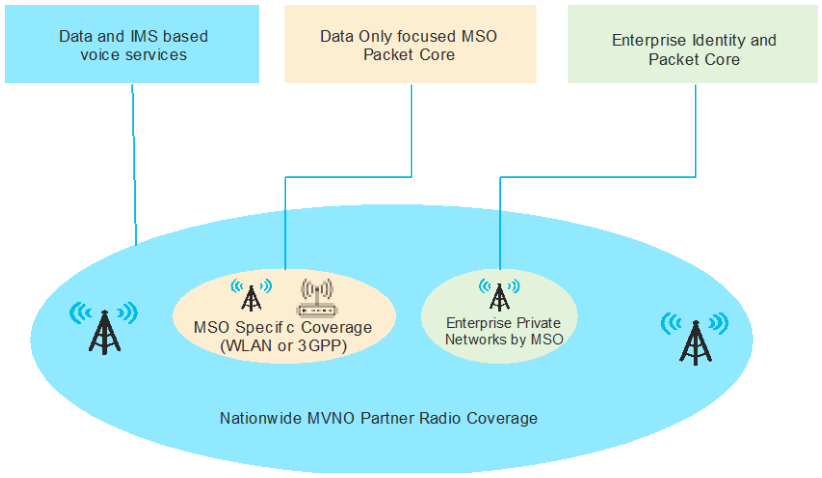


Figure 5: Focus on Private Networks and Limited MSO consumer coverage with nationwide MNO partnership

All these network options could exist by themselves as standalone or fully integrated. Introduction of additional capabilities could be in phases – as an example,

- 5G Fixed Wireless Access Network Deployment – Data offload only
- Voice Network introduction – migrate from MVNO Voice
- WiFi integration
- 5G Differentiator Features Deployment: Slicing; MEC; External APIs and Roaming supported
- 5G Core Advanced Features Deployment: Network Analytics (Artificial Intelligence (AI)/Machine Learning (ML)); Common Data Layer deployment

The sequence of the phases does not always have to be as depicted above but rather based on operator priorities.

While the 3GPP architecture seems simple on paper, what drives the deployment complexity is the service differentiation the operator chooses to provide towards the end user. As an example, a SMF/UPF combination depicted in the 3GPP architecture is supposed to deliver at a high level, the following base functionality for simple packet routing.

- Session Management and UE IP address allocation
- Configures traffic steering at UPF to route traffic
- Interfaces towards Policy control functions
- Lawful Interception and other regulatory requirements
- Policy rule enforcement, (e.g., Gating, Redirection, Traffic steering) and QoS handling
- DPI and application detection

In addition to the functionality defined above, due to operational requirements an operator will have:

- Local Redundancy in case of card or a process failure
- Geo-Redundancy in case of site failure or connectivity challenges
- Local Policy in case of temporary glitches in connectivity
- Local storage or accounting records and NAT Binding records

In certain scenarios, building out as a greenfield mobile network operator, the MSOs will be presented with choices to build a simplified network. However, it could end up being a competitive differentiation or a service parity with the MVNO network scenario which may end up forcing the MSOs to deploy an equivalent capability in the network.

As 5G network architectures evolve and deployment plans are formalized, the industry is driven not just by competitive pressures, but also based on evolving use cases which are heavily dependent on the availability of the device ecosystem that supports various 3GPP capabilities. The current expectation of the market trend as seen by the authors is depicted below.

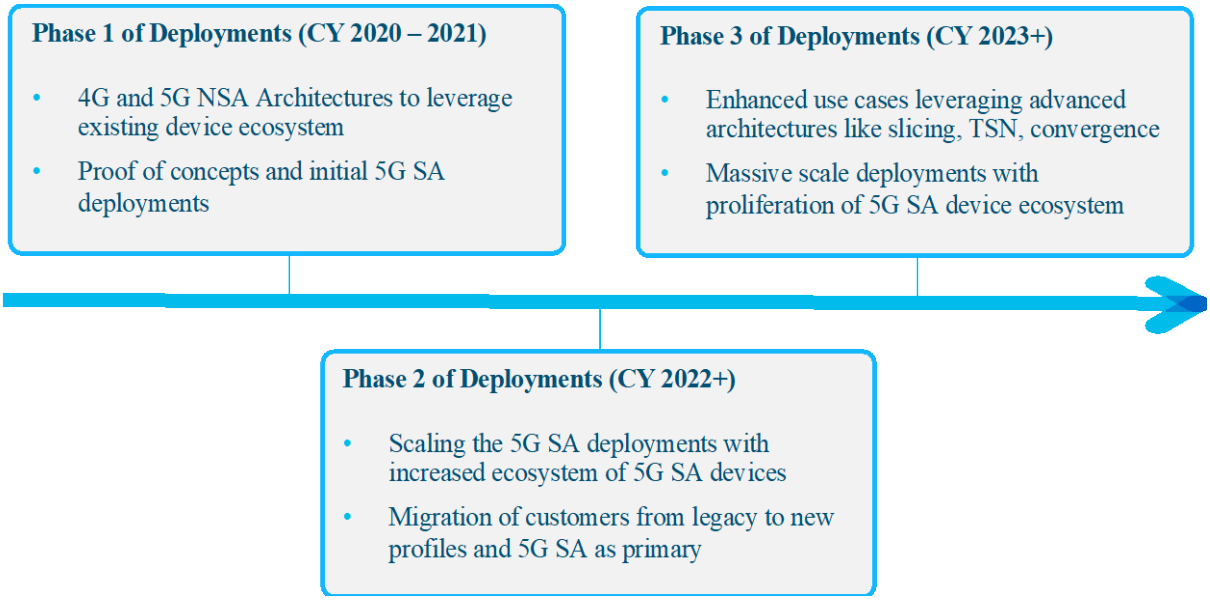


Figure 6: Likely timeline of deployments based on industry momentum

While there have been enough publications around what 3GPP network architectures in 5G offers for MSOs, technical capabilities, deployment models, this paper focuses on highlighting some of the key lessons learnt in existing Mobile network deployments – during 4G deployments and some initial 5G deployments. Intent of this paper is to offer solutions to common pit falls in mobile networks as it stands today – to account for these from the get-go and avoid any potential network issues and service interruptions. The authors believe that the lessons learnt during the initial 4G deployments, and the maturity achieved over the period of 10+ years in 4G networks should be leveraged as 5G networks are being built out.

4. Greenfield Network Considerations

4.1. Converged 4G+5G core Deployment

One of the important decisions an MSO faces would be to either deploy a 5G only RAN (also called as 5G Stand-alone (5G SA) or Option 2 RAN) or 4G 4G-and-5G non-standalone (5G NSA) or Option 3 RAN. The high-level architecture for the two options is shown in the Figure below.

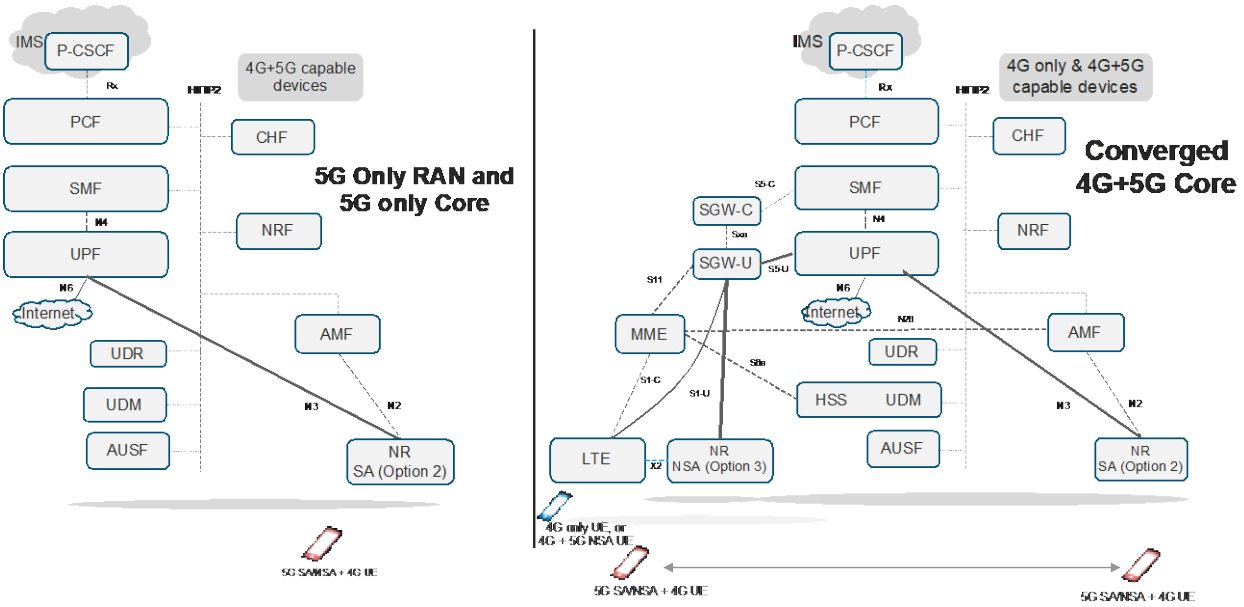


Figure 7: Converged 4G+5G Core architecture vs. 5G SA architecture

The left side of the Figure depicts the key network functions of 5GS system that only serves 5G RAN (NR) in a stand-alone mode. This would imply that in the gNB in NR band for the specific country is deployed and there is no 4G LTE deployment. Also, the mobile supports 5G SA mode of operations.

On the right side is a converged 4G and 5G core that also support 4G LTE radio and core network. The 4G related elements includes the MME and Serving Gateway (SGW). Also interworking interface between the MME and AMF (N26) is used to enable seamless mobility between NR and LTE. With this deployment the MSO can offer both LTE and NR in different frequency bands in the 150 MHz CBRS frequency spectrum, as an example. In addition, the mode where NR radio is added for additional downlink and uplink data rates with signaling going via the LTE cell. This configuration is called 5G NSA or Option 3. This requires the deployment of both LTE and NR i.e., eNB and gNB. The same NR radio can operate both in Option 3 and in Option 2 mode. With this deployment mode an operator can support legacy 4G UEs and the new 5G NSA and 5G SA UEs.

The main driving factor in such a decision is the availability of 5G SA capable devices. Though the number of devices that support 5G (NSA and/or SA) is increasing it is still very small compared to the 4G UEs (May 2021: approximately 800 device types that support 5G NSA or 5G SA compared to about 25,000 device types that support 4G technology[5]). Furthermore, the number of 5G devices that support 5G SA is a small fraction of the devices that support 5G (NSA or SA). Though the number of devices (phones and other form factors) that support 5G will increase, this number will still be smaller than those that support 4G. Depending on the timeline of deployment, it is possible for an MSO to deploy the converged 5G and 4G core. The operator should deploy this converged core also in comparison to deploying a separate 4G only EPC core (with PGW and PCRF instead of SMF+UPF and PCF), since EPC core cannot support 5G NR SA mobiles and EPC also is based on legacy telecom specific protocols like Diameter and GTP-C whereas the converged 5G and 4G core is based on protocols that have much wider deployment in the market, e.g., HTTP.

If the MSO chooses not to deploy 4G radio and no support for 4G only devices is required, the MSO can choose to deploy a SA core. However, given that MSO subscribers could potentially roam into a country that does not yet have 5G deployments or a partnership with a roaming provider with 4G only capabilities, it may be required to terminate the legacy GTP-C interfaces from the roaming partner. This would mean that the converged 4G+5G core will primarily be leveraged as a 5G SA Core when the device is on the MSO deployed RAN but when the device is connecting from a 4G only partner RAN, the converged core capabilities of exposing legacy GTP-C interfaces i.e., S8 interface could be leveraged. Note that in this case, N26 interface is not required unless a tight roaming relationship with handovers is planned.

4.2. MNO/MSO demarcation points

As stated in the previous sections, the most common deployment model mobile network deployment for an MSO could be one with partnership with an MNO, where the MNO provides nationwide coverage, and the MSO provides coverage in pockets (e.g., in urban dense areas). The network architecture diagram for such a deployment scenario is shown in the figure below where the MNO is providing 4G including 5G NSA coverage and the MSO has deployed 5G SA in NR bands in its pockets.

[5] https://gsacom.com/webinar/the-5g-story-so-far-5g-spectrum-networks-and-devices-in-1h-2021/

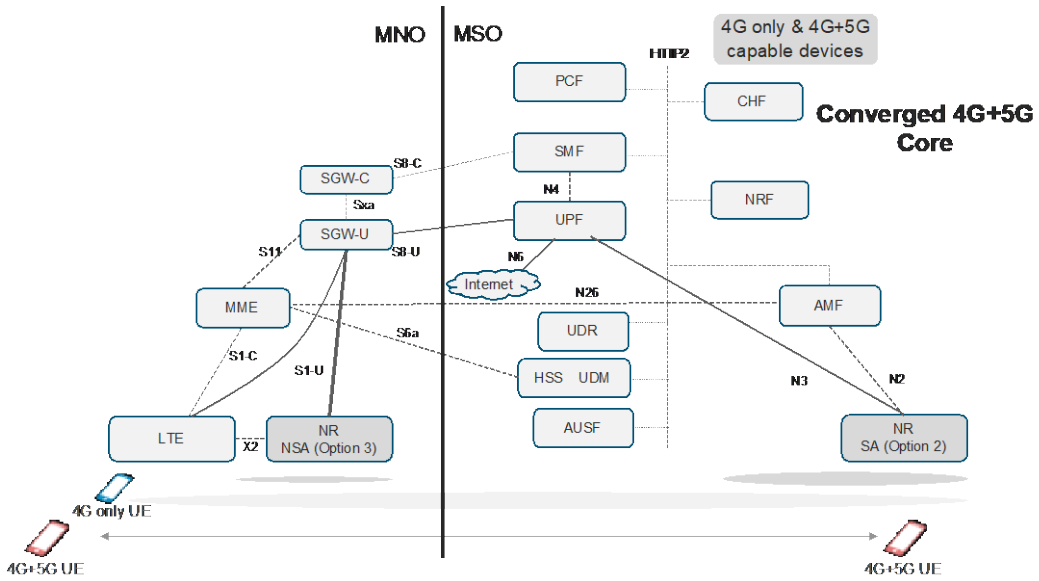


Figure 8: Demarcation of MNO and MSO networks in partnership when single-SIM is used.

The key interworking interfaces between the MNO and MSO are those that cross the thick black demarcation line between the MNO and MSO. The S6a interface between the MNO's MME and the UDM+HSS of the MSO is used to authenticate the MSO's UE when that UE is in MNO's coverage area. The S8-C and S8-U interfaces between the SGW-C and SMF and SGW-U and UPF respectively, are used to anchor the IP connection of the mobile in the MSO's SMF and UPF (this is also called home-routed scenario in 3GPP specifications). The reason for home-routing is that primary service policy and charging are provided by the SMF and PCF in the home-network of the UE, i.e., the MSO. The N26 interface between the MME and AMF is used to transfer contexts between the MNO and MSO network to enable seamless mobility.

Though 3GPP specifications exists that provides all the details of the call-flows to support the above architecture, the main challenge is in deployments of these inter-operator interfaces. The vendors providing the network functions on the two sides of the demarcation lines may be different and may require extensive inter-operability testing. Also, the interactions between the radio networks to handovers from the MNO network to the MSO network when a mobile is in the coverage of the MSO's radio network and vice-versa require quite a lot of radio planning (e.g., configuring neighboring cell information and updating these as the radio network of the MNO or MSO changes).

To mitigate some of these challenges, a DSDS based approach could be leveraged as described very well in Cable and Mobile Convergence, A Vision from the Cable Communities Around the World 6 and Section 4.1.5, “Dual-SIM”.

[6] https://www.nctatechnicalpapers.com/Paper/2020/2020-cable-and-mobile-convergence

4.3. Single vs. Multivendor strategy – Open Interfaces

While 3GPP provides a very well debated and industry vetted architecture with the operator, vendor and entire ecosystem contributing, it sometimes is unable to accommodate every single deployment scenario and use case. Especially, when it comes to requirements driven by MSOs, since the use case requirements and deployment models are not always aligned with established MNOs in the market, there could be a need to extend of enhance the capabilities for MSO deployments. Given that 3GPP and other SDOs have a set schedule for completing the documentation of specifications and priorities, it is likely that timelines for deployment of the operator may not align

This typically drives the MSOs to consider custom implementations. Take a scenario where an MSO would like to identify a customer not just via mobility identity but based on wireline identity or a cable identity. If leveraging 3GPP Release 15 specifications, this is not possible today without custom adaptation as 5WWC work item in Release 16 would add some of these capabilities. In these scenarios, an MSO may consider adding a Cable Line ID as a custom attribute across the nUDM service and on nChf service to identify the mobile user and tie into the Cable Service.

A similar capability was leveraged in 4G LTE networks extensively – especially on diameter interfaces. This was typically achieved via vendor defined AVPs or vendor specific AVPs. Additional details are captured in the Diameter RFC 3588, 5.3.3. Vendor-Id AVP. While this was a well understood mechanism, due to competitive scenarios, it is not always viable and possible to take this custom vendor defined AVPs to SDOs. In such scenarios, it may be required to continue supporting custom implementations which would require ensuring interoperability across vendor products.

Since 5G Service Based Interfaces support extensions or information elements that could added beyond the specifications, this could be leveraged by the MSO to achieve the use case and faster go to market option without having to wait for 3GPP or other SDO. Invariably, such custom behavior creates a challenge during software upgrades or interoperability across different vendor solutions.

We recommend that operators need to ensure that all deployment interfaces are open and compliant to 3GPP specifications. This is typically achieved via vendor-to-vendor interoperability testing either from a lab to lab or within the MSO labs. Any extensions or custom attributes implemented as described above by any vendor specific to the MSO network are recommended to be fully documented and more importantly, should be protected by a feature capability exchange and only utilized when the peer node indicates support.

- Option 1: Implement feature capability exchange: The recommended way is to ensure the initiator of the message that includes the custom information does so only once it has been established that the receiver can gracefully handle the custom information provided. This could be achieved in multiple ways in 5G network deployments.
    - In case a client / server relationship is established – leverage feature capability during initial connection establishment

- Leverage NRF capability exchange to communicate and register feature versions which include the custom attribute supports and only provide the producer information when the client specifically requests for this capability

- Option 2: Implement configuration options: This option provides an easy implementation choice by only including the custom attributes if specifically configured to do so. This removes the complexity of feature capability exchange or relying on NRF but poses an operational challenge of making sure every NF is appropriately configured and upgrades are sequenced and managed in a controlled manner
- Option 3: Ignoring uninterpreted attributes and default behavior: Alternate to a feature capability exchange or configuration is ensuring that the information passed through the custom attributes is ignored when a receiver is not able to interpret it and ignoring such information does not impact service. This is critical even if one of the additional options is implemented to avoid any service disruptions to minor changes in software releases or attribute parameters. The receiver should log an error so the issue could be debugged and addressed appropriately but while this is underway, network service is maintained without disruption, albeit with likely not all capabilities.

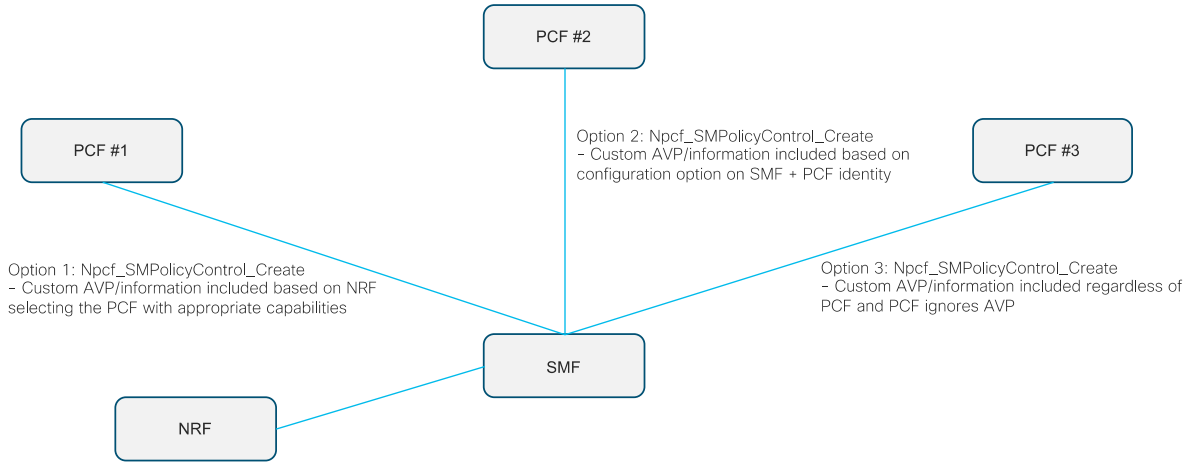


Figure 9: Options to achieve interoperability with custom implementation

We believe that at the end of the day – 3GPP compliant interfaces are the best way to deploy networks. The reason for recommending this is to ensure potential introduction of additional vendors in the mix. As an example, to achieve certain tracking area mapping capabilities, it was required in initial days of 4G LTE RAN ecosystem to implement custom behavior on the S1-AP interface between eNodeB and MME. However, it had eventually gone to a point where it was almost impossible to interwork an eNodeB from the RAN vendor with any other vendor's MME without impacts to service or capabilities which created a "lock in" problem for the operator and not being able to deploy best of breed network functions or a dependency on the operator introducing new services as the radio provider feature acceleration may not be at par with the operator requirements or a market specific mapping of RAN and MME i.e., tight dependencies cause operator to define a clear boundary of operation for radio and packet core

As an example, a radio vendor that has the radio infrastructure deployed in one region was able to only utilize MMEs from that vendor. During a system outage or migration or other similar scenarios, the operator is not able to leverage the additional capacity that is available from a peer market or region which is provided by another Radio vendor. Especially, the use of ASN.1 as a protocol between Radio and MME decreases the ability to interwork once custom attributes are introduced.

Similar rules apply to the non-radio deployment scenarios but due to the ability to extend diameter or GTP protocols without causing interoperability issues based on proposals above has resulted in somewhat of a manageable model within the packet core. Any deviations from common practices and not documenting could result in an interoperability issues and delay in service launch or disruption to existing services. Relying on open interfaces for scalable network architecture and debugging across network functions is critical for long term success of the solution.

4.4. Policy and charging architecture

The sheer number of different data-plans that are offered by MNOs is an evidence of the complicated policy and charging rules and traffic measurements that are performed in MNO networks. Specialized application charging rules, such as zero-rating of Netflix or Spotify traffic, require significant DPI and application detection capability in the network.

3GPP PCC architecture example for 5G System is depicted in the figure below and with all the options supported, the deployment could be daunting with ability to create policies across multiple service enablers in the network and ensuring policies do not contradict each other in this scenario. A well-defined PCC rule can help with detection of a service data flow and providing parameters for policy control and/or charging control and, for PGW enhanced with ADC, for application detection and control.

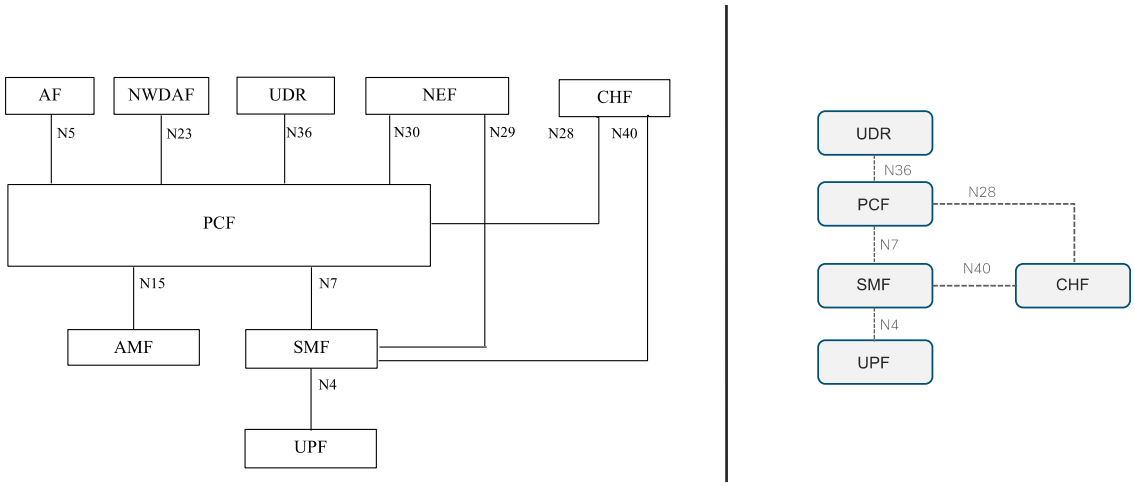


Figure 10: 3GPP PCC Architecture for 5G (left). Simplified PCC architecture for MSO (right).

There are two main drivers for the complex policy and charging infrastructure in MNO networks:

1. Requirement to support dynamic QoS, e.g., for providing periodic scheduling with low latency for voice traffic. (NOTE: This is not typically for charging perspective, since most MNOs are not monetizing voice traffic), and
2. Data caps and rate-enforcement required either for on-line charging (pay-as-you-go) or for offline charging (monthly payment) data-caps per month.

MSOs can simplify these by not supporting voice traffic natively. Not having to support voice or other traffic that requires dynamic signaling of filters and QoS rules at least does not require to support the N5 interface between the PCF and the application function (AF).

MSOs should look at only providing offline charging and not have to implement the more complicated on-line charging model. Typical MSO customers have at the minimum monthly subscriptions with them. Trying to support the more complicated on-line charging (pay-as-you-go) model will simplify the real-time requirements for policy and charging.

With the above simplifications, the only rules that an MSO needs to look at is enforcing are data-cap rules that are simpler to define and maintain. This design would align with similar billing and accounting capabilities offered within the Cable access for end users today in most networks. However, if the MSO chooses to offer more aligned capabilities as MNO, this will require similar PCC architecture and complex rule definition as done on the MNO network today.

Number of policies deployed could have impact on the overall performance of the system. When PCEF needs to perform Deep Packet Inspection or process policy while handling small packets, the overall system capacity could be impacted. Any DPI needs to be accounted for when deploying the static and predefined rules being installed on the gateways and only service / revenue impacting rules should be considered. Changes in policy based on specific locations require the PCF to be aware of the current location of the subscriber and in a mobility intensive environment, this could generate signaling all the way from the RAN, MME/AMF, SGW, PGW/SMF and eventually to the PCF causing massive signaling. Smart Phones tend to create Service Request in the network at an average of 2 mins and any location changes triggered during service request could generate unnecessary signaling.

Location based policy should be higher level granularity e.g., Presence Reporting Area or regional boundaries instead of individual RAN locations. Any rules installed by the PCF should be considered for the depth of the packet inspection required and impact to the system performance.

Designing for a simplified but robust policy, charging architecture and rule definition would ensure a seamless service experience for the end users and enabling new capabilities in the network with limited planning.

4.5. Device Behavior

When a network function or node is not able to handle the messages or is in maintenance mode, it is important that the downstream peer functions are aware of this. It is essential that the operators define a clear expected behavior within a client for error codes received from peer nodes and act in a graceful manner. As an example, if the PCF is unable to handle the message due to a session that is not found, the error code provided towards the SMF should provide enough information for the SMF to either initiate the session recovery procedures or gracefully disconnect the user and reestablish the session. While some of the network error behaviors could be gracefully handled or at least addressed with software updates, during the initial days of LTE launch, a critical lesson learnt was on the unexpected UE behavior.

As LTE networks were deployed, it became apparent that the error codes were not interpreted the same way, despite some documentation in 3GPP. This resulted in unpredictable behavior during error scenarios, but also resulted in non-uniform behavior in different markets or geographic regions of the same operator where different vendor solutions were deployed.

This was eventually addressed in LTE with much more prescriptive behavior expectations on error codes at the device (Reference: TS 24.301[13]) but also a massive effort to map error codes from every diameter interface to eventual error code mapping at the device. While some of these are based on lessons learnt from real deployments, some were mere clarifications of predefined behaviors in 3GPP.

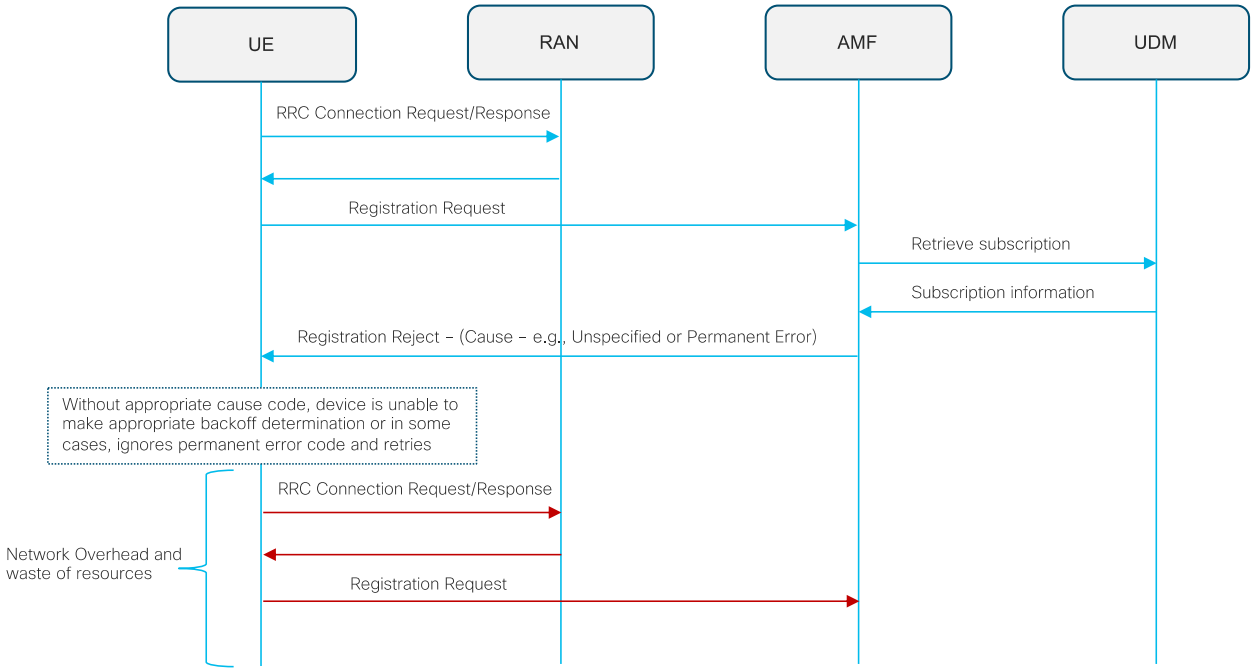


Figure 11: Call flow depicting UE errors and retries

Fortunately, most of this work has been forward ported into the 5G specifications already and similar capabilities as mature 4G networks exists already in 5G networks. However, it is imperative to plan for previously unseen behaviors in the network. Especially, new capabilities and interfaces in 5G defined around Service Based Architecture, slicing capabilities on devices and additional new functionality defined in 5G – it is very likely unforeseen scenarios will occur in the network.

It is critical to also note that the error scenarios are not just to be UE focused but rather a network focused approach as well. Some of these could be avoided by:

- Formal interoperability testing as described in previous sections or relying on statements of compliance to begin with
- Adding intelligence into the network to decommission or remove certain software functions in the network based on error trends
- Configurable error code behaviors

Accounting for unforeseen device behaviors and designing for graceful recovery mechanisms would help ensuring a reliable end use experience.

A couple of good examples of what has been achieved in 4G networks and positive outcomes of lessons learnt could be based on the overload scenarios or race conditions in the network.

Race conditions have typically occurred when the network and device are attempting conflicting procedures which are unfortunately not completely preventable but by adding message priorities and sequenced queues, the impact of these race conditions could be mitigated to a certain extend.

For overload scenarios, based on ecosystem and industry discussions, IETF work on overload control (Reference: https://datatracker.ietf.org/doc/html/rfc8582 [12]) or GTP-C load / overload control was introduced into 3GPP specifications. It should be noted that these are not widely leveraged today, partially due to the timing of availability of these capabilities in the specifications. By the time these specifications were available in 4G, there was an industry momentum within 5G and focus shifted towards 5G software development and implementation in the field. However, most of these capabilities have been made available with 5G specifications from day 1 and so could be leveraged as needed from initial deployments for MSOs.

One additionally way of ensuring device vs. network race conditions or duplicate messaging is avoided is by fine tuning the timers in the network and the message queues with validity timers / expiry tags for messages so that any messages stuck in queues are discarded after a certain time. In some rare scenarios, if a message from the queue is processed after the device internal timers have already expired and a procedure from the network is executed, this could result in out of band processing of the messages resulting in incorrect and unexpected behavior from the devices.

While there may be some reluctance to use some of these advanced capabilities in the specifications to begin with, we believe that having the capabilities planned in the network from the initial get go would be important.

5. Other topics for consideration

While not captured in detail in this document for brevity, it is critical for MSOs to take additional topics into consideration as listed below.

1. **Virtualization Stack**: Every MSO may choose to have a platform of choice – which may or may not be driven directly by the 5G deployment strategy. It is likely that some MSOs have already charted a path for virtualization of Network Functions in the existing deployment and service strategy. In some ways, the same platform could be the entry path for 5G Core cloud-native NF's or implementation of Open vRAN architectures that are critical for success of 5G. Depending on the capabilities of the MSO, operational focus within the company, a CaaS and PaaS capability may be under evaluation or already in deployment which could be leveraged or any existing investments on NFVI could be leveraged. Given that this is a choice for the MSO which could be influenced by other factors in the company – additional details are not discussed in this document.
2. **Convergence architectures**: 3GPP has defined the baseline architecture for convergence of wireline and wireless services as part of the 5WWC work item in Release 16 specifications. However, with additional dependencies being addressed in CableLabs and the detailed analysis and discussing in Cable and Mobile Convergence, A Vision from the Cable Communities Around the World[7], this document does not address convergence but could be a critical decision point for the MSOs before finalizing the 5G deployment architectures.
3. **Support of Voice Services:** Given the stringent performance and regulatory requirements for voice, in a deployment model where MSO partners with an MNO (see previous sections), it could be simpler for the MSO to have the MNO support voice services whereas the MSO provides internet connectivity as native voice support and client capabilities are made available specific to the MSO architecture. One challenge of achieving voice services delivered via the MNO partner with data services from the MSO native network is that either the device will have to support dual SIM capabilities or when the voice network of MNO is leveraged, the data is offloaded to the MNO partner network as well while the device is in a voice call
4. **Support for lawful interception**: Lawful interception is a required functionality by regulatory bodies in most countries where mobile networks are deployed. As cable operators consider their journey from MVNO to MSO, they will need to plan and deploy lawful interception. Most vendors who provide infrastructure equipment support lawful interception on their network function. Also, MSOs will need to create a system for handling lawful interception request from law enforcement agencies in their jurisdiction. Since information about targets of lawful interception needs to be guarded very closely in an operator's network, this involves only enabling a select few operation staff to have access to this information and for configuring and monitoring the system for lawful interception.

[7] https://www.nctatechnicalpapers.com/Paper/2020/2020-cable-and-mobile-convergence

5. **Roaming partnerships**: Roaming architectures are very well defined in 3GPP but most importantly a clear cross operator engagement model defined by GSMA to enable global roaming. It is recommended that MSOs leverage the existing guidelines of GSMA[8] for 5G being developed. One of the important considerations for the MSO as they finalize the deployment model is the any dependencies related to anchoring devices from 4G networks in countries where 5G is not yet available. Using a converged core architecture could simplify the operations and supporting different roaming models as described in Section 4.1
6. **RAN considerations**: This document does not cover RAN architectures in detail as there has been a significant amount of collateral and analysis on the various RAN deployment architectures. Please refer to Security Benefits of Open Virtualized RAN9 or additional standardization details as part of O-RAN alliance10. O-RAN could be discussed as a standalone document alone.
7. **SIM profiles and Device Management**: SIM profile development and device management capabilities could be driven by the operator choice of device partner ecosystem, compliance to OMA-DM specifications and partnerships with specific UICC provider itself. Given that the SIM profiles and UICC management are usually not disclosed publicly due to security concerns, this document does not specifically capture the lessons learnt from 4G deployments with respect to device management. However, this would be a critical area of investment for the MSO to ensure ability to configure new APNs/DNNs within the network and subsequently provisioning the device with the capabilities or being able to manage the MAPCON profiles on the device for WLAN vs. 3GPP selection. Note that with 5G, given the capabilities of ANDSF from 4G networks are now integrated with the PCF, we anticipate that once the devices start supporting the capability, network selection would become simpler to manage in 5G.
8. **Network Failures and Recovery**: Network failure scenarios are inherent in the regardless of the architecture and stability of the network functions and their underlying infrastructure. In the move to cloud infrastructure, this is even more prevalent as the virtualization infrastructure typically does not achieve more than 99.9% redundancy. To achieve a 5x9s service or greater, the redundancy of the network functions must be increased to support more common cloud infrastructure failures. Care needs to be taken to address:
    - Network Function Failures
    - Software Upgrades and Configuration Failures
    - Blast Radius of a Failure

[8] NG.113 5GS Roaming Guidelines v4.0 @ https://www.gsma.com/newsroom/resources/ng-113-5gs-roaming-guidelines-v4-0/
[9] https://www.cisco.com/c/dam/en/us/solutions/service-provider/pdfs/5g-network-architecture/white-paper-sp-open-vran-security-benefits.pdf
[10] https://www.o-ran.org/

6. Conclusion

We acknowledge and understand that every MSO has different reasons for investing in building a new Mobile Network. Whether it is to reduce the cost structure in offering mobile services by offloading as much traffic as possible from the MVNO partners network or to compete with differentiated services offering in the market, MSOs now have a unique opportunity to leverage new spectrum assets to do so. With 5G architecture defined in 3GPP and other SDOs enabling convergence capabilities in the future, investing in 5G to build a standalone network could be ideal in some cases but due to regulatory requirements or device ecosystem dependencies or roaming opportunities, there may be a need to support 4G / LTE networks as well.

As MSOs explore the option of natively building a greenfield network it is critical to ensure some of the lessons learned by MNOs while building the 4G networks or during the launch of VoLTE networks should be leveraged to ensure the new 5G networks by MSOs do not have the same challenges.

As described in the various sections, the following would be critical considerations, though not a comprehensive list.

1. Partnership scope with MNO partner or MVNO partner to ensure ability to influence policy across the network for seamless experience for the end user regardless of which network (MNO vs. MSO greenfield) the user device connects from
2. Augmenting voice and required regulatory compliance capabilities while the native coverage by MSOs is being expanded / built
3. Ensuring open interfaces across the network architecture and functions – to be able to leverage best of breed architecture vs. a single vendor strategy risking potential "lock in"
4. Reducing the complexity of policy and charging architecture but still being able to address inline service requirements to offer similar or same capabilities as the partner network
5. Expect that even with wide interoperability testing and device compliance, there will be a scenario where one device firmware/model may misbehave and with the volume of devices, this could quickly snowball into a massive network challenge. Preparing network with various mitigation capabilities when such device behavior is encountered could reduce subscriber impact and reduce operation costs
6. Design for various potential race conditions and network failure scenarios and fallback mechanisms from day 1 and ensure mobility is accounted for as devices connect from alternate access types or locations during a mass failure or pockets of failures in the access/core network

With the extensive knowledge that the MSOs have gathered over the years with existing networks and the interest to offer differentiated service to the end users, we believe that MSOs are at a cusp of changing the dynamics on wireless/wireline networks as a whole and the entire Service Provider model. Taking into consideration the variety of challenges that mobility and wireless pose and leveraging best practices will make this transition smoother and more successful.

7. Appendix A

A complete set of 3GPP network functions as defined in Release 15 and Release 16 specifications is listed below. The authors believe that while these capabilities are eventually required to be able to offer a wide set of use cases and completive solutions, not every function is required from day 1 of the deployment. It is possible to introduce most of the functions into an existing Greenfield network after launch without disruption of service or a major redesign. Some functions, like BSF or SCP may need some consideration before launch and planning for the future.

Table 1: Complete list of the 3GPP defined network functions

5G-EIR	5G-Equipment Identity Register
AF	Application Function
AMF	Access and Mobility Management Function
AUSF	Authentication Server Function
BSF	Binding Support Function
CAPIF	Common API Framework for 3GPP northbound APIs
CHF	Charging Function
ePDG	evolved Packet Data Gateway
GMLC	Gateway Mobile Location Centre
LMF	Location Management Function
LRF	Location Retrieval Function
N3IWF	Non-3GPP Interworking Function
NEF	Network Exposure Function
NR	New Radio
NRF	Network Repository Function
NSSF	Network Slice Selection Function
NWDAF	Network Data Analytics Function
PCF	Policy Control Function
(R)AN	(Radio) Access Network
SEAF	Security Anchor Functionality
SEPP	Security Edge Protection Proxy
SMF	Session Management Function
SMSF	Short Message Service Function
UDM	Unified Data Management
UDR	Unified Data Repository
UDSF	Unstructured Data Storage Function
UPF	User Plane Function

Abbreviations

3GPP	Third Generation Partnership Project
5GC	5G Core
5GS	5G System
AF	Application Function
AMF	Access and Mobility Management Function
AP	Access Point
APN	Access Point Name
AUSF	Authentication Server Function
CBRS	Citizen broadband radio service
CDR	Charging Detail Records
CHF	Charging Function
CPE	Customer premise equipment
CUPS	Control plane and user plane separation
DN	Data Network
DSDS	Dual-SIM Dual-Standby
eMBB	Enhanced mobile broadband
EPC	Evolved Packet Core
ePDG	Enhance Packet Data Gateway
EPS	Evolved Packet System
eSIM	embedded SIM
FWA	Fixed wireless access
LTE	Long Term Evolution
MCC	Mobile Country Code
MNC	Mobile Network Code
MNO	Mobile Network Operator
MSO	Multiple System Operator
MVNO	Mobile Virtual Network Operator
NEF	Network Exposure Function
NF	Network Function
NFV	Network Function Virtualization
NG-RAN	Next Generation RAN
NR	New Radio
NRF	Network Function Repository Function
NSA	Non-Stand Alone
NSSF	Network Slice Selection Function
NWDAF	Network Data Analytics Function
OCS	On-line Charging System
OfCS	Off-line Charging System
PCF	Policy Control Function
PCRF	Policy and Charging Function

PLMN	Public Land Mobile Network
SA	Stand Alone
SBA	Service Based Architecture
SBI	Service Based Interface
SMF	Session Management Function
SMSF	Short Message Service Function
SEPP	Security Edge Protection Proxy
RAN	Radio Access Network
UE	User Equipment
UPF	User Plane Function
UDM	Unified Data Management
UDR	Unified Data Repository
URLLC	Ultra-Reliable Low Latency Communication
VoLTE	Voice over LTE
VoNR	Voice over NR
vRAN	Virtual RAN
WWC	Wireless-Wireline Convergence

Bibliography & References

[1] 3GPP TS 23.501, System architecture for the 5G System (5GS)
[2] 3GPP TS 23.502, Procedures for the 5G System (5GS)
[3] “5G Implementation Guidelines: SA Option 2”, June 2020, GSMA
[4] Wireless and Wireline Convergence for the 5G system architecture, 5WWC, WI # 830050 - 5WWC
[5] 5G Wireless Wireline Converged Core Architecture Technical Report, CableLabs
[6] Jennifer Andréoli-Fang, PhD, CableLabs; John T. Chapman, Cisco et al, Cable and Mobile Convergence: A Vision from the Cable Communities Around the World
[7] The 5G story so far: 5G Spectrum, networks and devices in 1H 2021, Global mobile Suppliers Association
[8] NG.113 5GS Roaming Guidelines
[9] Eric Hanselman, Security Benefits of Open Virtualized RAN
[10] O-RAN Architecture Description, O-RAN Alliance
[11] O-RAN Use Cases Detailed Specification 5.0
[12] Diameter Overload Rate Control, RFC 8582
[13] 3GPP TS 24.301, Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3
[14] 3GPP TS 32.255, Telecommunication management; Charging management; 5G data connectivity domain charging; Stage 2

------ End of Document ------

Having the Whole Company in a Bag

Mediacom's Real-World Use of Automated Access Network Design and Optimization Technology

A Technical Paper prepared for SCTE by

Bill Wegener
Group Vice President, Engineering and Network Development
Mediacom Communications Corporation
1 Mediacom Way
Mediacom Park, NY
10918 USA
+1 845 443 2745
bwegener@mediacomcc.com

Mike Oja
Senior Manager, Access Network Design & Engineering
Mediacom Communications Corporation
1 Mediacom Way
Mediacom Park, NY
10918 USA
+1 845 419 6317
moja@mediacomcc.com

Ian Oliver
President
Versant Solutions Group Inc.
21 Radford Crescent
Markham, Ontario
L3P 4A2 Canada
+1 416 543 3360
ian.oliver@versantsolutionsgroup.com

1. Introduction

This paper presents findings and insights from Mediacom Communications' application of automated design and optimization (ADO) technology to its access network strategy and planning activities, and to its business-as-usual network design and implementation efforts. Mediacom's decision to adopt ADO technology is discussed in the context of having to make decisions on major capital investments in an environment that requires considering multiple, evolving network technologies and architectures as well as serving a variety of markets ranging from urban to rural, while having too little time to perform the necessary analyses using traditional manual methods.

This paper further describes Mediacom's strategic approach to managing such challenges and how it sees the use of ADO technology being transformative to its operations. The benefits of having 'the whole company in a bag' are discussed in detail. Such benefits include having the ability to ask and answer billion-dollar, footprint-wide questions in hours, as well as the ability to quickly react to network utilization issues and implement solutions on a node-by-node basis that are consistent with Mediacom's network evolution plan. The distinction between the use of actual network designs, as opposed to costing models and rules-of-thumb, in generating bills-of-materials and capital cost estimates, is described in terms of concrete benefits to the capital planning and budgeting process, as well as to construction planning and execution activities.

This paper also describes a unique situation in which Mediacom entered into a competitive bidding process to serve a mid-size US city with symmetrical Gigabit service and, using ADO technology, was able to quickly and confidently evaluate multiple N+X network architectures in terms of technical feasibility and capital cost. Finally, this paper presents and discusses the technical and operational path taken by Mediacom in evaluating and deploying ADO technology.

2. Business and Operational Challenges

Mediacom's network engineering group is constantly working to answer the deceptively simple sounding question: What must we spend on improvements to the access network to ensure it consistently satisfies our customers? Coming up with an answer to that question is fraught with challenges including:

- New and forthcoming transmission technologies that simultaneously offer excellent performance but complicate the capital expenditure decision-making process
- A network footprint that comprises relatively dense urban, medium-density suburban, and low-density rural environments, each presenting unique service requirements and design considerations
- The need to design, budget, approve and deploy network improvements in the shortest possible time to maximize the return on capital and to avoid losing customers to competitors
- Making the highest and best use of the capital funding available for network improvements, regardless of the challenges noted above
- Very long lead times for delivery of equipment and materials which is causing, and is expected to continue to cause, painful delays in network deployments
- The requirement to regularly (semi-annually) review and modify the long-term (5-year) network technology and architecture strategy to keep up with evolving transmission technology, growing customer demand, and competitive threats.

Mediacom, as a profit-making business, also must satisfy its investors that capital investments in the access network have been well planned and accurately budgeted. Similarly, as a telecommunications service provider to the public, Mediacom must satisfy its respective franchising authorities that, in fact, the access network will continue to reliably meet the service requirements of its customers.

3. Adoption of Automated Design and Optimization Technology

3.1. Pilot Project

In early 2020, Mediacom embarked on a program to determine and, over several years, deploy a new network architecture across its entire footprint. At that time, it was expected that most of the network would be upgraded to an N+2 architecture, though N+0 and FTTH architectures would be considered where customer demand warrants. Since then, Mediacom has adopted a network architecture strategy, dubbed Fiber-Deep, that does not call for a fixed maximum cascade length of 2 active devices, instead specifies a range of homes-passed per CMTS port – or Fiber-Deep node – as the determining factor for cascade length.

At that time, Mediacom undertook an evaluation of an ADO technology to determine if that technology could deliver:

- preliminary designs for N+2 architecture access network in accordance with Mediacom's design rules and equipment specifications
- optimization of locations of all new nodes to minimize node count
- fully calculated, technically valid RF plant design
- optimal routing of new fiber cable as required to connect new nodes to the existing fiber plant
- integration with Mediacom's existing plant engineering platform
- amplification of the effectiveness and productivity of Mediacom's existing network planning team
- savings of time and cost relative to manual planning and design methods

The results of the pilot project demonstrated that all of Mediacom's requirements were, or could be, satisfied by the ADO technology under evaluation. Mediacom subsequently decided to implement the ADO technology, commencing in early 2021 and continuing through Q3 of 2021.

3.2. Overview of ADO Platform

The ADO environment comprises three main components and two primary results, as shown in Figure 1. The three main components and the two primary results are:

- Mediacom's existing geo-spatial information GIS/network engineering database containing a complete model of the as-built access network
- a set of network design rules, one for each network architecture that Mediacom wishes to consider, which are created and maintained within the ADO platform, and
- the ADO engine itself, which reads in the as-built access network from the GIS/network engineering database and applies the design rules for a selected new architecture to the as-built network, resulting in
- preliminary designs and cost estimates for the new network architecture.

Mediacom initially utilized ADO to apply network design rules for its 'Fiber-Deep' architecture to its entire network footprint, which resulted in Mediacom having 'The Whole Company in a Bag'.

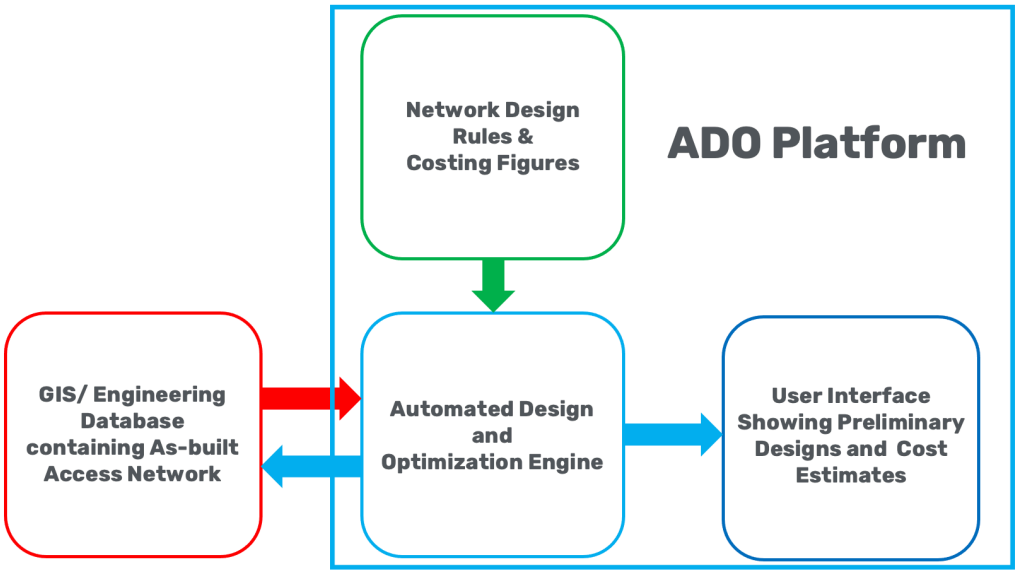


Figure 1 – Overview of ADO Environment and Platform

It is fundamental that the GIS/network engineering database be complete and accurate, otherwise the resulting preliminary designs and cost estimates will be incomplete and inaccurate. Mediacom, along with most major operators, has invested significant time and money in recent years in ensuring that its GIS/network engineering database is complete and accurate and, with the availability of ADO, is seeing a return on that investment beyond its previous expectations.

By serving as the repository for all of design rules for multiple network architectures, the ADO platform allows Mediacom to centrally develop, easily deploy and automatically enforce corporate design standards. This results in all preliminary designs and cost estimates being compliant with corporate standards, thus eliminating the typical variations that occur when multiple planners and network designers are each producing preliminary designs manually. The design rule sets in the ADO platform can be configured to allow the individual planner or designer to adjust certain design parameters, such as the use of express coaxial cable, and execute multiple designs runs to arrive at a preliminary network design that incorporates the insight of the planner or designer while remaining compliant with corporate standards.

The resulting preliminary designs can be viewed in the ADO user-interface in map view (geographically on a scaled land-base, including a satellite photo base) and schematically (in a simplified view that shows user-selected technical details, such as signal levels, active device cascade lengths, and cable lengths). The network designs are stored in the ADO platform's internal database for as long as required, which is typically until a new network architecture is decided upon and preliminary designs are then produced accordingly, or until a network deployment is decided upon and the preliminary designs are exported from the ADO platform to the GIS/network engineering database to be the basis for production of final design and construction drawings.

The cost estimates (and underlying bills-of-materials) are similarly stored in the ADO platform's internal database and are viewable in multiple formats via the user interface. They can also be exported for use by other software platforms within the network engineering group and within other groups in Mediacom.

3.3. Designs Versus Models

Traditionally, the operator's network planning personnel estimate access network deployment costs by manually producing a costing model whereby preliminary designs are developed for a relatively small portion of the network footprint in question, cost estimates are tabulated based on those few designs, then those cost estimates are extrapolated across the network footprint. Typically, the costing figures in the cost model are over-estimated by a significant amount to ensure that they do not prove to be unmanageably low. While this approach allows network planning personnel to provide information to support senior management in a timely manner, the cost model is necessarily subject to an undesirable degree of uncertainty, which in turn leaves the planners and senior management having to make major capital decisions with less-than-ideal information at hand.

The application of ADO technology offers the operator the opportunity to develop preliminary designs and cost estimates for multiple network architectures across the entire network footprint in the same, or less, time than required to manually develop sample designs and to extrapolate a cost estimate. By virtue of the speed of the ADO processing engine and of every single foot of the access network being processed to develop preliminary designs for a new architecture, there is no modelling required to develop the corresponding cost estimates. Rather, the ADO processing engine applies the operator's standard equipment, cable, and installation costing figures to the bills-of-materials for the preliminary designs to provide costing for every item of new equipment and cable, and for the construction labor to install same.

While it is understood that the preliminary designs, and resulting cost estimates, are subject to change during field validation and construction, that degree of change – or uncertainty – is much less and more easily managed than is the case with manually produced design and costing models.

Further, by having legitimate preliminary designs for the entirety of the network footprint in question, any questions as to the validity of the preliminary designs and the corresponding cost estimates can be very quickly addressed because all of the designs can be viewed on-demand both in map view (geographically on a scaled land-base) and schematically (in a simplified view that shows user-selected technical details, such as signal levels, active device cascade lengths, and cable lengths).

Ultimately, Mediacom expects that the application of ADO will simplify and expedite capital planning for access network upgrades by making better quality network planning information more accessible than previously possible.

4. Whole Company in a Bag and Why That's a Good Thing

Planning and budgeting for access network upgrades is more complex than ever before and has become a costly, full-time activity given the rate of development of new transmission technologies, competitive threats, and increasing customer demands. At Mediacom, the 5-year plan is updated semi-annually, which has required a significant allocation of personnel to carry out using traditional manual processes. The application of ADO to produce the Whole Company in a Bag, that is: A set of preliminary designs and cost estimates for Mediacom's current standard Fiber-Deep architecture, has given Mediacom a comprehensive and accurate baseline for use in planning and budgeting of access network upgrades - for the entire network footprint. Importantly, this capability obviates the need to invest time and money in manual production of preliminary designs which are often superseded by changes in technology by the time a given portion of the footprint is scheduled for upgrade.

Furthermore, by being able to produce, on demand, preliminary designs and cost estimates for alternative network architectures, Mediacom can quickly evaluate and compare each alternative network architecture in terms of equipment and cable quantities, and construction costs. This allows Mediacom to very quickly:

- update its long-term planning and budgeting by processing the entire footprint when needed, which satisfies the need of senior management for long-term planning information, and
- select and deploy the ideal network architecture for any portion the network footprint in question, which is particularly important given Mediacom's mix of urban and rural serving areas.

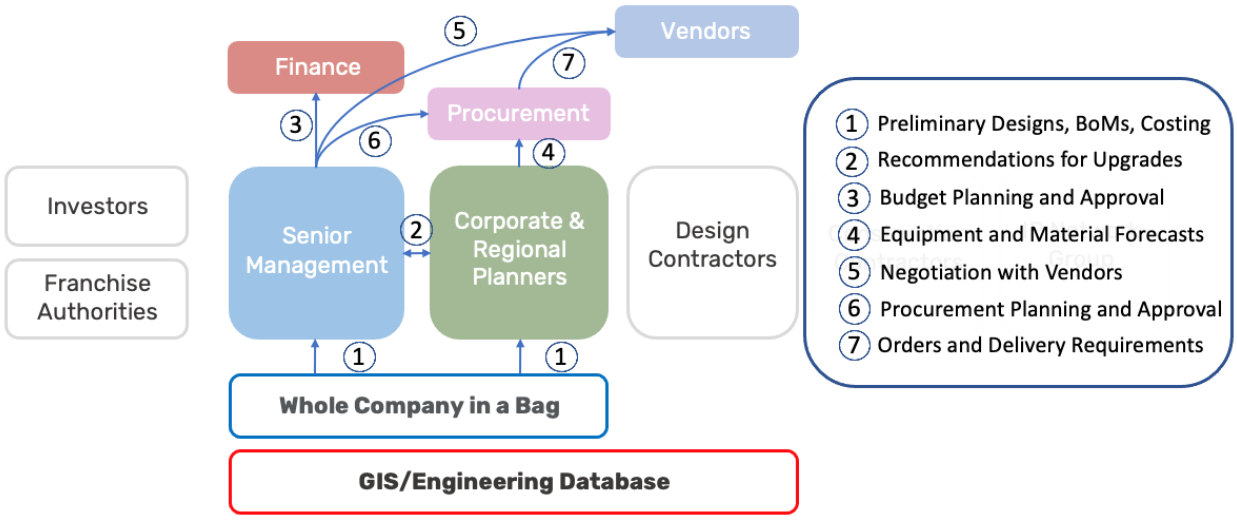


Figure 2 – Long Term Planning Activities

Traditionally, Mediacom has identified and upgraded single nodes to address congestion or other service issues. This approach has allowed Mediacom to keep up with customer demand but doesn't allow for any consideration of a more comprehensive and cost-effective approach to improving the surrounding larger portion of the network footprint.

The use of ADO allows Mediacom to do exactly that. For example, the entire serving area of a hub can be processed by the ADO platform practically as quickly as one or two nodes, allowing Mediacom to

identify opportunities to not only immediately upgrade problem areas, but to do so in a way that supports the eventual upgrade of the larger surrounding area and affords time and cost savings.

Mediacom expects this capability to become even more important when considered deployment of a Distributed Access Architecture (DAA) because the equipment complement in a hub will be reduced as DAA nodes are deployed in the access network. By using ADO to produce DAA preliminary designs for an entire hub serving area, Mediacom's access network planners can give their colleagues responsible for inside plant systems and technical facilities the information they need for their planning and deployment efforts. Similarly, Mediacom can then identify costs (and ideally overall cost savings) across its technical facilities, inside plant systems, and the access network when planning a DAA deployment.

An immediate benefit of having the Whole Company in a Bag is that serving areas that are well out of compliance with any current Mediacom network architectures can be easily identified when processed by the ADO platform against Mediacom's Fiber-Deep design rules (or, any other current architecture design rules) because the resulting cost estimates are higher than the average cost per node. This allows Mediacom to prioritize those network serving areas for upgrade or, if required, immediate maintenance work.

As noted in Section 3.2, the preliminary designs can be viewed in the ADO platform's user-interface in:

- map view, geographically on a scaled land-base, including a satellite photo base, and
- schematically, which is a simplified view that shows user-selected technical details, such as signal levels, active device cascade lengths, and cable lengths.

The various views allow Mediacom personnel to intuitively grasp the fundamental design characteristics of the proposed network itself and, by being able to see the network on a satellite photo base, assess its suitability to the physical topology of the serving area and distribution of the customer premises therein.

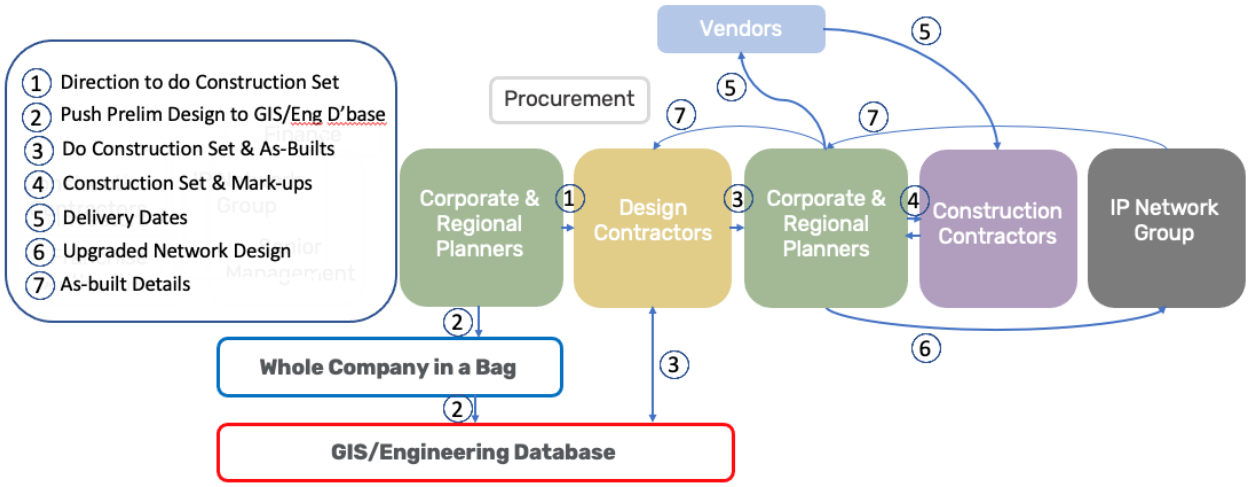


Figure 3 – Network Planning and Deployment Activities

Overall, Mediacom has seen a reduction in time for approval of network design for any area to be upgraded of two to four weeks. This is primarily due to having pre-approved preliminary designs for the entire footprint already 'in the bag'. Also, network planners at corporate and in the regions can see exactly the same designs at the same time, allowing them to collaborate very effectively to arrive at an agreement on exactly what network architecture to deploy by using the ADO platform.

It is noted that the aforementioned time savings do not include any time savings in the production of construction drawings expected by having the preliminary design from the ADO platform as an advanced starting point.

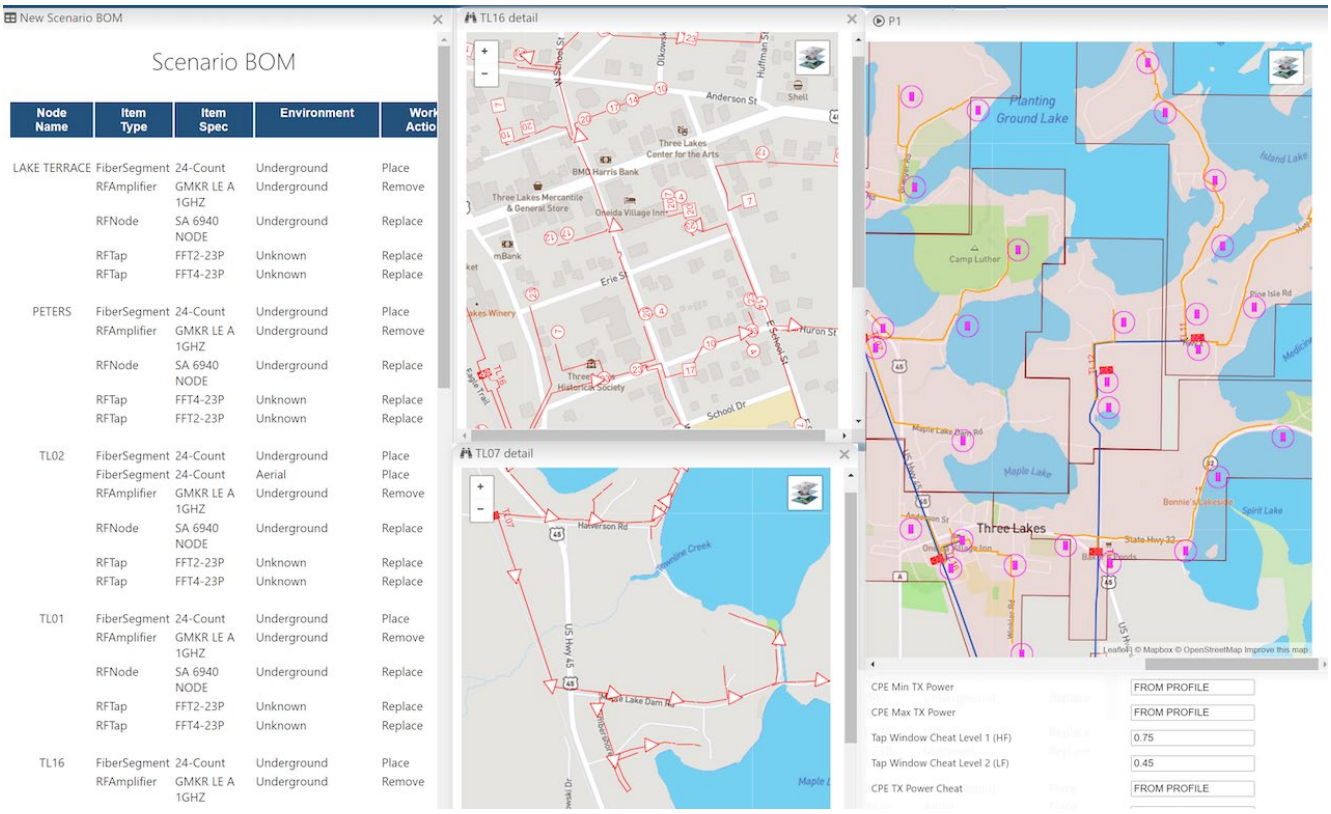


Figure 4 – Screenshot of User Interface – Network Map View

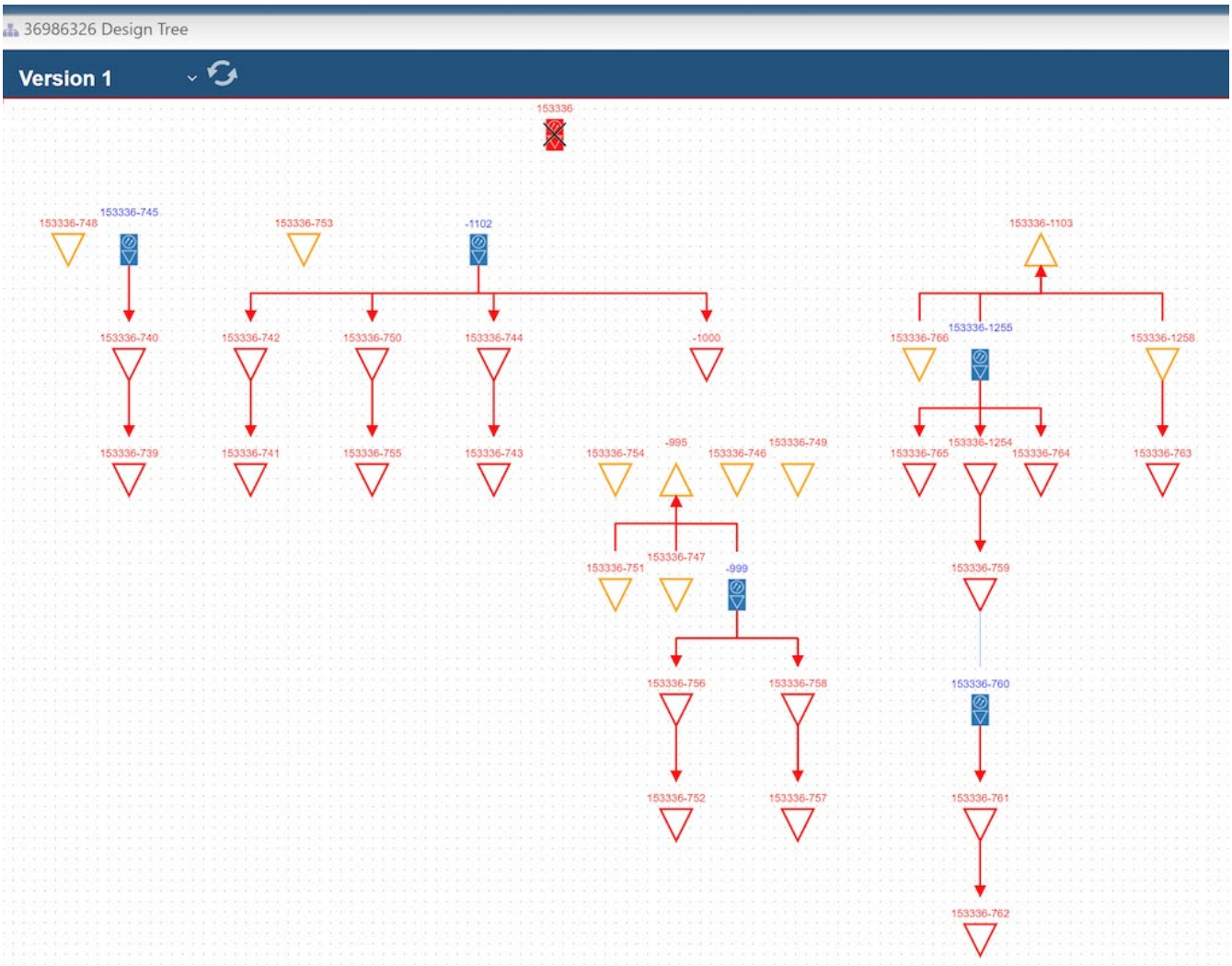


Figure 5 – Screenshot of User Interface – Network Schematic View

As much as Mediacom dedicates a great deal of time and money to ensuring its customers are very happy with their services, Mediacom must also ensure that its investors and its franchise authorities are confident that Mediacom is fulfilling the company's respective commitments.

By being able to provide investors with comprehensive insight into the company's capital expenditure and network upgrade plans, Mediacom can best maintain the confidence of the investor community, which is fundamental to ensuring access to capital funding.

Similarly, the franchise authorities must be confident that Mediacom is going to fulfill its obligations under the respective franchise agreements. Using ADO, Mediacom's ability to demonstrate (at an appropriate level of detail) its network upgrade plans and to, as a normal course of business, move increasingly quickly to maintain and improve the quality of the network.

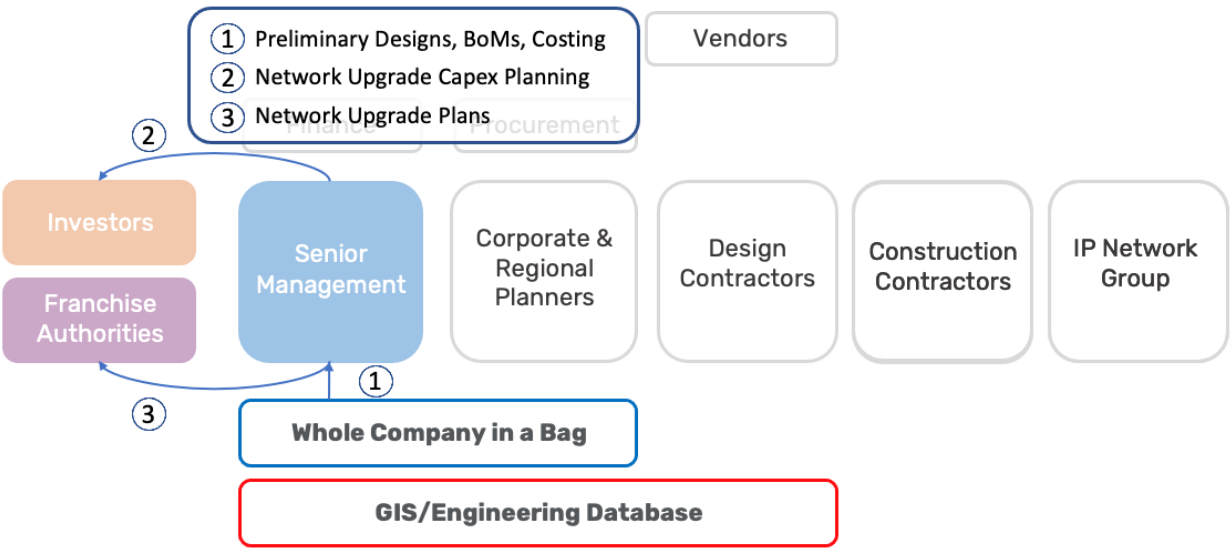


Figure 6 – Relations with Investors and Franchise Authorities

5. N+0 Planning – A Real World Example

In early 2021, Mediacom participated in a competitive bidding process that was set out by a mid-sized city wherein Mediacom is the incumbent service provider. The city solicited proposals from several service providers with regard to obtaining the highest possible bandwidth connectivity for all of the city's residents and businesses.

For the purposes of meeting the city's deadline for submitting its proposal, Mediacom utilized traditional network planning and cost estimation techniques for a N+0 network architecture. Subsequently, Mediacom utilized ADO to obtain more detailed and comprehensive designs and cost estimates in order to prepare for negotiations with the city. This was important given that the city was proposing to contribute funding to the network deployment effort. As a result, Mediacom was required to clearly demonstrate that its proposal would provide the most technically advanced service to its customers at the lowest possible cost.

In utilizing ADO, Mediacom was able to:

- quickly develop proposals for several different service levels, corresponding architectures, and respective costs thereby providing the city with a comprehensive range of options, instead of only one
- confidently compare (lower) capital costs of HFC upgrade options with the cost of building and FTTH network (as proposed by others)
- demonstrate that the deployment of its proposed network architectures would be much less physically disruptive to the community by virtue of maximizing the use of its existing access network
- be highly responsive to the city's requirements and strengthen the working relationship between itself and the responsible city officials.

Mediacom's application of ADO to the competitive proposal process demonstrated that having network designs and cost estimates for the entire access network footprint in the city, as opposed to only sample designs of parts of the footprint and costs extrapolated therefrom, allowed it to manage a significant competitive threat more effectively and confidently.

As noted previously, Mediacom operates under numerous franchises granted by the respective franchise authorities, which can prove challenging at times. In this situation, Mediacom was able to demonstrate a high degree of responsiveness to the franchise authority by providing a very detailed and well considered proposal which can only have a positive effect on Mediacom's working relationship with the franchise authority's personnel.

6. Deployment of ADO

Mediacom undertook a phased approach to the deployment of the ADO platform, with Phase 1 being to obtain 'the whole company in a bag' to support senior management in capital planning and other corporate-level activities, and Phase 2 being to implement ADO functionality for business-as-usual planning and design to support corporate and regional planning, design, and construction activities.

Phase 1 involved a number of activities including:

- development of Design Rules for Mediacom's Fiber-Deep architecture, which is Mediacom's default, or baseline, network architecture for purposes of corporate-wide capital planning and for considering any particular network area for upgrade,
- processing of Mediacom's by the ADO vendor of Mediacom's GIS/engineering database of the as-built network to produce Fiber-Deep preliminary designs and cost estimates for the entire network footprint, which were stored within the ADO platform's internal database,
- definition of Mediacom's requirements for querying the Fiber-Deep database, which were then implemented as functions within the user interface, and
- delivery of the user interface as a web-based portal into the Fiber-Deep database maintained on a server within the ADO vendor's IT infrastructure.

By pioneering the use of ADO technology, Mediacom was able to exert significant influence over the functionality and general design of the user interface and the Fiber-Deep database queries during the deployment process. This allowed Mediacom's senior management to obtain value from the ADO platform relatively early in the deployment process.

Phase 2, being focused on business-as-usual activities, involved:

- significant interaction with Mediacom's network planners at the corporate level and in the regions to ensure that the functionality provided by the ADO platform was truly useful in terms of time and cost savings in their day-to-day activities,
- implementation of an interface between the ADO platform and Mediacom's GIS/engineering platform so that the ADO platform
    - had on-demand access to the data therein describing the as-built access network configuration, thereby ensuring that any preliminary design produced in the ADO platform would be based on the best available information, and
    - can export selected preliminary designs to the GIS/engineering database so that network planners and design contractors could use those designs as the basis for completing detailed designs and construction drawing packages, thus saving a significant amount of time and effort in that process.

As with any major deployment of new technology, there were a few interesting, and some critical, factors to the successful implementation of the ADO platform, including:

- having to invest some time and expense to process the legacy computer-aided design (CAD) data within the existing GIS/engineering platform database into fully modelled data to ensure that the ADO platform could function optimally, which work was carried out in 2020 on the expectation that the ADO platform deployment would occur in 2021,
- working with the vendor of the GIS/engineering database to implement the interface between that system and the ADO platform so that the GIS/engineering database vendor could ensure that the ADO platform could read from and write to that database with no risk of data corruption,

- dedicating a reasonable amount of time on the part of a few senior Mediacom personnel on a weekly basis to working meetings and technical discussions to, most importantly, convey Mediacom's technical and operational requirements to the ADO platform vendor,
- revising the deployment plan at the kick-off meeting to include the N+0 planning work (christened Phase 0) described in Section 5 above, which involved developing N+0 Design Rules and applying them to the as-built network in the city in question to produce two sets of preliminary designs and cost estimates within 90 days
- engaging in regular discussion within Mediacom and with the ADO platform vendor to develop a working understanding of how Mediacom's internal processes will be improved, and
- communicating process changes to Mediacom personnel not originally involved in the deployment of the ADO platform but, nonetheless, affected by its deployment and adoption.

Phase 0, the N+0 planning work was executed in Q1 of 2021 while Phase 1 was executed in Q2/Q3 and Phase 2 in Q3.

Mediacom anticipates continued expansion of the ADO platform in terms of network design capability, e.g., FTTH, and supporting a broader group of users that would benefit by having access to the ADO platform. It will certainly take some time and effort to bring the broader user group onto the ADO platform, if only because people are naturally reticent to change the way they do their jobs, but the benefits to be gained by using ADO technology are expected to far outweigh the adoption costs.

7. Summary and Conclusions

Mediacom has found that the adoption of ADO technology does enable evaluation of the technical and capital costs characteristics of access network designs much more comprehensively and quickly than was previously possible.

While it was expected from the outset that ADO technology would immediately benefit those directly involved in network planning and engineering. It became increasingly apparent that the use of ADO technology – specifically, the data within the preliminary designs and cost estimates it produces so rapidly – has conferred benefits on practically every group within Mediacom involved in the funding, design, construction, and operation of the access network.

The increase in confidence and the saving of time afforded by having easy access to preliminary designs and cost estimates for multiple network architectures for network areas as small as a single node and as large as the entire footprint has allowed Mediacom to execute its access network upgrade strategy significantly more efficiently and effectively.

Interestingly, Mediacom's ability to work with external groups – investors, franchise authorities, and vendors – has also been enhanced by being able to provide clearly and consistently each of those groups with information they need to best understand and support Mediacom's access network upgrade strategy.

As Mediacom gains more experience with ADO technology, it is anticipating opportunities for automation of the design of power distribution for HFC networks and of FTTH networks.

Finally, it seems that, whereas 'having the whole company in a bag' was coined to refer to having at hand designs covering all of Mediacom's access network footprint, it more accurately means having designs for entire footprint in one place AND having everybody in the company who cares about the access network working together better than ever.

Abbreviations

ADO	automated design and optimization
CAD	Computer-aided design
CMTS	cable modem termination system
FTTH	fiber to the Home
GIS	geospatial information system
IP	internet protocol
IT	information technology
N+0	node plus zero active RF devices in cascade on a single coaxial cable leg of access network
N+2	node plus 2 or less active RF devices in cascade on a single coaxial cable leg of access network
N+X	node plus 0 or more active RF devices in cascade on a single coaxial cable leg of access network
RF	radio frequency
Q1	first quarter of calendar year
Q2	second quarter of calendar year
Q3	third quarter of calendar year
Q4	fourth quarter of calendar year

Helm: Self-Service Customer Data Platform

A Technical Paper prepared for SCTE by

Sriharsha Gangam
Principal Architect
Comcast Cable
1800 Arch St, Philadelphia, PA 19103
+1 215 583 8078
sriharsha_gangam@comcast.com

1. Introduction

For any organization or business, the importance of good customer experience cannot be overstated. Capturing and understanding customer interactions with the business is the first step towards empathizing with customers. For large organizations and conglomerates with millions of customers and multiple domains or business areas, typically the data is scattered across several business units. It becomes challenging to understand the holistic customer experience. For example, the data from domains such as billing, customer care, audit, and operational interactions could be generated by siloed business units. To understand customer experiences, it is important have these data points stitched together -- particularly when there are multiple products and services available to a customer.

These challenges are industry-agnostic and apply to industries beyond cable and telecommunications. For example, in the retail industry, it is important to understand customer interactions for products and services across multiple channels (e.g., marketing, orders, billing, shipment and call-center interactions). A customer might purchase certain merchandise online, exchange it with another product at a retail store, and subscribe to a new monthly payment installment plan. It is expected that all these interactions are captured and linked together to serve and assist customers. Similarly, in healthcare, it becomes important to have a comprehensive view of health records for all patient interactions, even as the health data could originate from several independent health care providers.

To address these goals, organizations build customer data platforms (CDP). A CDP consolidates and integrates customer interaction data from multiple domains to build and present a unified profile around each customer. At the heart of every CDP is a metadata management system to enable ingestion, governance, and access to the data. Once the data is available, it enables one to gain insights into customer experiences, troubleshoot problems, and interact with customers to improve their experiences.

While it is important to be data-centric, the dynamic nature and increasing variety of customer interaction datasets introduces new challenges. It is ever more important to manage customer data in a safe, secure, and self-service manner. While rich applications can be built on a CDP, data governance and ownership responsibilities can be overwhelming if they are not democratized and decentralized.

In this paper, we will present Helm, which is the name of one of Comcast's CDPs that manages data ingestion, processing, and consumption of anonymized customer data for purposes of advanced issue resolution. We share our experiences to understand what it takes to effectively build and maintain a CDP. The Helm platform ingests hundreds of event datasets from multiple domains, such as customer care, audit, billing, device activations and operations. To put the scale in context, the platform captures tens of billions of customer interactions from millions of customers every month. Internally, it is directly used by tens of thousands of employees to serve several hundred thousand customers, every month. An important component, Helm application programming interfaces (API), play a crucial role across several products in Comcast, serving several billions of API requests every month.

The paper is organized as follows: we will begin by introducing some of the challenges in building, maintaining, and governing CDPs. We then present an overview of the Helm CDP and its impact on customer experiences. This is followed by the platform architecture and design principles. We conclude the paper with a look at Helm’s journey so far, and possible future states.

2. Helm Overview

Everything discussed in this paper as “Helm” dates back to the 2012 timeframe, when a group of Comcast employees developed a “Lab Week” project, called “Timeline,” to help us to better stay in touch with our

customers who experience service issues. During its conception and the early stages, Helm was composed of bespoke data engineering extract, transform and load (ETL) steps. Each dataset had separate plumbing (data pipelines and code) to extract data from external database systems and apply custom business logic to transform the data. This was hard to maintain, as the data evolved over time. The Helm organization was responsible for data ownership, management, and governance of these datasets. The lack of domain expertise made it challenging for consumers to leverage the data. As the number of datasets grew in terms of domains and variety, data onboarding was hard to scale up, due to limited team capacities. This approach became untenable.

Helm Self-Service was built to address these problems. It represents the control plane or management layer of the Helm CDP. With it, data originating teams can define metadata and configurations about the datasets they publish into Helm. Data producer teams tend to be data business owners, data managers, and domain experts; the platform enables a seamless communication between the domain experts and data consumers. Dataset onboarding and governance deliberately do not require engineering effort by the Helm team. Instead, they are consistent, decentralized and democratized across business domains. As of this writing, a few hundred datasets are onboarded and made available in Helm. Several datasets are modified or added every week. Helm Self-Service user interface (UI) and APIs made this possible.

The Helm team has been in pursuit of continuously improving the Helm Self-Service user experience. Helm is integrated with Matomo (*https://matomo.org/*), InMoment (*https://inmoment.com/*), and Intercom (*https://www.intercom.com/*) to capture usage metrics and product net promoter score (NPS). Data onboarding and metadata management can have several configuration steps and occasionally require additional support from Helm administrators. The Helm success team leverages tools such as Atlassian (*https://www.atlassian.com/*) service desk, Slack (*https://www.slack.com/*) and Intercom for assisting Helm users.

Data's value is only as good as its quality. With Self-Service, data quality control responsibilities are shared with data producers. Helm is integrated with DataDog (*https://www.datadoghq.com/*) to expose data and platform observability metrics. Data availability (e.g., missing data) and quality (e.g., unexpected, or missing attributes) metrics are captured, summarized, and saved for historical lookups. These metrics are sent back to the data owners to resolve any data quality and availability issues. In addition to data quality, a metadata completeness score is captured to ensure that the metadata (e.g., attribute level documentation, tags) is up-to-date and accurate (see Figure 1). This way, Helm users can browse, understand, and consume data correctly as it evolves. Helm as a CDP platform publishes its own platform metrics, logs, and traces into DataDog to enable operational support and expose Helm's service level objectives (SLO).

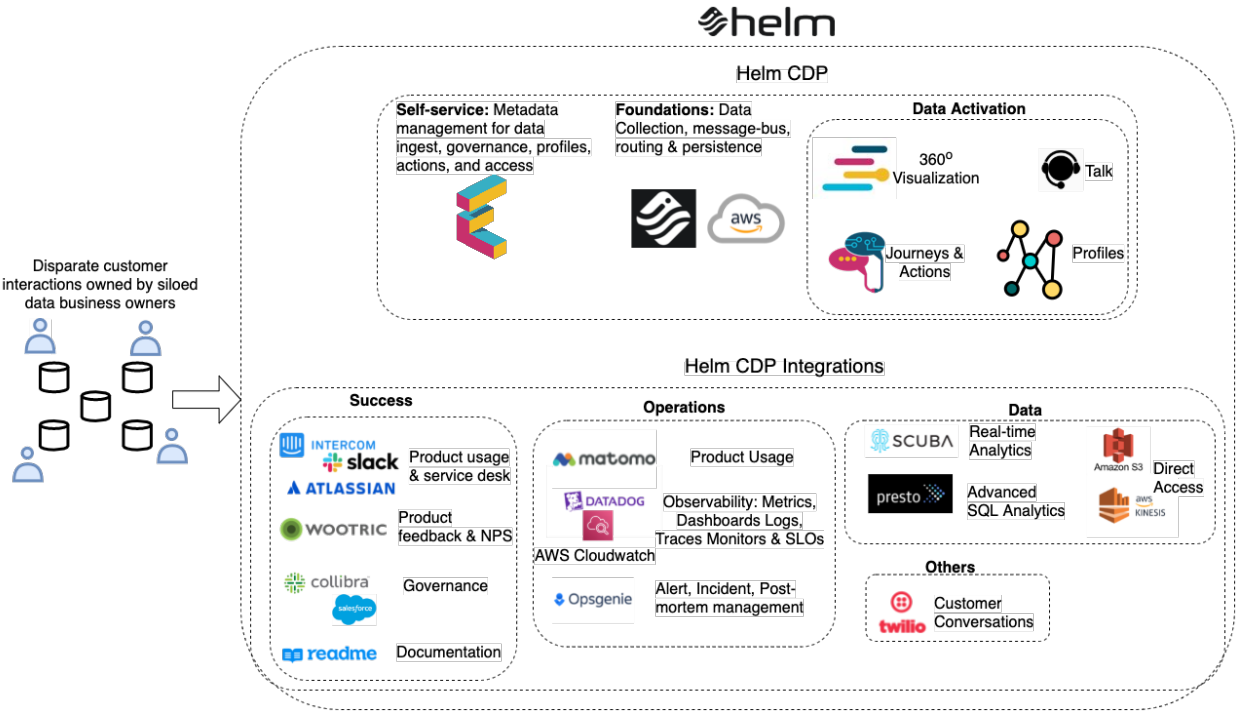


Figure 1 – Helm Overview

Once high-fidelity customer data is available, the next step is to enhance customer experiences and bring business value. Helm Applications is collection of packaged services that is included and integrated with the Helm CDP. The applications are a 360° Visualization application (timeseries visualization), Talk (customer conversations), and Analytics. They are designed to work together, draw out valuable data insights, and enable users (employees) to deliver the best customer experiences. To enable and support a variety of applications, data organization is foundational for CDPs. The next section introduces the Helm data model, which describes how data is stored within Helm. This is followed by an overview of Helm Self-Service for ingesting, managing, and governing the data.

Data Model

A profile (or entity) is the principal or thing around which timeseries event data is captured. For example, a "customer" profile is associated with a variety of customer interaction timeseries datasets such as product activations, billing data, or troubleshooting tickets. Such timeseries datasets could be originating in different business units within an organization. Each entry in these event datasets would have a unique customer identifier (the profile identifier) to link back to the customer profile, a timestamp to understand event chronology, and other attributes relevant to the specific dataset type. A product activation dataset could have additional attributes to capture a product activation experiences (e.g., attributes about the type product or service, lead time to activate the product, and activation failures).

The Helm data model supports different types profiles that may or may not have relationships with each other (see Figure 2). In addition to the customer profile, the "account" and "device" profiles are relevant in the cable and telecommunications domain. Every profile has a unique identifier and certain attributes. A customer's profile could have essential attributes to lookup, verify and assist a customer such as the name and contact information. An account profile could have attributes such the account activation date, service type, and a unique account number. If customers can open multiple accounts, this is represented by a one-to-many relationship between a customer profile and each opened account. Certain timeseries

event datasets, such as truck roll schedules, may apply to accounts rather than customers. For each account, multiple devices could be registered that generate device availability events. These datasets directly apply to devices (not customers or accounts). It is important to define and maintain profile relationships (*i.e.,* relationships between customers, accounts, and devices) to understand, navigate, and correlate events across datasets. These profile relationships are foundational in Helm and are represented by a graph – the profile graph. The profile graph along with timeseries datasets for each profile type constitute the Helm data model.

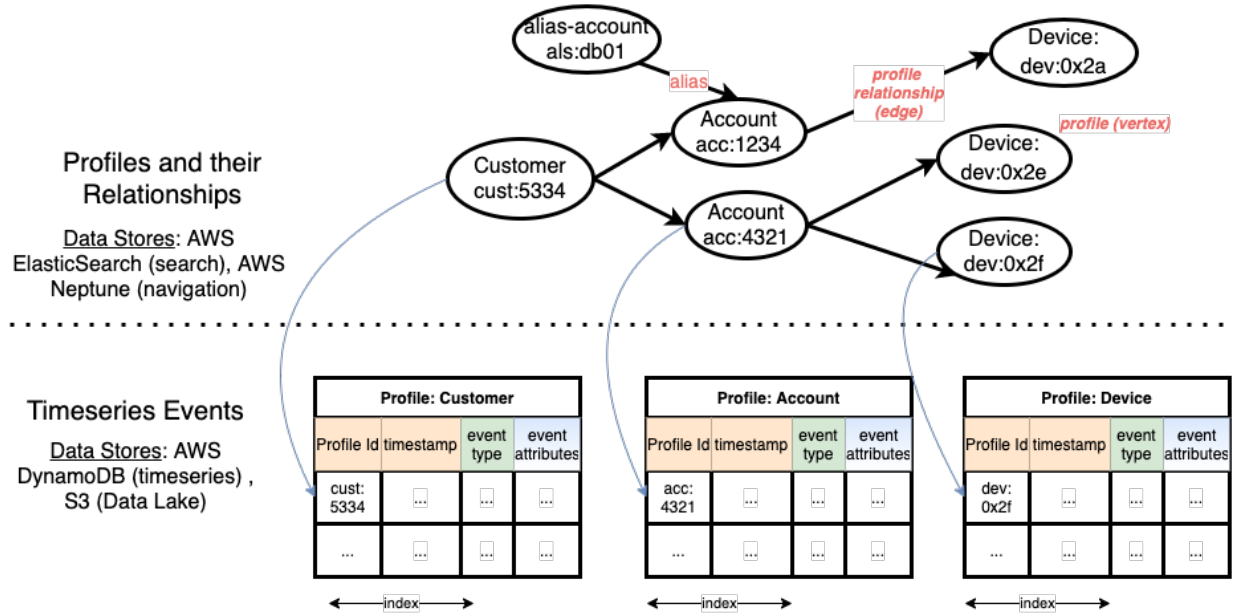


Figure 2 – Example Helm Data Model

2.1. Self-Service

First and foremost, a CDP needs to provide services to capture data effortlessly. Data owners across the organization should be able to publish data into Helm with minimal supervision. Data owners use Helm's Self-Service portal to define schema and other onboarding metadata. While the dataset schema is determined by the data producer and the business context, it must adhere to Helm's data model. It must have a timestamp (for timeseries data), a profile identifier, and additional attributes. Once a dataset is onboarded, the Helm platform will orchestrate the creation of specific infrastructure to enable data ingestion. Orchestration creates configurations to define and manage the data plane. For example, a dedicated hypertext transfer protocol secure (HTTPS) API endpoint is made available for the accepting the data. Additional application-specific orchestrations are executed when necessary.

As stated previously, Helm Self-Service represents the control plane or management layer of Helm CDP. In addition to incoming schema metadata, it enables metadata capture of transformations, routing, persistence, lifecycle, and access controls. The data producers configure the following metadata while onboarding a dataset (See figure 3):

Self-Service Metadata – Control Plane

- **Source Metadata**: Includes all the metadata relevant to a source dataset; e.g., security credentials, feed schema, and documentation. Security credentials ensure that only authorized users and applications produce data into the Helm platform. Data onboarding requires producers

to configure necessary credentials upfront. For real-time data ingest, Helm supports client authentication and authorization using OpenID connect (OIDC) and open authorization, version 2.0 (OAuth 2.0). The feed schema defines the structure of data published into Helm. It includes expected attribute names [JavaScript object notation (JSON) paths], expected data types, and expected values as per JSON schema specifications. The feed schema is used to validate ingested data and errors are communicated back to the data producer. The data producer will either need to fix the schema or the data being published. Every feed schema is accompanied by documentation that describes the dataset and attribute definitions with necessary business context. Documentation is important to build good data dictionaries. Helm users can search the dictionary to identify datasets of interest.

- **Transformation Metadata**: Includes all metadata definitions to transform the ingested data; e.g., normalization and encryption schemas. The encryption schema identifies sensitive attributes that require encryption. Helm encrypts these attributes while ingesting the data. Decryption of these attributes is restricted to privileged users and applications. The normalization schema defines naming and structural transformation specifications. It determines a target data schema that is persisted in Helm. Often times, data owners would like to change the name of an attribute or the structure of data stored in Helm without depending on the source dataset. This decouples data production from consumption and provides flexibility when data producers are unable to update their schemas. However, historical data processing can become tedious if the attribute names are updated over time. For example, analytical SQL queries would require multi-part queries, one part for each attribute name. Instead, the datasets flowing into the Helm data lake (for analytics use cases) are transformed. The transformation reverts attributes to their oldest (unique) registered names. This ensures consistent attribute names across historical data.
- **Routing, Persistence and Access Metadata**: The routing metadata for a dataset determines if one or more Helm applications (such as 360º Visualization, Analytics, or watchlists) are subscribed to the dataset. Not all datasets are suitable for every application and use case. The access metadata defines fine-grained access control policies for datasets and attributes based on their sensitivity levels. Persistence metadata defines the data lifecycle and retention goals for the dataset. These are influenced by an organization's data compliance and governance policies. Retention and lifecycle management features of underlying data stores are leveraged when applicable.
- **Application Metadata**: Helm is integrated with applications such as 360º Visualization, Talk, Watchlist, Profile Search & Navigation, and Analytics. Self-Service extends to these applications and enables users to configure them. For example, users can define icons for every data type in 360º Visualization. The profile schema metadata is applicable when the dataset is used to build a profile graph or used to define profile translations. It contains profile attributes, identifiers, and relationships between profile types. It also determines keyword-searchable attributes for a profile.

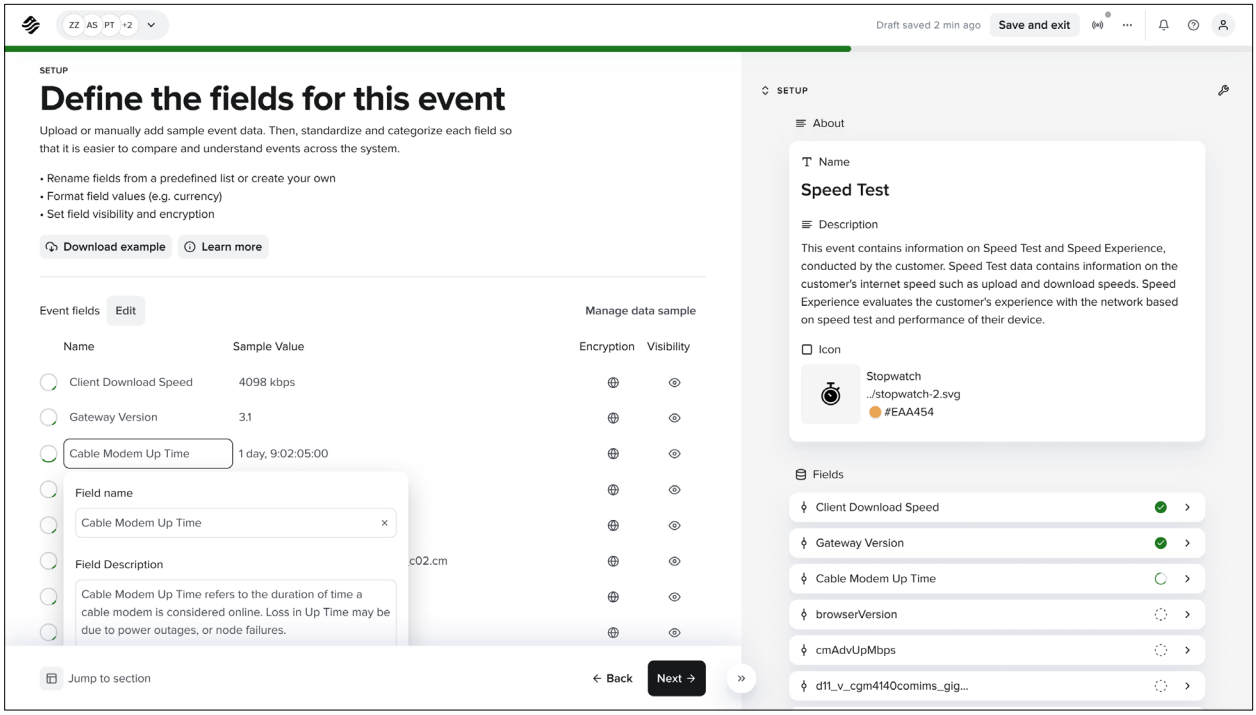


Figure 3 – Helm Self Service Data Onboarding

Ingestion & Persistence – Data Plane

After a dataset is registered in Helm Self-Service, the producer can publish the data into Helm. For ingesting real-time data into Helm, a dedicated HTTPS endpoint is created for each onboarded dataset. In this scenario, the ingested data is available for consumption in near real-time. For batch use cases (e.g., data is produced hourly or daily), data producers can directly push files to a dedicated Amazon Web Services (AWS) S3 bucket using temporary credentials generated by Helm's API. Unlike real-time ingestion, batch ingestion delays are less predictable due to their underlying nature (e.g., asynchronous and throughput-optimized).

For each event datum received from the producer, Helm performs a series of validation steps as per the security credentials and feed schema defined in Self-Service. The datum is then transformed in a sequence of steps such as normalization and encryption (if necessary). If any of these steps fail, the datum is rejected, and the data producer is notified. Helm provides live and historical observability of failures and rejected data. If all the steps execute successfully, the datum is accepted and is written to Helm's message bus for routing and consumption.

A datum's journey from the message bus is determined by the routing metadata configured in Self-Service. Helm applications consume this datum and may perform certain compute operations (e.g., sessionize, evaluate rules, aggregate, enrich, or store) specific to the application. Finally, the datum is persisted in datastores appropriate for the query use case. To enable profile keyword search and navigation based on profile attributes and relationships, the datum is stored in a modern graph database (AWS Neptune) and search database (AWS Elasticsearch). Similarly, to support real-time event lookups (timeseries data) for a profile, the datum is stored in a document store (AWS DynamoDB). For long term trends and aggregate analysis, the data is batched and written into the Helm data lake (AWS S3 and AWS Glue).

3. Data Activations

In the previous sections, we provided an overview of the Helm CDP and the importance of Self-Service in scaling the platform, and the underlying data model. In this section, we describe how employees engage with customers to enhance customer experience using Helm applications like Profile Search & Navigation, 360º Visualization, Talk, Journeys & Actions, and Analytics. Datasets that are onboarded once are reusable across these applications. This section describes how these applications complement each other and provide value that is more than the sum of its parts. Helm datasets are made available to non-Helm applications via data integrations and APIs.

3.1. Profile Search and Navigation

Once all the relevant datasets are onboarded and ingested into a CDP securely, the platform provides capabilities to query and serve this data. Helm's Profile Search & Navigation helps employees use search keywords to identify customers. For example, a care agent could use a customer's street address or a device's media access control (MAC) address to identify a customer profile. After identifying the correct profile, the account and device timeseries event data can be explored using 360º Visualization. A hypertext markup language (HTML) form accepts a keyword and returns all profiles that match the keyword. Auto-suggest results are provided for every keystroke, to enhance the search experience. The profile search application is served by APIs for programmatic integrations. The API accepts keyword requests to match with searchable profile attributes and returns a collection of matched profiles. It also accepts a profile identifier and allows one to navigate profile relationships and traverse to the profile graph.

The Profile API also aids in data ingestion, such as when a dataset does not have expected profile identifiers. For example, Comcast has several biller integrations. Depending on the services, customers could have multiple accounts, one for each biller. To serve customers effectively, it is important to transform biller-specific datasets to a common account identifier. Using profile relationships between biller accounts, the ingestion pipeline can transform the account identifiers of a dataset.

3.2. 360º Visualization

360º Visualization provides an intuitive and exploratory experience for understanding customer (profile) interactions. Customer interactions for a given customer's globally unique identifier (GUID) (i.e., profile identifier) occurring within a time range are presented in a single, chronological view. This "lifetime view" enables employees to serve customers more effectively, because they can see, sequentially, all current and previous interactions, without having to access disparate databases and systems. An employee can understand the customer's journey with historical data, as well as serve customers with real-time status and information. The visualization enables employees to seamlessly drill-down to a specific customer event of interest, from what can be many interactions associated with that customer. Figure 4 shows a screenshot of 360º Visualization, where events are represented by icons. In practice, clicking on these icons expands the event with necessary business context. The range-selector box (black) at the top of the screen helps users choose the timeframe of interest. The "swim lanes" and event filter selections enable one to hide irrelevant events. Visualization user preferences can be customized and saved.

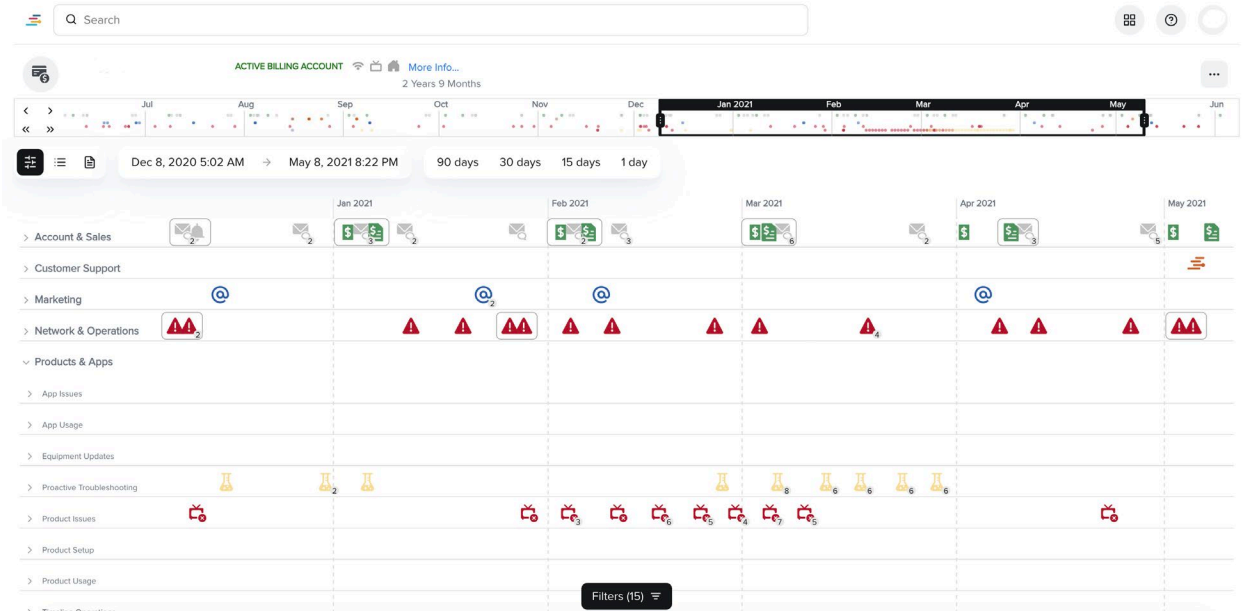


Figure 4 – 360º Visualization - Chronological View of Customer Interactions

Behind the scenes, the UI is backed by the Helm Events API. The Events API accepts a profile identifier, profile type, a collection of event types (optional), and a timestamp range (optional) and returns all events occurring within the time frame. These events are fetched from Helm's timeseries document store that is indexed on profile identifiers and timestamps. The Events API integrates and serves several critical Comcast applications external to Helm. At the time this paper was written, the Events API serves several billion API calls monthly, and the 360º Visualization UI serves several hundred thousand visits every month.

3.3. Talk

As employees troubleshoot customer concerns in 360º Visualization, they might want to reach out to the customer. Helm is integrated with "Talk", a collection of services that initiate short message service (SMS) messages, email, chat, or phone conversations with customers. This contact information is made available from Helm Profiles. No matter who or how many care agents communicate with a customer, all customer conversations are captured as events and ingested back into Helm. The 360º Visualization application allows one to review all these conversations in context of a customer's overall experience. Conversations with a shared context across channels and care agents are captured in a single place. Conversations are seamless, and care agents assisting the customer are literally all on the same page.

Helm Talk has played an important role with the NPS callback program at Comcast. Customer profiles with unsatisfactory NPS surveys are made available to certain employees (via Helm Analytics). The employees are required to proactively reach out to these customers to understand and address their concerns. With the integrated Helm Talk and 360º Visualization capabilities, employees can effortlessly review and initiate customer conversations.

Behind the scenes, Helm integrates with Twilio APIs (*https://www.twilio.com/*) to provide these communication capabilities. During the Covid-19 pandemic, Helm Talk capabilities were expanded to support video calls. This enabled our workforce to extend troubleshooting capabilities with audio and video communication without entering customer homes. This success story was feature in Comcast's corporate blog (*https://corporate.comcast.com/stories/digital-solutions-keeping-us-connected-covid-19*).

3.4. Journeys & Actions

Once customer troubleshooting is completed via the 360º Visualization and Talk applications, a care agent might wish to get notifications (triggered by specific events) about a customer, to ensure that there are no chronic issues. Similarly, care agents and customers might want to get notifications for planned network outages occurring within a certain neighborhood. These examples require configuration of event-driven rules and actions.

An event-driven rule defines the business criteria when a certain action (e.g., a customer notification) is performed. In the above example, the rule evaluation determines if an account belongs to a specific neighborhood and if the interaction type is a network outage. Information to evaluate a rule is typically available as attributes within the event and is enriched by external systems if necessary. In theory, actions could be defined to trigger any external service API call making them extensible.

The above examples do not consider a customer's historical events or prior context for evaluating event-driven rules. For example, a care agent might be interested in truck roll notifications that are re-scheduled more than twice. This requires information about a customer's journey or historical state to evaluate rules and execute actions. A "journey," in this context, is a snapshot of a customer's experience, built from a collection of seemingly disparate or distinct events occurring within a time period. A customer's truck roll journey state machine, for instance, could transition between the following states: Created, scheduled, completed, cancelled, or re-scheduled (see Figure 5). As individual events (e.g., "created," "scheduled") are observed, the workflow progresses, and the journey state is transitioned to a next one.

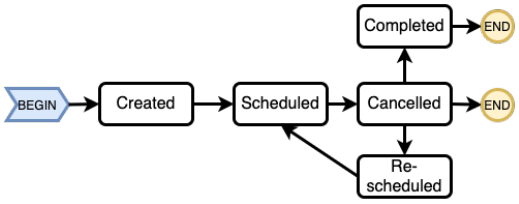


Figure 5 – An Example Truck Roll Journey

In 360º Visualization, care agents would need to mentally re-create a customer's journey by looking at individual historical events to understand the overall context. For complex workflows, this incurs a heavy cognitive load, every time. A journey state machine visualization aids this understanding and presents a customer's journey that could be understood with a single glance. The journey state machine is defined by business process flows and managed with workflow engines such as Flowable (*https://flowable.com/*). Each state transition creates an event in Helm representing the source and target states. By defining event-driven rules on journey state transition events, complex business rules and actions can be accommodated. Rule evaluation remains a stateless operation and is decoupled from journey state management.

Helm is integrated with 'Watchlist', an application that lets employees get notifications about customer events. This is achieved with a few clicks from the 360º Visualization experience. For example, a watchlist rule could be - "For new network outage events matching account numbers in the list [0003, 0004, 0005], trigger an SMS & email notification". The subscribed employees receive watchlist notifications via email or text (the action) along with the customer and event details that triggered the notification.

3.5. Analytics

It is valuable to review a single customer's experience in 360º Visualization or Journeys, but this can be hard to scale. With millions of customers, it becomes important to understand aggregated customer experiences across the population. This is accomplished with Helm Analytics, which provides both business insights and analytics capabilities. For example, it is useful to understand the population distributions at each step of a business workflow journey. This can be visualized with Sankey diagrams and infographics as shown in Figure 6. This helps business teams identify common problems across customers, and invest in areas that provide the most value. Product owners can understand the impact of their feature rollouts on customer journeys by correlating with product NPS. This fast feedback enables organizations to adapt and innovate with short lead times. Similar arguments apply for initiatives that aim to reduce an organization's budget expenses without impacting the customer experience.

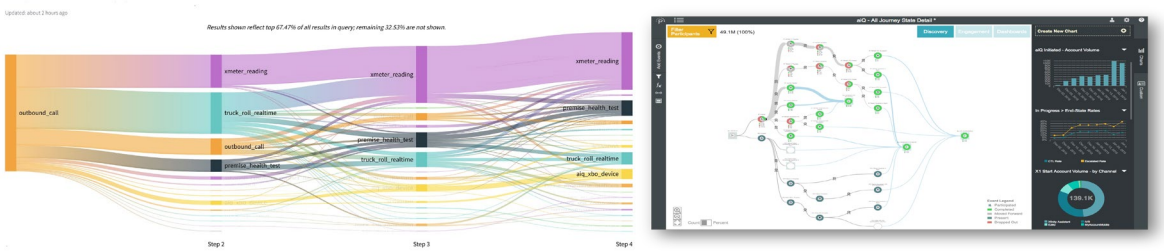


Figure 6 – Journey Analytics from Scuba and Pointillist

The NPS callback program at Comcast lets certain employees query Helm Analytics for customers with a low NPS. With a list of such customers, employees could review events in 360º Visualization, follow up with Talk, and query for customers with similar problems. These complementary applications provide the flexibility to switch context from an aggregate population view to individual customers and vice-versa, and demonstrate the immense value of Helm CDP at Comcast.

Datasets within Helm are processed and made available for analytical queries using modern data stores and analytical query engines. Users are exposed with a UI & API for interfacing with the underlying query engine. Like 360º Visualization, data access is managed using user groups and roles. At Comcast, Helm is integrated with Scuba (*https://www.scuba.io/*) and Pointillist (*https://www.pointillist.com/*) to provide near real-time analytical query capabilities. The platform serves tens of thousands of queries every month from thousands of users. The underlying datasets are partitioned and managed in Helm's data lake (backed by AWS S3andAWS Glue) to enable integration with other analytics products.

4. Helm Architecture & Design

In the previous sections, we provided an overview of Helm and showed how the applications activate data to improve customer experiences. In this section, we will present the building blocks of the Helm platform. This section details the design principles guiding the architecture, an architectural overview, and the platform's evolution over the years.

4.1. Design Principles

Below are some of the design principles that helped shape Helm's current architecture.

1. ***Metadata Driven, Domain Agnostic, and Multi-Tenant Capable***: Configurability was an essential architectural consideration, because it enables the re-use of functionality or services. Self-Service is intended to manage all metadata configurations per dataset. Observability metrics

are provided to understand the impact of configuration changes on the datasets. Platform configurability enables Helm to be domain- and industry-agnostic, and, as a result, to support multiple business verticals and industries. Profile types, their relationships, and associated time series datasets can be defined for each domain. Branding and user experiences (e.g., 360° Visualization swim-lanes) are configurable. Organization-specific integrations could be built with Helm SPIs (service provider interface), and as such support multi-tenancy by providing dedicated Helm instances in tenant namespaces. A Helm instance is created by cloning the infrastructure in a tenant namespace and applying tenant configurations. Helm is in the process of transforming into a shared-services, Software-as-a-Service (SaaS) model to reduce the time necessary to onboard new tenants.

2. ***Near Real-time Streaming with High-Fidelity***: Helm is designed for near real-time data ingestion and stream processing applications. Delays are unacceptable, from a usability standpoint, in applications like 360° Visualization and Watchlist (for troubleshooting and notifications). Certain other applications (e.g., Analytics) may be more delay-tolerant, especially if data efficiency is the priority, in which case data can be processed in batches. Data validation errors are immediately made available to the data producers. This enables data fidelity, creates a fast feedback loop with the data producer, and promotes dataset evolution.

3. ***Cloud Infrastructure and Performance Isolation***: Helm is designed to leverage cloud efficiencies, having been built using several managed and serverless AWS technologies (e.g., Beanstalk, Lambda, Kinesis, Firehose, S3, IAM, KMS, DynamoDB, Elasticsearch, Neptune, and RDS). Use of the public cloud as a technology choice is influenced by factors including corporate technology guidelines, Helm's business and strategic objectives, and personnel budgets.
Helm is designed to provide isolation and performance guarantees per dataset. For example, infrastructure orchestration (via Self-Service) creates a dedicated stream (AWS Kinesis) and compute infrastructure (AWS Lambda) for each dataset. The downside is the additional overhead for managing this infrastructure. The trade-off could be tuned by enabling shared infrastructure configurations. For example, a message bus with 's' streams could share 'd' datasets, where $1 \leq s \leq d$.

4. ***Delivery and Access Guarantees***: The Helm ingestion pipeline provides "at least once" delivery guarantee. "Exactly once" delivery requires every component in the data pipeline to guarantee it, and is generally hard to achieve. Assuming it is achievable, a data producer might still publish duplicate data into Helm. Instead, Helm provides "at least once" delivery guarantees and requires consumer applications to be idempotent. For 360° Visualization, the data store overwrites an event record if it receives another datum with the same idempotency key. The idempotency key uniquely identifies an event and is generated from a strong one-way hash function of event datum attributes. Similarly, for applications such as Journeys and Actions, the underlying state machines are designed to be idempotent. All aspects of Helm access, the product UIs, APIs, data, AWS services, and third-party integrations are secured by access controls with the principle of least privilege. All data is encrypted in transit and at rest. Additionally, sensitive attributes are encrypted at the application layer. Only authorized users and applications can decrypt these attributes.

4.2. Architecture Overview

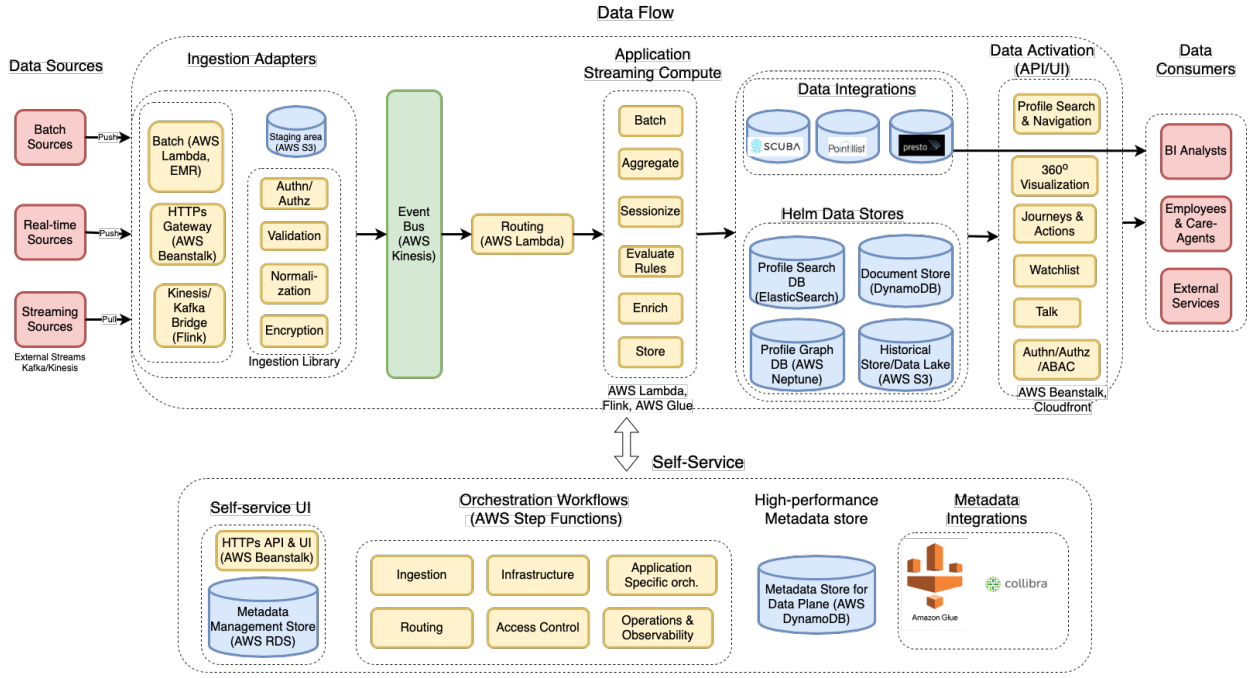


Figure 7 – Helm Architecture

Helm's high-level architecture, illustrated in Figure 7, starts with the foundational Self-Service control plane for metadata management. Self-Service is intended to manage datasets for all aspects of the data flow, from ingestion into Helm to consumption from Helm. Dataset registration in Self-Service triggers a workflow of orchestrations to enable ingestion, routing, infrastructure, persistence, access, and observability.

Ingestion adapters receive data from real-time, batch, and streaming sources. Data ingestion includes a sequence of steps such as authorization, normalization, encryption, and publishing to the Helm message bus. These steps are implemented as shared libraries and are reusable across different compute engines (AWS Beanstalk, EMR, Flink). A datum is said to be accepted in Helm only after it is written to the message bus.

Data from the message bus is consumed and written to Helm's historical data store (AWS S3). A dataset's routing configuration determines if it is consumed by one or more Helm applications. These applications and data integrations (e.g., Scuba, Pointillist) process, persist, and provide access to the data (via APIs and UIs). Applications consume the datasets independently using streaming compute (e.g., they transform, aggregate, batch, evaluate rules, sessionize, enrich, or store). For example, certain timeseries datasets are sessionized to reduce noise in 360° Visualization before persisting. For 'Watchlist', matching rules are evaluated to trigger notifications. Profile datasets are transformed into vertex and edge representations for storing in a graph database.

Occasionally, there are no data producers to publish profile datasets into Helm. In such cases, timeseries event data is processed to identify profiles missing in Helm. Profile-specific SPIs are used to fetch these profiles from source systems and ingest into Helm. These SPIs are also used to fetch profile updates from source systems. Similarly, certain timeseries datasets do not have data producers. Helm maintains a

collection of scheduled jobs to periodically fetch these datasets from source systems and ingest them into Helm.

4.3. Helm Evolution - Past and Future

Helm was initially conceptualized to support 360º Visualization. Back then, the platform was simply called "Customer Timeline". The Timeline web-application was originally built on Comcast's cloud infrastructure. It was designed to capture and visualize customer interaction (timeseries) data. Dataset onboarding required engineering new data pipelines to fetch data from source systems. It was time consuming to onboard new datasets. The 'Timeline API' enabled data access for external applications. Around this time, initiatives for migrating on-premise services to AWS public cloud had begun. Self-Service was introduced to catalog and manage datasets. Self-Service enabled dataset orchestrations and auto-created data pipeline infrastructure. Self-Service enabled the platform to scale using a shared responsibility model for datasets. However, day-to-day data operations were time consuming as the number of datasets scaled. Data producers could not get immediate feedback on the data quality. Customized dashboards were built for each dataset to monitor volume anomalies.

Helm was built to better understand customer journeys across every product and channel where they interact with Comcast. Data visualization tools and data APIs for sharing customer context with other applications are the cornerstone capabilities that enabled Helm to scale throughout Comcast. As high value datasets were onboarded, other applications such as Watchlist, Talk, Analytics, and messaging applications were introduced to leverage this data. Again, these applications had bespoke data pipes consuming from the Timeline's message bus. These applications required engineering work for every new onboarded dataset. Platform investments expanded Self-Service to support analytics and other applications. Dataset orchestrations that were originally intended for ingestion were expanded to support routing datasets to applications. At the same time, the applications evolved to consume these datasets and provide rich experiences. Eventually, the applications were integrated and managed with Self-Service and the platform was rebranded to Helm in 2019. Enhanced observability and process improvements enabled data producers to address the data quality and volume anomalies.

Over time, Self-Service and orchestrations expanded in complexity. Orchestrations were re-engineered to be modular, and workflow driven. Also, over time, Helm's operations, user onboarding, and Self-Service capabilities were hardened. Profiles were formally introduced to aid with domain-agnostic features. The platform was designed to scale, in terms of data variety and volume. With additional success, multi-tenancy had gained spotlight. Helm instances were spun up to support Comcast syndication partners and other tenants.

Current Challenges and Future Work: Spinning up new tenant and partner Helm instances demonstrates versatility. However, standing up a new Helm instance can take several hours. Our ongoing initiatives aim to reduce the time it takes to standup new instances. An end-goal is to host Helm in a SaaS model using a services-first approach. Additionally, certain aspects of Helm are not managed by Self-Service. For example, direct data access from Helm infrastructure (AWS Kinesis streams and S3 buckets) requires a code deployment. Similarly, observability dashboards are not externalizable, and teardown infrastructure orchestrations are yet to be integrated into Self-Service. With regards to data quality, statistical data quality checks over longer time ranges (days or months) are not managed in Helm. This is a future capability that can help identify data drift for analytics and machine learning use cases. We are also in the process of evaluating hot/warm storage techniques to optimize infrastructure costs.

5. Conclusion

In this paper, we introduced “Helm”, a CDP that enables employees to deliver best customer experiences by giving them a faster visual timeline of what’s going on, when our customers are experiencing service issues. We showed how CDPs like Helm provide value by consolidating customer experiences from siloed domains or organization verticals.

Data management and governance are the biggest challenges to any customer data platform. The paper highlights the importance of a Self-Service-oriented management layer to govern all aspects of data. As a direct result, the Helm CDP platform is modular and integrated with constituent applications, to enable data re-use. We demonstrated how these applications, including Analytics, 360º Visualization, and Journeys, work together to understand customer experiences and address concerns. We described the design principles that influenced Helm and described its architecture. Finally, we shared the architectural evolution of Helm and future work.

6. Acknowlegements

Neither Rome nor Helm was built in a day. Helm, as it stands today, has evolved several times since its inception as a “Lab Week” project at Comcast in 2012. It has come to fruition from the incredible work of a driven team of less than hundred members. I would like to thank and acknowledge all the engineers, managers, architects, and leaders who have made Helm successful over the years. I would like to particularly thank Leslie Ellis, Dave Torok, and Edward McLaughlin for their valuable inputs on the paper.

Disclaimer: All errors, viewpoints, and opinions in the paper are from the author. They do not necessarily represent the position of the author’s employer or affiliation.

Abbreviations

AWS	Amazon Web Services
API	application programming interface
CDP	customer data platform
ETL	extract, transform, load
GUID	globally unique identifier
HTML	hypertext markup language
HTTPS	hypertext transfer protocol secure
JSON	JavaScript object notation
MAC	media access control
OIDC	OpenID Connect
NPS	net promoter score
SaaS	software as a service
SLO	service level objective
SMS	short message service
SPI	service provider interface
SQL	structured query language
UI	user interface

Hidden Risk of Unpopularity in Open Source

A Technical Paper prepared for SCTE by

Chujiao Ma
Senior Security R&D Engineer
Comcast Cable Communications, LLC
Philadelphia, PA, USA
Chujiao_ma@comcast.com

Vaibhav Garg
Sr. Director Cybersecurity Research & Public Policy
Comcast Cable
Blacksburg VA
Vaibhav_garg@comcast.com

1. Abstract

Software development across the industry relies on the use of open source components (OSCs). Because these components are open-sourced, there is an assumption that these components are tested for security by third party researchers or open source communities. A vulnerability in a popular component can have ripple effects across the ecosystem. Consequently, more popular components are more likely to attract the attention of third-party researchers or the community. Less popular components are thus often left unexamined and potentially vulnerable. In this paper we propose a model to identify OSCs that create the greatest attack surface. Specifically, we propose a metric called relative popularity ratio and use it to risk-rank a set of JavaScript OSCs. We further refine the ranking using observable properties of code, such as number of lines of code. We then validate the efficacy of this metric by engaging third party university researchers to find vulnerabilities. Our results conclude that the hidden risk from unpopular OSCs is concentrated and can thus be addressed by small investments in the security analyses of OSCs.

2. Introduction

Black Duck's audit of commercial codebases in 2017 found that 96% of scanned applications contain open-source components [1]. According to Veracode software security report [2], 7 out of 10 applications contain flaws in their open source libraries on initial scan. 3 in every 10 applications have more flaws in their open source libraries than in the primary code base. Unlike commercial software, open source relies on voluntary contributors to identify vulnerabilities, build patches, and provide updates. Security researchers may spend more time finding vulnerabilities in popular components rather than those used less frequently. Thus, unpopular and consequently underexamined components may pose a hidden risk for their users.

A vulnerability in just one OSC can potentially impact software across multiple products and even across multiple companies. Consider "prototype pollution" [10]. It allows an attacker to modify an object in JavaScript, and its implications can range from injection attacks to denial of service. While fewer than 25 prototype pollution vulnerabilities were reported in 2019, according to a report from Synk, they impacted over 115,000 projects scanned [3]. For example, one high severity prototype pollution vulnerability was discovered in Lodash open source library. This library was used by 4.35 million projects on Github alone.

Once a vulnerability is found, the resulting exposure can take a long time to mitigate. Veracode's software security report notes that only 1 in 4 flaws are fixed in the first 32 days; 2 in 4 are still open after the first 6 months [2]. Simultaneously, a newly discovered vulnerability can turn into active exploits within days, leaving organizations with scant time to respond. A notable example is the 2017 Equifax data breach that affected 143 million customers. A vulnerability in the Apache Struts2 package made it possible for a remote attacker to send a malicious request that allowed them to execute arbitrary commands. 72 days after disclosure, the Equifax breach happened [4].

Thus, the hidden security risk of unpopular open source components cannot be ignored. One solution is for organizations to limit their software developers to an approved set of OSCs. Mature organizations may use tens of thousands of OSCs. Finding a secure and functionally equal OSC for each component is in itself a difficult, if not impossible, proposition. Another solution is to conduct continuous assessment of all the OSCs being adopted in an organization's software supply chain. Unfortunately, an exhaustive analysis would likely be cost prohibitive.

How, then, should organizations and development shops address the potential hidden security risk of the open source libraries? In this paper, we present a three-step approach to address this question:

1. Define a framework to risk rank OSCs based on factors such as popularity
2. Apply the framework to OSCs to produce a list of OSCs and identify areas of concentrated risk
3. Conduct security analysis of the high risk OSCs to identify vulnerabilities

Section 3 starts off by introducing different indicators of security risk for OSCs. In section 4 we describe how these factors can be combined to operationalize a *relative popularity ratio* and then used to determine areas of concentrated risk. Section 5 presents a case study of the security analysis of the OSCs identified in Section 4. Section 6 concludes with a discussion of key findings.

3. Risk Indicators

There are many direct and indirect indicators of the security risk posed by OSCs. They can be categorized as: 1) security status, 2) code characteristics, and 3) OSC popularity. Table 1 provides a list of potential measures for each category. Different language or package managers collect different information and rank OSCs differently. Even the metadata and statistics available also differs. Even if an organization has a complete inventory of all the OSC it uses, the details may be incomplete. Thus, there may be no one size fits all solution to using these OSC security risk indicators. Instead, each organization may select the indicators that suit their development environment. Here, we discuss the three categories of OSC security risk in detail.

Table 1. Possible risk indicators.

Security Status	Code Characteristics	Popularity
▪ Reported vulnerabilities from CVE and NVD ▪ Severity of vulnerabilities based on CVSS ▪ Security of the language ▪ Typical time to remediation ▪ Number of open issues	▪ Lines of code ▪ Complexity of the code ▪ Number of versions ▪ Time of creation ▪ Number of libraries they use ▪ Where/how it is used	▪ Number of contributors ▪ Number of Github stars ▪ Number of forks and pull requests ▪ Number of subscribers ▪ Number of downloads ▪ Number of dependencies

3.1. Security Status

Indicators of security status can include the number of reported vulnerabilities from CVE and NVD databases as well as severity of reported vulnerabilities based on CVSS scores. However, not all flaws have CVEs. The percentage of flaws or vulnerabilites that have a CVE vs those that don't can vary widely depending on the language. In Veracode's analysis of open source, 61.8% of Javascript vulnerabilities don't have a CVE while it is only 10.5% for PHP [5].

Furthermore, different repositories may have distinct level of security exposure depending on how security conscious the owner is and how many independent researchers have analyzed the OSC. They may be captured by the number of open issues against a repository as well as the mean time to remediate.

The language in which the code is written may provide additional indicators. Some languages provide more in-built protection against specific security threat, such as type safe languages. Each language is also more susceptible to specific attacks and less vulnerable to others [11]. For example, JavaScript unlike Ruby is more susceptible to deserialization.

3.2. Code Characteristics

Characteristics of the OSC code can also be indicative of the probability of risk. Prior research has found that code complexity and code size both may correlate with the probability of finding a vulnerability [6] [7]. The coding language used also matters, since almost half of PHP packages contain vulnerabilities while the number is much smaller for other languages [2]. The number of libraries used by the OSC also matters. A report by Snyk found that 86% Node.js vulnerabilities are from indirect dependencies, 81% for Ruby and 74% for Java [3].

In addition to the characteristics of OSC itself, characteristics of the application using the OSC also matters. Depending on where and how it's used, a vulnerability in the OSC can have widespread impact that's difficult to fix or be easily fixed with a compensating control.

3.3. Popularity

Another way to evaluate the risk of open source is by looking at its health or popularity. If a project has low contributor attraction or retention, there's a high chance of it dying out and used by attackers in typosquatting or social engineering. However, the total number of contributors does not correlate to whether a project survives, experience level of the contributors matter as well for the quality of the project [8].

With that said, there are many smaller libraries with just a handful of total contributors and 1-2 active maintainers, so it's important to look at activity level as well instead of just contributors. Node package manager (npm), for example, has a popularity rating that takes into account the number of stars, forks, subscribers, contributors, dependents, downloads. They also have a maintenance score looking at ratio of open/total issues, time takes to close issues, most recent commit, commit frequency, release frequency [9]. While the popularity of the project is important, it is also important to account for the stability of the project as well as distinguish between a huge pull request vs. multiple smaller ones.

4. Risk concentration

Current security solutions aim to identify known vulnerabilities and corresponding patches for OSCs used within an organization's software development lifecycle (SDL). Discovery of new vulnerabilities and verification through exploits necessitates manual analysis. If the organization uses thousands of OSCs across different applications, an exhaustive analysis would likely be cost prohibitive.

It might be difficult to get information on some OSC security risk indicators due to scale or incomplete inventory. Furthermore, it may be difficult to compare these indicators as distinct package managers collect different types of information. Thus, each organization must begin with a set of indicators for which they have the most complete information. These indicators can be combined into a relative popularity ratio to address the bias introduced by one single metric.

Based on this ratio organizations can identify areas of concentrated risk. This allows prioritized security analyses of OSCs that have a greater probability of containing hidden risks as well as for whom

vulnerabilities will result in greater organizational impact. In this section we discuss this process with an example from Comcast.

4.1. Relative Popularity ratio

The goal of the relative popularity ratio is to identify OSCs that are more popular within the organization (high impact) than externally (under evaluated) and is calculated as:

$$Relative\ Popularity\ ratio\ = \frac{Internal\ popularity}{External\ popularity}$$

By using a ratio of two measures, rather than a single measure, we are able to reduce biases specific to a single dataset. Furthermore, relative popularity captures both the under evaluation of an OSC by third party security researchers, and therefore the hidden risk, and the possible impact for the concerned organization. Table 1 notes the various indicators of popularity, e.g. number of forks.

For Comcast, internal popularity was measured by the number of dependents, as this information is easily assessable through Software Component Analysis or SCA tools. External popularity was measured as an average of: 1) dependents (impact), 2) weekly downloads (current popularity), and 3) Github stars (long term popularity). These were selected based on the specific information available for npm, the specific repository that hosts the specific OSCs under analysis.

$$Popularity\ ratio = \frac{\%\ of\ Comcast\ dependents}{(\%\ of\ npm\ dependents + \%\ of\ github\ stars + \%\ of\ weekly\ downloads)/3}$$

4.2. Risk Concentration

While ranking OSCs according to their relative popularity can help prioritize the list of OSCs, there are still have thousands or tens of thousands of OSCs for each language. The next step then is to determine areas of concentrated risk. For example, consider the main languages used for development at an organization. At Comcast most of the development is limited to three: 1) C, 2) Java, and 3) JavaScript.

For each language, developers may use hundreds of OSCs. An individual library may be used just once or in up to thousands of applications. Figure 1 plots the OSCs for each language by percentage of dependents. Observe that the top 50 OSCs covers 26.2% of dependents for C, 37.7% of dependents for JavaScript and 25.9% of dependents for Java. In fact, top 200 OSCs provide more than 50% coverage for each language, i.e. 52.2% for C, 59.6% for Java, and 59.5% for JavaScript.

Figure 2 shows the number of dependents or applications using each OSC for the top 1000 JavaScript for 2019, 2020 and 2021. The top 50 OSCs cover 37.7%, 23.7% and 37.9% of the risk for 2019, 2020 and 2021 respectively. The top 200 OSCs covers 63%, 45% and 60% of the risk for 2019, 2020 and 2021 respectively. Thus, as more development teams adopt SCA tools and more OSCs are inventoried, the general trend continues to hold true from year to year.

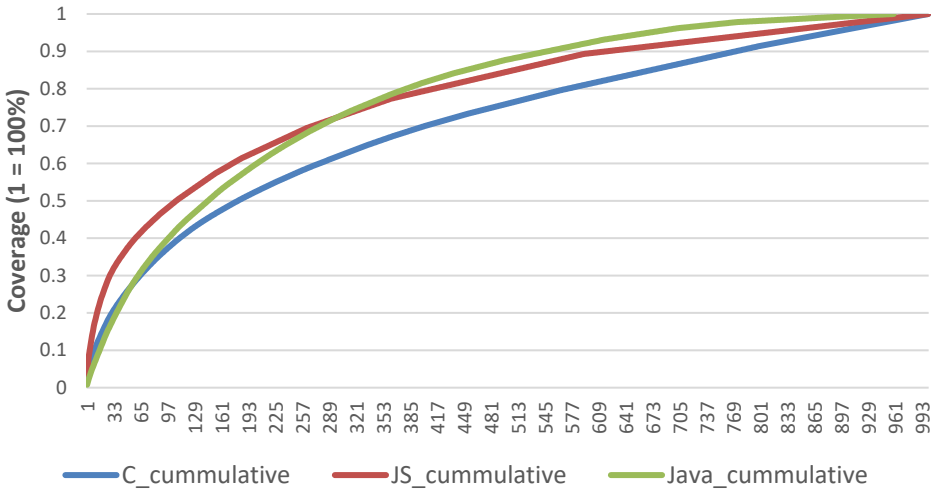


Figure 1. Percentage of dependents covered by the top 1000 OSC for C, JavaScript and Java

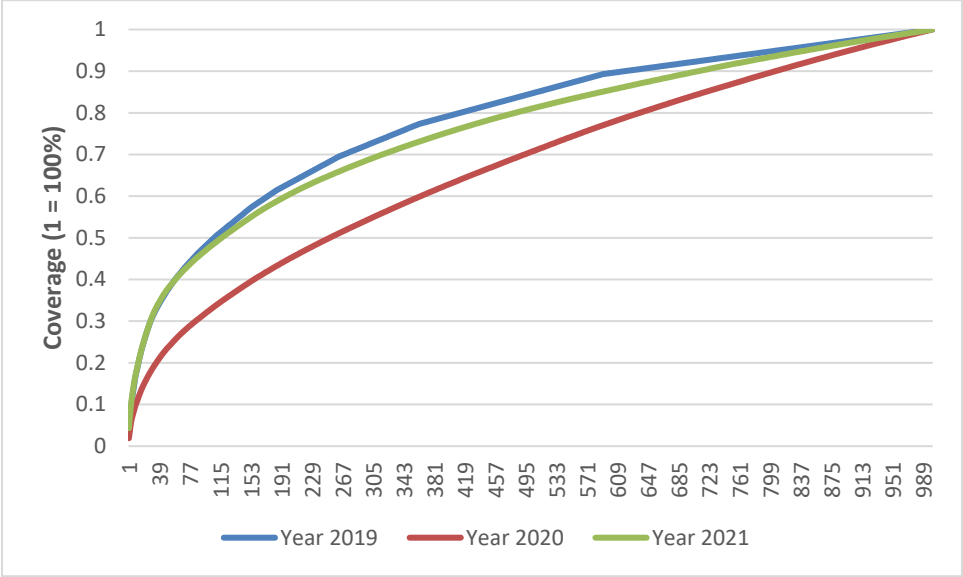


Figure 2. Number of Dependents for top 1000 JavaScript OSCs for 2019, 2020, and 2021.

Thus, the security risk from OSCs is concentrated. Analyzing the top 50-200 components will provide coverage for majority of the risk. There is then no need to conduct an exhaustive security assessment of all components. Instead, the organization can focus its resources on OSCs with greater number of dependents. Here we focus on the top 50 components in JavaScript, which covers approximately a third of the risk.

5. Case Study

As we note in Section 4, the security risk from OSCs is often concentrated. Thus, a detailed security assessment of a small set of OSCs provide coverage for much of the security risk. In this section we discuss an example of such an assessment. We began by selecting the OSCs. Based on the trajectory of usage of OSCs within Comcast, we decided to analyze top 50 JavaScript OSCs. We further narrowed our scope based on the relative popularity ratio.

In terms of manual analysis, there are many ways to attack and test the OSCs. OWASP top 10 and SANS top 25 are a good place to start. In addition, according to a report by Veracode [5], there are four categories of flaws that represent 75% of all flaws found in the open source libraries: access control, cross-site scripting, sensitive data exposure and injection. We decided to focus on the flaws within the source code rather than due to access control or other external factors, and those that are server side rather than client side.

5.1. Methodology

The first step of the analysis is to collect a list of OSCs. The information for OSCs used within Comcast are collected from WhiteSource, thus the list only includes OSCs used by applications of teams using WhiteSource. There is a limited amount of information available for all OSCs so the list only contains the name of the component, and the number of dependents (applications using it with all versions combined together). For this experiment we focused on JavaScript.

Once the internal list of popular OSC is collected, we ranked them according to the number of dependents, and external data is then collected for the top 50 OSCs that are used by most applications within Comcast. External data used for popularity ratio calculation such as number of dependents and weekly downloads are collected from npm repository. The number of stars is collected from the Github repository. The popularity ratio is then calculated for each OSC, where a higher number indicates it is more popular within Comcast than externally, while a lower number indicates that it is more popular outside of Comcast than internally.

The list of top 50 JavaScript OSCs was given to third party university researchers at Comcast Center of Excellence for Security Innovation at the University of Connecticut for security analysis. The analysis focused on identifying previously undiscovered code injection vulnerabilities and verifying that they are exploitable.

5.2. Results

Figures 3, 4, and 5 lists the 50 OSCs under analyses based on external popularity, internal popularity, and relative popularity ratio respectively. Security analyses found new code injection vulnerabilities in 3 of the 50 OSCs: jade, depot and prototype. These components are color coded red. The problems came from embedded dependencies; newer versions of the packages have fixed the vulnerabilities. Separately, one package was found to be deprecated, and colored grey.

When ranked according to either internal or external popularity, as seen in Figure 3 and 4, the vulnerable packages are scattered throughout the top 50 list. In Figure 5 however, when the top 50 OSCs are ranked according to the popularity ratio, where higher ratio indicates high popularity within Comcast but low popularity externally, we can see that the vulnerable components are concentrated in the top third of the list. It also makes sense that the deprecated OSC would rank high on the popularity ratio since there would

be almost no external usage (thus the external popularity is only from a few GitHub stars) while it was still used internally. The OSC has been phased out of use since then.

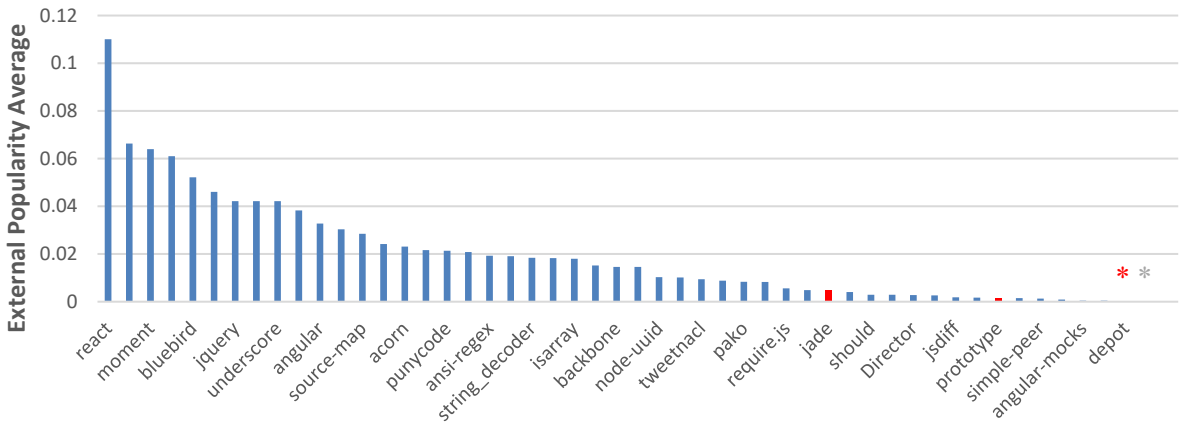


Figure 3. Top 50 OSCs ranked by external popularity. * denotes location of vulnerable/deprecated components difficult to see.

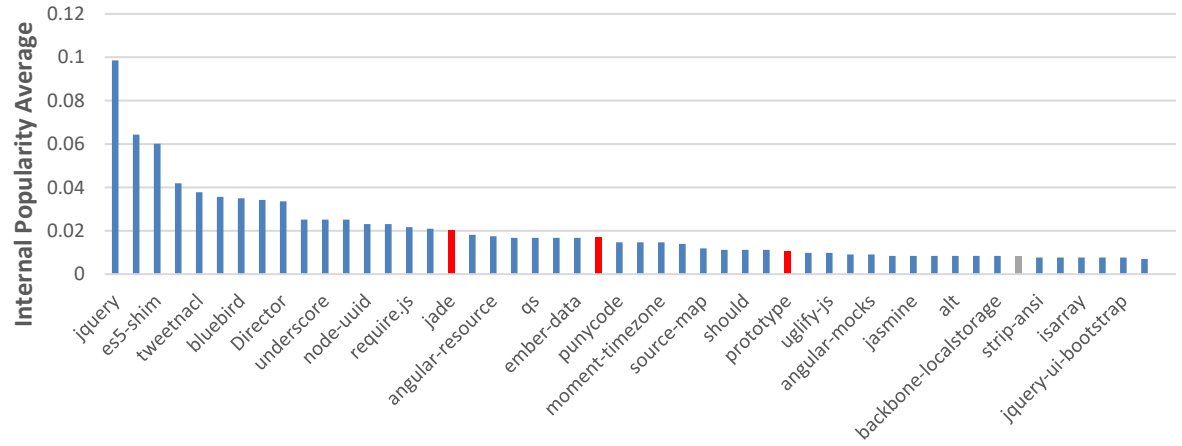


Figure 4. Top 50 OSCs ranked by internal popularity.

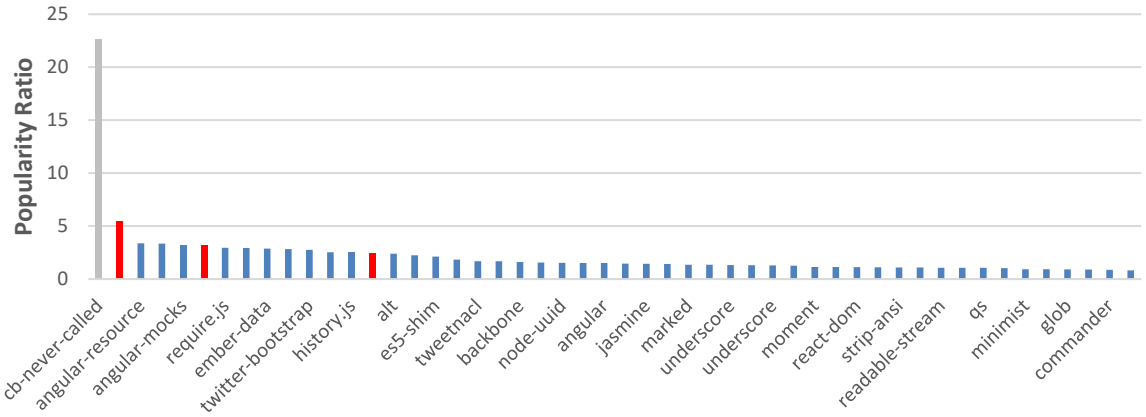


Figure 5. Top 50 OSCs ranked by popularity ratio.

The three vulnerable OSCs covers 1.78% of dependents or applications used by the teams in the list and the analysis was only done for code injection vulnerability. From this case study, we have verified our hypothesis that less popular OSCs contain hidden risk.

6. Conclusion

The first step toward managing the risk from open source is to identify the OSCs used across the organization. Using software composition analysis helps with both taking inventory of OSCs used and whether there are any existing vulnerabilities as well as corresponding patches available. It also automates the discovery of old or deprecated versions of an OSC being used. For many OSCs no significant vulnerabilities may be publicly known. This may be because these OSCs are secure. Alternatively, it is possible that these components are not popular enough for third party security researchers to analyze them. This creates a hidden security risk. As most organizations use hundreds, if not thousands, of OSCs an exhaustive security assessment would be both cost prohibitive and impractical.

In this paper we provide a roadmap to address the hidden security risk from such OSCs. First, we describe the various indicators that can be used to approximate the security status of an OSC. Then we provide a way to use these indicators to risk rank OSCs via a relative unpopularity ratio and identify areas of concentrated risk that can be prioritized for security assessments. We then present an example case study. Our results indicate that a majority of the hidden risk is concentrated in a minority of components. For us, the top 50-200 components cover 30-50% of the risk. Thus, small investments in security assessments of OSCs can address any potential unknown vulnerabilities.

This paper presents a specific instantiation of the relative popularity ratio. However, organizations may choose their own indicators based on their operational context, information sources, and data completeness. Furthermore, the relative unpopularity ratio is only intended to provide a way to prioritize security assessments. It is unlikely that there will be a quantitative correlation between the ratio and the number of vulnerabilities found. As the ratio can be constructed in different ways, the specific ranking of an OSC will differ in each distinct instantiation of the formula depending on the indicators used. Additionally, assessors (and assessments) themselves may be better at finding certain types of vulnerabilities over others. Regardless, as we show in this paper this qualitative approach is still effective at identifying areas of concentrated risk and addressing them. Thus, it is possible to shine a light on the hidden risk of unpopularity in OSCs.

Abbreviations

CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
npm	Node Package Manager
NVD	National Vulnerability Database
OSC	Open Source Component
OWASP	Open Web Application Security Project
PHP	Hypertext Preprocessor

Bibliography

[1] Z. Zorz, "The percentage of open source code in proprietary apps is rising," 2018. [Online]. Available: https://www.helpnetsecurity.com/2018/05/22/open-source-code-security-risk/.

[2] "State of Software Security v11," Veracode, 2020.

[3] A. M. a. S. Zitzman, "The State of Open Source Security Report 2020," Synk, 2020.

[4] C. Fearon, "Examining Apache Struts remote code execution vulnerabilities," 3 October 2017. [Online]. Available: https://www.synopsys.com/blogs/software-security/apache-struts-remote-code-execution-vulnerabilities/.

[5] "State of Software Security: Open Source Edition," Veracode, 2020.

[6] R. a. W. J. a. H. A. a. J. W. Scandariato, "Predicting Vulnerable Software Components via Text Mining," *IEEE Transactions on Software Engineering,* 2014.

[7] Y. a. M. A. a. W. L. a. O. J. A. Shin, "Evaluating Complexity, Code Churn, and Developer Activity Metrics as Indicators of Software Vulnerabilities," *IEEE Transactions on Software Engineering,* 2011.

[8] N. Eghbal, "Methodologies for measuring project health," 2018. [Online]. Available: https://nadiaeghbal.com/project-health.

[9] A. C. a. A. Duarte, "npms - About," 2021. [Online]. Available: https://npms.io/about .

[10] Kim HY, Kim JH, Oh HK, Lee BJ, Mun SW, Shin JH, Kim K. DAPP: automatic detection and analysis of prototype pollution vulnerability in Node. js modules. International Journal of Information Security. 2021 Feb 13:1-23.

[11] “Most Secure Programming Languages,” WhiteSource. 2021.
https://www.whitesourcesoftware.com/most-secure-programming-languages/

Hitchhiker's Guide to Quantum Key Distribution

A Technical Paper prepared for SCTE by

Vaibhav Garg
Sr. Director Cybersecurity Research & Public Policy
Comcast Cable
Blacksburg VA
Vaibhav_garg@comcast.com

Walter Krawec
Assistant Professor
University of Connecticut
Storrs, CT
Walter.krawec@uconn.edu

Pete Quesada
Sr. Principal Engineer
Comcast Innovation Labs
Denver, CO
pete_quesada@cable.comcast.com

Tony Tauber
Distinguished Engineer
Comcast Cable
Boston, MA
Tony_Tauber@cable.comcast.com

Aman Satija
Research Engineer
Purdue University
West Lafayette
asatija@purdue.edu

1. Abstract

Quantum Key Distribution or QKD offers a quantum safe mechanism to establish an encrypted and authenticated communications channel. QKD is theoretically secure against a computationally unbounded adversary; this contrasts with classical key distribution systems where security is contingent on the assumptions made about the computational capacity of the adversary. These systems are being deployed in operational test environments across the globe. Most commercial systems require expensive proprietary technology, where the marginal cost of deployment is proportional to the capital investment. This means that the cost of adding an additional user is proportional to the cost of the original system. Thus, some experts have argued that the security assurance of QKD systems is not adequate to justify a transition from current approaches, other than for niche or otherwise narrow use cases. This paper provides an overview of current QKD systems, provides insight into the economics of deployment, and discusses the potential for commercial applications.

2. Introduction

The promised advent of quantum computers has focused on the negative impact of quantum technologies for security. Large quantum computers can void the security of current public key cryptosystems, necessitating a transition to Post Quantum Cryptography (PQC), i.e. cryptography that would be secure against know quantum cryptoanalysis. Less attention has been paid to security enabling solutions such as Quantum Key Distribution (QKD) [1]. These solutions, unlike quantum computers, are currently being tested in operational settings across the world. South Korea Telecom, which announced the first 5G smartphone enabled with Quantum Random Number Generator (QRNG), is now testing ID Quantique's QKD system in its 5G network [2]. British Telecom partnered with Toshiba Labs and is now testing a QKD-enabled link between two research sites in Bristol, United Kingdom [3]. Stateside, Verizon is piloting a QKD network in the Washington D.C. area [4].

The key security advantage of QKD systems is that they are theoretically secure against a computationally unbounded adversary [1]. Thus, unlike RSA or Elliptic Curve cryptography, the security of these systems is not based on the hardness of solving a mathematical problem. Simultaneously, unlike AES and post-quantum cryptography, their security is not based on the computational limits of the adversary. Instead, the security of QKD is based on the laws of quantum physics. This means that any communications secured using QKD can be recorded by an adversary and the communicating parties can be assured that, regardless of any future technological advances, the messages will be secure in perpetuity [5]. Critics, however, will point out that adversaries typically do not rely on cryptoanalysis or breaking the math. The easier solution is to find a vulnerability in the software [6].

In the United States, the NSA has supported the critics' view by stating that national security systems should not be secured with the use of quantum cryptography, including quantum key distribution [7]. Yet, the Bureau of Industry and Security continues to restrict the dissemination of QKD technology as part of export controls [8]. Internationally, ETSI [9], ITU [10], as well as ISO/IEC [11] have continued their standards efforts for QKD, advancing the view of supporters of QKD.

If you are new to QKD and all this confuses things, DON'T PANIC. This paper is an effort to provide a representative overview of QKD, the arguments for and against this technology, and the possible future implications. We begin with an introduction to QKD in Section 3. Section 4 provides an overview of the commercial landscape and deployment architectures. In Section 5, we discuss the benefits and limitations of various solutions as well as potential use cases. Section 6 concludes with a summary of key points.

3. An Introduction to QKD

QKD is a subset of quantum cryptography technologies [12]. These solutions are based in large part on the quantum physical property of *no cloning*, i.e. it is impossible to measure a quantum state without changing its properties [13]. This property has led to many quantum cryptography solutions, such as quantum coin tossing [14], quantum oblivious transfer [15], quantum zero knowledge proofs [16], quantum bit commitment [17], quantum secret sharing [18], and QRNGs [19]. While many of these solutions have been of theoretical interest to cryptographers, some, like QRNGs and QKD, have gotten significant commercial traction [1].

The first and perhaps the most well-known QKD protocol is BB84, which was designed by Charles Bennett and Giles Brassard in 1984 [20]. Every subsequent protocol is to a degree an adaptation for BB84 for a specific technical design under distinct constraints. BB84 has two main stages. In stage 1 Alice sends a series of quantum states to Bob over a quantum channel. These states are randomly constructed in either the X or Z basis by Alice and similarly arbitrarily measured in either basis by Bob. If the states are measured in the same basis as the one in which they are constructed, the outcome is deterministic, otherwise it is random and destructive.

In stage 2 Alice reveals her basis to Bob through a classical channel and Alice and Bob keep the states in which Bob chose the correct basis. The channel for classical communication is public but authenticated, i.e. Eve can eavesdrop on this channel, but not tamper with any message sent on it. Due to the no cloning theorem of quantum mechanics, Eve must actively attack the quantum signal (she cannot store the quantum data to attack at some future time). Finally, since Eve does not know the basis choice that Alice used, she cannot deterministically extract any information with certainty. In fact, any attempt by Eve to extract information from the quantum data will cause the quantum state to become disturbed, which may be detected by Alice and Bob.

The BB84 protocol was designed to be used with single photon emitters. However, these can be difficult to build and operate and thus add expense to the system. Researchers have addressed this by proposing decoy state BB84, which can be implemented using weak coherent lasers [21]. Others have proposed measurement device independent or MDI QKD protocols, which protect against side channel attacks [22]. Finally, a recent addition has been semi-quantum QKD protocols [23]. In these protocols, some of the participants need to be able to prepare or measure quantum states, whereas others can simply reflect them.

The cost of QKD systems is driven by the photon detectors, i.e. Single Photon Avalanche photoDiodes (SPADs). These may require special cooling equipment, resulting in higher capital costs, increased operational costs, greater operational complexity, as well as more cumbersome form factors. A good SPAD should always click when a photon hits it (quantum efficiency), should not click when a photon is not present (dark count), have a low reset time (dead time), should not result in multiple clicks for a single photon (afterpulsing), and be accurately able to ascertain when a photon was detected (timing jitter).

The physics of single-photon detection leads to fundamental trade-offs between the various desired characteristics. For example, an increase in quantum efficiency is achievable with a larger detector area. However, that leads to a greater uncertainty in the incident location of the photon which translates an increase in timing jitter. Similarly, it is desirable to lower detector temperature to reduce dark noise however that increase the probability of after-pulsing (as well as the cost).

QKD systems can operate over fiber or in free space. Systems that operate over fiber are typically limited to a range of 50-100kms. This is a fundamental difference between classical and quantum communication. One cannot use a traditional fiber amplifier on single-photons due to the *no cloning* theorem [13]. Thus,

QKD systems use trusted nodes to extend their range. This inherently limits the security guarantees offered by QKD, i.e. the overall security of the system becomes dependent on the trustworthiness of the trusted nodes. Another limitation is the hold-off period of SPADs applied to prevent after-pulsing after an avalanche breakdown, i.e. the detector needs a time gap between detecting two distinct photons. Without this gap, or hold-off period, the detector may click multiple times even though there is only one photon in the fiber. The hold-off period of Indium Gallium Arsenic or InGaAs SPADs, for example, is 1-10 μs. This limits communication speed to 100 kbps-1 Mbps.

Free space solutions, i.e. earth to space or space to space, may overcome the distance constraints of fiber-based systems. They are instead constrained by line of sight. Furthermore, their efficiency may be impaired by atmospheric conditions such as rain. In fact, QKD systems have a range of variations in implementations, each with distinct advantages and disadvantages. For example, QKD systems that form the base of quantum communication use entangled photons. Some QKD systems use discrete variable implementations, while others use continuous variable alternatives. An exhaustive discussion is beyond the scope of this paper and is covered elsewhere [1].

4. Systems in the Wild

QKD systems are increasingly being tested in a variety of operational setting for various use cases. Some of these are being deployed by the industry alone, others in collaboration with academia and government. There are a range of commercial solutions being offered by traditional technology companies as well as startups of varying levels of maturity. In this section we aim to provide a representative overview of these systems.

4.1. ID Quantique

ID Quantique's flagship product is the Cerberis XG QKD system [24]. The system is designed to be operated over 50km with a secret key rate of 1.4 kbps. The quantum channel can operate over dark fiber or via channel multiplexing with the quantum channel around 1310nm (O-band). The system can be integrated in a diversity of network topologies, including point-to-point, relay, and ring, as well as hub-and-spoke. The system complies with the Advanced Telecommunications Computing Architecture (ACTA) to allow for easy plug and play into existing physical infrastructure.

4.2. Toshiba

Toshiba systems use both proprietary detector, self-differencing single photon detectors, as well as a proprietary protocol, T12 [25, 26, 27]. This allows Toshiba to offer 1 Mbps key rates over 50km deployments, with the equipment operating at room temperature. Toshiba offers two commercial solutions. The first uses multiplexing to allow data and QKD on the fiber. The second operates on dark fiber for long-distance applications. Toshiba has been awarded the contract to deploy and manage the QKD infrastructure of the National Institute of Information and Communications Technology in Japan. The company has partnered with British Telecom to lead the first industry deployment of QKD in UK. In the U.S., it is partnered with Verizon and Quantum Xchange to pursue QKD demonstrations in operational environments.

4.3. Quantum Xchange

At the time of writing, Quantum Xchange does not produce its own QKD systems. Instead, it has partnered with other QKD vendors to provide a key management solution that works with and without QKD. Additionally, Quantum Xchange also has a proprietary trusted node technology developed in collaboration

with Battelle [28]. This is being used to build a QKD network from Boston to Washington, D.C., over which it plans to offer QKD as a service. Its systems operate over dark fiber to allow for better key rates and longer-range deployments [29]. The flagship product, known as Phio QK, is available standalone or as a managed service to allow for an easier transition. At the time of writing (summer 2021), the company's primary focus is the financial sector, although, as noted, it is conducting pilots with Verizon.

4.4. Qubitekk

Qubitekk's solution uses entanglement-based QKD rather than prepare and measure systems [30]. Its systems are specifically designed for industrial control systems, and run over standard optical fibers with an offering of 100kbps over 20 kms. Qubitekk's focus is on the energy sector, as a participant in the Department of Energy's Quantum Grid initiative. In one instance, it demonstrated trusted relay based QKD in collaboration with the Electric Power Board in Chattanooga, TN. The demonstration used three distinct QKD systems: 1) a COTS system with BBM92, 2) a research system with BB84, and 3) COTS systems with SARG04. The overall key generation rate was slower than the slowest QKD system [31].

4.5. Quintessence Labs

Quintessence systems are based on Continuous Variable QKD (CV-QKD) vs. Discrete Variable QKD (DV-QKD) [32]. The company's flagship product is qOptica™ 100. Quintessence is focusing on free-space QKD, where CV-QKD may offer additional advantages. The latter does not use SPADs, which are sensitive to atmospheric conditions. Instead, CV-QKD relies on homodyne detectors that may operate unimpeded under daylight conditions, without the need for spatial filters or spectral filters.

4.6. VeriQloud

VeriQloud's flagship product, QLine, takes a different approach to QKD architecture [33]. It does not require each client to both prepare and measure the state of a qubit. The product QLine then requires only one photon source and detector for a set of clients on a single linear fiber. These clients have neither the ability to measure nor prepare qubits. However, they can transform the state of the photon using optical modulators. Its architecture may be used with either DV-QKD or CV-QKD. At the time of writing (summer 2021), no commercial product is available.

5. The Case for QKD

The unconditional security of QKD comes at a cost, which impacts the availability of potential use cases. In this section we will discuss how the economics of QKD may change with technological evolution and thus open the technology to additional use cases.

5.1. Economics of Deployment

QKD systems often use specialized hardware, have more usable keys rates over dark fiber, and may also require special cooling equipment. Given these requirements, a QKD system connecting two nodes may cost on the order of ~$100,000. Simultaneously, as these systems are designed to operate point to point. Thus, the marginal cost of adding another node is the same as that of the initial capital investment. This does not account for the cost of adding a trusted relay, which can further impinge costs for deployments over longer distances. Fortunately, there are both academic as well as commercial solutions emerging that may redefine the economics of deployment.

Many commercial solutions are now ACTA-compliant to allow the use of existing chassis. Commercial products, such as IDQuantique's Gen 4 Cerberis, work in a variety of network topologies, e.g. star, spoke and hub, etc. New experiments from Toshiba have extended the range of fiber-based QKD from 50-100km to 600km [34]. This reduces the cost of long-range deployments by a factor of six (or perhaps even 12). The distance limitations of fiber-based QKD can also be compensated by satellite-based QKD. In Canada, Loft, in partnership with QEYSSat and the Canadian Space Agency, is looking to build small, inexpensive satellites for QKD [35]. Singapore-based startup Speqtral has already demonstrated entanglement in space with ira nano satellite SpooQy-1 CubeSat [36].

Another promising solution is the potential for auto-compensated QKD systems that use Faraday Mirrors in combination with semi-quantum protocols. These essentially act as mirrors for optical signals and simply reflect them back down the fiber. As early as 2002, IDQuantique conducted a demonstration of one such solution over 67km with a bit rate of 50bps at 1550nm [37]. This was followed up with another auto-compensated system as part of the SECOQC initiative where key rates were closer to 1Kbps [38]. VeriQloud's QLine requires but one set of sender and receiver for a set of clients. With these solutions the capital cost of deployment may be high, but the marginal cost of adding another client will likely be orders of magnitude lower. For example, Faraday Mirrors (FM) used in auto-compensated QKD are on the order of $100 (vs. a single detector which runs on the order of $25,000). The corresponding end point systems may be on the order of ~$1000 for Faraday Mirrors vs. ~$100,000 for detectors.

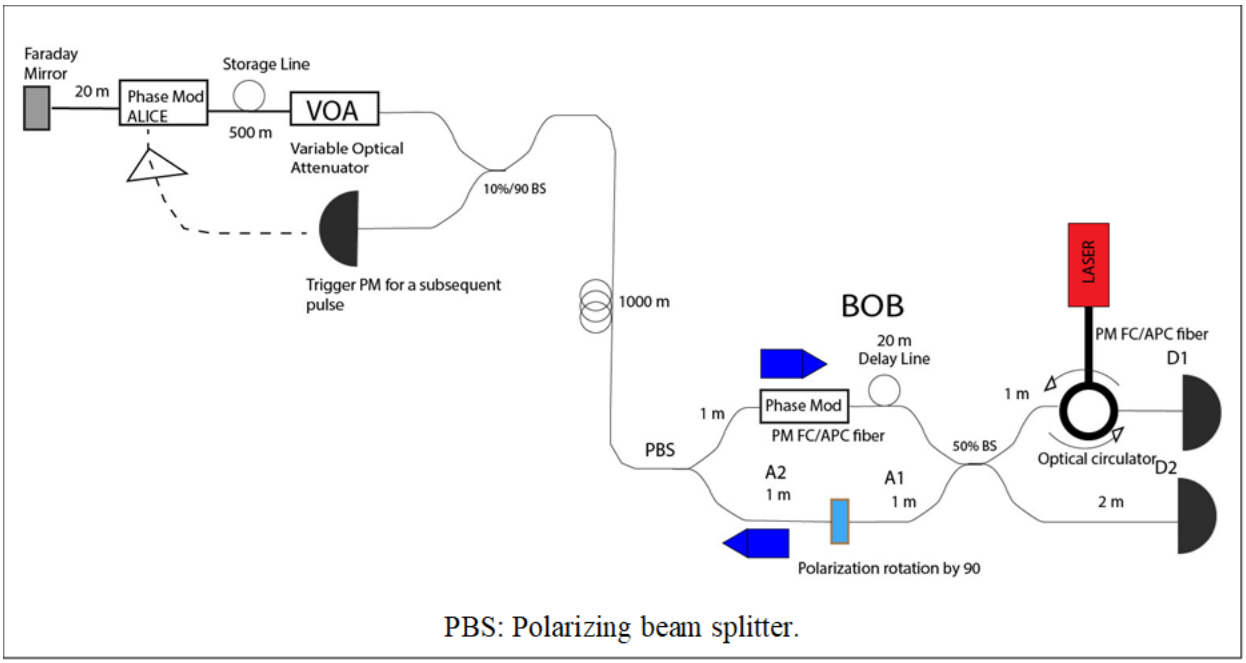


Figure 1 - Example Auto-Compensated DV-QKD System

At its core QKD is simply a mechanism for securely distributing keys. These keys can then be used in any of the various contexts that currently use PKI for key distribution. For example, as far back as 2010, ETSI published a white paper noting that the keys generated from QKD can be integrated in any part of the network stack from Data Link Layer to Application Layer [39]. On the Data Link Layer, QKD keys can be used be used to support the encryption component of Point-to-Point Protocol (PPP). They can be combined with a link encryptor to support a Virtual Private Network (VPN) [40]. Fortinet, for example, in partnership with IDQuantique ,will begin to offer a QKD-based VPN platform [41].

It has been argued that any protocol that requires a pre-shared secret can use QKD keying material instead [42]. Thus, the use of QKD keys can be extended to Network Layer protocols such as IPSec [43], Transport Layer Protocols such as TLS [44], and application layer protocols such as Kerberos [45] or Single Sign On [46]. For example, IPSec uses Internet Key Exchange (IKE) protocol for key management. A QKD key may be used in place of the Diffie-Hellman shared secret used by IKE. This can be done in conjunction with an agreed upon cipher suite, or alternatively, the QKD key may be used as an One Time Pad (OTP). The second option will ensure security in perpetuity, whereas the security of the first option will be contingent on the security of the cipher suites.

Thus, QKD keys can be integrated in any part of the infrastructure. The specific use cases will depend on the cost of the QKD end-points vs. the benefit of the security offered. There are many sectors, such as defense and health, where data may need to be protected for 100 years, and perhaps more, in the case of genetic data. However, at $100,000 per QKD-end point, the solution may be cost prohibitive. Over fiber, QKD requires trusted nodes over long distance, which impacts its security assurance.

One solution is to focus fiber use cases for smaller distances [33]. Smart buildings, smart campuses, data centers etc. all require key distribution over short distances. Furthermore, many current approaches assume that both the sender and the receiver need to be able to prepare and measure quantum states. Another alternative is to use auto-compensated systems, which require only the sender to have this capability [37]. The receiver need only manipulate the phase of the photon. The receiver then needs to have a phase modulator along with a Faraday Mirror.

The sender, Bob, can be treated as a central key distribution authority interacting with multiple Alices using polarization division multiplexing [47] (or wavelength division multiplexing [48]). Assuming a distance limit of 50km based on current commercial systems, this system on average can provide coverage for 7850 sq kms. (=3.14 x 50^2). An average U.S. city is 338 sq kms, whereas the average county is 2911.4 sq kms. Thus, many city-wide and metropolitan networks may also satisfy the distance constraints of current QKD systems without trusted repeaters [49].

This can be used to connect local government offices, to secure electronic voting, or even to offer highly secure VPN for remote workers. Advances in the use of QKD with drones may allow for secure key distribution in hard-to-reach areas, disaster zones, etc [50]. With an auto-compensated system using a Faraday Mirror + Phase Modulator, the receiver will offer a more usable form factor both in terms of size and weight for drones.

6. Conclusion

In the past two decades, QKD systems have gone from an academic pursuit to being commercially viable. Implemented correctly, their security is based on the laws of physics rather than the computation ability of the attacker. This guarantee comes at a cost that is partly driven by the need for specialized hardware. However, crypto has often required specialized hardware -- from Hardware Roots of Trust to Hardware Security Modules. Furthermore, there are certain classes of data, such as health data, that may need to be secured for long time, e.g. 100 years. There are no current crypto solutions, including PQC, which can guarantee this. The only solution for these classes of data is QKD, as it does not make assumptions about the computation capability of the attacker, which may evolve with time [51]. (Although it does require QKD key rates that allow for OTPs.)

Perhaps unsurprisingly, market analysts sized the QKD market in 2019 at $2472.4M, with expected growth to $8562.7M by 2026 [52]. Currently, the cost of deployment has limited use cases to high-risk environments such as defense. However, as technological evolution reduces the cost of deployment, it is

likely that new use cases will emerge to further increase the size of the market. This potential revenue stream is driving investments in both startups across the globe, and operational proofs of concepts (PoCs) in a range of sectors, from energy to communications.

Finally, entanglement-based QKD is critical to securing a Quantum Internet. QKD, then, is a technology that will increasingly become mainstream. While it will never be the answer to life, the universe, and everything cybersecurity, it will certainly have its use cases. This is reflected in the efforts to standardize architectures and interfaces to ensure interoperability [10,11]. QKD itself comes in many different flavors, i.e. CV-QKD vs. DV-QKD; entanglement vs. prepare and measure; fiber vs. free-space, etc. As with any emerging technology, it is unclear which solution will become the default option. It is possible that different versions may be applicable for distinct uses cases. Thus, this paper was intended to provide a representative overview of relevant technical, operational, and economic considerations, with an eye toward practical QKD.

Abbreviations

ACTA	Advanced Telecommunications Computing Architecture
AES	Advanced Encryption System
BB84	(Charles) Brassard + (Giles) Brassard (19)84
COTS	Commercial Off the Shelf
CV-QKD	Continuous Variable Quantum Key Distribution
DV-QKD	Discrete Variable Quantum Key Distribution
ETSI	European Telecommunications Standards Institute
FM	Faraday Mirror
ISO/IEC	International Standards Organization / International Electrotechnical Commission
IKE	Internet Key Exchange
MFA	Multi-Factor Authentication
NSA	National Security Agency
OTP	One Time Pad
PoC	Proof of Concept
PPP	Point-to-Point Protocol
PQC	Post Quantum Cryptography
QKD	Quantum Key Distribution
RSA	Rivest-Shamir-Adleman
SECOQC	SEcure Communication based On Quantum Computing
SPAD	Single Photon Avalanche (photo)Diode
TLS	Transport Layer Security
UK	United Kingdom and Northern Ireland
USA	United States of America
VPN	Virtual Private Network

Bibliography & References

1. Amer, O., Garg, V. and Krawec, W.O., 2021. An Introduction to Practical Quantum Key Distribution. *IEEE Aerospace and Electronic Systems Magazine*, *36*(3), pp.30-55.
2. Press Release, 2020. ID Quantique and SK Broadband selected for the construction of the first nation-wide QKD network in Korea. [weblink]
3. Dunn, J., 2020. BT Is Using Quantum Technology to Secure Gigabytes of Sensitive Data Sent Between Two Industrial Sites In The UK. Forbes. [weblink]
4. Ashraf, C., 2020. Verizon achieves milestone in future-proofing data from hackers. [weblink]
5. Stebila, D., Mosca, M. and Lütkenhaus, N., 2009, October. The case for quantum key distribution. In *International Conference on Quantum Communication and Quantum Networking* (pp. 283-296). Springer, Berlin, Heidelberg.
6. Paterson, K.G., Piper, F. and Schack, R., 2007. Quantum cryptography: a practical information security perspective. *Nato Security Through Science Series D-Information and Communication Security*, *11*, p.175.
7. NSA, 2019. Quantum Key Distribution and Quantum Cryptography. [weblink]
8. Bureau of Industry and Security, 2017. Encryption and Export Administration Regulations. [weblink]
9. Alléaume, R., 2018. Implementation Security of Quantum Cryptography: Introduction, challenges, solutions. *ETSI White Paper*, *27*, p.28.
10. Technical Report, 2020. Security Considerations for Quantum Key Distribution Networks. Telecommunication Standardization Sector of ITU. [weblink]
11. ISO/IEC CD 23837-1. Security requirements, test, and evaluation methods for quantum key distribution. [weblink]
12. Wiesner, S., 1983. Conjugate coding. *ACM Sigact News*, *15*(1), pp.78-88.
13. Wootters, W.K. and Zurek, W.H., 1982. A single quantum cannot be cloned. *Nature*, *299*(5886), pp.802-803.
14. Döscher, C. and Keyl, M., 2002. An introduction to quantum coin tossing. *Fluctuation and Noise Letters*, *2*(04), pp.R125-R137.
15. Bennett, C.H., Brassard, G., Crépeau, C. and Skubiszewska, M.H., 1991, August. Practical quantum oblivious transfer. In *Annual international cryptology conference* (pp. 351-366). Springer, Berlin, Heidelberg.
16. Kobayashi, H., 2008, March. General properties of quantum zero-knowledge proofs. In *Theory of Cryptography Conference*(pp. 107-124). Springer, Berlin, Heidelberg.
17. Hillery, M., Bužek, V. and Berthiaume, A., 1999. Quantum secret sharing. *Physical Review A*, *59*(3), p.1829.
18. Song, Y. and Yang, L., 2020. Semi-counterfactual Quantum Bit commitment protocol. *Scientific reports*, *10*(1), pp.1-12.
19. Herrero-Collantes, M. and Garcia-Escartin, J.C., 2017. Quantum random number generators. *Reviews of Modern Physics*, *89*(1), p.015004.
20. Bennett, C.H. and Brassard, G., 1984, August. An update on quantum cryptography. In *Workshop on the theory and application of cryptographic techniques* (pp. 475-480). Springer, Berlin, Heidelberg.
21. Lo, H.K., Ma, X. and Chen, K., 2005. Decoy state quantum key distribution. *Physical review letters*, *94*(23), p.230504.
22. Xu, F., Curty, M., Qi, B. and Lo, H.K., 2014. Measurement-device-independent quantum cryptography. *IEEE Journal of Selected Topics in Quantum Electronics*, *21*(3), pp.148-158.
23. Krawec, W.O., 2015. Mediated semiquantum key distribution. *Physical Review A*, *91*(3), p.032323.
24. IDQuantique. Cerberis QKD System. [weblink].

25. M. Lucamarini, K. A. Patel, J. F. Dynes, B. Fröhlich, A. W. Sharpe, A. R. Dixon, Z. L. Yuan, R. V. Penty, and A. J. Shields. Efficient decoy-state quantum key distribution with quantified security. *Optics Express*, 21(21):24550, October 2013.
26. Z. L. Yuan, B. E. Kardynal, A. W. Sharpe, and A. J. Shields. High speed single photon detection in the near infrared. *Applied Physics Letters*, 91(4):041114, July 2007.
27. L. C. Comandar, B. Fro¨hlich, M. Lucamarini, K. A. Patel, A. W. Sharpe, J. F. Dynes, Z. L. Yuan, R. V. Penty, and A. J. Shields. Room temperature single-photon detectors for high bit rate quantum key distribution. *Applied Physics Letters*, 104(2), January 2014.
28. Hadley, T, 2018. Quantum Xchange Launches the First Quantum Network in the United States to Provide Quantum-Safe Encryption Over Unlimited Distances. BusinessWire. [weblink]
29. Quantum Xchange, 2018. Quantum Xchange Selects Zayo Group for Dark Fiber to Deploy First Quantum Network in the United States. [weblink]
30. Mink, A., Frankel, S. and Perlner, R., 2010. Quantum key distribution (QKD) and commodity security protocols: Introduction and integration. *arXiv preprint arXiv:1004.0605*.
31. Alshowkan, M., Evans, P., Peters, N., Earl, D., Grice, W., Mulkay, D., Jones, K., Morgan, T., Morrison, S., Newell, R. and Peterson, G., 2021. Field Demonstration of a Multiple Trusted Node Quantum Key Distribution on an Electric Utility Fiber Network. *Bulletin of the American Physical Society*.
32. Lance, A., Leiseboer, J., and Symul, T., 2020. Quantum Key Distribution Systems Compared. White Paper – ID 3676. Quintessence Labs. [weblink]
33. Kaplan, M., 2020. Building Small-Scale Quantum Communication Networks. VeriQloud. [weblink]
34. Pittaluga, M., Minder, M., Lucamarini, M., Sanzaro, M., Woodward, R.I., Li, M.J., Yuan, Z. and Shields, A.J., 2020. 600 km repeater-like quantum communications with dual-band stabilisation. *arXiv preprint arXiv:2012.15099*.
35. Winder, D., 2020. Meet the Scrappy Space Startup Taking Quantum Security Into Space. Forbes. [weblink]
36. Villar, A., Lohrmann, A., Bai, X., Vergoossen, T., Bedington, R., Perumangatt, C., Lim, H.Y., Islam, T., Reezwana, A., Tang, Z. and Chandrasekara, R., 2020. Entanglement demonstration on board a nano-satellite. *Optica*, *7*(7), pp.734-737.
37. Stucki, D., Gisin, N., Guinnard, O., Ribordy, G. and Zbinden, H., 2002. Quantum key distribution over 67 km with a plug&play system. *New Journal of Physics*, *4*(1), p.41.
38. Peev, M., Pacher, C., Alléaume, R., Barreiro, C., Bouda, J., Boxleitner, W., Debuisschert, T., Diamanti, E., Dianati, M., Dynes, J.F. and Fasel, S., 2009. The SECOQC quantum key distribution network in Vienna. *New Journal of Physics*, *11*(7), p.075001.
39. Quantum Key Distribution: Use Cases. ETSI GS QKD 002 v1.1.1. 2010. [weblink]
40. Aguado, A., López, V., Martinez-Mateo, J., Peev, M., López, D. and Martín, V., 2018, March. VPN service provisioning via virtual router deployment and quantum key distribution. In *2018 Optical Fiber Communications Conference and Exposition (OFC)* (pp. 1-3). IEEE.
41. ID Quantique, 2020. Partner Fortinet Commercializes a Quantum-Safe VPN Solution. [weblink]
42. Piotr K Tysowski, Xinhua Ling, Norbert Lütkenhaus, and Michele Mosca. The engineering of a scalable multi-site communications system utilizing quantum key distribution (QKD). *Quantum Science and Technology*, 3(2):024001, 2018.
43. Mink, A., Frankel, S. and Perlner, R., 2010. Quantum key distribution (QKD) and commodity security protocols: Introduction and integration. *arXiv preprint arXiv:1004.0605*.
44. Mohamed Elboukhari, Mostafa Azizi, and Abdelmalek Azizi. Improving TLS security by quantum cryptography. *International Journal of Network Security & Its Applications (IJNSA)*, 2(3):87–100, 2010.
45. Fatima, S. and Ahmad, S., 2021. Quantum Key Distribution Approach for Secure Authentication of Cloud Servers. *International Journal of Cloud Applications and Computing (IJCAC)*, *11*(3), pp.19-32.
46. Dai, G. and Wang, Y., 2014. A Non-entanglement Quantum Single Sign-On Solution. *International Journal of Theoretical Physics*, *53*(4), pp.1143-1149.

47. Park, B.K., Woo, M.K., Kim, Y.S., Cho, Y.W., Moon, S. and Han, S.W., 2020. User-independent optical path length compensation scheme with sub-nanosecond timing resolution for a 1× N quantum key distribution network system. *Photonics Research*, *8*(3), pp.296-302.
48. Woo, M.K., Park, B.K., Kim, Y.S., Cho, Y.W., Jung, H., Lim, H.T., Kim, S., Moon, S. and Han, S.W., 2020. One to Many QKD Network System Using Polarization-Wavelength Division Multiplexing. *IEEE Access*, *8*, pp.194007-194014.
49. Wonfor, A., Dynes, J.F., Kumar, R., Qin, H., Tam, W.W.S., Plews, A., Sharpe, A.W., Lucamarini, M., Yuan, Z.L., Penty, R.V. and White, I.H., 2017. High performance field trials of QKD over a metropolitan network. *Quantum Cryptography (Qcrypt)*, p. 467.
50. Conrad, A., Isaac, S., Cochran, R., Sanchez-Rosales, D., Wilens, B., Gutha, A., Rezaei, T., Gauthier, D.J. and Kwiat, P., 2021, March. Drone-based quantum key distribution: QKD. In *Free-Space Laser Communications XXXIII* (Vol. 11678, p. 116780X). International Society for Optics and Photonics.
51. Lovic, V. Quantum Key Distribution: Advantages, Challenges and Policy. *Cambridge Journal of Science and Policy, 1* (2. e8410270193)https://doi.org/10.17863/CAM.58622
52. Market Watch, 2021. Quantum Key Distribution (QKD) Market 2021: Analysis of Key Trends, Industry Dynamics and Future Growth 2026 with Top Countries Data. [weblink]

How Cox Communications Implemented an Expert System for Service-First Autonomous Operations

A Technical Paper prepared for SCTE by

Dave Norris
Sr Director, Video Engineering
Cox Communications
6205-B Peachtree Dunwoody Rd
Atlanta GA, 30328
dave.norris@cox.com

1. Introduction

At the beginning of the current video evolution to deliver live IPTV streaming to subscribers, the Cox Video Engineering team recognized they were entering a period of unprecedented change and growth. There would be many challenges coming fast to high-capacity video delivery. We knew we'd need to absorb and incorporate new methods of compressing and delivering video without dropping the preexisting systems. At the same time, we were approaching end-of-life status on our national video compression systems. All these factors drove the need for and timing of deploying a new and more flexible video compression platform.

In 2017, the Cox Video Engineering team evolved our national video encoding systems into a virtual platform. For agility, flexibility, API controllability, and ultimately economy, we selected a software video encoder running on commodity servers in dense blade-center chassis. One of our chief concerns was that deploying a virtualized video encoder might be less reliable than the legacy hardware-based appliances. Any negative impact to video stream reliability would be highly visible, driving subscriber frustration, calls, and expense.

We decided to close any performance gaps by implementing a closed-loop Expert System solution to manage all system day-1 build and day-2 operational functions.

This paper will present some of the reasons that drove the evolution to software video encoding, some unexpected motivation that expanded our scope, a few details about the phases of the implementation, and the results.

2. Motivation and Drivers

At the end of 2015, we recognized that in the next 18-24 months we'd need to replace our aging national video encoders. We realized that Cox would be expanding our delivery of MPEG-4 to commercial and residential customers, as well as growing our offering in OTT-style IPTV live streaming for both in and out-of-home uses. We opened an RFP and began the search to find our new video encoders. At that point, no single vendor offered everything we were looking for.

It quickly became apparent that the software-based vendors had the most up-to-date feature sets, and the least disruptive pathway for upgrades. Our requirements included things like: time-aligned multi-bitrate ("MBR") encoding for adaptive ("ABR") playback; accepting HEVC contribution streams from content-providers and eventually outputting content compressed in HEVC; ATSC 3 features supporting new audio CODECs and metadata; etc. The hardware-based compression vendors could meet many of the requirements but were challenged when addressing the path and timing for incorporating new features. The software-based compression vendors provided a much faster and more efficient path for incorporating new technologies and features. Thanks to Moore's Law, CPUs (and GPUs) were closing the density gaps that had stifled software video encoding for years.

As our path to software-based compression became clear, we were increasingly intrigued by the prospect of leveraging APIs and scripting to provision both the environment and the video streams on the platform. While server stand-up automation is mature, automating video

compression software was new territory. Video compression devices, whether hardware or software based, traditionally use GUIs for configuration. The operator must manually enter a lot of parameters to define the inputs, outputs, customized filters and processing of the audio and video streams, ad-insertion messaging, and any other necessary metadata. This manual interface interaction is very repetitive and error-prone, particularly when done at a high volume. We recognized we needed to bring automation into the process.

3. Platform Evolution

With the decision to move to software-based encoding, it was time to select hardware for our new platform. We decided to purchase and host our own Cisco UCS blade server chassis and run them as bare metal for video encoding. Bare metal saved us CPU processing overhead vs VMs, and running warm spares enabled the fastest methods of video stream failover. We leveraged PXE boot and Spacewalk to take our hundreds of bare-metal servers from boot-up through the installation of CentOS.

4. Selecting an Automation Strategy

While it was expected that we would be scripting the API calls to load the video streams onto the encoders, we also decided to engage companies offering more advanced automation solutions. We spoke with a vendor offering an “Expert System” automation platform and quickly became motivated by the idea. Instead of just executing a script referencing a database that defines the systems configuration, the Expert System can be programmed with rules and exceptions. It’s also a form of a state machine and can take “expert” actions based on a pre-programmed understanding of the operational states of the components within its scope of authority.

To explain the differences in strategies, consider a self-driving car. It can be programmed to deliver its occupants and cargo to a destination. Scripting can do that in a closed and controlled environment like at a proving ground—just define the destination and route to move the car from here to there. But if you want that car to navigate on real roads, alongside the rest of us and arrive safely, advanced steps must be taken to adapt to the real-world issues it will encounter along the way. Advanced algorithms are necessary to account for the dynamic environment consisting of traffic, adverse road conditions, construction zones, weather, flat tires, etc. The Expert System self-driving car is programmed to adapt and react to each variable like a human expert would. Each of the many various task algorithms are employed as part of the autopilot system as needed to avoid that box in the road, stop for the school bus, or slow down over the rough roads, etc.

We saw a ton of potential in the Expert System approach. Properly designed, this could enable capabilities well beyond simply provisioning thousands of streams automatically. It could become an autonomous operator of the platform after that provisioning. We selected the StratOS Expert System from Sea Street Technologies, and partnered with them in the development process.

5. Development

We created a flat database file that included every parameter needed to build our hundreds of video streams into the encoders. For efficiency, we utilized look-up tables for the highly repetitive audio and video settings based on classes of outputs in MPEG-2 and 4 like SD, HD 1080i, HD 720P-60, single program AC3 audio, primary and AC3 audios, AAC audio, etc.

That database would be loaded into the Expert System to kick off the automatic build activities. Any future additions, changes, or deletions would be conducted in the same way. Once read into the Expert System, it automatically modifies whatever necessary to reflect any changes to the latest version of the database.

At Cox we use the TeleStream (formerly Tektronix) Sentry product family as video probes. The Sentries monitor and alert on issues that impact audio and video quality. These can be conditions like audio silence or excessive loudness, loss of stream, frozen video, tiling or blocky video artifacts, among many other conditions. When any Sentry sounds an alarm condition, SNMP traps are sent to collectors in the Network Operations Center and are then forwarded to the NOC personnel for triage and escalation to the correct fix agents.

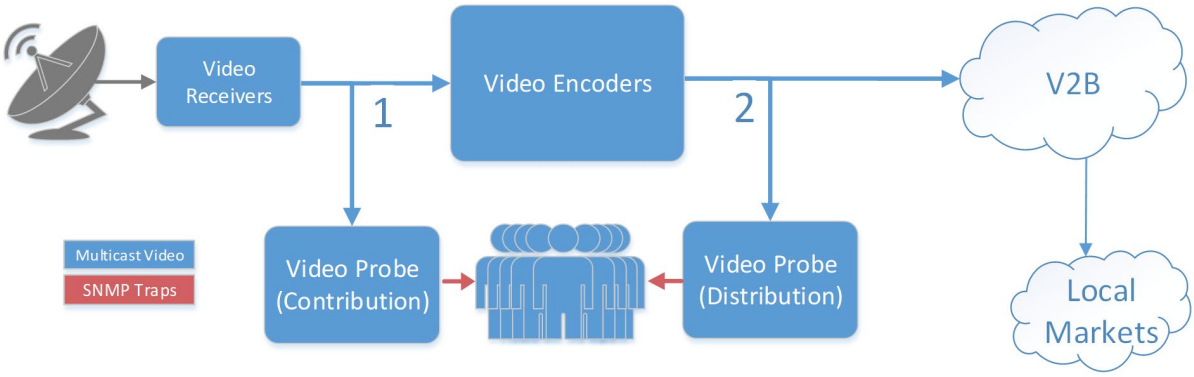


Figure 1 – Humans respond to SNMP traps from the Video Probes on the Contribution and Distribution sides of the video encoders

We integrated the Contribution and Distribution probes SNMP alerting capabilities into the Expert System by adding it as an additional destination for traps and then modeling how it should respond as they arrive. This allows the system to be aware of the health and status of the streams at both the input and output sides of the video encoders.

We also documented the steps a human expert fix agent would take based on the good or bad status reports from the video probes. The Expert System was programmed to monitor for traps on the distribution streams, and when an alarm condition exists, to check for the same traps on any

or all the available redundant contribution streams. With knowledge of stream health on both sides of the video encoders it's controlling, our Expert System automatically makes informed "expert" decisions on what to do, in sequenced steps, to restore from the outage or impaired condition.

To define the actions an expert human operator would take during break-fix activities, we created a decision matrix consisting probe-error feedback that defines the problem conditions, and the first few steps the expert would take under those circumstances to restore the stream. For example:

- Output bad, all inputs good→ restart the encoding service
  - If still bad→ mute the primary output and rebuild the stream in spare encoder pool, then remove the stream instance from the primary encoder
- Output bad, primary input bad, secondary input good→ toggle input
  - If still bad→ restart the encoding service
  - If still bad→ mute the primary output and rebuild the stream in spare encoder pool, then remove the stream from the primary encoder
- Output bad, all inputs bad→ do nothing, let the alarms flow to the NOC Operators

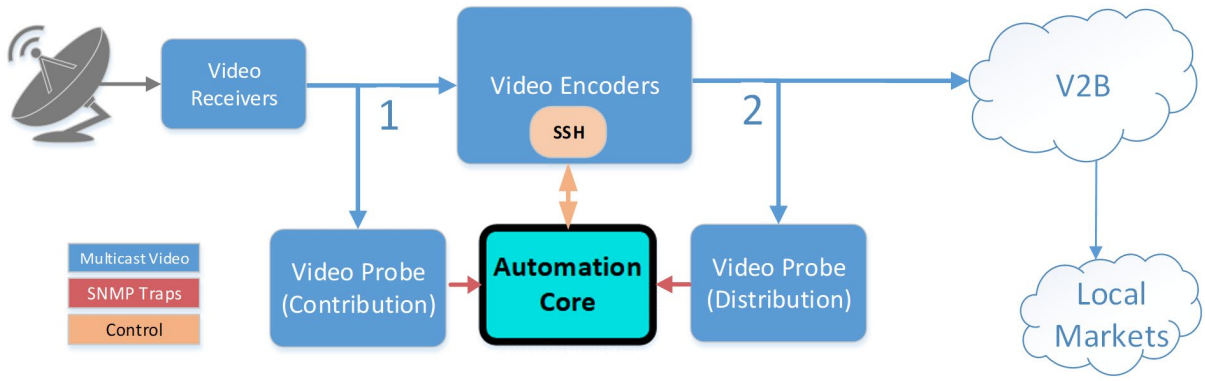


Figure 2 – The Expert System Automation Core Responds to SNMP Traps

6. "Service-First" Self-Assured Autonomous Operations Platform

A Service-First solution is one that focuses on the health of the service over the health of the individual elements. Frequently, automation platforms focus on ensuring each element is healthy but aren't aware of the health of the service itself end-to-end. Working with Sea Street Technologies, we used their StratOS platform to create an Expert System that focuses on the health and remediation of the service as the first priority, before it takes steps to fix the elements.

For example, in our solution when a fault is detected with a video stream, the Expert System will first focus on fixing it by restarting the encoder service or moving it to a spare encoder within seconds of the fault. Only after the stream is reported to be healthy again does the Expert System return to fix the faulted encoder.

The Expert System is designed to manage each video stream from initial creation to daily operations. When new video streams are brought online, the Expert System will autonomously create and configure each one and then start to monitor and independently operate each one indefinitely. This is done through "Objectives," which are continuously running objects that collect health data, execute business logic, and take actions. Instances of an Objective are created for every encoded stream and the instances all run in concert with each other, working to ensure the health of all streams and underlying elements. This method of self-assurance has enabled us to improve the reliability of our video streams, significantly improve MTTR, and reduce the need for additional staffing.

7. Launched!

In 2017, we launched our new software encoder platform under Expert System control. We entered the launch maintenance activity with all the servers online and provisioned only with CentOS. The platform-defining database was loaded into the StratOS Expert System. StratOS interpreted the database and took over the maintenance, loading and licensing the video encoding software before it began to automatically provision approximately 650 streams in the first data center.

The full configuration was completed in a less than an hour. A manual configuration of the same platform would have taken a few months of maintenance-window activities. Since launch, our Expert System autonomous operation platform has been routinely restoring errored streams in seconds with machine speed and accuracy, typically before the NOC is even alerted to the service impact.

8. Continous Improvement

The modular architecture of our Expert System makes the process more straightforward for adding features or modifying parameters. For example, we were able to easily add repair-retry logic to the system for use after extended contribution stream outages.

We've also added a second video encoder vendor's software to our Expert System. This has proven the operational environment to be independent of any single video encoding solution. Our preexisting business and operations logic was simply inherited by the new vendor-specific API resource driver module. Should Cox decide to modify our video encoding automation business rules or operations logic, those changes would immediately be applied to all the Objectives, regardless of the underlying software encoding platform.

9. How's it Working?

After years of successful operation, this has been a trailblazing and fantastic success! Bell Labs was hired to evaluate the state of automation in use at Cox and described our Expert System autonomous operation platform as the most advanced of its kind they had ever seen. They conducted an evaluation of some of the benefits and concluded the following:

- Reduced Eng effort: Fewer Maintenance Windows
  - **.4 Full Time Employee ("FTE") saved (efficiency) per year**
- Reduced Eng effort: Incident management
  - 100 in-scope incidents per week
    - **38 minutes saved per incident (T2)**
    - **15 Minutes saved per incident (T3)**
    - **2.2 FTE saved (avoidance) per year**
- Reduced Eng effort: Platform Process savings
  - **Cut time to process by 90%**
  - **1.9 FTE saved (efficiency) per year**
- Reduced Eng effort: Software release upgrade
  - **.5 FTE saved (efficiency) per year**
- Team avoided FTE growth
  - **5.0 FTE saved (avoidance) per year**

Per Bell Labs calculations, our Expert System has saved Cox over $10 Million, and continues to provide savings of $2 Million per year.

Abbreviations

ABR	Adaptive Bitrate
API	Application Programming Interface
GUI	Graphical User Interface
HEVC	High Efficiency Video Coding ("MPEG-5")
MBR	Multi-Bitrate Encoding
MTTR	Mean Time to Repair
NOC	Network Operations Center
SNMP	Simple Network Management Protocol
V2B	"Video to the Backbone" -Video distribution via the Cox backbone
VM	Virtual Machine
GUI	Graphical User Interface

How Network Topology Impacts Rf Performance: A Study Powered By Graph Representation of the Access Network

A Technical Paper prepared for SCTE by

Maher Harb
Director of Data Science, Next Generation Access Network
Comcast
maher_harb@comcast.com

Karthik Subramanya
Senior Engineer, Next Generation Access Network
Comcast
karthik_subramanya@comcast.com

Ramya Narayanaswamy
Senior Manager, Next Generation Access Network
Comcast
ramya_narayanaswamy@cable.comcast.com

Sanket Walavalkar
Executive Director, Next Generation Access Network
Comcast
sanket_walavalkar@comcast.com

Dan Rice
VP, Next Generation Access Network
Comcast
daniel_rice4@comcast.com

1. Introduction

We have recently embarked on a project with the aim of capturing all the building blocks of the access network, their relationships, and their properties in a graph representation encompassing vertices and edges. This representation is to be available in a high performance and scalable graph database that allows access to the data through application programming interface (API) endpoints and in batch. The graph database mirrors the dynamic nature of the network by getting updated as customers get connected & disconnected, optical nodes get segmented, network equipment gets commissioned & decommissioned, as well as the happening of any other impactful network change. Having all the relational information in one source, and combining the physical & logical elements in a single view allows analyzing the access network at any level of network topology (e.g., service group, fiber node, amplifier, tap, drop) on a use case basis. The graph database technology also allows enrichment of the data with ease by overlaying device telemetry, Cable Modem Termination System (CMTS) telemetry, and maintenance data on top in order to implement algorithms for business intelligence and troubleshooting (e.g., root cause analysis).

Building the graph database required combining and reconciling data across many different sources of the organization without identified primary keys (ids) and creating algorithms to automate inference of connections. In this paper, we share Comcast's journey into this process that is currently scaled to cover ~20% of our footprint. We present the very first use case of utilizing the network graph to study the effects of the amplifier cascade length on radio frequency (RF) performance in the upstream (US) and downstream (DS). The interest in this investigation falls within a broader question on the operational effort required to maintain nodes with large cascade of amplifiers (both in terms of depth and breadth). Our findings reveal that, predictably, longer amplifier cascade lengths exhibit degraded RF signal-to-noise ratio (SNR) -- yet with no significant impact on quality of service, likely due to the mitigating impact of the profile management application (PMA) system currently deployed in Comcast.

We acknowledge contributions to the project from former team members Doga Keresteciogl u, Matt Lord, and Athanasios Tsiaras.

2. Constructing the Graph

The Comcast access network data platform is an enterprise-scale graph data platform that maps the entire access network from the CMTS to the Customer Premise Equipment (CPE) while incorporating all of the physical and logical elements that form part of the network. At a very high level, the graph platform brings together site/headend topology, computer aided design (CAD) physical plans, telemetry, and billing systems together to construct the access network graph as one connected entity (see illustration in Figure 1).

While the individual data sources that form the building blocks of this graph data platform are highly mature with rich offerings, they have varied goals and are managed by various teams. There are no common keys that connect the boundaries between these data sources to form one connected access network graph. Hence, the need for such a graph data platform assumes significant importance to drive various data science applications and serve as a source of truth for deriving relationships between various access network components.

The platform supports a rich set of use cases from anomaly/network deterioration detection, triangulation, capacity planning [1] and various aggregation use cases. In addition, the platform uses several statistical

and data science techniques to bridge boundaries between individual data sources by inferring relationships and provides a single unified view of the access network.

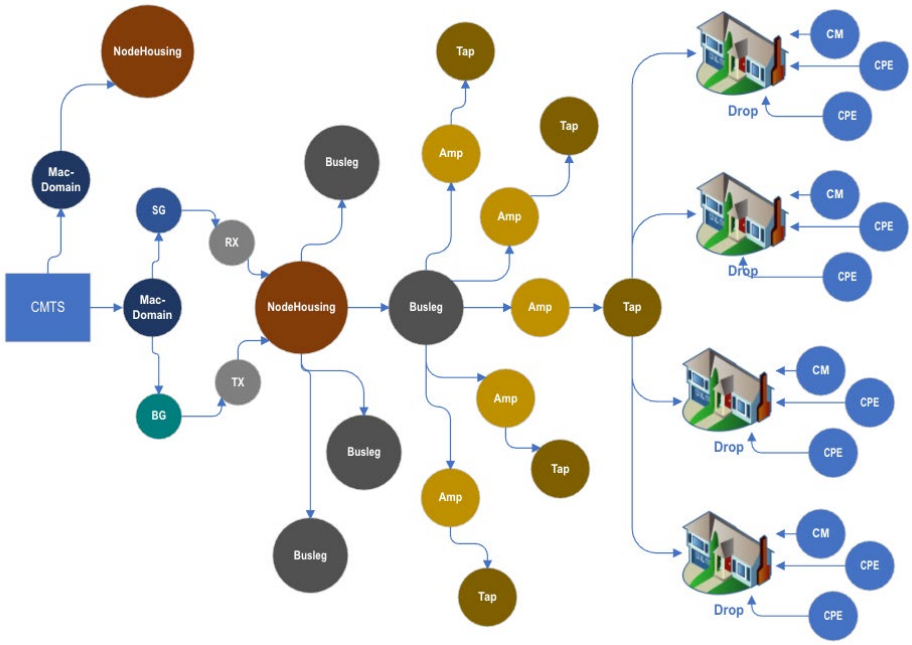


Figure 1 - Schematic representation of some of the main entities captured in the graph database.

By the virtue of being a scalable graph platform that is being purpose-built within the cloud, we're able to build multiple data pipelines to ingest various data sources and rapidly experiment and iterate through various algorithms, data processing, and refreshing techniques that feed into the end-product. The graph defines vertices and corresponding attributes for CMTS, nodes, RF equipment such as bus legs (port on a node), amplifiers, passives, taps, drops, and customer entities such as household, devices entitled, and so on. Each of these vertices are connected by edges, and in the case of equipment, these edges define the attributes of RF cables that form the connection (e.g., cable length). This graph platform allows us to query and traverse the network in either traffic direction and start at the most granular level (i.e., the CPE) and go all the way up to the CMTS or start at the CMTS and terminate at the CPE.

One of the many functions that the graph platform realizes is to diagrammatically connect physical address drops that are defined in CAD documents, to the appropriate households, which are logical elements defined in billing systems. Since there are no common keys connecting these 2 elements, the platform performs address standardization, further employs various Natural Language Processing (NLP) techniques and coordinate-based proximity to identify the right households and match them to the appropriate drops. The telemetry data set that consists of online cable modem inventory and CMTS configurations helps cross-validate these inferred relationships using reported MAC (Media Access Control) domain, bonding group, service groups, and so on.

A significant engineering challenge that was solved during the development of this platform is to account for graph refreshes. The access network consists of components that independently refresh at different rates. Hence, we needed a decoupled solution that independently refreshes various sub-graphs, even within the footprint of a single CMTS. We use property-graph architecture to implement our data model and use Apache Gremlin Tinkerpop to perform traversals, look ups, or aggregation queries. We often

notice that the depth of traversals for our queries is more than 40 layers deep, indicating the extent of cascading and density of our access network. These traversals would never be effective for any relational or NoSQL database. In contrast, through the use of large-scale data processing platforms, we're able to complete the required traversals or queries on the graph platform in a matter of minutes, for the entire footprint.

Figure 2 is a screen capture from a graphical interactive tool used to explore the graph database. It allows retrieval of attributes for vertices and edges by clicking through the graphical interface.

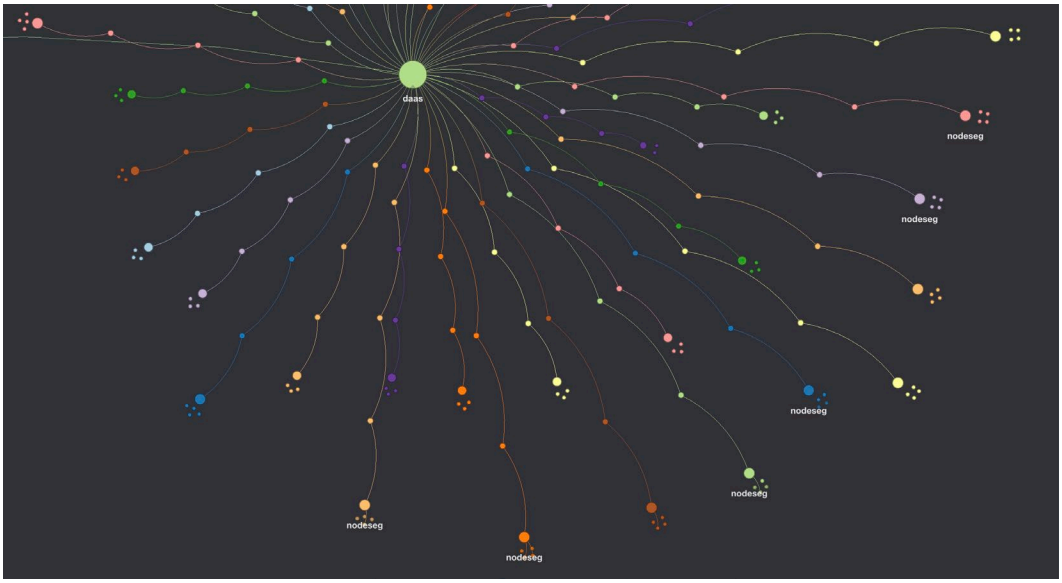


Figure 2 - Screen capture from Graphistry—an interactive graph map that allows browsing attributes of the different vertices and edges by hovering over the element with the mouse tip.

3. Visualizing the Graph

The ability to visualize the network in the form of a graph, which encompasses vertices and edges, serves a multitude of purposes. These include validating the graph construction algorithms by thorough visual inspection of the outcome, uncovering errors & inconsistencies where they may appear, and supporting use cases related to troubleshooting network issues by layering key telemetry data on top of the basic topological view. There exists a host of libraries for the purpose of graph visualization—both in the form of stand-alone proprietary software as well as open-source packages that integrate with Data Science programming language such as R and Python. For this analysis, the R packages *tidygraph* and *ggraph* were adopted for creating graph visuals. They both follow the data principles established by the popular *tidyverse* family of R packages used for data wrangling and visualization [2], thus, allowing enrichment of the graph data with ease.

The graph database at Comcast establishes relationships that span thousands of CMTSs, hundreds of thousands of optical nodes, and millions of households. For visualization purposes, the task is limited to producing a visual for a very small subset of the full network at a time and on demand. Typically, the interest is in plotting the tree-like structure that connects devices to the same fiber optics node (also referred to as a busleg/port on a node clamshell). The pathway from the device to the node includes elements such as amplifiers, passives, taps, and drops. This view is very relevant to troubleshooting of access network problems because issues originating in elements under a node may impact multiple

customers (this is especially true in the upstream direction in which noise is known to funnel). In contrast, customers connected to different nodes are usually isolated from each other. Figure 3 shows an example of a graph visual for a single node. The power of the graph is manifested in its ability to combine physical elements and logical elements in the same view. For example, the customer's billing address (shown as black square) is included in the graph and hierarchically positioned between the drop and the cable modem (tree leaf). Notice that some customers have multiple cable modems under the same billing address. These include the household internet gateway and one or multiple DOCSIS video set-top boxes.

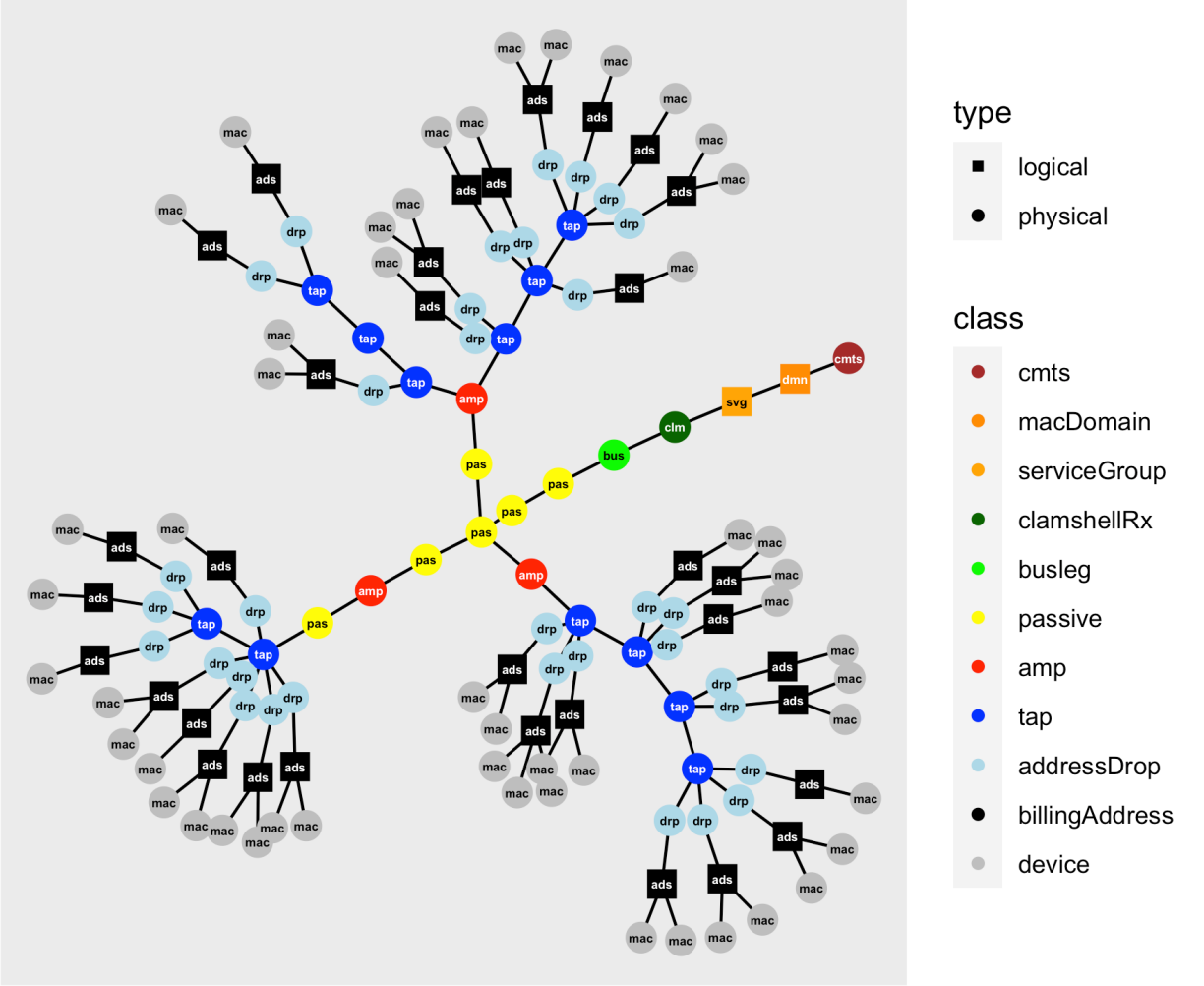


Figure 3 - Example of a basic graph visual for a relatively small size node. The tree trunk (brown circle) corresponds to the CMTS, and the leaves (gray circles) correspond to the cable modems/IP devices. The pathway connecting the two traverses multiple physical and logical elements as shown here.

Figure 3 represents a basic view of the graph that can be enriched in many ways. A few illustrative examples are considered henceforth. In the first example, shown in Figure 4, the thickness of the edge is utilized to designate the traffic volume as measured by the number of devices that transmit/receive traffic through that link. In the same figure, the number of devices is also annotated on the graph vertices. Alternatively, the weights can be easily adjusted to correspond to actual traffic volume rather than number of connected devices.

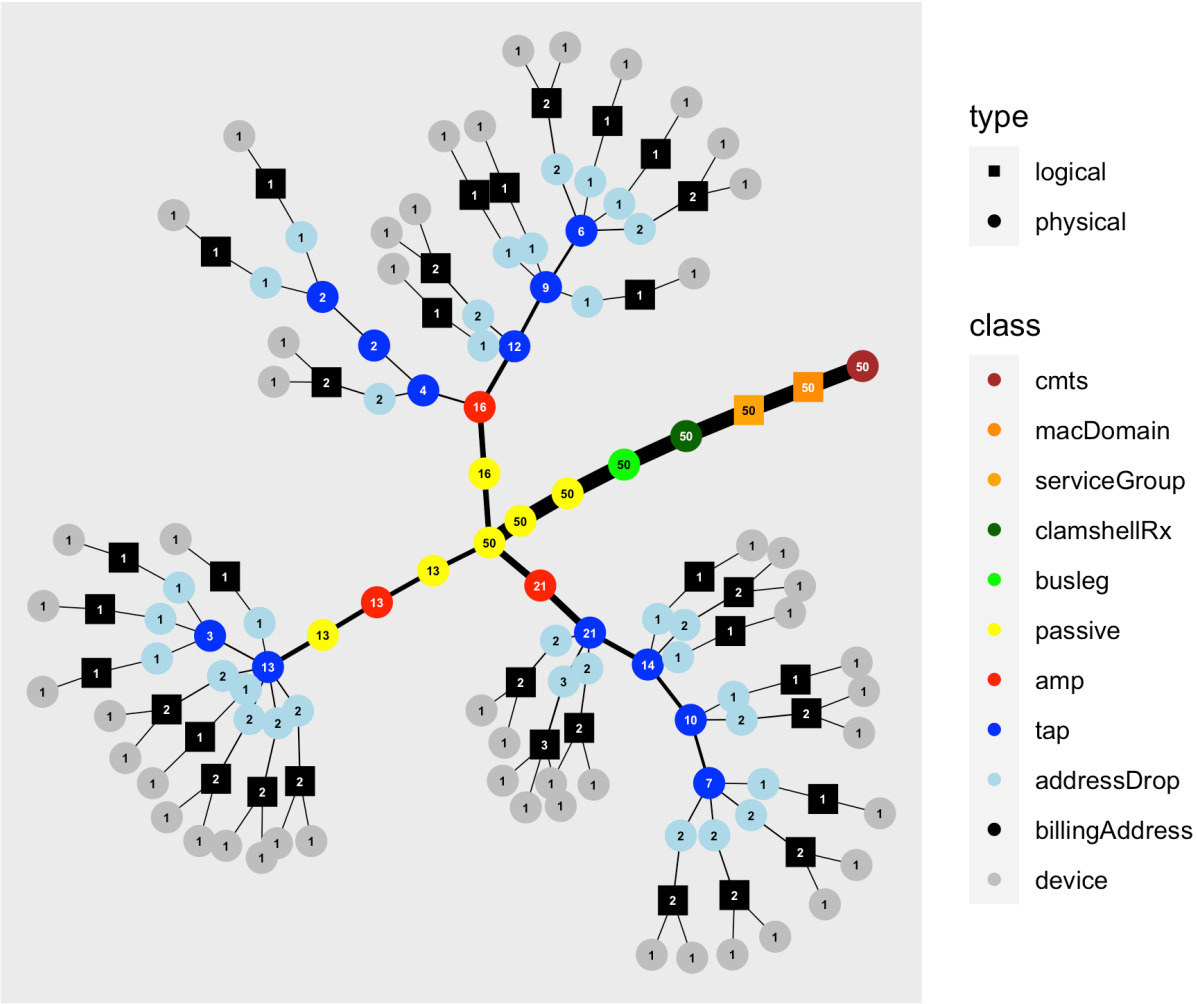


Figure 4 - Example of graph visual in which the thickness of the link (edge) is weighted by number of cable modems connected through the link. The label on each vertex represents the total number of cable modems that are hierarchically located underneath the vertex.

The example shown in Figure 5 overlays telemetry data onto the topological view. In this example, the US device-level SNR is coded as the color of the device symbol. Such view is useful for visual identification of "hot spots". These could be a cluster of devices suffering from the same impairment and in which the graph visual provides a clue to the problem root cause.

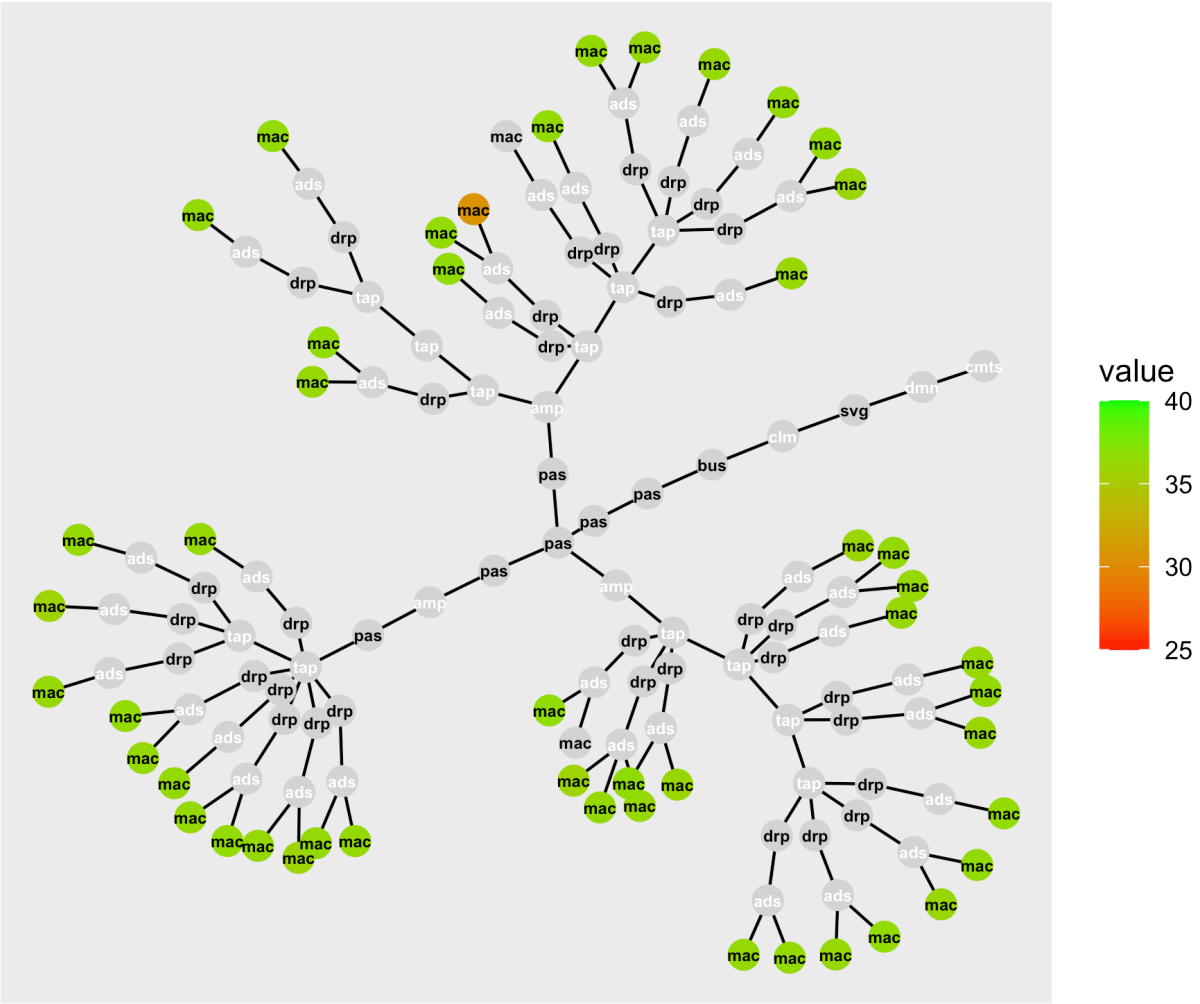


Figure 5 - Example of a graph visual in which the color of the cable modem symbol indicates the upstream SNR on the 25-to-40 dB scale shown in the adjacent color bar. The one cable modem colored gray had missing telemetry data for that polling time sample.

4. Amplifier Cascade Length Analysis

A question that arises frequently in discussions around network design best practices is the impact on the customer experience of a large cascade of amplifiers. From an RF design perspective, it is accepted that amplifiers introduce distortions to the signal—there is no such thing as an "ideal amplifier". Albeit there is an implicit assumption that distortions are tolerated if they are deemed to be within an acceptable design range. However, the reality is that network growth is dictated by customer demand and geography, and often deviates from original plans, causing limitations on node size to be exceeded. In this context, no hard written rule exists on the maximum allowable amplifier cascade length. Furthermore, there is no straightforward approach to quantifying the impact of the amplifier cascade length on customer experience. In fact, examining this question was the very first use case of the network graph database.

The analysis dataset contains the ~20% of our footprint's CMTSs that are fully captured in the graph database at the time of writing this paper (summer of 2021.) The basic idea behind the analysis is to correlate the node size with key telemetry data and explore this relationship in depth. The very first task within the analysis requires defining what is meant by "node size" in relation to the RF amplifiers. It was decided to explore the following two features:

- **Amplifier cascade length**: This is a device-level feature that represents the total number of amplifiers traversed in the path between a cable modem and the node (see Figure 6).
- **Total number of amplifiers**: This is a node-level feature that represents the count of all amplifiers within the node.

The rationale behind the choice of cascade length and total amplifier count was to accommodate the different noise accumulation behaviors between downstream and upstream paths. In the downstream, noise does not funnel between devices. Therefore, what matters from a distortion perspective is the number of amplifiers in a device's path. Whereas in the upstream, noise funneling causes all amplifiers to potentially contribute to the distortion for any given device in the node.

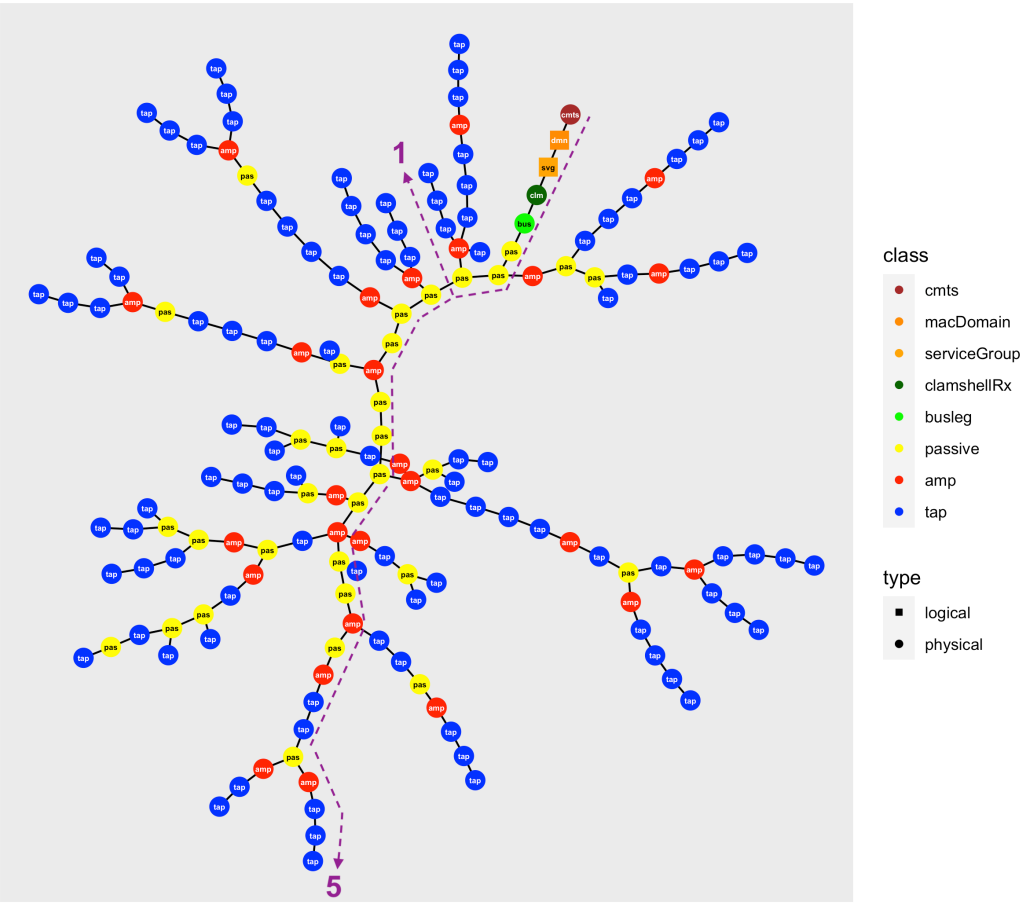


Figure 6 - Visual example of a large node in which devices, billing addresses, and drops were removed for clarity. Two paths from CMTS to tap are highlighted in purple. One path traverses a single amplifier and the other traverses 5 amplifiers. Devices attached to the latter are expected to experience larger amplifier distortions in the downstream path.

Figure 7 shows the distribution of the two features across the ~20% of CMTSs mapped in the graph database. The amplifier cascade length distribution is unimodal with a peak at 2, meaning that the majority of devices connect to the node via a cascade of 2 amplifiers. The distribution in the figure was intentionally cut off at 10, even though the data contains outliers with cascade lengths that exceed this value. The total amplifier distribution has a peak at 1 and gradually diminishes at ~40 amplifiers. It may be surprising to see ~3000 nodes served by a single amplifier. These are small size nodes in terms of either the number of customers or the geographical extent of the node (or both).

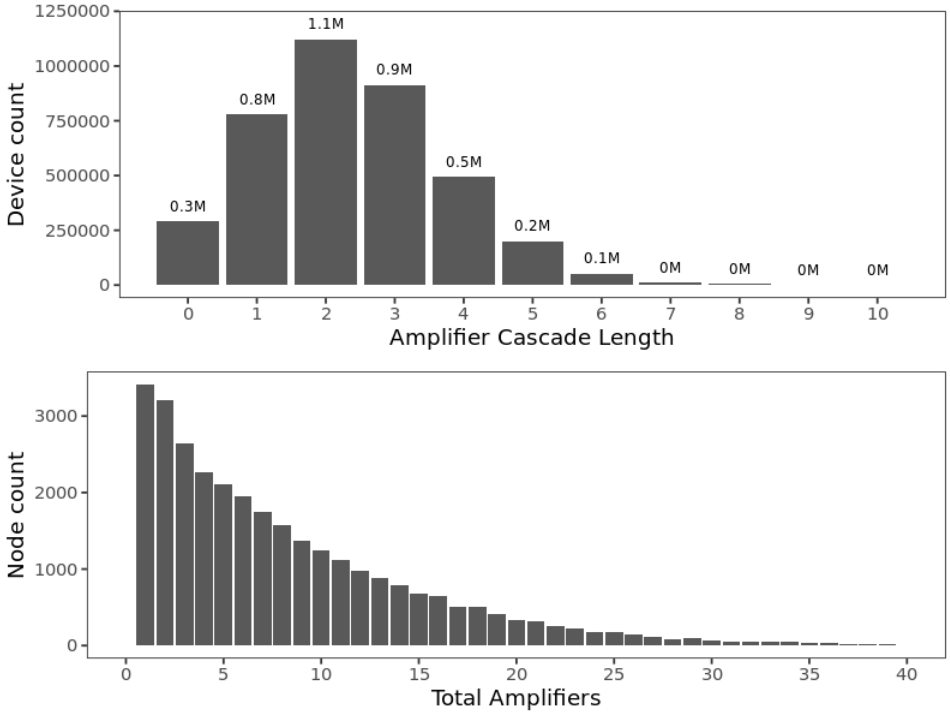


Figure 7 - Distribution of amplifier cascade length across devices (top panel) and distribution of total number of amplifiers across nodes (bottom panel). This data was generated by querying the graph database.

With the features derived from the graph database extracted, the next step involved identifying and examining telemetry data in relation to the graph features. The following variables were considered in the analysis:

- **Upstream Signal-to-Noise Ratio (US SNR)**: 10th percentile of a device's upstream SNR samples (i.e., data is collected from multiple time polls and aggregated)
- **Upstream Forward Error Correction (US FEC)**: The percentage of time a device polled when transmitting upstream experiences an uncorrectable codeword error rate > 0
- **Upstream Partial Service**: The percentage of time a polled device goes into partial service with respect to an US channel
- **Upstream Power**: mean device power level
- **Downstream Receive Modulation Error Ratio (RxMER)**: A device's OFDM channel RxMER samples (data resolution by 50 KHz subcarrier)

The correlations between graph features and the variables above were explored and are shown in Table 1. One glaring outcome was the lack of a correlation between graph features and the FEC metric. The

absence of correlation with FEC is attributed to the mitigating effect of Comcast's Profile Management Application (PMA). PMA was deployed for both DS D3.1 (OFDM) and US D3.0 since early 2020 [3,4]. Under PMA, channels that exhibit degraded spectrum get assigned a configuration that ensures that devices continue to use the spectrum without experiencing unacceptable levels of uncorrectable errors. While the algorithms are quite different between PMA for DS D3.1 vs. US D3.0, the result is the same: for those degraded channels, capacity is traded off for robustness. Hence, the lack of correlation is an assuring sign that PMA is doing its intended job.

Table 1 - Correlations between the node features and device telemetry.

	US SNR	**US Rx Power**	**US FEC**	**US Partial Service**	**DS OFDM MER**
Amplifier Cascade Length	-0.04	0.03	0.01	0.00	-0.12
Total Amplifiers	-0.05	0.04	0.02	0.01	-0.07

The strongest correlation exists between amplifier cascade length and OFDM RxMER. In Figure 8, two views supporting this trend are presented. The first shows different aggregating RxMER percentiles vs. amplifier cascade length. In all of these, there exists a trend of decreasing RxMER with increasing amplifier cascade length. Trend lines are included in the plot, which one can use to estimate an effect of ~2 dB for every 10 amplifiers added to the path (a more proper calculation, based on regression model, is presented below). The second view shows the distribution of modulation assigned to each subcarrier based on the standard D3.1 RxMER-to-modulation mapping for DS [5]. In this view, no aggregation is done as every subcarrier contributes to the distribution. Once again, there's a clear effect that is most visible in the diminishing proportion of 4096-QAM as the amplifier cascade length increases.

A question that remains is to quantify the effect of this relationship on customer experience. Given that PMA is mitigating the impact of low RxMER by dynamically managing OFDM profiles, we restrict the definition of "customer experience" to the "capacity" dimension. Below, a model is introduced that quantifies the impact of the amplifier cascade length on available capacity. In the first step, a linear regression model is fitted to that RxMER data as presented in Figure 9. The linear model yields a statistically significant relationship (with a p-value $< 2 \times 10^{-16}$) between RxMER and the amplifier cascade length. The relationship is outlined in the following equation:

$$\text{RxMER} = 37.7\ \text{dB} - 0.37 N_{Amps}\ , \tag{1}$$

in which N_{Amps} is the length of the cascade. In other words, every additional 10 amplifiers reduce RxMER, on average, by 3.7 dB.

An interesting feature of the overall distribution of RxMER is that the 38 dB line falls through the middle of the distribution (notice that the linear fit intercept is 37.7 dB). This level happens to be the PMA threshold for assigning a modulation of 4096-QAM, the highest possible under the current vendor implementation of D3.1. The threshold was deliberately set to be 3 dB lower than the recommended value under the D3.1 specification [5] (i.e., it is more aggressive). This means that for more than half the population (at and below this level), reduction in SNR due to increasing amplifier cascade length may

cause demotion to lower modulations. Indeed, this explanation agrees with the trend shown in the lower panel of Figure 8.

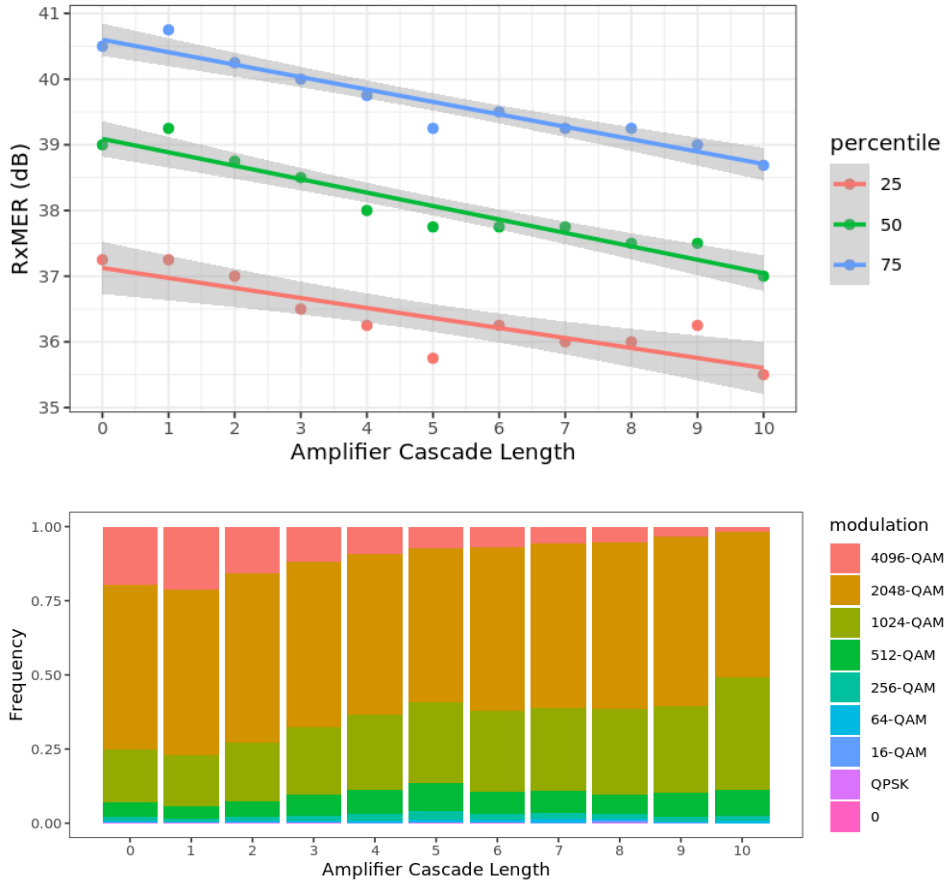


Figure 8 - The top panel shows the relationship between RxMER and amplifier cascade length for different aggregating percentiles (25th, 50th, and 75th). There's a clear downward trend, as highlighted by the linear best-of-fit lines. The bottom panel shows the corresponding modulation distribution. Once again, the trend is visible by the decreasing ratio of 4096-QAM as the amplifier cascade length increases.

To translate the results from the linear model into a capacity impact figure-of-merit, we turn to the Shannon capacity theorem [7]. In a previous SCTE contribution [3,4], we argued that D3.1, with its low-density parity check (LDPC) error correction algorithm, operates close to the Shannon limit. The Shannon theorem can be approximated in the large SNR regime as follows:

$$C \approx 0.332 \cdot B \cdot \mathrm{SNR(dB)}\,, \tag{2}$$

where C is capacity, B is bandwidth, and SNR is measured in dB. The derivative of the equation above yields:

$$\Delta C \approx 0.332 \cdot B \cdot \Delta\mathrm{SNR(dB)}\,, \tag{3}$$

which provides a recipe for translating small changes in SNR to changes in capacity. Based on the above, for a standard 96 MHz wide OFDM channel, the results from the linear model translate into ~118 Mbps reduction in capacity for every 10 amplifiers traversed in the node-to-device DS path.

As demonstrated in this analysis, the effect is measurable and statistically significant. However, it is not impactful in the context of DS capacity for several reasons. First, the 10 amplifiers represent the upper bound of the distribution, i.e., this is an extreme scenario, as the bulk of the population falls below a cascade length of 5 amplifiers. Second, the spread in RxMER values within each "bucket" is much wider than the effect of increasing cascade length. This implies that there are other pressing issues one can solve before turning attention to cascade lengths. Third, even with loss of ~100 Mbps of capacity, D3.1-capable devices have access to sufficient D3.0 and D3.1 spectrum to support our highest speed tiers.

The same analysis was conducted for the US, exploring the relationship between US SNR and the total number of amplifiers in the node. Once again, the linear regression model shown in Figure 10 reveals a statistically significant relationship. However, the effect of increasing the number of amplifiers is even less impactful on customers compared to the DS, because of the large safety net that is intrinsic to a D3.0 US: while the highest possible modulation under D3.0 US is 64-QAM, the distribution of SNR lies ~10 dB above that level.

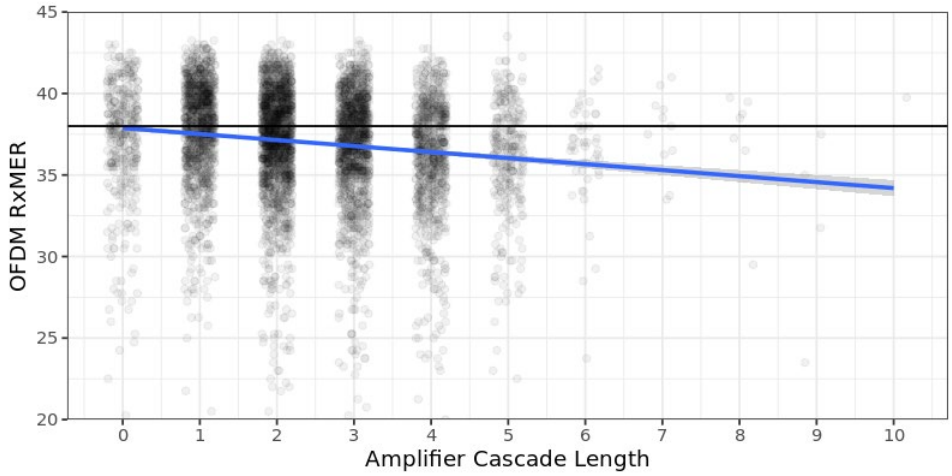


Figure 9 - Linear Fit to the RxMER vs. amplifier cascade length data (blue line with gray confidence band). The horizontal black line, at 38 dB, is the current threshold in PMA for assigning a 4096-QAM modulation.

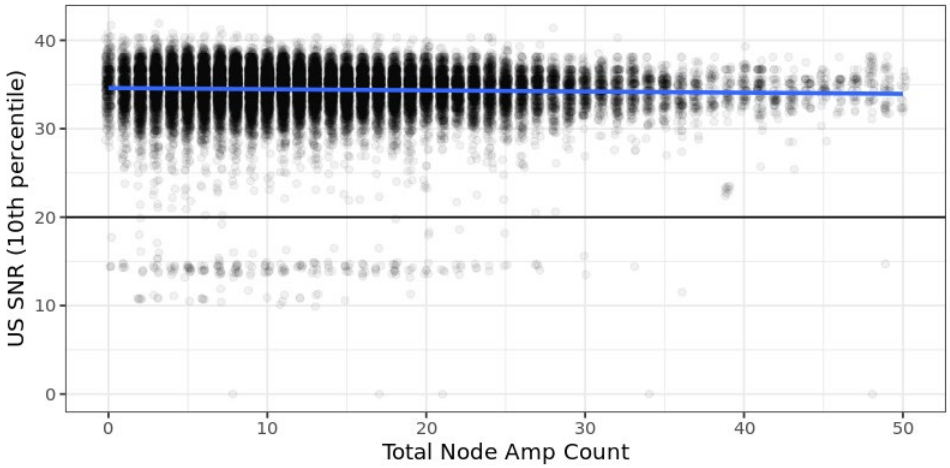


Figure 10 - Linear Fit to the US SNR vs. total number of amplifiers (blue line with gray confidence band). The horizontal black line at ~20 dB is the current threshold in PMA for assigning a 64-QAM modulation.

5. Conclusion

This paper highlights the power of the graph database in getting a better understanding of our network, and especially to track all physical and logical elements within the network. Using the relevant information, we were able to analyze and measure the impact of the amplifier cascade length on RF performance. The outcome was in the form of guidance on the effect of increased cascade length on capacity. While the impact on customers is assessed to be minimal today, future evolution of the network will bring service offerings that push the physical bandwidth to its limit (e.g., symmetrical Gbps service). These developments will make it critical to have a thorough understanding of the impact of network topology on customer experience—beyond the particular "amplifier cascade length" feature.

The graph database is being scaled up to cover our entire service footprint. In addition to the use case presented in this paper, there is much excitement about utilizing the graph database to build algorithms for root cause analysis and noise triangulation. These examples constitute a "holy grail" for us and the industry, as they promise to significantly cut down the labor-intensive troubleshooting processes involved in locating sources of noise/ingress. Given that machine learning techniques that utilize graph data representation exist today and are mature enough to support the intended use cases, the remaining bottleneck to conquer is to complete the construction and maintenance of the high-quality graph database presented in this paper.

Abbreviations

API	Application Programming Interface
CMTS	Cable Modem Termination System
CPE	Customer Premise Equipment
D3.0	DOCSIS 3.0
D3.1	DOCSIS 3.1
DOCSIS	Data Over Cable Service Interface Specification
DS	Downstream
MAC	Medium Access Control

NLP	Natural Language Processing
OFDM	Orthogonal Frequency Division Multiplexing
PMA	Profile Management Application
RF	Radio Frequency
RxMER	Receive Modulation Error Ratio
US	Upstream
SNR	Signal to Noise Ratio

Bibliography & References

1. *"Access Capacity Planning: Staying Well Ahead Of Customer Demand Helped Ensure Stability During COVID-19", B. Baker, C. Bou Abboud, E.Neeld. NCTA technical paper, 2020.*
2. *"R for Data Science", H. Wickam,G. Grolemund, O-Reilly, 2017.*
3. *"A Machine Learning Pipeline for D3.1 Profile Management", M. Harb, J. Ferreira, D. Rice, B. Santangelo, and R. Spanbauer, NCTA technical paper, 2019.*
4. *"Full Scale Deployment of PMA", M. Harb, B. Santangelo, D. Rice, J. Ferreira, NCTA technical paper,2020.*
5. *"Data-Over-Cable Service Interface Specifications DOCSIS 3.1, PHY Layer Specificaiton,CM-SP-PHYv3.1-I18-210125", Cable Labs, https://community.cablelabs.com/wiki/plugins/servlet/cablelabs/alfresco/download?id=f00df402-7367-4f86-a35c-5c22a2bfbaed*
6. *"Practical Lessons from D3.1 Deployments and a Profile Management Application (PMA)", NCTA technical paper, 2019.*
7. *"A Mathematical Theory of Communication",C.E.Shannon,The BellSystem Technical Journal,vol27,pp. 379-423,623-656 July,October1948*

How To Not Pop the Balloons —

Migrating The Analog Headend for The Digital Broadband Future Facility

A Technical Paper prepared for SCTE by

Benjamin Strunk
Senior Director – Power and Infrastructure
Comcast Cable – Northeast Division
1131 South Duke Street Lancaster Pa 17601
717.505.1410
Benjamin_Strunk@comcast.com

1. Introduction

The transformation of the legacy broadband buildings into Broadband Digital Nexus (BDN) structures that enable the digital connections to millions of people and businesses daily, is at a crossroads. At a fast-approaching time horizon, all legacy analog sites will need to change from housing an analog-facing technology to a digital infrastructure. These structures will gain more useful life in the cable broadband ecosystem. This paper will show in detail the transformative steps achievable in hundreds of locations simultaneously or individually. All these steps can be accomplished while maintaining a limited or negative amount of critical infrastructure equipment. When comparing energy consumed in an analog facility versus a digital facility, the initial data is showing an energy reduction of 200%. Coupling these transformations to adopting sustainable building technics and energy reduction options, the space and power of Broadband Digital Nexus structures is presenting opportunities hidden within physical constrictions setting operators up for rapid growth and digital transformations in the field.

2. Pump up the Volume

To understand the start of the Community Cable Television Antenna (CCTA) systems, we should look back into the ways of the telephone. The switching phone equipment and operator system needed to be housed within a structure. It sounds obvious today, but think about what that required for the telephone and telegraph systems. Engineers needed to evaluate and string telephone wires along with the power conductors of the time. The brick and block buildings that housed this function made possible the birth of the telephony exchange or Central Office (CO). These simple repeatable structures also were built for vertical expansion. Countless buildings were built on a first floor and then an expansion to a second floor and upwards. The telephony buildings were, in fact, the beginnings of the world's internet.

As the usage of the Bell and ATT telephony systems grew, the physical network equipment in the COs were forced to accommodate these new switches. These digital exchange switches were classified as Class 5 switches. These switches came in a vast variety. There was the Bell Technology Lucent 5ESS. The General Telephone Equipment GTD5. Siemens even had their own version I called the Brown Monster. All these large main frame switches ushered in the dawn of the electronic switching age. With the scope, growth and size of these switches, many telephony buildings needed a way to provide energy backup and 24/7 operations. The solutions were the adoption and full usage of DC batteries and power conversion units provided by AC to DC conversion systems. This architecture enabled the telephone powered network. This powering system allowed for the uninterruptable services that society began to rely on for communication, comfort, and business.

As we follow the communication timeline to the 50s, we have the start of the satellite age with the launch of Sputnik and the broad licensing of over the air broadcast television. Amplified Modulation and Frequency Modulation. The world at this time was "wired" for power and telephone and "wireless" for entertainment/news. Communication infrastructure is poised to get another player into the content and via for eyeball time of the population. This brings us to the need for video infrastructure.

The video ingest of the time was via an antenna. The tower would have multiple antennas attached facing a tuned direction to pick up the broadcaster's signals. These signals were then received and attached to the coaxial cable within the building. Encoders and other devices allowed the RF signals to be embedded and modulated upon the coax and to leave the building. The stringing of the coax cables to the subscriber base enabled the community part of the cable systems to fledge. The

greater the height of the tower, the better the signal or more signals were able to be received. More receive signals means more channels and more equipment needing space in the buildings.

Figure 1: Image of CATV headend-image by author – circa 03/2016

3. Cup…..String….Cup

The available bandwidth on the coaxial cable was being utilized to its known capability. The MegaHertz signals were determined by the thickness of the coaxial cable, distance between amplifiers and pure decibel signal loss over the distance to the customers. The invention and applications for fiber optical cable allows for the manufacturing and production of radio frequency nodes to be connected into the field. These nodes allowed for the usage of Headend-located transmitters, situated into equipment cabinets. These transmitters received a RF signal over coax and converted the signals to analog optical light that would be transmitted along the glass until it landed at an RF node in the field. At the receiver housed within the field node, the receiver injected the optical analog signal and converted the light amplitudes back to a radio frequency to be embedded upon the RF coax cable to be run down past the homes it was going to serve. This "Cup-String-Cup" methodology allowed the distance of unpowered optical plant to stretch for kilometers. The longer the distance achievable by the lasers, the closer the node could be placed to the end customer.

The lasers began as large form factor units, sized at approximately 6 Rack Units (RUs) per transmitter or receiver. With a standard rack being 19" wide, that's a transmitter-to-receiver density of approx. 3500 cubic inches (19w x 10.5" tall X 18" deep). Given the usage of the transmitter, receivers and the adoption of combining on the RF forward and the RF return decombining… the spatial needs of the headend started to move into not just active equipment, but a new category of passive equipment. The need for passive equipment driven by cost savings, testing necessity, and optimization.

Cost savings on the transmitters happened with the adoption of service groups. Take 1 Transmitter and split it to feed 2 or 3 or 4 field nodes. This service group allows the cost (transmitter) to be distributed

into the field connecting to multiple nodes. In this application, the transmitters will still communicate, but their power output needs to be modified. These headend amps will take an input power level and increase it to optimize the dB receive signal in the field.

Every node still needs to transmit the upstream data and information to the headend. The new 2-way communication require new devices to be placed within the Headend. The receiver allows for the receipt of the signal, and new decombining equipment allows for the separation of the return signal. The receive signals are broken down for multiple applications: Conditional Access for video authentication; signal maintenance by the outside plant teams and communication to the cable modem termination station (CMTS). So, when we look at the new spatial needs for these analog touching devices, the Headend space is rapidly being consumed at only the third generation of Headend equipment -- generation 1 equipment being simple receiver and encoders; generation 2 equipment being 300 MHz systems operating with conditional access equipment that controlled the set top boxes, which ensured that only the video programming purchased was made available for consumption. Generation 3 equipment is true bi-directional communication between the Headend equipment and the customer premise equipment (CPE).

If you were keeping count so far, we have:

Table 1: Video Analog Equipment

Receivers	Encoders	Transmitters	Receivers
Signal testing equip	Controllers	Amplifiers	Signal Combining
Signal Decombining	Test ports	EAS equipment	

All the above equipment works as a homogeneous group of standalone equipment needed to send 1 defined signal down the coax line to the homes.

We are approaching a situation where power to these devices is critical. Two developments within the electronic space are driving to what is being called clean power and/or conditioned power. The electronics have more and more solid-state components needing a clean and uninterruptible sine wave to bridge the gaps on an alternating cycle. The original equipment handled the under-voltage section of the sine wave by embedded capacitors that would hold the electrical current charge to the device for the other ½ cycle. The adoption of the Alternating Current Uninterruptable Power System (UPS) would allow for not only the conditioned, clean power but also the opportunity for backup power. The UPS systems are currently a simple dual power conversion unit that takes AC power in. Rectifies the power and allows some of the energy to be stored within batteries. The System then takes the direct current energy and feeds it into an inverter system. The output of the converter system them passes out of the unit and feeds into the Headend electrical outlets to power the headend signal generating equipment. But everything comes at a cost. The space needs for the UPS would take up rack space or floor space. For a typical 22kva UPS, you could or would lose about 2 racks of space within the equipment 4 walls of the Headend.

With the wide usage of UPS systems taking hold within the industry, there is another push to protect the UPS system. This came with the adoption of on-site power generators and Automatic Transfer Switches (ATS). The ATS serves as a large power selector switch. During 99% of the operation of the building, the Utility power grid provides the power to the building. For the 1% of the time when the utility power company cannot provide power to the building, a signal is sent to the generator to start. The ATS then sees a reliable voltage on source 2 emergency input to the switch. The ATS will then select the generator

voltage over the lack or no voltage coming from the power utility. This transfer from utility failure to Generator start and transfer is usually less than 12 seconds.

The physical nature of the ATS is that it is a basic electrical component. It is placed within a UL listed enclosure based on its size and configuration. Typical for this time within the history of cable, the units are wall mounted and typically less than 12 inches in depth. Since these unit are not on the floor, we have some level of relief from losing rack space. However, since the access to the unit is front hinged door, we have the need to leave what is called a safe working border to the front of the unit. This will typically require 36” of free access space from a grounded location. The ATS can be placed outside the building on a concrete slab or a wall mount location at the desire of the engineering agent. This spatial offload allows for some space relief within the building but comes at a cost of reliability and capabilities. We want to always strive to keep this equipment within the conditioned space of a building.

4. Can I phone a Friend

Bidirectional communication between a customer and the communication facility begins to usher in the age of coaxial cable television voice or Voice over IP (VoIP). This meant that, unknown to the customer at the time, their experience at home was tied directly to the successful operation of the building supporting the services. The functionality of the legacy CO joined the video content delivery subscription services of the time: The voice over internet opportunity needed space in the Headend.

Voice over Internet Protocol was the third leg of the triple play stool pitched by many cable operators. The equipment in the headend was identical to the voice soft switches that began to replace the legacy Mainframe Class 5 switches of the CO. The advancements of the CPE equipment to be able to inject a common telephony phone input and ring service output connected the equipment platforms in the homes. The other end of the wire at the Headend removed the voice modulation signals from the RF carriers and sent the signals out to the world telephony network.

This advancement was not without difficulty. This new internet-based voice system was built commonly as a modernization effort for the big phone companies, that have been operating on a 48v platform of power for the network/telephony equipment for scores of years. This legacy power platform is rigid, hardened by years of OEM investments and backup systems that would cause a massive shift in facility architecture and mindsets of the Cable TV Facility Engineers. The video side of content was content with the platform being supported by only utility voltage (120v). With the onboarding of a UPS, the equipment was still predominantly AC powered. So emerged equipment co-habitation issues, with specifically different powering needs. Would the equipment be AC powered or DC powered?

Figure 2: An analog-only two-way Headend. Image by Author. Circa 12/2018

5. Rock, Paper, Scissors.

In that timeframe, the battle for power supremacy was fought on a square foot by square foot basis. The telephony camp was locked-in on the operation of the power input to the power supplies to be 48v nominal. This allowed for the most typical and standard power supplies available to be procured with the telephony devices. The uptime argument between the cable TV guys and the Telephony Central Office guys entered the regulated space in the Code of Federal Regulations (CFR)47. If the companies entering the telephony service provider space wanted to compete and market a reliable, capable, and cheaper service than the legacy land line carriers, they had to prove a reliable and available network. The answer: Cable companies would use the same equipment as the telephone carriers. DC powered equipment backed up by batteries for 4 hours and generators for 24 hours. The OEMs continued to provide DC power supplies to make the power delivery reliable and enable the headend to run under multiple failure conditions -- first, a failure of utility power; second, a failure of the utility generator. Operational survivability on batteries via direct current power to the network devices.

With the supportive adoption of the DC power input to critical network routers, optical transport nodes and by default the CMTSs that communicate with the cable modems in the home. With the supportive adoption of the DC power input to critical network equipment, the stage is set to run and operate (2) independent power delivery systems to the multiple devices within a headend space. The adoption of telecom grade 2 volt batteries would also drive a rapid consumption of network-required space allocation.

At the worst, I have seen where the power and batteries needed to deliver back up energy and distribution to the network consumed as much space as the network equipment. In completing over 8 years of planning for the space and power within the headend, indications are that spatial consumption distribution from a per sq foot perspective leaves only about 40% of the total under roof space available to the network engineers to populate revenue generating equipment. Chewing through more footprint for multiple power providing equipment would only continue to diminish the spatial ratio.

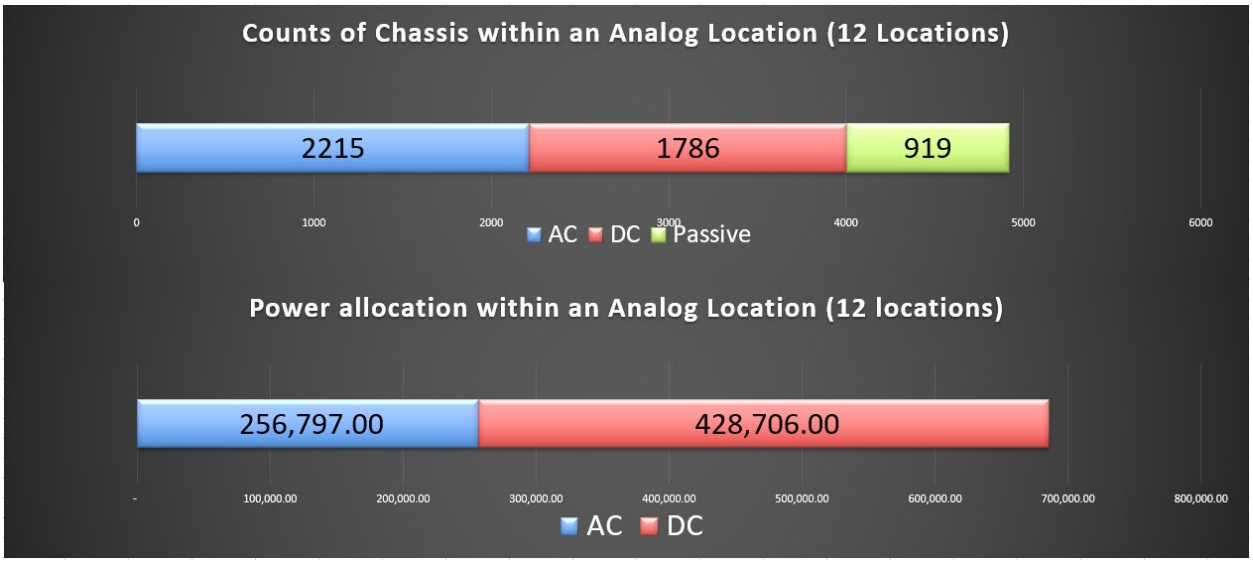


Figure 3: AC vs DC energy consumption relative to RUs consumed

6. "I Want It Now"

With the enhancements of the cable spectrum and capabilities, the industry was able to launch video on demand as a service. This service required the adoption of something within the headend not typically used… Servers. Servers allowed for the VOD content to be stored as close to the customer location as we could get based on needs and caching library. The server farms were typically located within the datacenters of the blossoming Internet world. The easiest ways to get a lot of content to the masses of people was to centralize the content and release the content when requested by the user. This would require massive transport builds from locations spread out across the nation. Over time and the adoption of Video on Demand (VOD), the scalability of the transport networks showed a level of weakness. This weakness was only resolved by placing local content caching as close to the packet build of the time. This location was called the headend. Eventually, the thoughts of putting servers everywhere were scaled back to placing it in a multi transport hop location so that multiple local edge sites can grab the content from 1 hop or two. Terabits of storage at this point was being located as close to the customer as possible to support VOD and the fledging over the stop streaming services content.

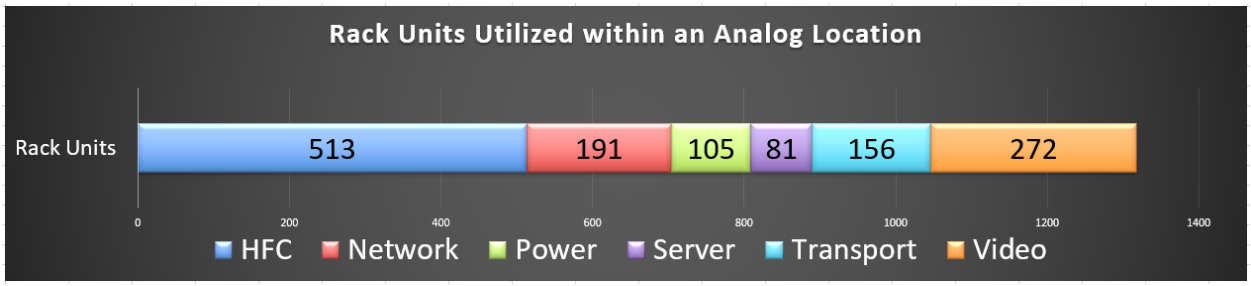


Figure 4: Rack unit allocation within 1 headend

The headend was almost at the peak of physical loading the in the early 2010s, but there was still a growing customer base needing fiber optic connections and speed tier offerings. Small, Medium and Enterprise businesses were growing rapidly. This rapid opportunity for revenue was based on the premise of providing a direct fiber connection to a place of business. This Metro-ethernet connection afforded customers the opportunity to choose a path of speed they could choose to fund. These full fiber connections required a multitude of new equipment to also be located within the Headend. These were also the most robust, largest, power hungry and feature rich devices companies could engineer and code. As more and more of these customers loaded up on the boxes, services and support needs also grew at the physical layer to supplement what was provided at the customer locations accompanying virtual command and control. These enterprise scale routers, with the drive for redundancy, stretched the footprint consumption within the smaller Headend to the point of costly expansions, utility upgrades hampered by abandoned technology not ready to go to the grave.

7. If The Shoe does not Fit…Stretch the Shoe

For those keeping track, we are now at the below list of items and services within the small cable TV building at the top of the hill:

Table 2: Active Equipment Groups in an Analog Headend

Receivers	Encoders	Transmitters	Servers
Signal testing equip	Controllers	Amplifiers	Signal Combining
Signal Decombining	Test ports/gear	EAS equipment	UPS
Routers	DC Power Plants	Batteries	Battery Distribution
Commercial Edge Gateways	MetroE Switch	OLT	VOD QAM chassis
Linear Video Edge Qams	Fiber Muxes	Optical Panels	Pilot Carrier equipment

The buildings were at a max capacity, providing most of all traffic via analog linear video spectrum, with a minor amount (one to two 6 MHz channels) of HSD digital spectrum.

The advent and adoption of the virtual cable modem termination station (vCMTS) and the potential to generate RF quadrature amplitude modulation (QAMs) at the node has opened a new world. This ability of sending 0s and 1s to the node and for the node to change digital into RF signals to extend the useable life of the existing RF plant will drive even more equipment into analog buildings as they move to be Digital Broadband Nexus structures.

The vCMTS is made up of many different components likened to the "medusa" boxes of the legacy VOD-connected CMTS. Eventually, the medusa boxes collapsed into a converged cable aggregation platform (CCAP). There are two key items to the new vCMTS. Routing and video…. Sounds like the entire buildings of the 1990s. The difference is that this new vCMTS platform is function built for the digital world. For the bargain price of servers, layer 1 switches, and a robust routing table and chip set, the first part of the vCMTS will allow the sites to send broadband digital traffic to and from the cable modems. The second part of the platform is the video generation units. This is a must carry due to the legacy needs of the existing customer base, the customers who are on a legacy RF only tuning set top boxes that are still on network. The adoption of an IP only set top was not universally given nor forced to the customers. The business decision to continue allowing analog devices in the home will maintain a legacy of analog equipment base operating withing the headend that would inevitably need to move to the RF strand or leave the home.

8. Now Boarding At Central Station

We have the base platform needed to accommodate the digital goodness in the HE and the nodes in the field with the capability of the remote physical layer conversion (RPHY). Where do we place all this equipment? This raises two additional questions that are repeatedly is addressed by CEOs and CTOs across the industry. Do we place all this equipment in one location, like a data center, or do we disperse it into the network like a neural feed to all the buildings, and keep the equipment at the edge?

Using the Centralization theory:

- Less infrastructure to maintain at a "five 9s or better" status
- Fewer locations to staff
- Fewer sites needing "big iron"
- Already probably exists with a heavy loading of fiber.

Using the Distributed theory:

- Smaller "blast radius"
- Shorter latency
- Cheaper Transport costs
- Already inhabited by legacy analog CMTSs

The platform architecture is the origination of the new digital packet train that transfers data from site A to Site K. In the physical layer, we have a laser or transmitter that will send a packet of information down the fiber pipe. This pipe has only 1 entrance and 1 exit. In the analog world, you could take a signal from the Master Headend and channelize the signal onto an analog transmitter. This allowed you to send this signal to multiple hubs. At these Hubs, more RF signals could be injected into the coax stream and allow for combining locally as needed. In the binary world of digital signals, packets need to be assembled at

the origin and be understood at the destination. The packet build can only occur at a vCMTS location. This will inevitably require a centralization effort of all legacy dispersed servers to cohabitate with the vCMTS. Pulling to a centralized location will eliminate the edge equipment located at the existing locations. This edge equipment we have identified above refers to mostly caching servers and edge VOD qam devices. These two items were needed in the legacy analog world to support the demand for less transport and more edge services. As the VOD QAMs also deprecate with the general adoption of IP content delivery, there is less need for edge linear video equipment. However, the product still needs to get to every customer. Pushing out would refer to the location and placement of the vCMTS into the edge locations. There are multiple benefits and detriments to address.

Pushing the vCMTS to the edge effects the operational efficiency of the vCMTS and the transport layer to support it. In the first assumption of decentralization, deployment and efficiency of equipment modernization will be key cost determining items. The assembly of the vCMTS is completed in a centralized location akin to an assembly line or production center. This assembled platform requires a level of assembly control through the physical build to configuration and onto deployment and original testing on network (OTON). Pre-configurations loaded for the end location will drive very tangible and cost savings opportunities. Speed to deployment, speed to growth, and value-added remote access from a centralized configuration support center. This will all require building to scale effectively with the nodes attached to the physical location. The existing and total node counts need to be reviewed before anyone can make a best utilization decision on vCMTS locations.

A vCMTS is connected to the node via a chip set on the server/ line card within the platform. A Hub may have 400 nodes. A HE may have 90 nodes. A master Headend may have 1000 nodes. We need to understand the math of the connections to validate the options for locational placement of the vCMTS. If your vCMTS is capable of 50 nodes to 200 nodes, the ideal location may be in the edge location. If the vCMTS is cable of supporting 200 to 500 nodes per cabinet, it would be better to centralize so the disbursement of compute would be efficient to the subtending Broadband Digital Structure (BDS)s.

9. The Slinky Effect

Do you remember as a child playing with a Slinky? Standing at the top of the stairs, starting a slinky, which by the forces of gravity and nature crawls down the stairs, much to our amazement. The other thing we did as young engineers, of course, was pull the slinky apart. Hopefully, you never pulled the slinky too far apart and then it would never compress back to original again. The effect of stretching a slinky too far is deformation. The spring never returns to an original dimension, and the original dimensions are never recovered. Stretching the cable network will have good and bad effects on the facilities. The optical limits of fiber and the effective usefulness of transmitters and receivers allow us to stretch digitally for 80 KM from the switch originating the packets to the node. This allows the facilities to range in homes served from a modest amount to a large amount based on the location and serving density.

Space in a legacy analog site has (4) classifications: AC powered network equipment, DC powered network equipment, passive equipment, and non-functional units. In every facility, operators should understand the requirements and dependencies of these assets. AC powered equipment is anything that is powered from a UPS, inverter, or standard utility power. The power distribution to these locations typically is a power strip in the back of the cabinet or rack. No RUs within the cabinets are consumed by these power strips. DC powered equipment is from a fuse alarm panel (FAP)s/ Distribution Circuit Breaker Panel (DCBP) or direct feed from a power system. These FAPs and DCBPs are typically located in the top of equipment cabinets, taking space away from the active equipment. Passive equipment is anything that does not consume energy but has a purpose within the building. These are typically fiber

pre-terms and RF combining on the upstream or the downstream. Non-functional units are a great level of opportunity when we talk about the transition from analog to digital.

The power capabilities at a facility are among of the easiest to understand, trend, project and manage. The service entrance ampacity is the single energy gate and can measure what I have, what I have capacity to add, what will I have left after I add it. The Electrical service utility entrance is a single energy gate. This allows for an Engineer to measure what they have, what energy capacity they can consume, and project what is left after that decision is made. This measurement is typically in amps or in kilowatts. Simple meters and measuring tools have been exhaustively added to the buildings, allowing for the consistent trending and tracking of this metric. We would also look at the capacity of energy conversion at the DC plant level and UPS level. The final power draw component is the HVAC consumption measurements.

Cooling energy is a necessary evil, as it is called: Energy must be spent, in order to cool energy. Since most energy processes are consumable in nature, we traditionally see a 1 to 1.6 relationship on energy consumed, to energy needed to cool. These energy metrics and options generate our "Facility Spring". If we work within the facility spring, we can survive the transition to digital without a major influx of capital (the "deformation," in Slinky terms.)

10. I Think I can Fit One More Thing

The start of a journey begins with a plan. Our planning takes place with the kickoff called Blueprint-531. The B-531 is a planning assisting workflow that allows space planning managers and engineers to understand the destination. Once you decide where to go, it is easy to map your journey. The B-531 looks at the finite data available to be understood at the facility level. The physics around the transmitter/SFP distance limitations, lambda restrictions in a piece of glass and the availability to mux and demux optical paths are defined parameters we can engineer solutions from.

The other parameter feeding into the blueprint is the fiber density in and around the physical building. This fiber density looks at all services available to be rendered from a facility in way of commercial customers and SMB connections. Fiber consumption take rates for the embedded customer bases and designs the number of households passed (HHPs) per node split within the digital transformation end state. The equation is:

$$C = D_{nt}\,(@\text{Xhhp density}) + (O_n{*}d) + S^{2ch} + V + U + F_p$$

Equation 1: Cabinets needed = total digital nodes at "X" HHP density + Degrees of Optical transport nodes + number of redundant chassis needed per service + Video + utility + passive fiber terminations.

The above would indicate the total cabinet need of a site with 22,000 homes passed and commercial customers requiring PEG channel origination would equal 1+3+2+4+1+1 = 12 cabinets needing approximately 190 square feet for cabinets and aisle space walkways. With the average size of any BDS being 300 to 500 square feet, building additions based on space end state are very rare. At the time of this writing (summer 2021), after evaluating over 218 locations, only 9 sites would have no path to sustained footprints when a proper execution path is achieved.

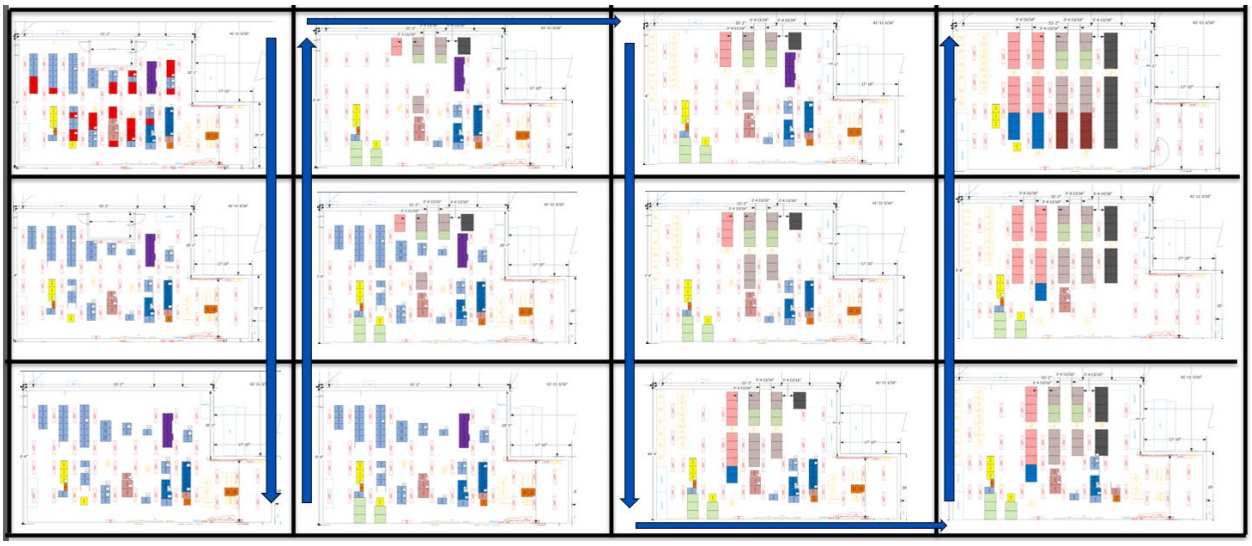


Figure 5: Facility Equipment location migration plan for a vCMTS location

What comes into the building first to facilitate a digital conversion is an effort of trial and error. Due to capital intensity, modernization efforts and customer needs consuming items at the port level, multiple B531 start points are to be "war gamed." The first most obvious path to free space is the conversion of the legacy analog CCAP cable CMTS to a vCMTS platform. As we evaluate the opportunities for space conversion, a decision of transition process needs to be understood. Original equipment Manufacturer (OEM) vendors have the hardware available to utilize a vCTMS Digital to analog shim called Shelf Physical (SPHY) Layer. This shim allows for the installation of multiple RF output ports to tie to existing transmitters and receivers allowing for the removal of all supporting analog hardware.

The method of deploying the SPHY equipment is a major decision point for energy and customer impacts. If the decision is made for equipment velocity, the result is a fragmented mess of analog abandoned equipment running in a very inefficient space and powering plan. Legacy best methods of collocating combining with TX/RX may leave the building with no recovered space until a defragmentation plan is enacted. This will impact the customer experience twice: One time to cut over to vCMTS, and the second time on the transmitter centralization plan to free up needed space. In the world of hardware modernization, velocity usually trumps efficiency.

Table 3: Dual Platform BDS - Analog and Digital

Receivers	Encoders	Transmitters	Servers
Signal testing equip	Controllers	Amplifiers	Signal Combining
Signal Decombining	Test ports/gear	EAS equipment	UPS
Routers	DC Power Plants	Batteries	Battery Distribution
Commercial Edge Gateways	MetroE Switch	OLT	VOD QAM chassis
Linear Video Edge Qams	Fiber Muxes	Optical Panels	Pilot Carrier equipment

Receivers	Encoders	Transmitters	Servers
DAAS Switch	SPHY chassis		

The sister step to the transition from CMTS to vCMTS involves the deployment of Distributed Access Architecture Switches (DAASs). The isolation of the DAAS to a separate cabinet may seem a gratuitous consumption of space, but it layers in a "forever rack". These forever racks allow for the long-term build and spatial reservation for nodes to terminate to mux shelves and switches. Existing density architecture allows for the build of over 554 nodes within 1 cabinet space consuming 15.83 sq feet. When addressing the legacy CMTS spatial consumption of over 285 sq foot via a spatial consumption of 95 for the same digital node, a savings of over 300% is realizable. But that requires taking the first step, which is installing the Digital DAAS cabinet.

The next parameter to understand is time. The time horizon to a location conversion to vCMTS and SPHY will drive the downstream part 2 conversion plans -- commercial chassis modernization and transport updates will inevitably conflict with the time horizon on the digital node conversions in large sites. The actual upgrade of transport chassis and routers to absorb the dedicated pipes running from the DAAS switches to the Broadband Digital Primary (BDP) centralization center will potentially drive more chassis before a hardware refresh can begin. The growth of the commercial side of the business may require separate routing networks that will scale differently than the routers for residential traffic. When the site has been converted to an entire SPHY connected DAAS, the beginning of the digital transition can begin and end. Within the table shown below, the highlighted equipment can at this point be removed from the edge locations.

Table 4: Analog Items to remove post vCMTS connectivity

Receivers	Encoders	Transmitters	Servers
Signal testing equip	Controllers	Amplifiers	Signal Combining
Signal Decombining	Test ports/gear	EAS equipment	UPS
Routers	DC Power Plants	Batteries	Battery Distribution
Commercial Edge Gateways	MetroE Switch	OLT	VOD QAM chassis
Linear Video Edge Qams	Fiber Muxes	Optical Panels	Pilot Carrier equipment
DAAS Switch	SPHY Chassis		

11. Do Not Stop At Go

As we reflect to where we are in this long process of removing the first round of analog equipment, we may have left a path of inefficiency in our wake. Leaving the existing Tx and RX in their existing locations will cause a level of effort that in some sites will be minor. In other sites, it will be immensely

difficult. On the level of easy, I have injected an image of a location that has left the analog world behind but still has the leftover equipment sparsely located with the HE:

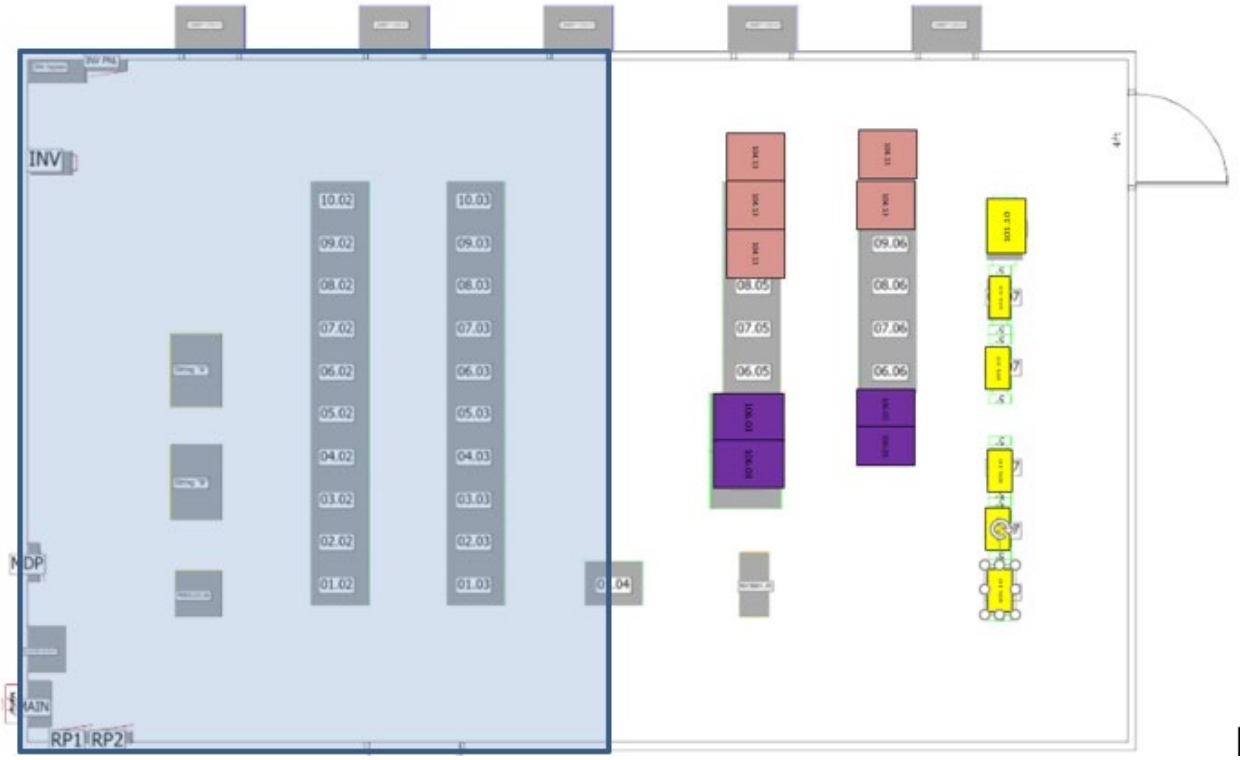


Figure 6: Active Digital Cabinet surrounded by legacy analog equipment

The leftover spare locations are inhibiting not only the ability to make the site as efficient as possible, but also leaves a complex legacy of old transmitters and receivers that no longer have a meaningful return in investment to be upgraded. At this point, a cost benefit analysis for the replacement of legacy RX and TX instead of a forced conversion to digital nodes in the field and connections to only the Digital DAAS. This will cause conflict and discussions and analysis and meetings and more conflict in priorities. As you see in figure 6, the blue highlighted area is capable of 100% vacancy. The SCTE champions the energy conservation of the MSO industry. Leaving so much analog active equipment burning through the HVAC airflow problems will be a counter argument to the reliability plans and limited touch operation that drives our strategic objectives. We as an industry need to understand there is a tradeoff between digital improvement opportunities, energy efficiencies and the customer experiences and that they all should be in balance.

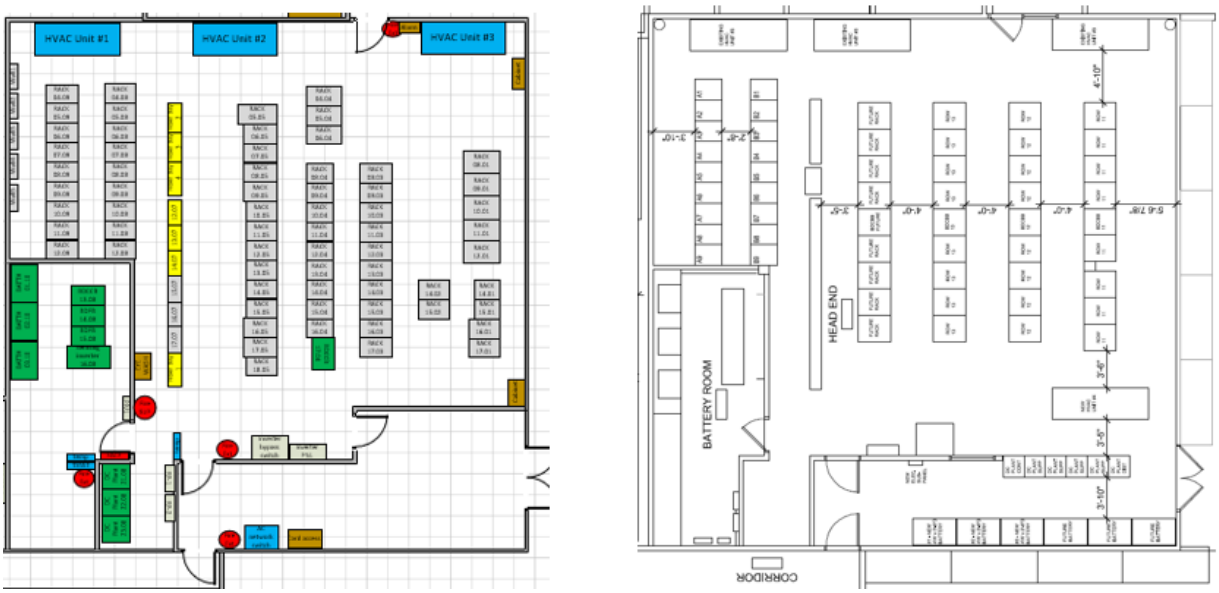


Figure 7: Before and after planning showing a path to all Digital

The final transition to digital will require the outside plant to do a significant amount of work to support the final transition. This may or may not require the replacement of actives downstream of the node by either replacing the amplifiers and or the nodes and split the service groups to a manageable amount of cable modems on each of the remote physical devices (RPDs). This activity and business cases need to take many items into consideration. The acceleration into a permanent standard cabinet and layout will accelerate the potential year over year growth opportunities. Speed to complete future activities can be achieved via this spatial and energy consolidation efforts. As efforts have been realized, the below items need to be understood as to the level of need for the final conversion to digital:

- Energy savings to be achieve by the digital conversion
- Recovery of energy capacity within the former edge sites
- Recovery of space within the former edge sites
- Operational improvements from the conversion to digital
- Additional bandwidth available due to the conversion to digital
- Elimination of OEM Analog equipment due to the conversion to digital
- Better OSP efficiency from operating digital nodes in the field
- Eliminating infrastructure equipment within the headends and hubs due to the conversion to digital equipment
- Right sizing of space needs within leased locations due to the conversion to digital operations
- Optional removals of underperforming sites within the networks due to the optic reach from digital small form factor pluggable (SFP)s to the nodes in the field
- Low-cost expansion into rural markets due to the range and options digital provides.

Table 5: BDS items left after analog leaves

Receivers	Encoders	Transmitters	Servers
Signal testing equip	Controllers	Amplifiers	Signal Combining
Signal Decombining	Test ports/gear	EAS equipment	UPS
Routers	DC Power Plants	Batteries	Battery Distribution
Commercial Edge Gateways	MetroE Switch	OLT	VOD QAM chassis
Linear Video Edge Qams	Fiber Muxes	Optical Panels	Pilot Carrier equipment
DAAS Switch	SPHY Chassis		

Within the above list of tangible improvements realized from the full conversion to digital, the 1 item that provides a meaningful tangible capability is the recovery of usable energy capacity at the edge locations. If we take a review of the imbedded energy opportunities within a former edge site, we can calculate the available energy and available space at a today cost to build the same surplus. In a presentation produced back in 2017, an MSO had the MW of embedded energy capacity at the edge locations of 73mw, of which only 44mw was being consumed. Since then, the overall capacity of energy has grown to approximately 87mw and a consumed amount of approx. 55mw. With the known advantages of a digital conversion to eliminate 80% of energy consumption at the edge locations, the asset value of unused and provisioned energy capacity can be measured in megawatts of network available energy for growth and new edge technology. When our industry looks back of the growth opportunity at the edge locations and the almost-free cost of connectivity to energy to serve customers, the low cost of expansion and opportunity will be plentiful, if digital realities are strategically aligned to the opportunities within the facilities we operate.

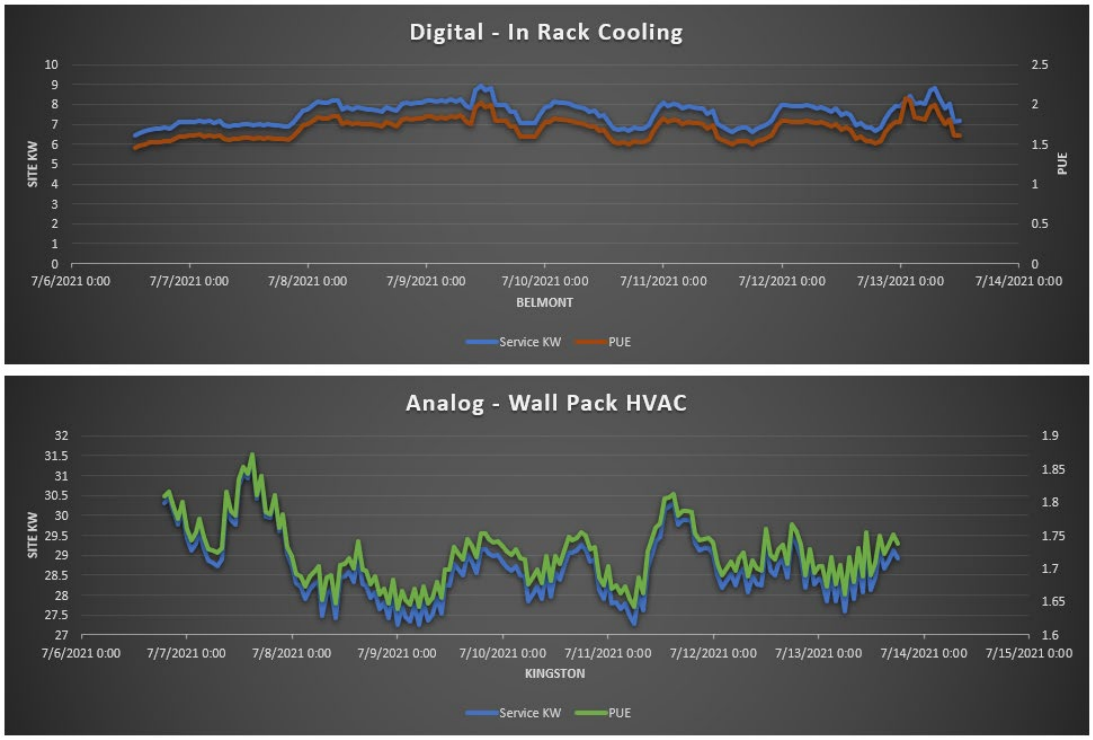


Figure 8: Digital vs Analog

12. Conclusion

The name "Cable TV" is becoming synonymous with antique, legacy, old and clunky. The names Headend, Master Facilities and Signal Buildings are becoming just as legacy. We are in the Information Age as a sea change event since 2003. The brick and mortar headend of the 60s, 70s, 80s and 90s have a long roadway available to them. Buildings we have built since 1963 to enable the analog age are still able to grow the Information Age and the upcoming Age of Digital Expansion. Innovations of the future will inevitably need to be connected to a physical layer; the cloud will always need the ground. This paper reviewed the opportunities to plan the layout and location of items within a building and how to phase plan for all activity. By following a pre-determined playbook of tech refresh, capacity augments and right sizing infrastructure, you also will achieve a building able to sustainably grow with the network that demands its support. The Age of Digital expansion will connect the future world within the bricks laid in the past.

Abbreviations

AC	Alternating Current
AC UPS	Alternate Current Uninterruptible Power Supply
ATS	Automatic Transfer Switch
BDNS	Broadband Digital Nexus
BDS	Broadband Digital Structure

BDP	Broadband Digital Primary
CA	Conditional Access
CATV	Community Antenna Television
CCTA	Community Cable Television Antenna
CFR	Code of Federal Regulations
CMTS	Cable Modem Termination Station
CO	Central Office
CPE	Customer Premise Equipment
DAAS	Distributed Access Architecture Switch
DC	Direct Current
DCBPs	Distribution Circuit Breaker Panels
FAPs	Fuse Alarm Panels
HE	Head end
OEM	Original Equipment Manufacturer
OTON	Original Testing on Network
QAMs	Quadratic Amplitude Modulation
RF	Radio Frequency
RPD	Remote Physical Device
RPHY	Remote Physical Layer
RUs	Rack Units
SFP	Small Form Factor Pluggable
SG	Service Group
SPHY	Shelf Physical
SSD	Solid State Drives
UL	Underwriters Laboratory
vCTMS	Virtual Cable Modem Termination Station

VOD	Video on Demand
VoIP	Voice over Internet Protocol

How to Optimize TCO and QoE in a Cloud Environment Using a Context Adaptive Delivery Solution

A Technical Paper prepared for SCTE by

Patrick Gendron
Director, Innovation
Harmonic
2590 Orchard Parkway
San Jose, CA 95131
U.S
+1.800.828.5521
Patrick.Gendron@harmonicinc.com

Thierry Fautier,
Vice President of Video Strategy, Harmonic

1. Introduction

OTT delivery is increasingly becoming a primary solution for the consumption of live video content. With OTT, the QoE provided to users should be at the same level as traditional broadcast TV.

OTT has become so mainstream that even live content is now available from video streaming providers. In the U.S., between Sling, DirecTV Now, Hulu, YouTube and Sony Vue, there were more than 9 million OTT subscribers at the end of 2018, according to a Fierce Video report. Yet, quality is sometimes an issue, and that's a problem because consumers expect the same video QoE for OTT as they've experienced with broadcast TV.

While the experience is expected to be the same or better, there are many technical differences between OTT and broadcast delivery. OTT targets a variety of devices (i.e., smartphones, tablets, desktop computers, connected TVs, game consoles) and delivers the content over a variety of networks (i.e., xDSL, fiber, radio 4G and now 5G).

To address these issues, many efforts have been made to define technical solutions for OTT streaming. One example is lowering the latency for the most popular streaming protocols HLS[1] and DASH[2]. This was presented in the 2019 SMPTE conference paper "How OTT Services Can Match the Quality of Broadcast"[3], but there are still some problems to be tackled when it comes to achieving massive at scale viewing of live events.

This paper will examine the different solutions that can be deployed in an OTT environment, comparing the technical merit, the integration aspects in an open ecosystem, the need for standardization and the overall impact on total cost of ownership (TCO) and QoE. It will provide suggestions on how scalability can be achieved to deliver high-quality live video to millions of subscribers on every device at any time, even during peak hours.

2. Current State of the Art for the Video Streaming Industry

There have been a lot of new streaming protocols and formats popping up over the past decade, but when you observe the current OTT delivery landscape for video on demand (VOD) and live content, it's clear that a vast majority of content is distributed using either HLS or DASH formats. Both use HTTP[4] as the underlying transport protocol and are based on adaptive bitrate (ABR) technology, which makes it possible to deliver video over unmanaged networks with variable available access bandwidth. The other formats have either reached obsolescence (e.g., Microsoft Smooth Streaming) or should be reserved for more specific use cases (e.g., ultra-low latency) as they come with some additional constraints. For example, WebRTC[5] can provide low latency but relies on peer-to-peer and thus has some serious scalability issues.

Since the focus of this paper is on large-scale content delivery, it is not a debate that the streaming industry will rely on two dominant streaming technologies for the next few years: HLS and DASH both using CMAF[6] as a common format for the delivery segments.

While there's been a growing demand for live content over the past few years, the vast majority of OTT consumption has always been and is still VOD content. This brings some additional constraints to the delivery workflow.

Several content delivery optimizations have already been deployed but most, though not all, are dedicated to VOD asset distribution or suffer huge limitations:

- Several years ago, Netflix introduced its per-title encoding [7], then per-scene encoding to provide a better video quality at a given bitrate. It also included a new paradigm of adaptive ladder, removing the profile when it doesn't bring any value to the end user from a video quality standpoint. These innovations improve the viewing experience, but they have been implemented for VOD assets with no real-time constraints on the processing.
- During the BEITC 2019 conference, Brightcove presented [7] some techniques to dynamically adapt the profile ladder based on the network conditions. These are a promising path but not yet available for live content distribution.
- Storage of recorded assets (e.g., cDVR applications) can be optimized by an offline profile curation removing the nonessential rendition in the profile ladder. The criteria to remove the nonessential rendition is based on a perceived QoE by the end user. This approach can work to optimize the storage volume and therefore the costs but again this is not yet something that can operate in real time for live content.
- Current multi-CDN strategies are based on static cache allocation. For large-scale events like the Olympics or FIFA World Cup, a major CDN would need to book physical resources up to 12 months in advance.
- Finally, as required for any closed-loop optimization client/CDN analytics may be collected in real time (which sometimes occurs at the end of the viewing session) but are generally not processed in real time to build actionable insights to optimize QoE.

When a popular event must be delivered live over a variety of networks to a variety of devices, it is a lot of work to make sure everything goes smoothly. For such an event or for regular peak audience, most popular services experience some QoE issues like rebuffering and long start time, when it is not a total impossibility to connect to the service.

The next section will discuss how the scalability issues and generally how a service operator can improve the delivered QoE by moving from a fixed, non-optimal workflow to a much more flexible workflow that is adaptative to the external context.

3. Motivations to Move From Static Workflows to Dynamic Delivery of Content

Depending on the business model for the service provider, the most critical metrics it needs to monitor and improve are:

- The acquisition of new subscribers, at a reasonable cost
- The churn ratio (or the ratio of new subscribers) for a subscription-based model or the total viewing time for an ad-subsidized model
- The cost of service (that can be approximated by the TCO value).

The first bullet will not be addressed in this paper, as the acquisition of new subscribers is mostly linked to the content proposed and service feature package, more than to the quality of the delivery.

The metrics reflected in the second bullet are the result of multiple factors, technical and nontechnical (typically based on the content offering and service features). However, for the technical side, which is what the platform can offer, the metric that is prominent is the QoE perceived by the end user. The QoE metric itself is a combination of multiple factors, including the video startup time (VST), video start failures (VSF), rebuffering ratio (CIRR), end-to-end latency and the perceived video quality. This metric can be estimated by collecting telemetry directly on the user devices or by deducing from other telemetry collected on the network.

The TCO includes any costs needed to run the service. This can encompass hardware costs and the cost of operation in the case of an appliance-based service or the cost of service invoiced when operated in SaaS mode. In any case, this total cost covers the headend and delivery (e.g., CDNs).

Any evolution of the current workflows should therefore be considered keeping these two goals in mind: improving the QoE and reducing or keeping the TCO under control.

Looking at the whole delivery workflow from the content capture to the end device, there are several areas where the processing needs to be flexible to get the optimum use from resources (i.e., computation resources, bandwidth, storage, etc.). In the video compression space, it has been well known for decades that the codec engine should enable the most appropriate mode depending on the content nature. This is a supported capability in any video codec since fixed QP approaches have been enriched by many other techniques driven by the content nature. Video codecs have dynamically adapted to the content characteristics for years, providing compression improvements. But things are now getting more complex with ABR distribution. Content is now made available in several bitrates, several resolutions and several frame rates. Even if one can anticipate general rules linking these parameters (for example, when the bitrate is reduced, at some point, it becomes more efficient to reduce the resolution), these thresholds are highly dependent on the nature of the content. The same applies for the frame rate. It is commonly agreed on that for sports content a high frame rate will provide a better result at a given bitrate, which is not the case for other types of content.

The delivery side, compared with the "good old simple broadcast era," is also much more complex. Delivery is made via multiple types of networks, all of which have limited resources that should nevertheless cope with the unicast paradigm used in OTT. Moreover, the QoS of the delivery network is highly variable over time and locations (on the open internet QoS is not guaranteed as it is on a cable network, for example). Dealing with these variable parameters on the network side can be achieved by overprovisioning resources. For example, putting more edge caching in the CDN. But this approach carries a significant cost that might impact the profitability of the service. If considering the worst case in a static configuration is not a viable option, then the alternative is to adapt the delivery to the current condition in order to find the sweet spot that will, at any time, give the best compromise and maximize the end-user satisfaction. All this is even more complex when consumption is on mobile networks, meaning there are even less predictable network conditions.

There is a clear motivation for the service provider to maximize the perceived QoE and keep TCO under control in a fast-moving environment. This multi-variable equation cannot be solved efficiently in a rigid, static delivery workflow. Moving from a static to a more dynamic approach can greatly impact the entire delivery workflow.

4. Industry Trends and Research

On the content preparation side, compression technologies have evolved and become more complex. In the past few years a new paradigm called content-aware encoding (CAE) has emerged. CAE embraces different technologies that are highly dependent on the encoder vendor but overall the codec decisions are more driven by an on-the-flow content analysis. More recently, artificial intelligence (AI) and machine learning (ML) technologies were added to cloud-based solutions, making these tools economically viable. Thanks to a big push by industry leaders like Netflix, AI entered the game to propose per-title and then per-scene encoding. Here the video content is not only analyzed in real time to extract the relevant feature but, then, a prediction model is created offline using a large database of content to select the best codec configuration. More details can be found on this topic in the SMPTE 2019 conference paper [9]. All these techniques, from CAE to more advanced AI-based processing, were first used in production for VOD

assets but now are starting to be deployed for live content, with the even greater challenge of matching real-time operation. This is the next step the industry should take, introducing some new dimensions and new flexibilities to create an optimal profile ladder at any time. These new areas will be described in the following sections.

On the delivery side, the situation is a bit less mature, as this part of the global workflow has been a moving target in the past decade. Nevertheless, the global trend on the delivery network side is to transition to a more flexible software-based architecture. On the one side deploying new network elements (software based) according to demand is an option. On the other side, providing some hints to the end-user player so that it can make smarter requests to the network is also a possibility. There has been a significant amount of academic research on building some models of the various delivery networks. Much of the research is focused on the mobile network where data consumption (mainly driven by video content) explodes and will continue to grow in the coming years while QoS is still an issue, especially in crowded areas and at peak hours. Stanford University has done work on network optimization using deep reinforcement learning [10]. On the client side, MIT has developed research around an improved ABR algorithm using reinforcement learning to improve the player behavior in difficult network conditions [11].

As explained in Cassie Tolhurst's blog [12], deep learning algorithms can help secure highly demanding content like UHD delivered at a large scale. We are seeing AI spread more and more across the workflow from content preparation to network delivery to enhance the end-user experience. The different functions, part of these new dynamic workflows, are presented in the following sections.

5. Moving to More Dynamic Workflows

5.1. Introduction

Dynamicity of the workflows must be done in relation to external context, as explained above. Taking a holistic view of the global situation for live video streaming, we came to a conclusion about where it's important to build a new way to distribute video. This analysis leads, therefore, to the creation of adaptive workflows taking into account all possible contextual sets of information:

- **Content characteristics:** Live video is per nature changing over time. Encoding and packaging it with fixed configuration (i.e., bitrate, resolution, frame rate) as done today is not optimal to ensure the best QoE the delivery network can give at any time.
- **Content consumption:** Having the same encoding, same packaging (same profile ladder) for all live channels is sub-optimal. This should be dynamically adjusted over time using feedback from the network (i.e., player, CDN, access network).
- **Content importance:** Premium content with high value attached to it will have to be encoded with a higher quality than less valuable content. This should be dynamically adjusted over time according to the operator's preference.

Moving from the traditional static (set and forget) approach to a workflow where many configurations can change over the time, sometimes with a high dynamicity, is not a simple task. This covers many aspects summarized in Table 1.

Table 1 - OTT evolution from static to dynamic workflows

Static workflow	Dynamic workflow	QoE improvement	Cost improvement
All parameters are fixed	Variable parameters	✓	✓
Fixed resource allocation	Variable resource allocation	✓	✓
Fixed architecture/ maximum TCO	Usage-based architecture / Optimized TCO		✓
Siloed approach	End-to-end approach	✓	✓
Deterministic approach	AI-based approach	✓	✓

As illustrated in Table 1, the different areas moving from a static paradigm to a dynamic workflow may impact either the end-user QoE, the service provider costs or both.

The move to a dynamic approach implies that the decisions made should be driven by various criteria linked to the contextual sets of information mentioned above (i.e., content characteristics, content consumption and content importance).

The various dynamic actions across the delivery chain can be split into two categories:

- Actions and tools aimed at optimizing the production on the profile ladder, either for the purpose of improving the QoE or reducing the operation costs
- Actions or tools aimed at optimizing the delivery path to improve the QoE (mitigation of network congestion during peak audience)

The next sections give a description of the various tools, most being guided by AI, which are part of the Context Adaptive Delivery solution embracing the two categories of actions.

5.2. AI-based Encoding/Content Aware Encoding

One of the ways service providers are battling QoE issues for OTT is through advanced compression methods. Content Aware Encoding (CAE), a per-title encoding technique currently used by Netflix, is one such method that supports both VOD and live applications.

CAE assesses the video complexity in real time and adjusts the encoding parameters to provide the best picture quality. It works similarly to VBR for statistical multiplexing, except that only one program is encoded, and the video quality measurement is more refined since it is based on the Human Visual System (HVS) model. In order to have a more accurate video quality measurement, the CAE live system is trained offline using artificial intelligence technologies. For more details, see Harmonic's technical guide on EyeQ [13].

Over the past few years, CAE has made a real change in video compression and is now backed by Apple, Netflix, and the Ultra HD Forum, which has demonstrated a consistent savings of 40% vs. CBR for UHD ABR using CAE in 2018.

The next step of video compression improvement using AI is what Harmonic calls "Dynamic Encoding Style." We leveraged our first research in AI to embed prediction models in the encoding engines to feed the compression engines with the most appropriate set of parameters. As with many ML-based solutions, we train a prediction model offline using a large database of assets in order to find the best compromise among the huge set of encoding parameters that can be fine-tuned. Then, on the live system, the prediction model uses the video characteristics extracted in real time by the first blocks in the encoding

pipeline and matches it to the “optimum” set of encoding parameters. Dynamic Encoding Style is a natural complement to the CAE approach and shares a lot of common principles, including the video quality assessment technique used to build the prediction models in the Human Visual System Model.

Dynamic Encoding Style uses an AI based two-step approach, depicted in Figure 1.

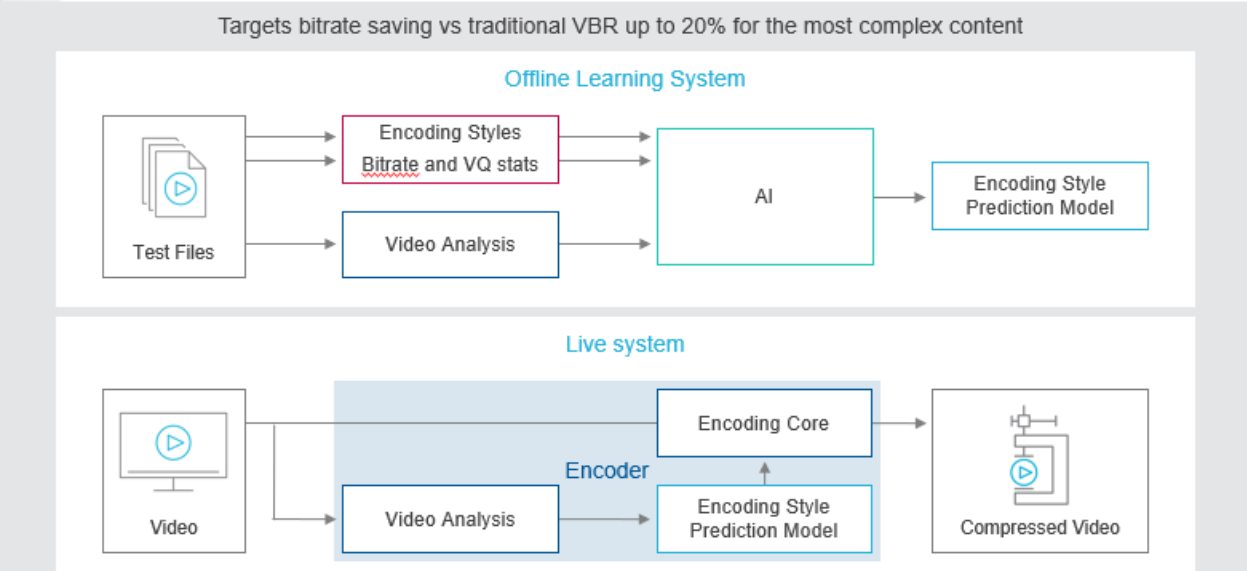


Figure 1 – AI-based Encoding

5.3. Elastic Encoding

Elastic encoding is the last tool mentioned in this paper that focuses on the compression core for live content distribution. The general approach is to use feedback from the network on the content popularity in order to allocate variable CPU resources for the transcoding. The new generation codecs have a very large toolbox that could, if they are all used for every content, have a huge impact on the solution density (number of transcoding instances that can run in parallel on a given cloud resource). All the live encoder vendors are therefore making compromises to find the sweet spot between quality, bitrate and CPU resources. This compromise can be different when the distributed content is very popular. Allocating more CPU cycles will lead to lowering the bitrate at a given video quality which, in turn, will reduce the CDN costs. This is very interesting when the CDN egress is high for popular content.

As it is sometimes difficult to predict which event will be very popular, having a flexible solution that can adapt dynamically when the live event is being distributed is very important.

5.4. Dynamic Resolution Encoding

This tool (and the next one) is different from the previous ones, as the AI is not used to modify the configuration of the compression algorithm but is used to select the optimal resolution of the encoded video. It is well known that there is a link between the representation bitrates and resolutions in an OTT profile ladder. For low bitrate representations, it is more efficient to reduce the content resolution before encoding to get the best QoE. As the threshold to change the resolution at a given bitrate is highly dependent on the content nature and evolves dynamically over time, a solution based on a ML prediction model is a good choice to estimate the best resolution before the encoding processes.

Like the previously mentioned tools, Dynamic Resolution Encoding uses an AI-based two-step approach, as depicted in Figure 2.

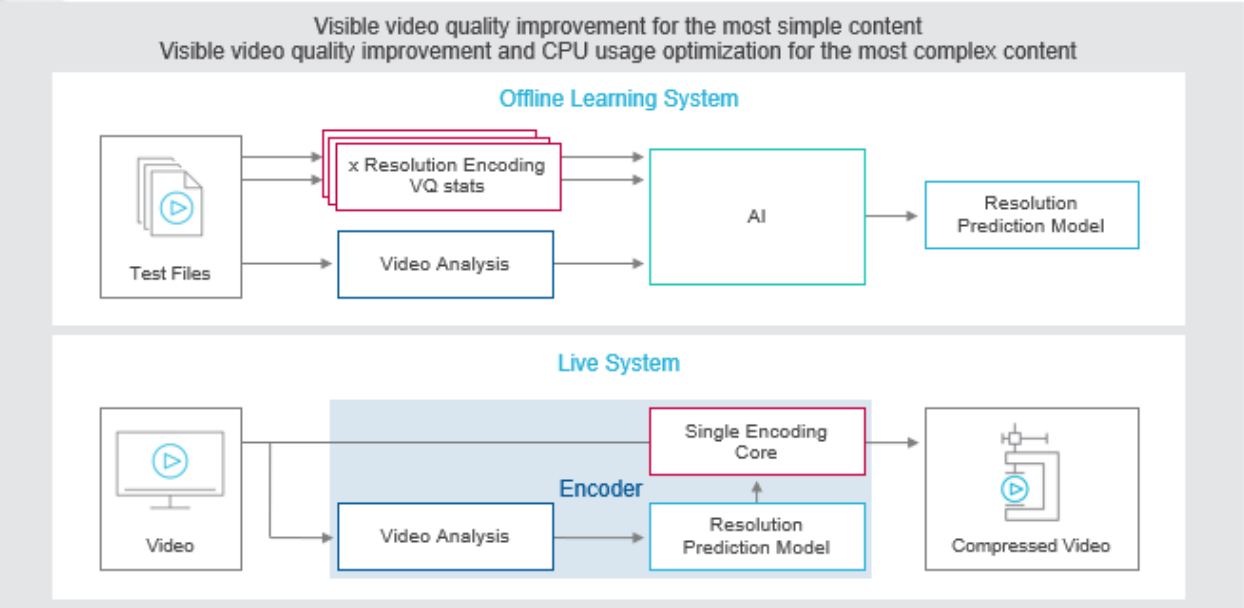


Figure 2 - Dynamic Resolution Encoding

Dynamic Resolution Encoding allows operators to improve the QoE by enhancing the picture quality perceived by the user. It also improves the density of the solution, thus the TCO, as less profiles need to be used for OTT.

5.5. Dynamic Frame Rate Encoding

Dynamic Frame Rate Encoding works on the temporal activity of the content. Relying on known properties in conjunction with what the user can perceive (all of this is described in the HVS model), this tool will determine the optimum frame rate for a given piece of content, making temporal decimation when a full frame rate is not required.

The value brought by this decimation is that the encoding core will not encode all the frames, therefore saving CPU cycles that can be used either to achieve a better bitrate or to improve the network reach (i.e., less stalling, less rebuffering when the content bitrate is lower). Another use for the CPU savings is to get a denser architecture and/or reduce power consumption, and this adds to the TCO for a service. Dynamic Frame Rate Encoding uses an AI-based two-step approach, as depicted in Figure 3.

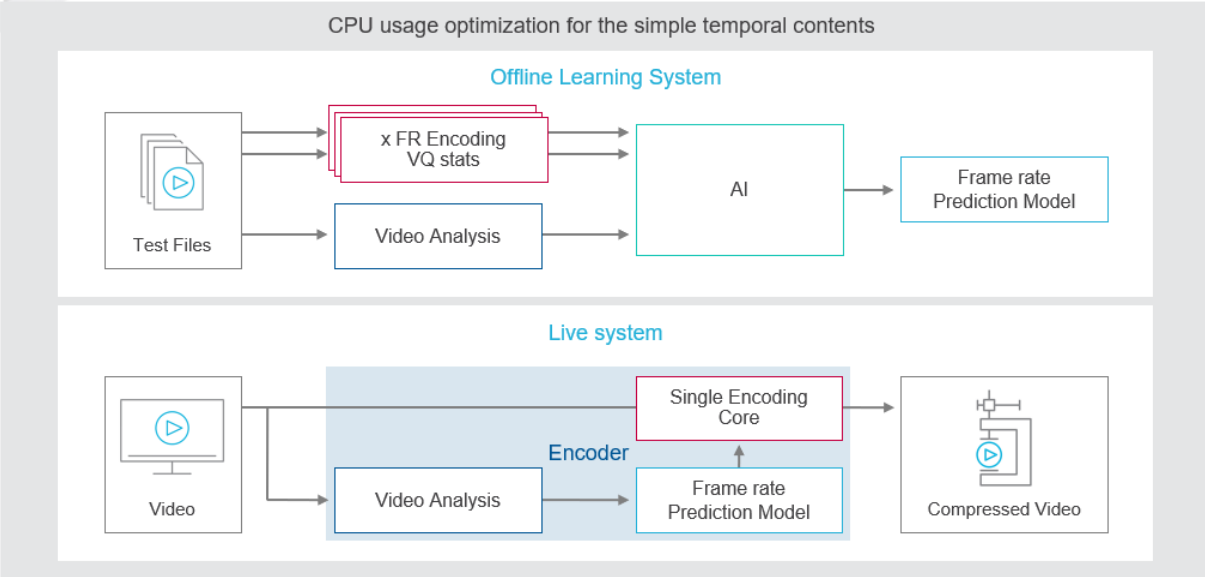


Figure 3 - Dynamic Frame Rate Encoding

5.6. Delivery Optimization

The different tools presented above are aimed at selecting the best encoding configuration depending on the content characteristics. Using these approaches, or a combination, one can create an adaptive content preparation workflow that should optimize the profile ladder based on the content itself or its popularity. But, at this point, another important aspect needs to be considered. The content will be sent over various networks from a core network to the edge and then to the delivery network with a lot of different situations depending on whether the user is on a fixed or radio network, and depending on if it is in a geographic area where this content is very popular compared with other areas (think about a sports match between team A and team B where the audience will be higher in regions A and B compared with the rest of the eligible territory).

With a traditional broadcast paradigm, the service provider delivers one single stream to all the users, making sure that the signal-to-noise ratio will be good enough to ensure a reliable reception on the covered geographic area. On the other hand, a modern OTT distribution platform needs to cope with multiple CDNs, multiple devices, and adapt to much different situations. It seems very ambitious or may be suboptimal to define one single strategy to dynamically adapt the delivery workflows to all these situations. Therefore, flexible architectures will be easily tunable to find the best configuration at a given time.

Whatever the strategy for this delivery optimization is, there are some basics that need to be met on the network. Collecting information from the different elements in the networks is necessary to understand, in real time, how the network is behaving and what the actual QoE is for end users.

A typical analytics collection architecture is illustrated in Figure 4.

In many deployed systems, analytics are collected for the purpose of offline marketing dashboarding but not for real-time usage on a feedback loop.

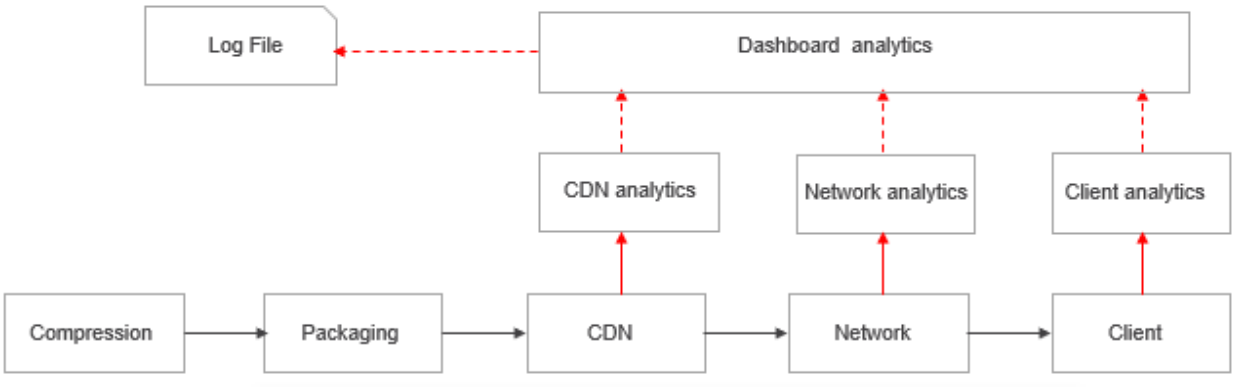


Figure 4 - Typical analytics collection architecture

As presented in the first category of tools that will dynamically adjust the profile ladder, all these decisions may be influenced by the current situation in the delivery network.

The generic architecture for a full Context Adaptive Delivery (CAD) workflow, including the profile ladder optimization and the network path optimization, is illustrated in Figure 5.

Many variants can exist, but the high-level idea is to use the raw data collected on the network to feed a decision engine that will trigger some actions on:

- The compression engine (tools mentioned in the previous sections)
- The packager/origin
- The network path controller

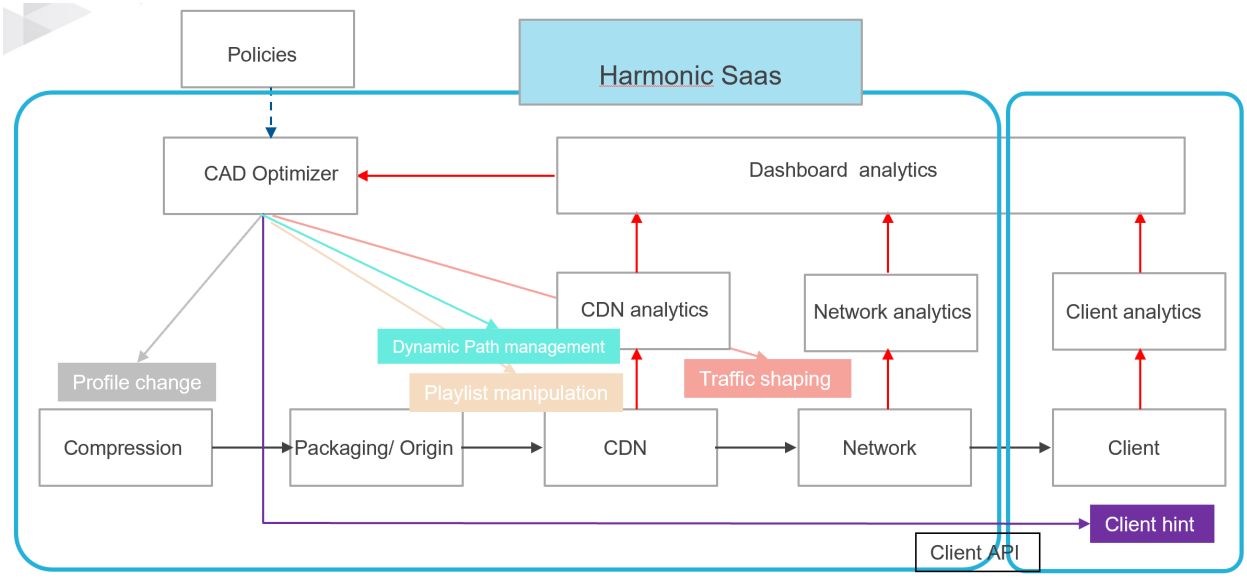


Figure 5 - Context Adaptive Delivery (CAD) generic architecture

The decision engine located in the "CAD optimizer" box above can be as simple as logical switches triggered on fixed thresholds of any of the raw data. But a more forward-looking architecture is to bring AI into this system to enable action before the network crash occurs. Building an accurate and reliable

model of the complete delivery network, from the core network to the last mile, is still a research topic but modern AI approaches look very promising. Because the network conditions are evolving over the time, the AI should adapt automatically based on the current network situation. To achieve this goal, reinforcement learning[14], an area of machine learning, provides some interesting tools. Through a rewarding mechanism, the decision engine gets instantaneous and continuous feedback from the network on the actions taken. It can then adjust the decisions to find, at any time, and under any conditions, the optimum configuration.

As mentioned above, different strategies or scenarios to optimize the delivery at scale for live events can be imagined using the raw information collected in the network. The aggressiveness of the scenario depends on the policy the service provider wants to use to prevent or reduce the breakdown in case of peak audience for a given event.

The global optimization scenario will therefore combine, with a holistic view, the optimization that can be made on the profile ladder and the class of actions to:

- propose a dedicated manifest to some category of player (through manifest manipulation approach),
- dynamically change the delivery path by selecting the most appropriate CDN or delivery nodes

As the second part of the contextual information that the CAD should use, the content consumption is translated into the network load that can be heterogenous, typically based on geographical distribution. To estimate this network load in real time, the system should collect information telemetry from different points in the delivery network. This includes client-side telemetry as well as CDN analytics and network traffic measurements. The collection and capability to perform real-time processing on this information is critical in order to have a timely answer (feedback loop) to any significant change in the content consumption.

Below are some possible scenarios that the CAD optimizer can implement using the network telemetries combined with the service provider policies:

- Based on the reported network load, either global or in some geographic areas, the Customer Management System (CMS) can decide, when a given threshold is reached, whether to prevent any new subscribers from connecting to the service.
- Based on reported network load, either global or in some geographic areas, the traffic can be routed to one CDN or another (when the problem doesn’t come from the last mile).
- Based on reported network load, either global or in some geographic areas, the manifest generator can be instructed, when a given threshold is reached, to remove one (the top one) or several high demanding representations in the manifest either to all or a subset of subscribers. This decision can also be influenced by some business rules to give higher privileges to premium customers.
- Some geographic areas can be isolated and treated with a particular scheme if the network indicates that something is wrong in this area.

These scenarios can be seen as a reaction to a given situation but should bring more value if, thanks to an accurate prediction model, the action anticipates and therefore avoids a future crash.

As depicted in Figure 5, the loop-back action can be directed to different elements in the delivery workflow:

- This can be on the encoder where an action can be decided based on content consumption information. This is the elastic encoding tool presented above.
- This can be at CMS level where new subscribers to the service are rejected when network capacity is exceeded.

- This can be on the packager/origin where playlist manipulation can be done to present the best profile ladder to all or a subset of end-user players.
- This can be on the path management system that can dynamically move the delivery from one CDN to another based on reported consumption or risk of overload on the delivery path.
- This can be on the player itself where some instruction or guidance can be delivered in real time to make sure it will request the most appropriate resource (this may include some hints to help the ABR decision algorithm, for example)

In summary, depending on the contextual set of information on the content characteristics, its consumption and its importance and on scenario choices, Table 2 gives an overview of the different actions triggered by the CAD optimizer. This comes together with the content characteristics-only related tools mentioned in the previous sections.

Table 2 - Delivery optimization summary

Tool category	Usage	QoE improvement	Cost improvement
Profile change	Adjustment of the profile ladder based on delivery network status	✓	
Playlist manipulation	Provide different playlist/manifest to groups of users based on network congestion, user category, device groups	✓	
Dynamic path management	Optimize the distribution between several CDNs or private delivery nodes based on reported consumptions and business rules	✓	✓
Traffic shaping	Set business rules to limit the traffic (enhanced zero rating approach)		✓

6. Conclusion

Context Adaptive Delivery is a new paradigm that takes the end-to-end video delivery workflow to the next level in order to address the two most important aspects for an OTT service provider: delivering better QoE to end users while reducing or keeping the TCO under control. Taking into consideration the dynamicity of the content's consumption over a variety of delivery networks is the next step now that OTT technologies are ready for the main screen. Moving from today's rigid configuration to much more adaptive workflows should be seen the same as the transition from hardware- to software-based solutions. With Content Adaptive Encoding, Dynamic Resolution Encoding, and Dynamic Frame Rate Encoding, the content preparation can be much more flexible and adapt to the content type itself as well at its popularity. All these tools should be used to prepare the optimum profile ladder at any time. Then, once the profile ladder is optimized, using network optimization in an adequate scenario allows one to provide the best delivery to users, no matter what their location, device or subscription is. This will quickly create opportunities to improve the user experience, leveraging the value that AI-based processing and cloud-native deployments bring into this landscape.

With more adaptive workflows for the delivery of large-scale live events, the end-user experience can be dramatically improved to reach the expected level and match broadcast services. In addition, service operator profitability can be improved. Even better, we can imagine the introduction of new services and new user experiences that are not possible today.

Abbreviations

ABR	Adaptive Bit Rate
AI	Artificial Intelligence
CAD	Context Adaptive Delivery
CAE	Content Aware Encoding
CDN	Content Delivery Network
cDVR	Cloud Digital Video Recorder
CIRR	Connection Induced Rebuffering Ratio
CMAF	Common Media Application Format
DASH	Dynamic Adaptive Streaming over HTTP
HLS	HTTP Live Streaming
HTTP	HyperText Transport Protocol
MIT	Massachusetts Institute of Technology
ML	Machine Learning
OTT	Over The Top
QoE	Quality of Experience
QoS	Quality of Service
QP	Quantization parameter
SMPTE	Society of Motion Picture and Television Engineers
TCO	Total Cost of Ownership
UHD	Ultra High Definition
VOD	Video on Demand
VSF	Video Start Failure
VST	Video Start Time
WebRTC	Web Real-Time Communication

Bibliography & References

[1] HTTP Live Streaming 2nd Edition draft-pantos-hls-rfc8216bis-07. https://datatracker.ietf.org/doc/draft-pantos-hls-rfc8216bis/

[2] ISO/IEC 23009-1:2019, Information Technology — Dynamic Adaptive Streaming Over HTTP (DASH) — Part 1: Media Presentation Description and Segment Formats (4th Edition), https://www.iso.org/standard/79329.html

[3] T. Fautier, "How OTT Services Can Match the Quality of Broadcast," SMPTE 2019 Annual Technical Conference & Exhibition.

[4] Wikipedia, "HTTP," https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol

[5] WebRTC, https://webrtc.org/

[6] ISO/IEC 23000-19:2020, Information Technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media, https://www.iso.org/standard/79106.html

[7] "Per-Title Encode Optimization," Netflix, https://netflixtechblog.com/per-title-encode-optimization-7e99442b62a2

[8] Y. Reznik, et al, "Optimizing Mass-Scale Multi-Screen Video Delivery," Brightcove, 2019, http://reznik.org/papers/ReznikY_BEITC2019.pdf

[9] J.L. Diascorn, "How AI Technology is Dramatically Improving Video Compression for Broadcast and OTT Content Delivery," SMPTE 2019 Annual Technical Conference & Exhibition.

[10] S. Chinchali, et al, "Cellular Network Traffic Scheduling With Deep Reinforcement Learning," Stanford University, 2018, http://asl.stanford.edu/wp-content/papercite-data/pdf/Chinchali.ea.AAAI18.pdf

[11] H. Mao, R. Netravali, M. Alizadeh, "Neural Adaptive Video Streaming With Pensieve," MIT Computer Science and Artificial Intelligence Laboratory, http://web.mit.edu/pensieve/content/pensieve-sigcomm17.pptx and http://web.mit.edu/pensieve/content/pensieve-sigcomm17.pdf

[12] C. Tolhurst, "Deep Learning Algorithms Could Secure the Future of 4K Streaming," Venture Beat, 2017, https://venturebeat.com/2017/10/26/deep-learning-algorithms-could-secure-the-future-of-4k-streaming/amp/?__twitter_impression=true

[13] Harmonic, "EyeQ Achieving Superior Viewing Experience," https://info.harmonicinc.com/technical-guide/achieving-superior-viewing-experience/?hsLang=en

[14] Wikipedia, "Reinforcement Learning" https://en.wikipedia.org/wiki/Reinforcement_learning

How VCMTS Paves The Way For 5G Over DOCSIS

Exploring Software-centric Solutions for 5G Xhaul and FMC

A Technical Paper prepared for SCTE by

Brendan Ryan
System Architect
Intel Corporation

Ed Dylag
Marketing Development Manager
Intel Corporation

Thushara Hewavithana,
Wireless Systems Architect
Intel Corporation

Eric Heaton,
Platform Solutions Architect
Intel Corporation

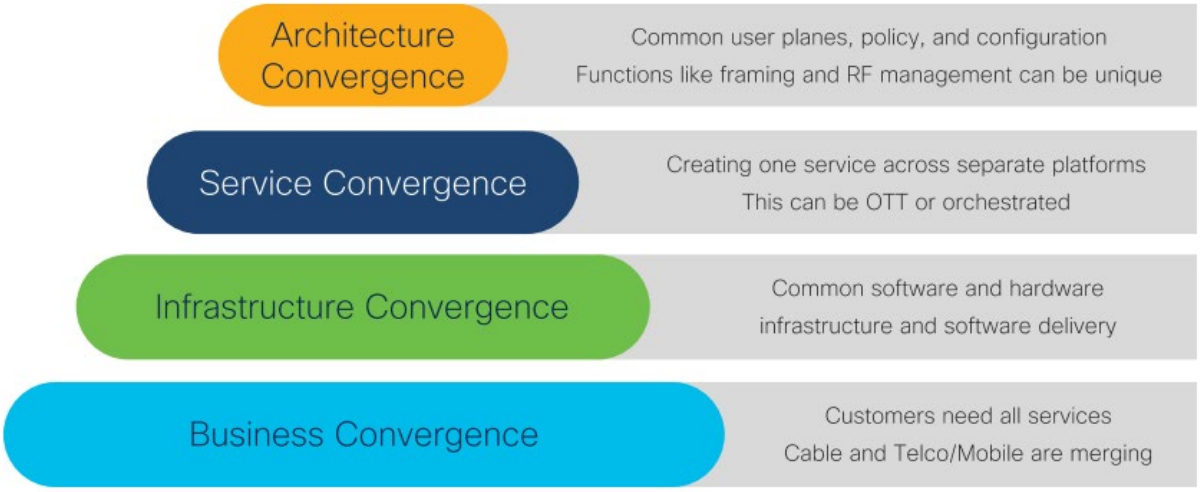
1. Introduction

Network convergence has been gaining a lot of attention in the telecoms industry recently with the architecture shift in both wireline and wireless communication networks to be more centralized in terms of core and baseband processing, and more distributed in terms of the analogue and RF functions.

This has also led to the disaggregation of the hardware and software functions of the network. Furthermore, the standardization of the interfaces in this network, for example the O-RAN, 3GPP wireless wireline convergence (WWC), Broadband Forum fixed mobile convergence (FMC), cable distributed access architecture (DAA) and the more recent flexible mobile architecture (FMA) initiatives enables wider participation of ecosystem vendors leading to competitive solutions for operators.

Centralized functions of the network such as 5G-Core for wireless and cable modem termination system (CMTS) for data over cable system infrastructure specifications (DOCSIS) are well suited to being implemented in software and to run on common off the shelf (COTS) hardware, giving better scalability over time and adaptability for different deployment scenarios.

There are many aspects of network convergence as shown below (from reference [1]), at business, infrastructure, service, and architecture level.



Source: Cisco & CableLabs, "Cable and Mobile Convergence, A Vision from the Cable Communities Around the World"

Figure 1 – Four Levels of Convergence

This paper will focus on infrastructure and architecture convergence and specifically, how the advent of 5G has created a great opportunity to leverage the existing DOCSIS network for two key purposes:-

- Xhaul of 5G small-cell traffic over DOCSIS
- Fixed mobile convergence (FMC) using the Cable access network

It will be shown how the flexibility of a software-based vCMTS makes it much easier to adapt the Cable access network for these new 5G use-cases and to evolve over time.

This paper will explore the rationale outlined above as to why Cable MSOs should prioritize vCMTS deployment in preparation for 5G xhaul and FMC, while identifying some gaps that need to be addressed by the Cable industry to prepare for 5G mobile convergence.

2. Cloud-native Infrastructure Convergence

Infrastructure convergence as shown in Figure 1 in the previous section is already happening in terms of common hardware platforms and software stacks being used to host wireless and wireline network functions as shown in Figure 2 below. Network function virtualization (NFV) and the subsequent progression to adoption of cloud technologies for communications network functions started in the Mobile core and is now gathering pace for central-office deployments of access network functions such as virtual broadband network gateway (vBNG), and the virtual cable modem termination system (vCMTS) in cable head-ends.

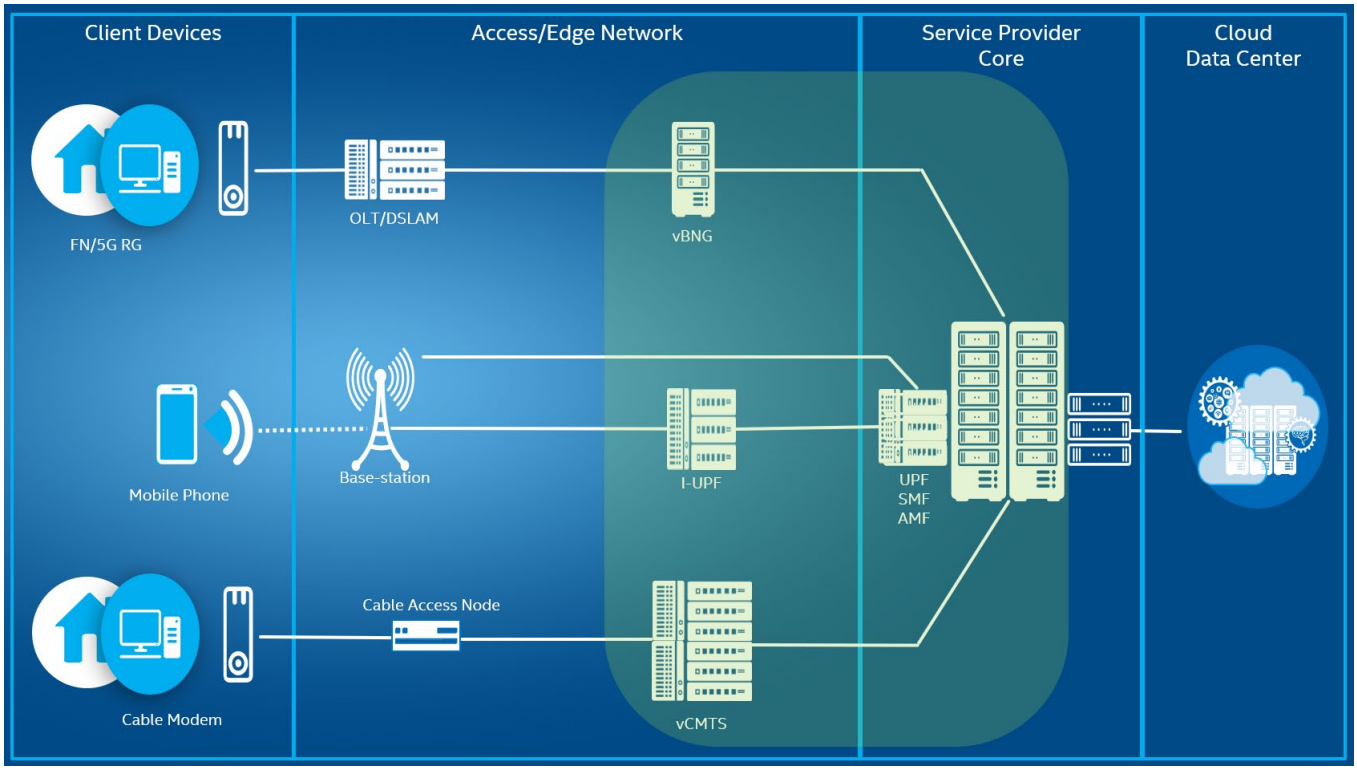


Figure 2 –Telecoms Network Infrastructure Convergence

Cloud-native is a term that has come to be used to describe the use of cloud technology for network function deployment on COTS hardware platforms. Significant work has been done in several mainstream open-source projects to enable high performance network function data-planes to be deployed on a cloud-native platform. Open-source packet-processing frameworks such as DPDK and FD.io/VPP (see references [2] and [3]) make the required data-plane performance possible in software while key features have also been added to open-source cloud computing platforms such as Kubernetes to enable network applications based on these packet-processing frameworks to be deployed with the scalability, flexibility, and observability of a cloud environment.

2.1. Cloud-native deployment of vCMTS

Significant effort has been invested in open-source projects to advance cloud-native network function virtualization by providing reference software which is highly optimized to run on general purpose COTS hardware platforms while also leveraging the benefits of cloud technology such as automation, scalability, flexibility and observability.

A Container Bare-metal Reference Architecture as shown in Figure 3 below (see reference [4] for details) is provided by Intel which combines general purpose hardware and open-source software components to provide a reference cloud-native platform which may be configured appropriately with Ansible playbooks to host network functions at various network locations - mobile core, central-office, access network edge and on-prem.

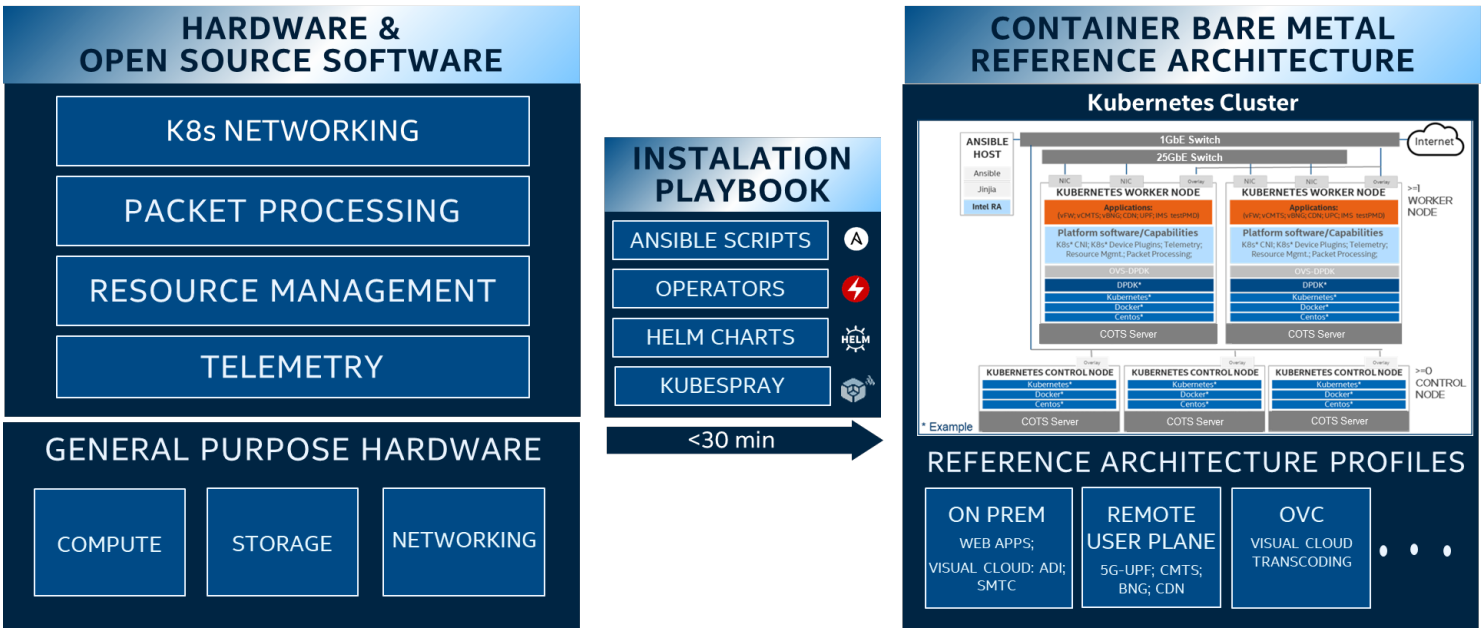


Figure 3 - Container Bare-metal Reference Architecture

A cloud-native software stack such as this is key to preparing the cable access network for mobile xhaul and FMC, as both require new functionality to be added. The network function adaption needed for these new DOCSIS use-cases are greatly simplified by having a modular extensible cloud-native vCMTS software architecture with rich telemetry and observability already in place.

A reference vCMTS data-plane implementation based on DPDK as shown in Figure 4 below has also been provided by Intel (see reference [6] for details). Such an implementation demonstrates the software performance capability of this key component of the cable access network on a cloud-native platform. DPDK forms the foundation for any high-speed packet processing software. And for this reference DOCSIS MAC data-plane, 80% of the code comes from DPDK, with each function in the DOCSIS MAC pipeline leveraging existing CPU-optimized DPDK library functions.

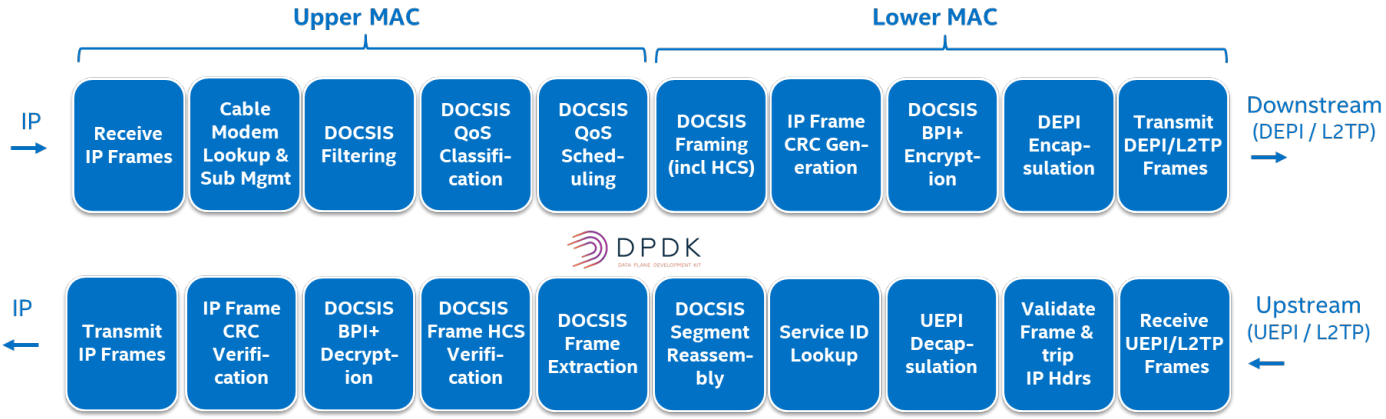


Figure 4 –vCMTS Data-plane based on DPDK

An example of a cloud-native software stack based on open-source components for deployment of vCMTS is shown below. Such a stack may be deployed in a lab environment for performance and TCO analysis using Intel's container BMRA (reference [4]) and vCMTS reference data-plane (reference [6]).

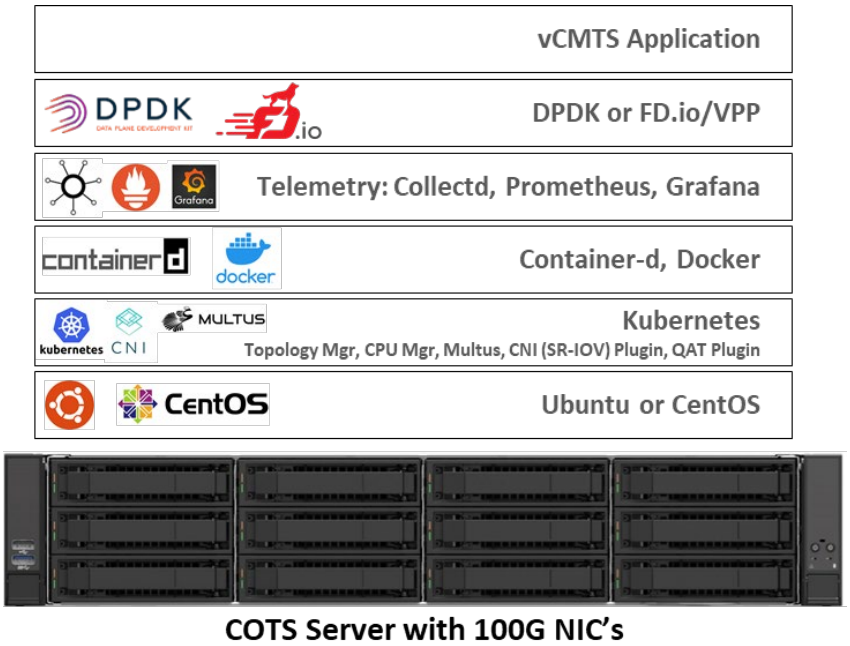


Figure 5 - Sample Software Stack for vCMTS Cloud-native Deployment

A cloud-native vCMTS runtime environment is shown in Figure 5 below. The vCMTS software architecture shown below is an example of how a monolithic CMTS may be decomposed into multiple Pod's (for example one per cable service-group) and each Pod decomposed into containers running distinct functional parts. This is based on the cloud concept of micro-services which enables optimum extensibility and resilience while simplifying the maintenance and upgradability of network function deployments.

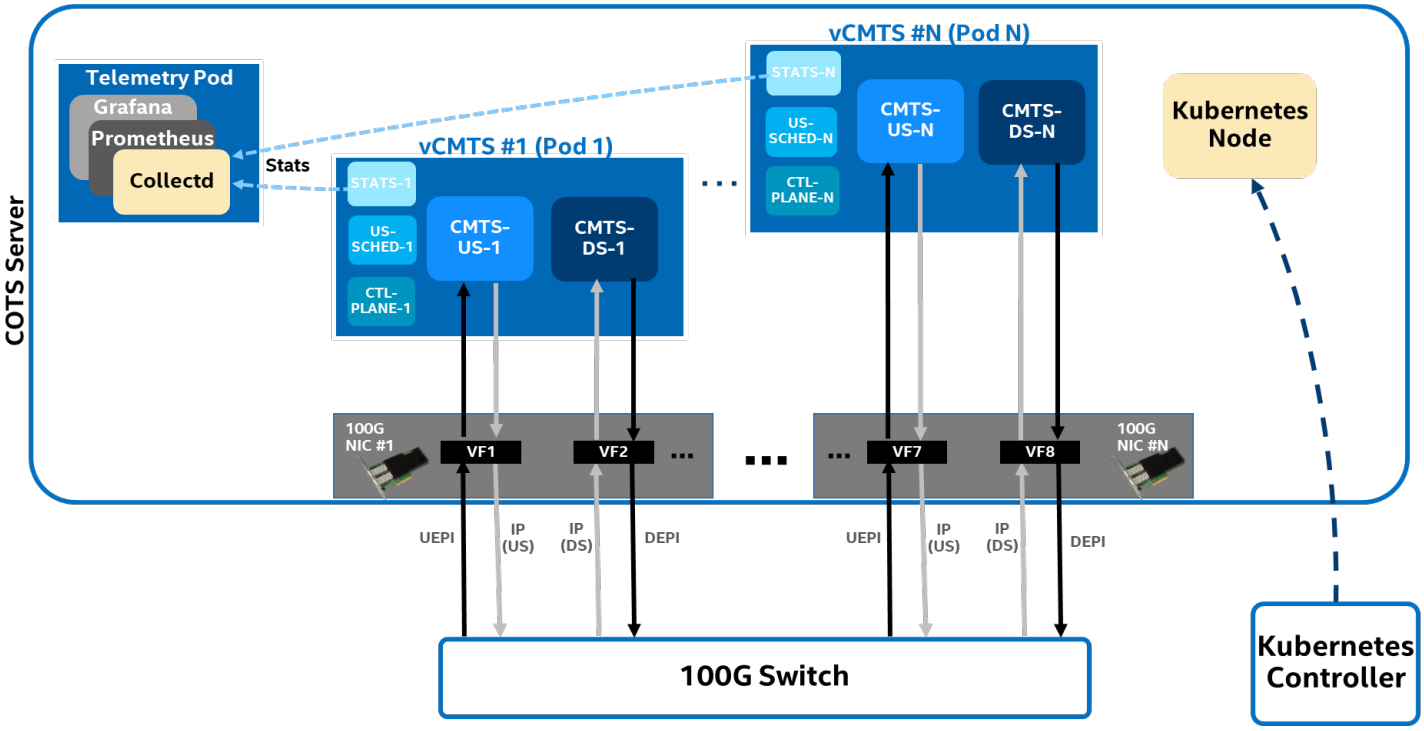


Figure 6 - Cloud-native vCMTS Simulation Environment

Reference [5] provides details of the performance capability of a cloud-native vCMTS data-plane. Empirical performance measurements such as shown in Figure 6, taken using the Intel vCMTS reference data-plane, prove that a cloud-native vCMTS is more than capable of achieving the throughput required for future DOCSIS 3.1 and 4.0 capabilities, and indeed to support transport of 5G small-cell traffic.

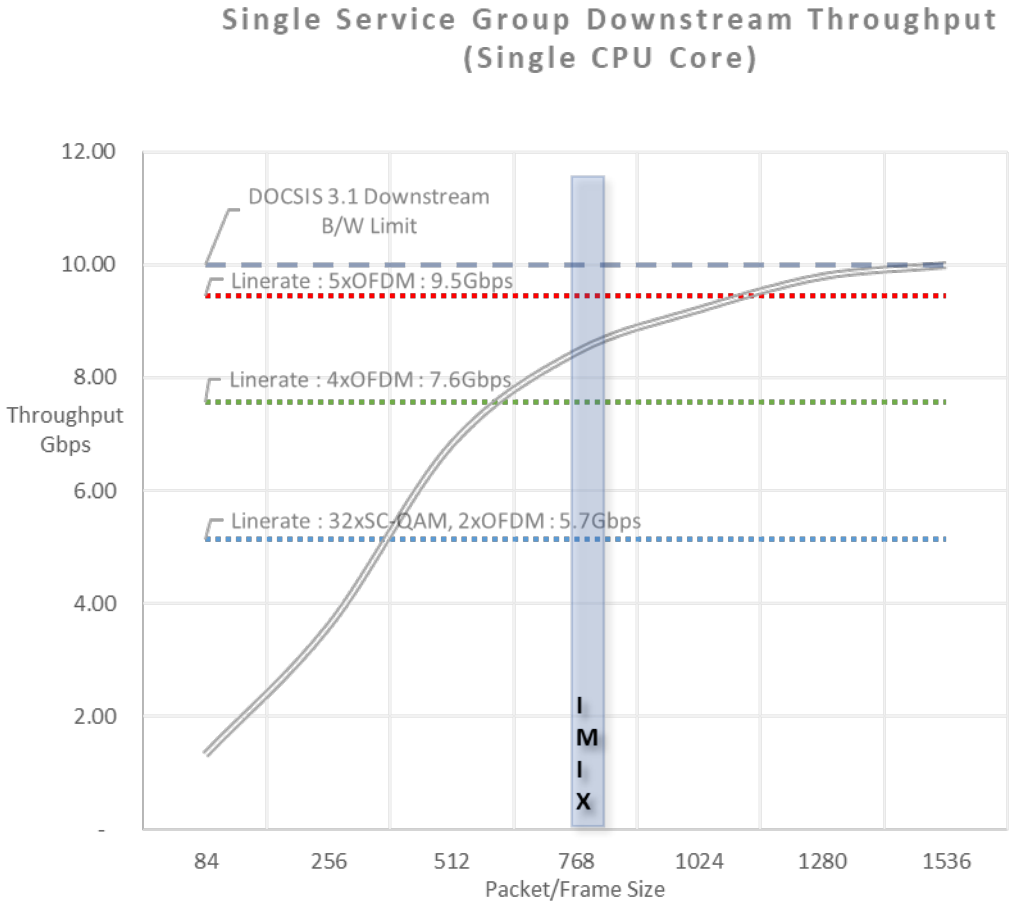


Figure 7 –vCMTS Data-plane throughput capability of a single CPU core

Scalable performance is a key aspect of vCMTS that directly paves the way for 5G xhaul and FMC use-cases. Each have their own particular performance requirements which will be covered in the respective sections that follow.

3. Extending vCMTS to support 5G Xhaul

Multiple systems operators (MSOs) are rolling out 5G deployments for a number of different business drivers and using a number of different convergence approaches. Operators like Shaw Communications and Videotron have acquired or built wireless assets to become mobile network operators (MNO). Charter, Comcast and Cox have entered into mobile virtual network operator (MVNO) agreements with MNOs such as Verizon and T-Mobile. In an MVNO agreement, the MVNO pays the MNO a fee to carry subscriber traffic over the MNO network. Still other operators like Vodafone began as an MNO and later acquired fixed network assets.

One common element across these operators is that there is a certain cost associated with carrying subscriber traffic and there is a goal to minimize that cost, thereby increasing operating margin. Minimizing cost to a large extent means leveraging existing assets to their full extent. A key existing asset in the cable industry of course is the hybrid fiber coax (HFC) network. Charter and Comcast plan to use the HFC network to build a network of 5G small cells to offload MVNO traffic, thereby reducing service cost. They and other operators will also market their HFC network to carry traffic for other operators planning to deploy their own small cell networks.

Three possible 5G functional split options are shown in Figure 7 below. Throughput and latency requirements placed on the HFC network will vary depending upon the split. These requirements are categorized as fronthaul, midhaul and backhaul with fronthaul having the most stringent bit-rate and latency constraints and backhaul the least.

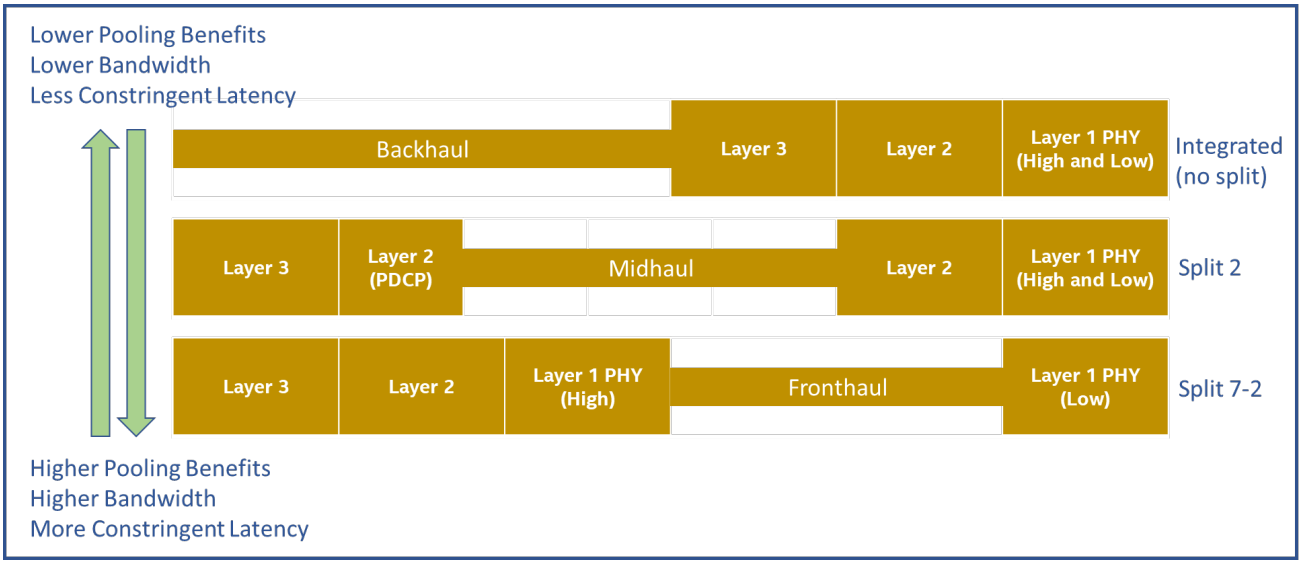


Figure 8 – Possible 5G Functional Split Options

Because HFC networks reach over 90% of North America and many parts of Europe, MSOs are uniquely positioned to quickly and cost-effectively roll out 5G if the HFC network can meet these latency and bit-rate requirements. Unfortunately, a 'plain vanilla' DOCSIS 3.1 network is hard pressed to meet all but basic mobile backhaul requirements.

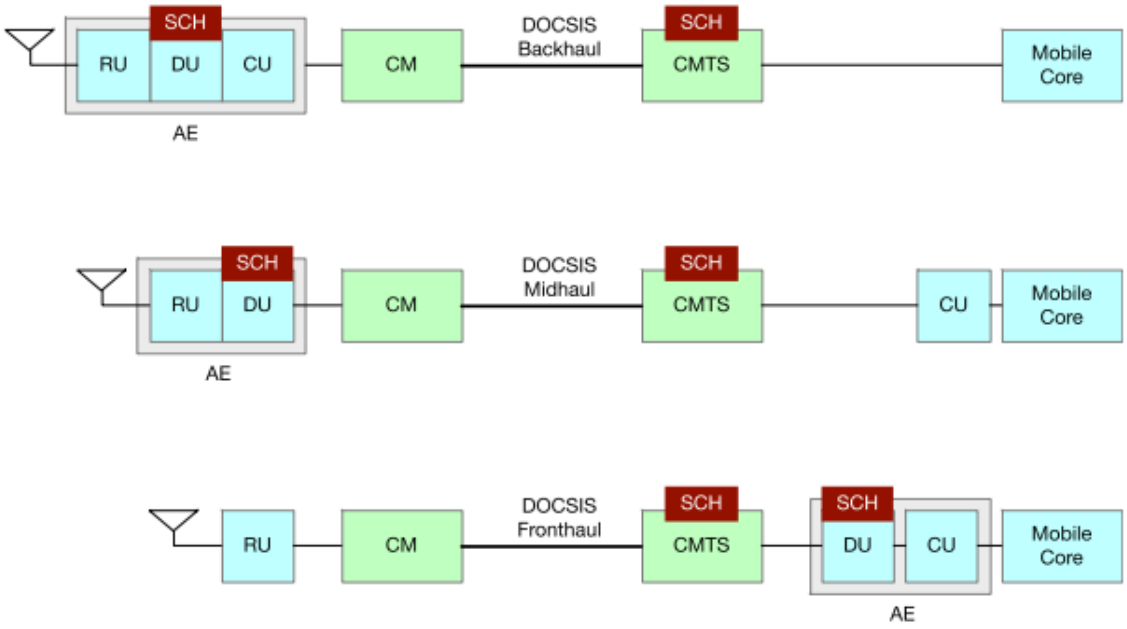
But several recent technology advancements will greatly enhance the ability of the HFC network to meet 5G xhaul requirements - namely:

- The recent release of the DOCSIS 4.0 standard enabling multi-gigabit capacity in both uplink and downlink
- Distributed access architecture (DAA) which enables a software-based DOCSIS MAC to be deployed at the Cable head-end or remote HFC node
- The development of flexible MAC architecture (FMA) which enables flexible deployment of DAA options
- Low latency DOCSIS (LLD) which provides latency-sensitive traffic management
- Low latency xHaul (LLX) which streamlines upstream scheduling across 5G and DOCSIS schedulers, further reducing latency
- Finalization of the generic access platform (GAP) hybrid fiber coax (HFC) node specification which enables operators to deploy modern CPUs in nodes to implement evolving access protocols and optimization techniques in software much more cost effectively

In the remainder of this section, we will focus on LLX and GAP to showcase how implementing DOCSIS in software (i.e. implementing vCMTS) paves the way for 5G xhaul.

3.1. How to achieve low latency for 5G xhaul over DOCSIS

Figure 8 below shows how a DOCSIS access network may be used as a transport for 5G backhaul, midhaul and fronthaul cases. This is an example of infrastructure convergence which includes the transport network as well as common hardware and software.

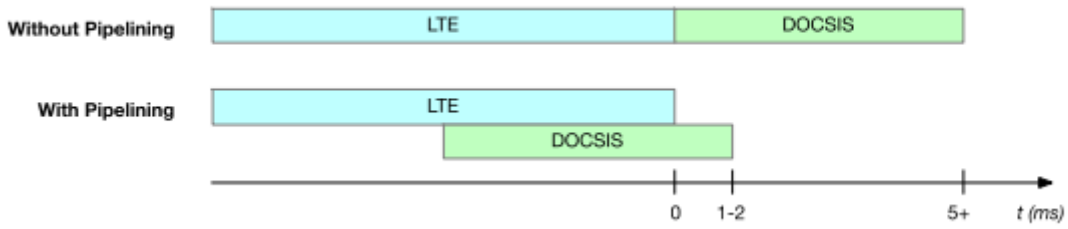


Source: CableLabs, "Low Latency Mobile xhaul over DOCSIS Technology"

Figure 9 – Access Entity Definition for Mobile Xhaul

As both DOCSIS and 5G are shared-medium networks, an upstream scheduler (SCH in the diagram above) must be present to ensure that traffic from subscribers does not collide. These respective schedulers introduce latency because a subscriber must ask and wait for permission to transmit through a request-grant message sequence. Without any optimization, the delay is compounded since each network has its own scheduler cycle.

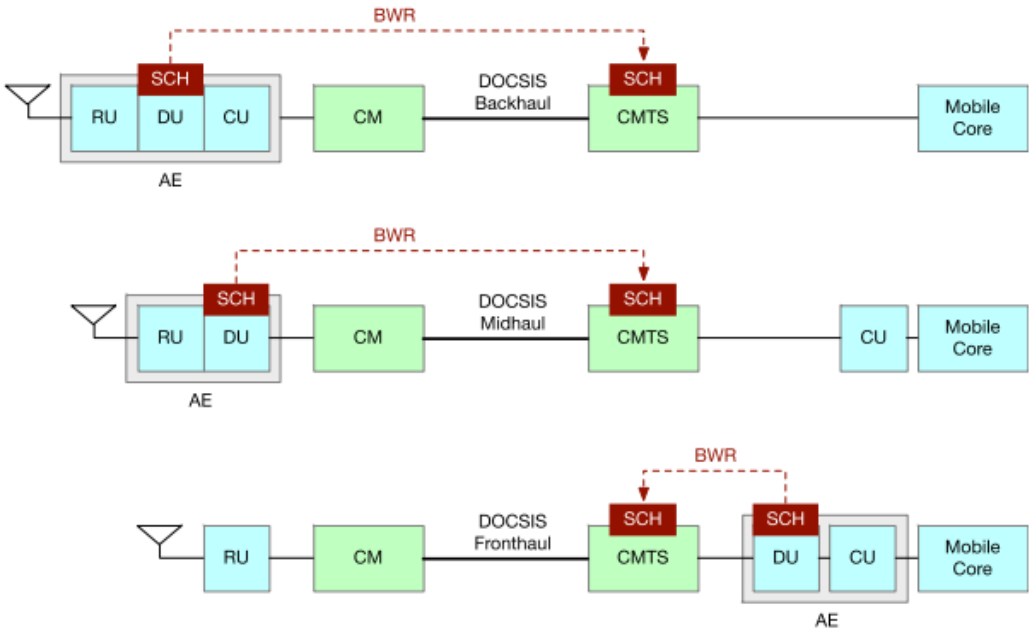
Through use of LLX (as specified in reference [8]) the DOCSIS upstream bandwidth grant request-response delay is hidden behind the equivalent upstream traffic scheduling function for wireless access as shown below (noting that while the diagram shows LTE, the same principles apply 5G)



Source: CableLabs, "Low Latency Mobile Xhaul over DOCSIS Technology"

Figure 10 – Hiding DOCSIS Upstream Sheduling Latency behind Mobile

This is achieved by adding the bandwidth report (BWR) capability to the 5G DU and the DOCSIS Upstream scheduler in order to pre-determine the upstream grants needed on the DOCSIS network for Mobile xhaul.



Source: CableLabs, "Low Latency Mobile Xhaul over DOCSIS Technology"

Figure 11 – BWR Transits Across Xhaul

As shown below, a cloud-native vCMTS software architecture can be relatively easily adapted to support the LLX BWR feature required for low-latency transport of 5G small-cell traffic. This is because the upstream scheduler is simply another algorithm implemented in software. As such, only software in the vCMTS system needs to be updated to implement LLX – and specifically only the software algorithm that implements the scheduler. Contrast that with a proprietary CMTS consisting of several sub-systems engineered to a specific set of requirements. Any change in the target requirements means that each sub-system needs to be carefully evaluated and re-tuned – which could include updates to proprietary hardware – a long and expensive undertaking.

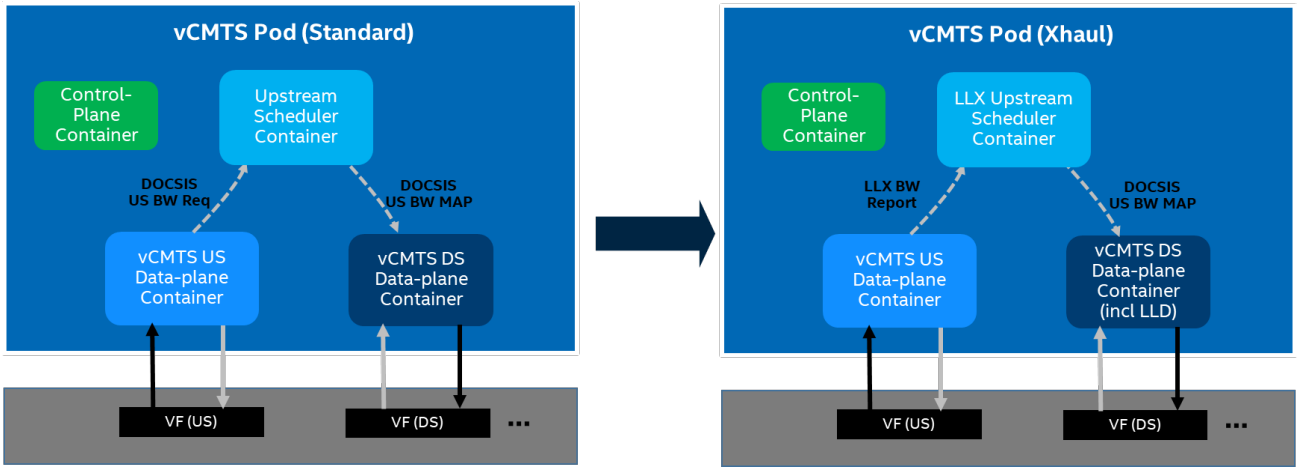


Figure 12 – Cloud-native Software Architectiture for LLX

But does implementing vCMTS (i.e. DOCSIS processing in software) actually introduce significant latency?

It can be seen in Figure 12 below from a CPU cycle count breakdown of the vCMTS upstream and downstream data-plane pipelines in software that it is possible to achieve ultra-low packet-processing latency in software with a single core run-to-completion pipeline based on a high-performance packet-processing frame-work such as DPDK or FD.io/VPP.

Considering average CPU cycles per packet for upstream and downstream vCMTS data-plane pipelines, based on measurement using the Intel vCMTS reference data-plane (reference [6]) with a 2.2 GHz CPU clock, software can perform DOCSIS MAC processing on upstream and downstream packets at an average of 615 nano-seconds and 967 nano-seconds respectively.

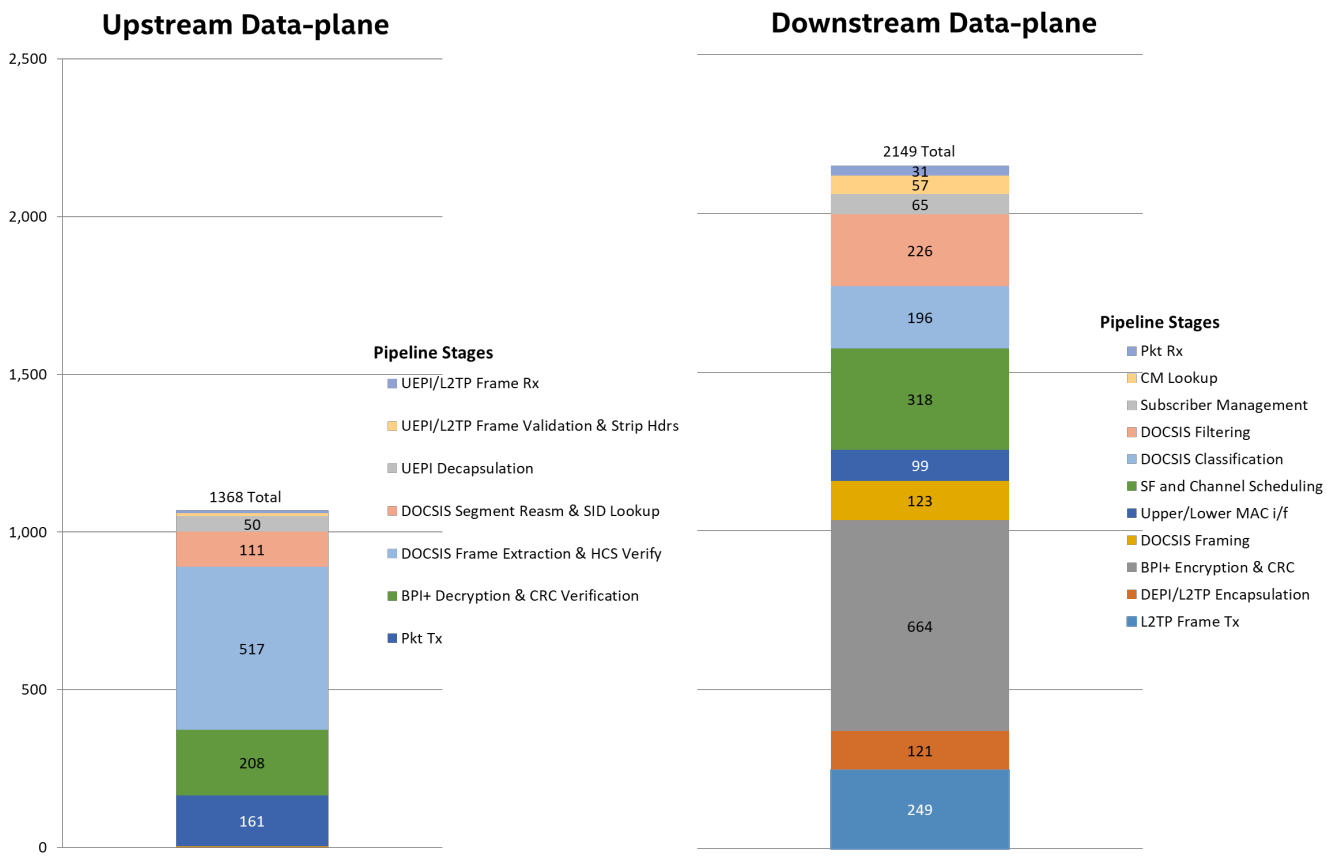


Figure 13 – vCMTS data-plane CPU cycle-count breakdown

However, latency is also added by packet buffering which occurs at three points in a DOCSIS MAC data-plane software pipeline - NIC Receive, Transmit and Downstream QoS scheduling. However, based on latency analysis using the Intel vCMTS reference data-plane, an optimally tuned system, can achieve sub 100 micro-second average latency per frame with a DOCSIS MAC data-plane implementation in software.

Low latency DOCSIS (LLD) functionality as described in a white-paper from CableLabs (reference [9]) may be used to reduce the impact of this by extending the standard Downstream Service-Flow QoS scheduling component to prioritize latency sensitive flows such as 5G signaling when vCMTS is being used for 5G xhaul over DOCSIS. LLD will most likely be needed to achieve the ultra low latency required for 5G fronthaul over DOCSIS.

The combination of an ultra-low-CPU-cycle software pipeline and the low-latency extensions to the DOCSIS standard described earlier should enable DOCSIS to meet the latency requirements for 5G backhaul and it should even be possible to meet the more stringent latency requirements for midhaul by combining LLX with proactive upstream scheduling.

Further work is needed to prove that DOCSIS is sufficient for 5G fronthaul, which has latency requirements in the order of hundreds of microseconds. At this low level of latency, even the PHY layer

latencies become critical and need addressing. As for the DOCSIS upstream schedular related latency, LLX could potentially be augmented with AI models for better traffic pattern prediction resulting in even smaller request-grant cycles and less grant waste. Perhaps the solution lies in a custom grant-request cycle just for small cells, or a combination. Whatever the future solution brings, vCMTS is just a software upgrade away from implementing that solution which makes vCMTS a logical first step in 5G over DOCSIS and network convergence.

Not only is vCMTS more easily adapted to new innovations, but it can be readily scaled down and distributed as needed.

3.2. How FMA and GAP can help with 5G xhaul scenarios

As shown below, the portability of software enables the same DOCSIS MAC software to be used for all three flexible MAC architecture (FMA) scenarios being specified by CableLabs (see reference [10]).

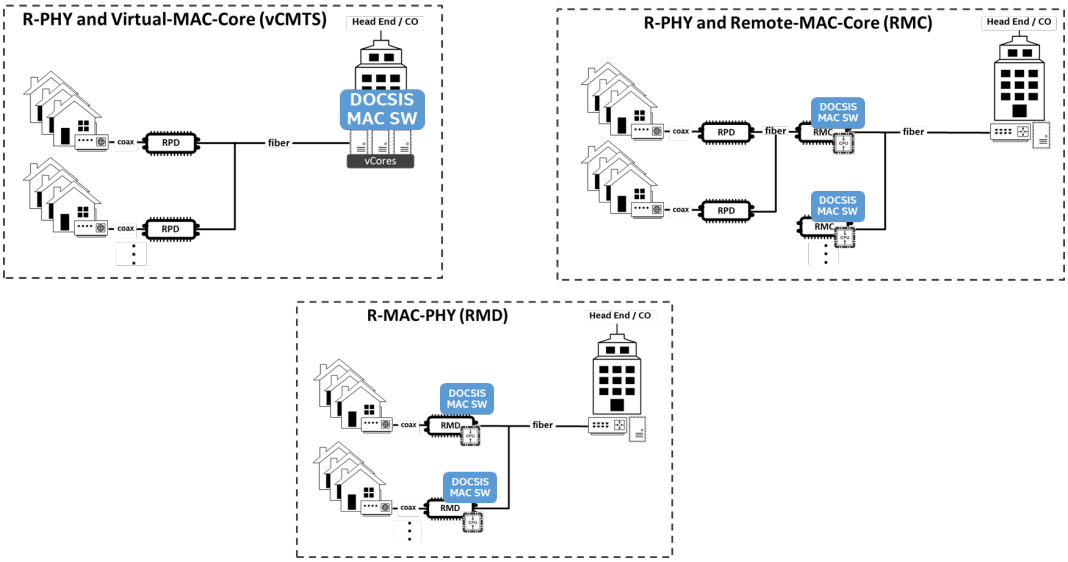


Figure 14 – Portability of Software for Flexible MAC Architecture Scenarios

Overlaying these three FMA options with the RU, DU, CU split options shown in Figure 7 produces a rather large matrix of deployment options for 5G over DOCSIS.

The industry is still in the process of evaluating these deployment options. While not all options ultimately will get deployed, it is most likely that at least several will get deployed because operators all have different assets and aspirations.

Analogous to the Intel vCMTS reference data-plane (see reference [6]), Intel has developed vRAN reference software (see reference [12]) for the upper PHY portion of a 5G base band unit (BBU). The industry eco-system adds layer 2 and layer 3 components to implement a full software based vRAN stack. This means both DOCSIS and RAN access implementations that are almost entirely software-based exist in the market today. In the case of vCMTS, the PHY is implemented in hardware and in the case of vRAN, the lower half of the PHY is implemented in hardware.

Because vCMTS and vRAN are both software based and can be scaled and moved as needed, implementing the matrix of FMA and RAN splits in combination becomes a much more tenable exercise. And as technology evolves both in the RAN and CMTS stacks to make them work better together, implementation of that new technology is a software upgrade away.

Of course, a suitable CPU host is required to run these software stacks. In the headend, the host will be a standard off the shelf rack mountable server with one or more x86 processors. But until recently a suitable CPU host in the HFC node did not exist.

The SCTE has recently released the generic access platform (GAP) specification. This specification defines the mechanical, electrical and thermal requirements of a modular, standardized strand mount HFC node enclosure. GAP includes the specification of a high speed PCIe/KR backplane which means that GAP is a suitable host for modern day x86 and other processors.

AOI, ATX, Charter, Cisco, Intel and Silicom demonstrated an early working prototype of a general purpose compute module at CableTec Expo in 2019 (see reference [13]). Coupled with an RPD module, the prototype was an early implementation of a software based RMD device. By adding the appropriate vRAN BBU components, a GAP RMD node becomes a key enabler for 5G over DOCSIS. While out of scope for this paper, by adding a pluggable PON SFP, a GAP node can similarly be used to drive 5G small cells over PON.

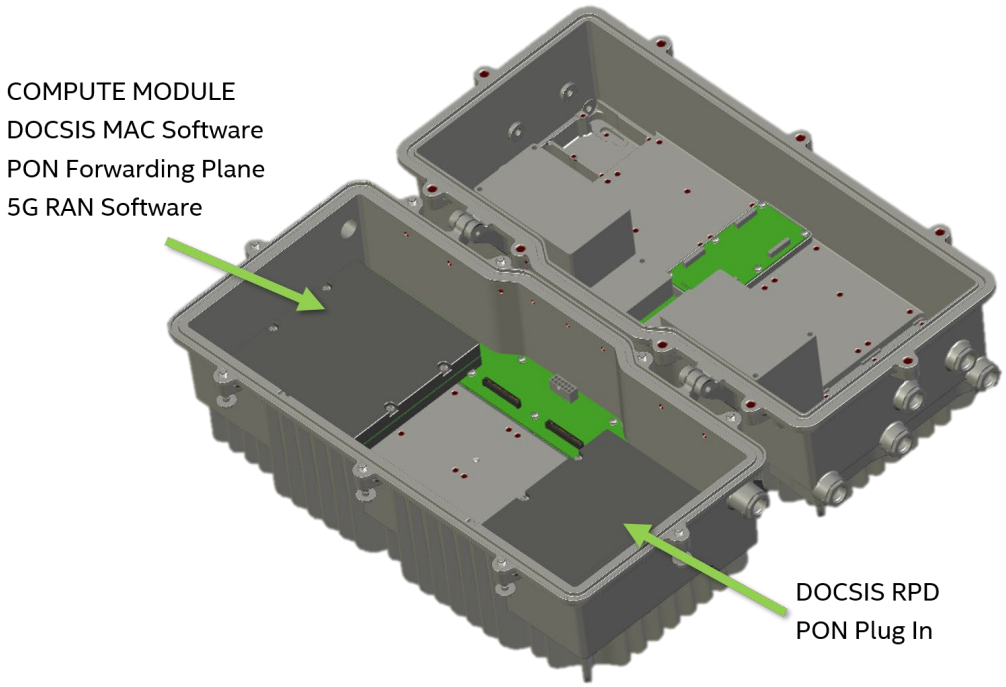


Figure 15 – GAP Configuration for RMD

Other Benefits of a software-based access network include enabling the highest degree of flexibility as it can be written once and re-targeted onto different platforms such as central-office servers or edge nodes.

But there are other more compelling reasons to move to software-based access that revolve around the availability of a tremendous amount of readily available software tools and packages.

Examples include:

- Telemetry frameworks and tools for faster fault detection and correction built around modern interfaces such as Yang models.
- Management and orchestration tools for easier and faster provisioning as well as 'self-service' customer provisioning.
- Network slicing applications to automate the creation of customer defined networks with customer specific characteristics around throughput, latency and quality of service.
- AI and Deep Learning frameworks to build sophisticated network optimization algorithms.

While it could be argued that all of the above are POSSIBLE to implement on networks with proprietary access nodes, software-based access nodes make the above much more PRACTICAL to implement. The reason is simple economics. Standard tools and utilities are developed for a broad range of markets and industries. This means that the cost of development is amortized over a much larger footprint and in the case of open source, that development cost is shared across several companies. Contrast that to a proprietary CMTS as an example where a single vendor historically had to develop 100% of these tools – including custom command line interface (CLI) protocols tasked with managing several internal sub systems such as ASIC based line cards. The latter is costly, time consuming to implement, costly to maintain and largely not re-usable.

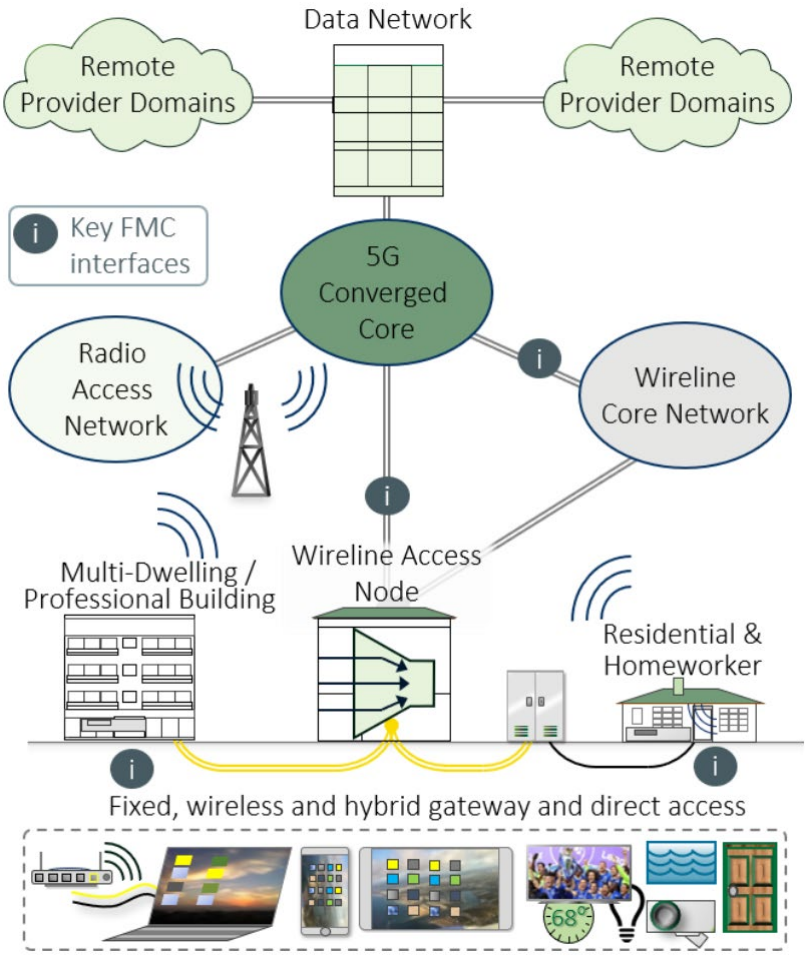
4. Adapting vCMTS for Cable FMC

The goal of fixed mobile convergence (FMC) is to have a common core across wireless and wireline access networks. The term wireless and wireline convergence (WWC) is used by 3GPP to describe this convergence which may also be applicable to Cable access; for the purpose of this paper FMC will be used to describe this, whether for Cable or BBF Wireline.

As described in reference[16], the drivers for this include:

- Seamless multi-access connectivity and simplified service experience for end-users
- More effective use fixed and mobile assets for service providers
- Opportunity for service providers to properly integrate emerging technologies such as Connected Home, Virtualization/Cloud and High-speed Broadband Access

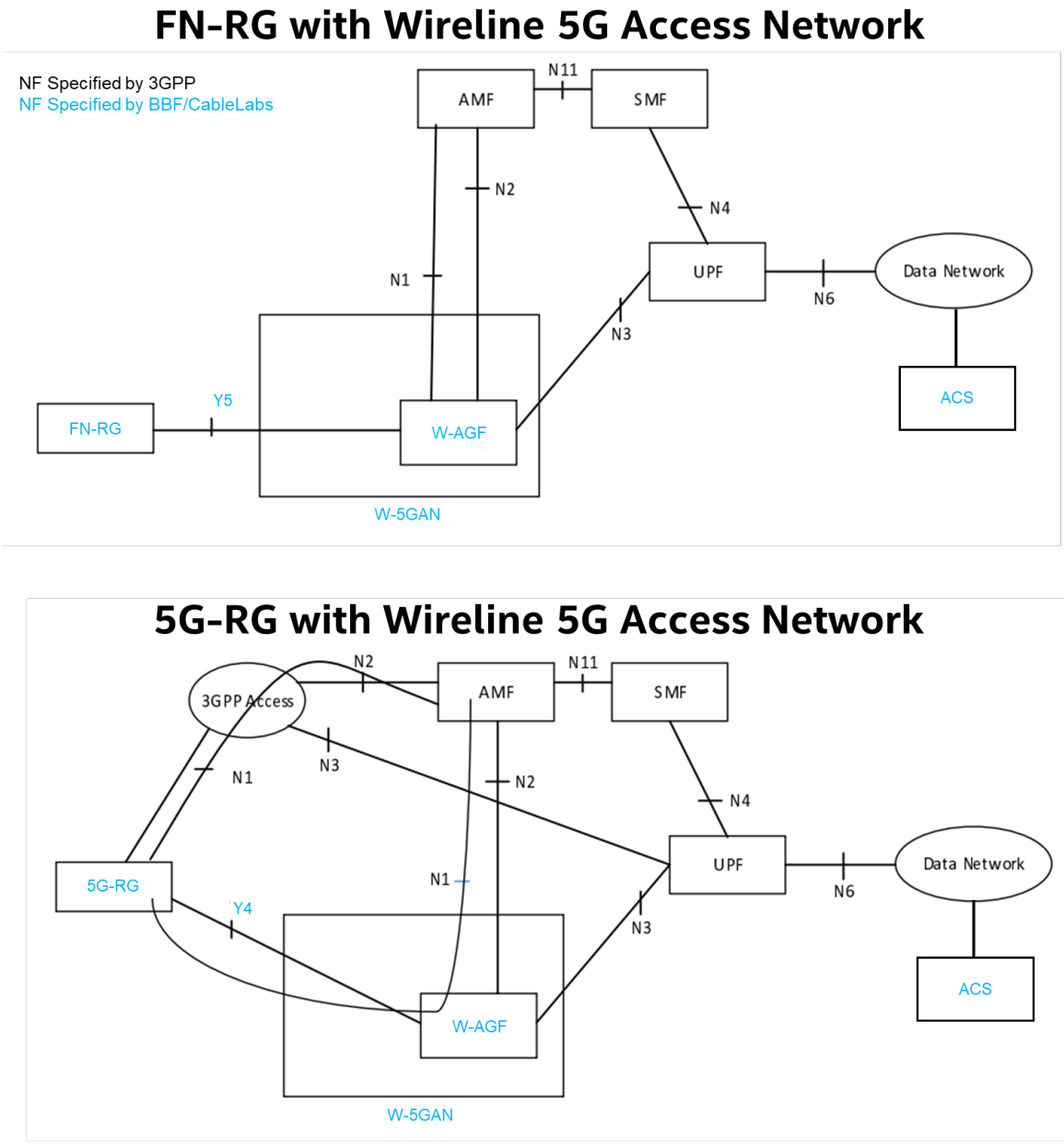
Figure 15 shows the desired end state for WWC/FMC – both wireline and wireless access served by a single 5G core.



Source: Broadband Forum, "5G Fixed-Mobile Convergence – Marketing Report"

Figure 16 – 5G Fixed-Mobile Convergence

As stated in the 3GPP TS 23.316 specification for wireless and wireline access support (reference [17]) the key network function required for convergence is the Wireline AGF (W-AGF). This is a network function that mediates between the wireline access network and the 5G Core Network. In addition to supporting existing residential gateways (RGs), it is assumed that a new residential gateway (5G RG) will be introduced which must also be supported in the future (through the AGF). Both scenarios are shown below.



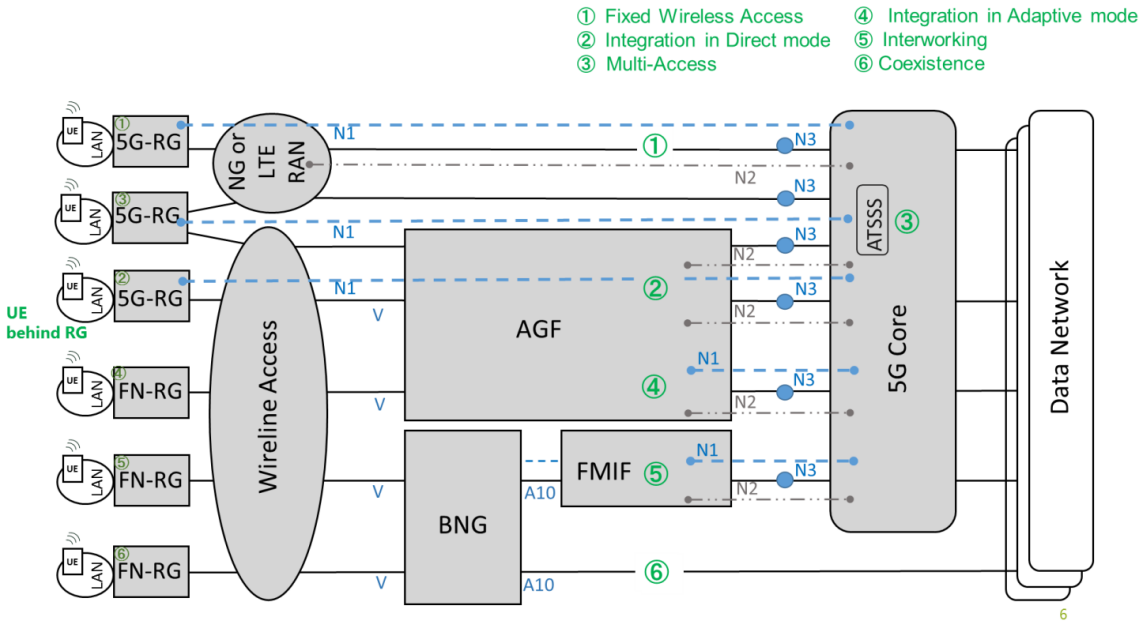
Source: 3GPP, "23.501 Rel 17: System architecture for the 5G System (5GS)"

Figure 17 – 5G Converged Core Architecture for FMC

While CableLabs was involved in the 3GPP Wireless Wireline Converged Core working group and published a technical report in relation to this (see reference [15]), the Broadband Forum (BBF) has made significantly more progress by providing FMC and AGF specifications for PON and DSL wireline access.

The requirements for initial convergence with existing RG's/Cable-modems should be similar to BBF: the integration of vCMTS functionality into a 5G-Core access gateway function (AGF) will be required to enable the use of the cable access network for FMC. At this point 5G and cable access networks become architecturally converged, the top level of convergence as shown in Figure 1.

Figure 17 below shows the five different FMC deployment scenarios/modes as defined by the BBF in TR-470.



Source: Broadband Forum, "TR-470: 5G Wireless Wireline Convergence Architecture"

Figure 18 – FMC Deployment Scenarios

The scenario most likely to be used for initial BBF convergence is **Integration in Adaptive Mode** with FN-RG (i.e. existing residential-gateways/modems). It is proposed that a similar approach also makes most sense for Cable FMC.

In this case the RG (aka modem) is connected over the wireline access network and the AGF mediates layer 2 traffic with the 5G core network based on N2 and N3 interfaces. However, FN-RG does not support N1, so the AGF acts as end point of N1 on behalf of the FN-RG. The AGF is said to integrate access sessions in "adaptive mode".

In the case of "adaptive mode" for cable access the AGF will perform the same functions as the vCMTS, on the Cable access network (aka CIN) side, while providing the following additional functionality to integrate with a 5G-Core on the other side:

- Control Plane function:
  - map wireline information into 5GC information
  - map 5GC information into wireline information

- proxy N1 and N2 on behalf of the FN-RG and generate all the relevant N1 / N2 signalling for an FN-RG that has been identified / connected via the wireline access network
- map data-plane L3 connections from the Cable access network to PDU sessions in the 5GC and provide relevant signalling on the N1, N2, N3 interfaces
- manage access specific resources based on 5G QoS profiles

- User Plane functions:
  - for Uplink: encapsulate incoming FN-RG traffic with a GTP-U Header (mapping to the appropriate GTP-U Session), apply an appropriate DSCP value to the outer packet and forward it to the UPF
  - for Downlink: decapsulate incoming UPF traffic by removing the GTP-U Header, mapping the incoming traffic to the appropriate RG service-flow (using the GTP-U header, QoS flow indicator (QFI) and DSCP value if provided) and forward the extracted packet to the target FN-RG

Below is a high-level view of an AGF for DOCSIS, with control and user plane split, connecting cable residential gateways to the 5G core through the cable access network.

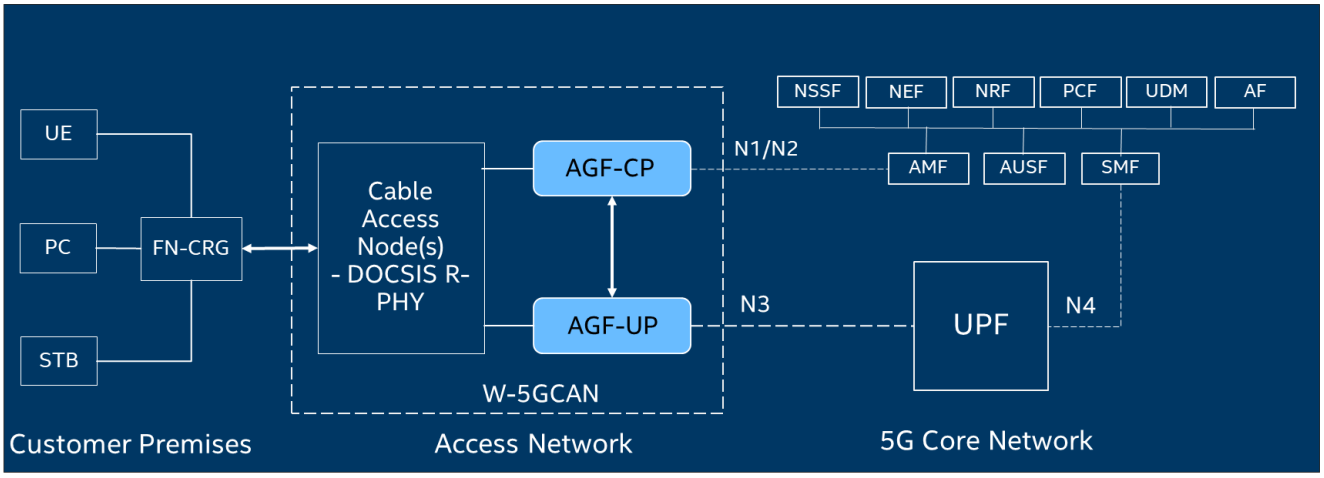


Figure 19 – Cable Residential Gateway Integration with 5G Core (in Adaptive Mode)

The existing software-based vCMTS network function will essentially need to re-purposed to become the vAGF network function as shown above. This will involve re-using vCMTS components such as DOCSIC MAC upstream and downstream data-planes, DOCSIS upstream scheduler and control-plane and integrating these with some new components for the 5G Core interface such as N2 control-plane, GTP-U encap/decap for 5G user-plane conversion and mapping of 5G QoS flows to DOCSIS service flows.

It is recommended that the Cable AGF adopt a control-plane/user-plane separation (CUPS) architecture which may require some re-architecting of an existing non-CUPS vCMTS architecture to deploy control-plane and data-plane functions in separate Kubernetes Pods on the same server or on separate servers in the same Kubernetes Cluster. Indeed, such an architecture has the added advantage of allowing control-plane and data-plane functions to be deployed at separate locations in the network. For example, it may be desirable to place the control-plane function in the cloud.

In terms of performance requirements for a Cable vAGF, data-plane throughput and latency are the critical KPI's that need to be considered. Upstream scheduling latency may be addressed in the same way

as for 5G xhaul as shown in the previous section. The impact of adding additional processing stages for 5G user-plane adaption to data-pane also needs to be considered for FMC.

Below are the additional packet processing stages that would be added to a standard vCMTS data-plane to re-purpose it to a Cable vAGF data-plane processing pipeline. The additional stages are highlighted.

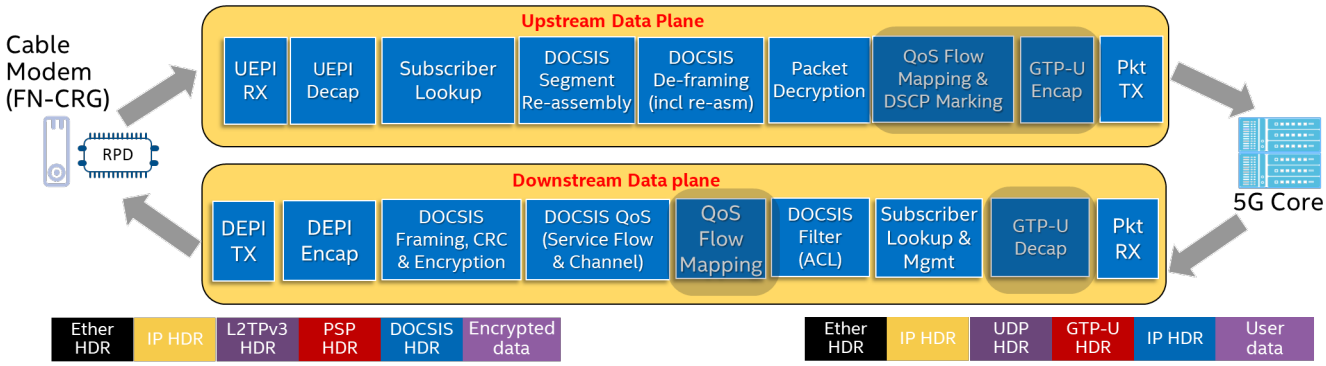


Figure 20 – Comparison of vCMTS and Cable vAGF data-pane processing pipeline

It is estimated that the additional processing shown above for 5G user-plane conversion adds 15 to 20% more CPU cycles per frame when compared to standard vCMTS data-plane processing. Given the performance capability of modern CPU architecture shown in Figure 6 and Figure 12 the required packet-processing performance for Cable access FMC is considered feasible in software.

It remains to be seen how broadly applicable AGF will be to cable operators; it may be a more desirable approach for mobile operators rather than cable operators. Furthermore, there are other viable approaches to fixed mobile convergence. For example, it could be said that MSOs such as Comcast have achieved this by deploying Wi-Fi Access Gateways in their cable network. This coupled with a high density of Wi-Fi hotspots allows subscribers to seamlessly switch between mobile and cable access networks.

However, it is clear that when starting with a cloud native vCMTS software architecture which is modular, extensible, and scalable, implementing FMC is greatly simplified versus starting with a proprietary hardware based solution.

As was shown for the xhaul case, software offers the flexibility to enable deployment at the pace of innovation and standardization of FMC while advancements in CPU architecture and the accompanying high-performance packet-processing frameworks ensure that throughput and latency requirements can also be satisfied. Such flexibility will be key to MSO's leveraging their Cable access networks to take advantage of advantage of such a major new use-case as 5G FMC.

5. Conclusion

As shown in this paper, the flexibility of a software based vCMTS makes it much easier to adapt for new 5G use-cases such as xhaul and FMC and to evolve over time.

It has also been shown that a software-based vCMTS is more than capable of meeting the throughput and latency requirements to transport 5G small-cell traffic by leveraging new low-latency DOCSIS methods. These should be used in conjunction with high-performance packet-processing frameworks such as DPDK and FD.io/VPP which are highly optimized for modern CPU architectures, which in turn are continuously improving gen-on-gen to meet ever-increasing network processing performance demands.

A software-centric architecture also has the benefit of leveraging open-source and the thousands of engineering person-hours of optimization and feature-development in mainstream projects such as Kubernetes, DPDK and FD.io/VPP. Additionally, there exist off the shelf tools for telemetry, management, orchestration, and AI which can be leveraged to improve network performance and resiliency.

Ultimately it is clear that starting with a modular, extensible, and scalable cloud-native software based vCMTS greatly reduces development effort and time-to-market for the two major new 5G use-cases of Xhaul and FMC as well as future-proofing for further standardization and innovation.

Abbreviations

3GPP	Third generation partnership project
5GC	5G core
BBU	Base band unit
BNG	Broadband network gateway
Bps	Bits per second
BWR	Bandwidth report
CM	Cable modem
CMTS	Cable modem termination system
CNF	Cloud native function
CO	Central office
COTS	Common-off-the-shelf
CP	Control plane
CPE	Customer premise equipment
CPRI	Common public radio interface
CU	Central unit
CUPS	Control and user plane separation
DAA	Distributed access architecture
DC	Data center
DL	Downlink
DOCSIS	Data over cable system interface specification
DP	Data plane
DPDK	Data plane development kit (open-source project)
DS	Downstream
DTP	DOCSIS Time Protocol
DU	Distributed unit
eCPRI	Enhanced CPRI
eMBB	Enhanced mobile broadband
eNB	eNodeB
EPC	Evolved packet core
FD.io/VPP	Fast Data I/O (open-source project)
FMA	Flexible MAC architecture
FMC	Fixed mobile convergence
FTTH	Fiber to the home
FWA	Fixed wireless access
gNB	gNodeB
HFC	Hybrid fiber-coaxial
LAN	Local area network
LLX	Low latency xhaul
LTE	Long term evolution
MIMO	Multiple in multiple out
MNO	Mobile network operator
ms	Millisecond
MSO	Multiple system operator
MVNO	Mobile virtual network operator
NFV	Network function virtualization
NGA	Next generation access

OLT	Optical line termination
Opex	Operating expense
O-RAN	Open RAN
O-RU	O-RAN RU
OS	Operating system
PON	Passive optical network
PGS	Proactive grant service
PTP	Precision time protocol
QoE	Quality of experience
QoS	Quality of service
RAN	Radio access network
RF	Radio frequency
RG	Residential gateway
RGW	Residential gateway
RMACPHY	Remote MAC and PHY
RMD	RMACPHY device
RPD	Remote PHY device
RPHY	Remote PHY
RTT	Round trip time
RU	Radio unit
SCTE	Society of Cable Telecommunications Engineers
SDN	Software defined network
SG	Service group
SP	Service provider
TCO	Total cost of ownership
UE	User equipment
UL	Uplink
UP	User plane
UPF	User plane function
US	Upstream
URLLC	Ultra-reliable and low-latency communications
vCCAP	Virtualized CCAP
vCMTS	Virtualized CMTS
vEPC	Virtualized EPC
VM	Virtual machine
VNF	Virtual network function
VPP	Vector packet processing (project)
vRAN	Virtualized RAN
WWC	Wireless and wireline convergence

Bibliography & References

[1] *Cable and Mobile Convergence: A Vision from the Cable Communities Around the World*, Cisco & CableLabs, https://www.nctatechnicalpapers.com/Paper/2020/2020-cable-and-mobile-convergence

[2] Data Plane Development Kit project, https://www.dpdk.org

[3] FD.io VPP project, https://fd.io/vppproject/vpptech

[4] *Container Bare-metal Reference Architecture,* Intel Corporation, https://networkbuilders.intel.com/solutionslibrary/container-bare-metal-for-2nd-3rd-generation-intel-xeon-scalable-processor

[5] *Maximizing vCMTS Data Plane Performance with 3rd Gen Intel® Xeon® Scalable Processor Architecture,* Intel Corporation, https://builders.intel.com/docs/networkbuilders/maximizing-vcmts-data-plane-performance-with-3rd-gen-intel-xeon-scalable-processor-architecture.pdf

[6] *Intel vCMTS Reference Data-plane project,* Intel Corporation, https://01.org/access-network-dataplanes

[7] *View on 5G Architecture*, 5G PPP Architecture Working Group, https://5g-ppp.eu/wp-content/uploads/2019/07/5G-PPP-5G-Architecture-White-Paper_v3.0_PublicConsultation.pdf

[8] *Low Latency Mobile Xhaul over DOCSIS Technology*, CM-SP-LLX, CableLabs, https://www.cablelabs.com/specifications/CM-SP-LLX

[9] *Low Latency DOCSIS: Technology Overview,* CableLabs, https://www.cablelabs.com/technologies/low-latency-docsis

[10] *Flexible MAC Architecture Specification,* CableLabs, https://www.cablelabs.com/specifications/CM-SP-FMA-SYS

[11] *Generic Access Platform(GAP),* SCTE, https://www.scte.org/generic-access-platform

[12] *Intel FlexRAN Project,* Intel Corporation, https://github.com/intel/FlexRAN

[13] *Can cable bridge the GAP?,* LightReading Article, https://www.lightreading.com/cable/can-cable-bridge-the-gap-/d/d-id/754620

[14] *What is network convergence and why do we need it?*, BT (IEEE BMSB 2018 Key-note), https://www.mcg.upv.es/wp-content/uploads/2018/06/IEEE_BMSB_2018_Keynote_Day2_MariaCuevas_Public.pdf

[15] *5G Wireless Wireline Converged Core Architecture Technical Report,* CableLabs https://www.cablelabs.com/specifications/WR-TR-5WWC-ARCH

[16] *MR-427: 5G Fixed-Mobile Convergence Marketing Report*, Broadband Forum (BBF) https://www.broadband-forum.org/download/MR-427.pdf

[17] 23.316 *Rel 17*: Wireless and wireline convergence access support for the 5G System, 3GPP https://www.3gpp.org/ftp/Specs/archive/23_series/23.316/23316-h00.zip

[18] *23.501 Rel 17: System architecture for the 5G System (5GS),* 3GPP, https://www.3gpp.org/ftp/Specs/archive/23_series/23.501/23501-h00.zip

[19] *TR-470: 5G Wireless Wireline Convergence Architecture*, Broadband Forum (BBF), https://www.broadband-forum.org/technical/download/TR-470.pdf

How Working and Schooling from Home has now Driven a Change in How we View Home Connectivity and Networking

What We Did During the Pandemic and How it Could Define New Products and Services

A Technical Paper prepared for SCTE by

Charles Cheevers
CTO Home Networks
CommScope
3871 Lakefield Dr, Suwanee, GA 30024
678-473-8507
Charles.Cheevers@CommScope.com

1. Introduction

During the last 18 months many things have changed for many people. One indelible change seems to be how important our home is for working, schooling, and staying connected. While we cannot fully predict what the real ramifications of virtual working from home , schooling from home are going to be – there are many predictions of significant changes in our workforce locations from new permanent at home work increases to flexible working being introduced more by employers. As well as the work from home changes there are likely to be other permanent changes in how we view and use connectivity and digital connection from our home.

- Our consumption and use of OTT and IP video services has increased and adoption of its use by even more 'traditional telly' focused consumers has increased
- The number of connected IoT devices we have has increased – because of more online consumer adoption. The current silicon shortages have been the result in this rapid change in consumer purchasing behavior of electronic devices for the home.
- The rapid acceleration of adoption of Telemedicine services including virtual doctor visits that is likely to change the way we engage with our medical resources.

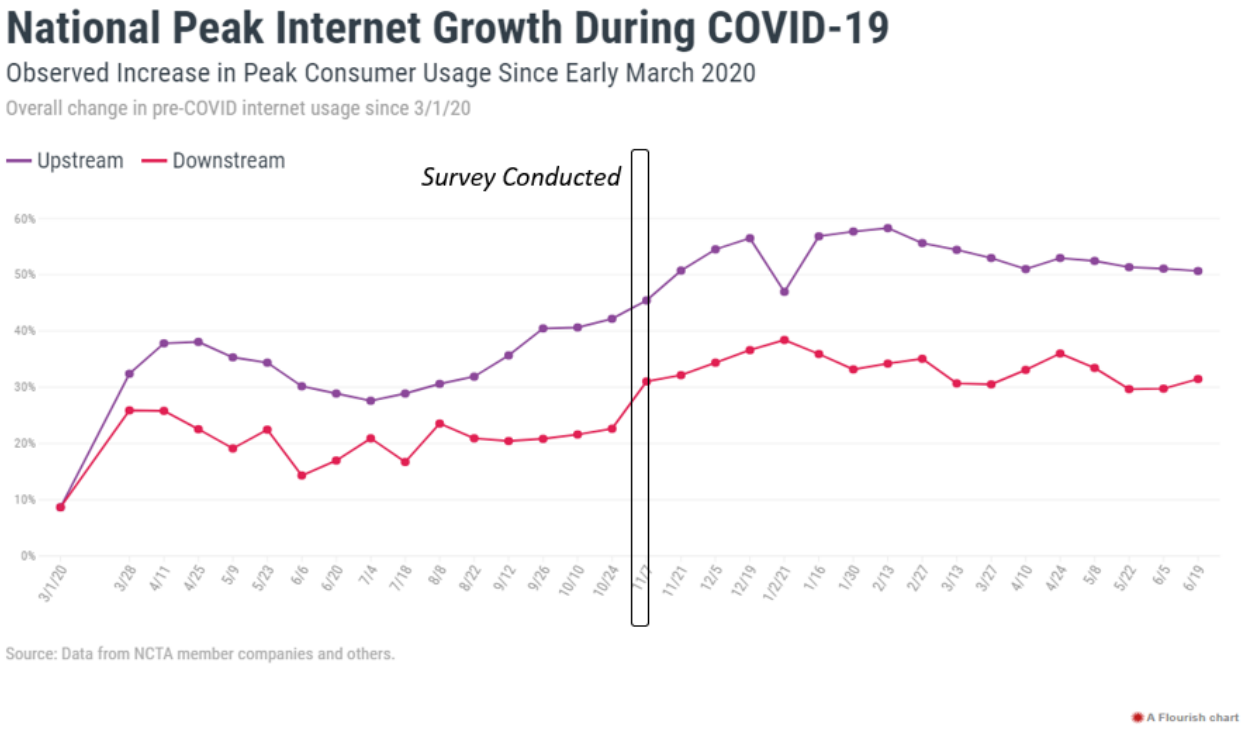
Because of this step function and acceleration in behavior change driven by being at home for 12-18 months there is now also likely potential for new and improved Service Provider services that offer to cater to raised awareness and desire from Consumers to improve elements of their connected digital home. In particular

- Reliability of home network connection
- Latency of home network connection
- Improved performance of Wi-Fi
- Separation of work environment at home from residential environments

And potential opportunity for Service providers to bring the Enterprise, School or Doctors office to the home via new services, devices, and technologies.

- WFH and Corporate services to the home
- Telemedicine and Aging in Place services
- Increased Privacy and Security
- Increased Reliability

This paper is the result of a poll of ~2,200 CommScope employees during the height of the pandemic last year in November 2020 when according to the NCTA Covid-19 dashboard – home downstream/upstream connectivity growth had not yet hit its peak (which seemed to happen during Christmas/Holiday Season end of December and early January 2021). The poll sought to understand the connectivity drivers and conditions of consumers during the pandemic and their robustness of connectivity and performance ratings. The paper will use some of this feedback to try and put the case forward that there are potentially new requirements for the Broadband delivered home. Based on the survey analysis it will then try and propose the potential for new pandemic home architecture and services potential and outline some of the new device and system elements that could play a new role for Cable Operators to continue to own the digital connectivity and services for their consumers.



2. November 2020 Survey of Working and Schooling at home

Given the drastic change in all our lives, 2020 presented a unique opportunity to survey people on their new environments of lockdown and working, schooling and effectively trying to do everything from home and on-line. This change of

- Number of end points on Broadband networks concurrently active
- Number of concurrent active devices in the home
- Frequency and time of use of Broadband
- Increase in upstream traffic from Video and Audio conferences
- Number of IP connected devices in the home
- Rooms and new locations in home being used for high value and high capacity services like laptop driven Video conferencing

has been described as almost 2 years of growth of capacity and services in a 6 month period. It did however accelerate how people create their home connectivity platform, resolve their niggling issues that were tolerable when they only had entertainment services primarily at home, improve range of connectivity as new rooms were occupied with WFH and Schooling activities.

Consumers views on reliability, camera resolutions, microphone quality and latency were all sensitized more because of the higher performance of being able to work effectively and to not be that person upsetting everyone else on Video conferences. Because of these changes in our home environment the thesis put forward was that the existing set of Broadband, Video and IoT services may

- Not be good enough to cope with the demands for service the pandemic has highlighted

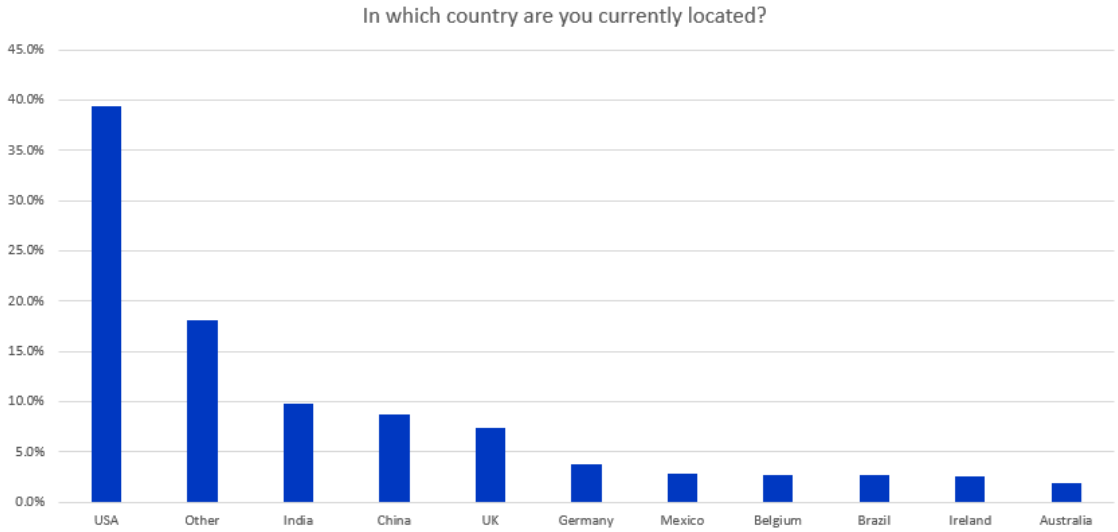
- Offer opportunity to improve services with new features that may either be sticky for consumer retention or drive new business opportunities
- Be driven by acceleration of new technologies like products developed using new 6GHz spectrum allocations, improvements in IoT like the new Matter solution.
- Be driven by price decreases in technologies like Cameras, Smart Assistants, Microphones, IoT devices which make them more accessible or potential for Operators to bundle into traditional broadband and Video services
- Be driven by acceleration of areas like Telemedicine which even as people resume some normality has reached a point in areas like Virtual Doctor visits where this has become more acceptable and the benefits of not having to wait in Doctors office for minor issues may move this faster to mass adoption
- Be driven by our adoption and acceptance of Video conferencing as a more normal and regular means of connecting with people generally and from the home
- Be driven by new niche areas that have been highlighted as capable of being addressed with technology solutions like Aging in Place – where during the pandemic innovative ways had to be created to keep elderly loved one's safe and connected

With the goal of trying to open up a discussion on new opportunities to capitalize on this technology acceleration and mind shift in consumers having changed habits pre-pandemic to more use of digital services this paper will first explore how over 2,000 consumers in different countries and with different

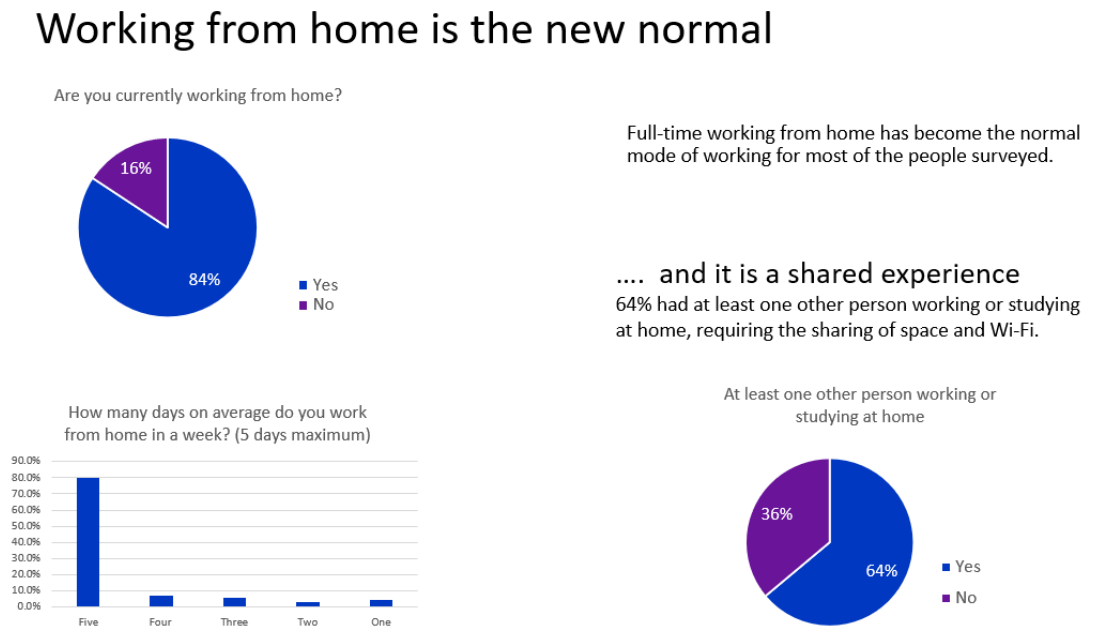
- Service Providers
- Home Sizes and constructions
- Home Family and occupancy differences
- Demographics

all coped and shared their insights to a series of questions that were asked to gage what they valued and what they will value going forward to ensure their homes are not just for after 5pm during the week for digital excellence.

Almost 40%/~1,000 of our respondents were from the US with the following spread of other non-domestic respondents in the table below. For the sake of this paper the answers to the questions are presented in a blended fashion as generally the differences created by the pandemic and the needs were similar in all countries.



The sample of respondents had 84% working from home. 80% of these were working full time (5 days) from home. Of course, the key issue with the working from home experience is that many homes during the pandemic had multiple people trying to work and school from home. In our sample 64% of respondents had others working or schooling with them while 36% were on their own.



Exploring the respondents working conditions in more detail, 23% of them had to share a working space in the home. Over 40% had 2 people in a room and 10% had 3 people trying to work and school in a

single room. While its is expected that many people will start moving back to their workplaces towards the end of 2021 many companies are now trying to find the best balance to

- Investment in Office real estate
- Creating collaboration face to face for innovation
- Maintaining flexibility for the employees to work more from home
- Allowing employees to work from home more permanently where it has made sense

To create and sell services to the needs of the height of the pandemic is the big question that is still hard to answer. However, there is a much more acute awareness of the reliability and performance of the home environment for connection to the internet and corporate office that is likely to remain, particularly for those employees who will remain working flexible conditions from home. As of the time writing this paper it does look like schooling will resume for the majority back in the classrooms at the start of the 2021/2022 school calendar. The worry about new developments in strains of the Covid-19 virus will keep consumers vigilant on their home connectivity performance but consumers will always balance their investment to their budget and additional performance/reliability features offered may not be invested in ahead of need.

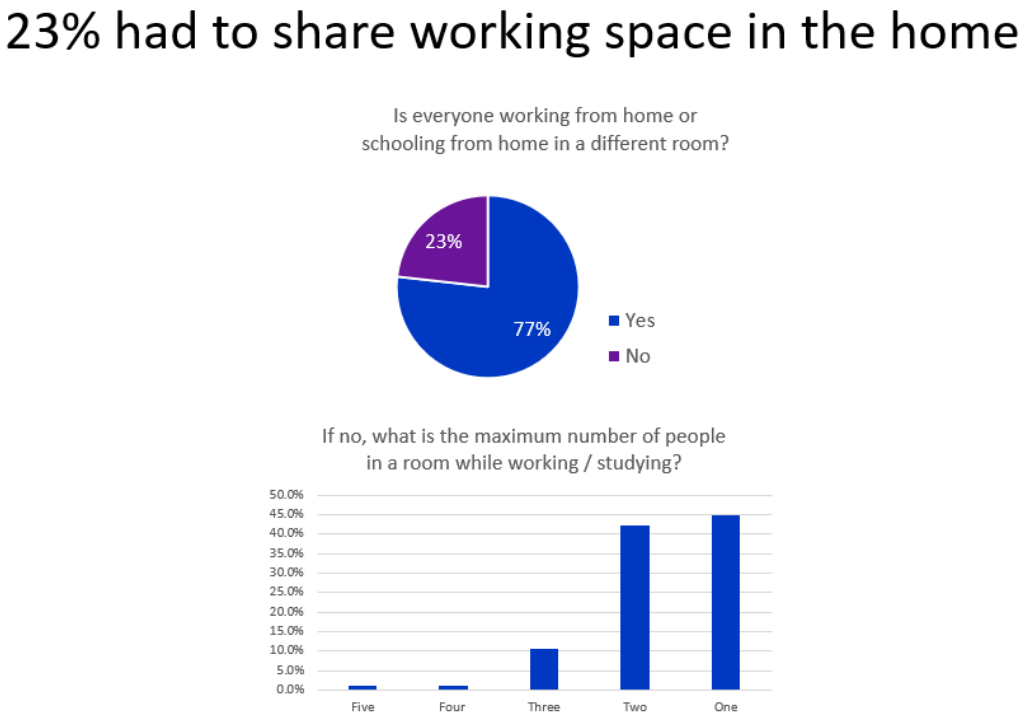


Figure 4 How many people had to share a room to work and school

Of those respondents working from home 42% of them had another person working with them and typically working 5 days per week. This has been a key stress point for both the psychological wellness of the workers from home as well as their connectivity performance.

Phycological wellness has been challenging for many households

- Parents working from home and schooling and minding children

- Multiple people trying to work simultaneously in small homes
- Audio and background interference during working and schooling sessions
- Problematic connectivity on important schooling or work video or audio conferences
- Additional group meetings to stay connected

Connectivity performance and its issues was also higher and more exacerbated in homes with multiple people working and schooling at the same time. The stress of problematic calls and connectivity would generate household questions like

- Was someone watching streamed video content while others where trying to work – and it is causing problems
- Could someone in the house drop their sessions to see would it help others
- Could someone having problems get closer to the Wi-Fi router and work in that room for a while
- Could everyone stop what they were doing so Gateway or Access Point could be rebooted to see does it resolve the problem
- Trying to resolve an outage with your Service Provider in time for that important call
- Giving schooling children priority during exams even at the expense of issues with Work tasks and meetings

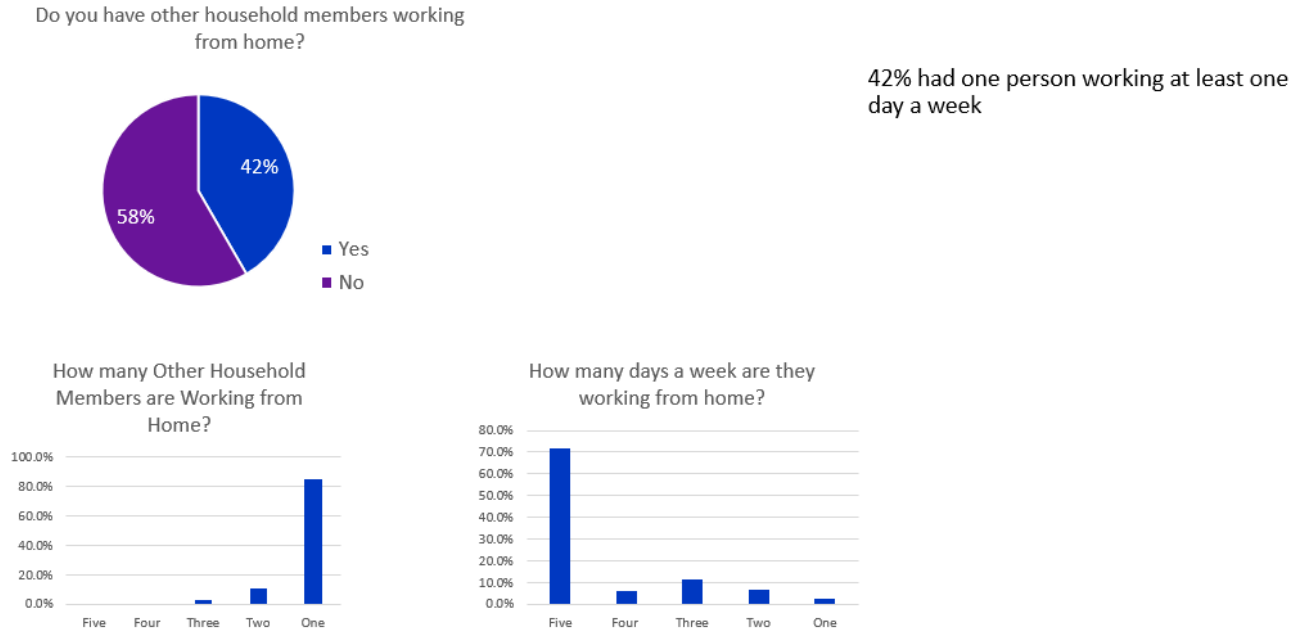


Figure 5 People working at home - did they have others working with them

Schooling from home was also important during the pandemic. In this survey 43% of the respondents had at least one child schooling at home with them. Over 50% of students were doing one day at home but most children were still attending school for the most part of the week. It is expected that most children or University students will go back to in school tuition in the coming 2021/2022 year. However, there is probably a marked rise in future homeschooling which in the US has typically been about 3% of children pre-pandemic but set to rise by potentially 5% based on effects of pandemic

(https://ifstudies.org/blog/homeschooling-is-up-thanks-to-covid-but-will-the-increase-last , https://www.census.gov/library/stories/2021/03/homeschooling-on-the-rise-during-covid-19-pandemic.html?utm_campaign=20210322msacos1ccstors&utm_medium=email&utm_source=govdelivery).

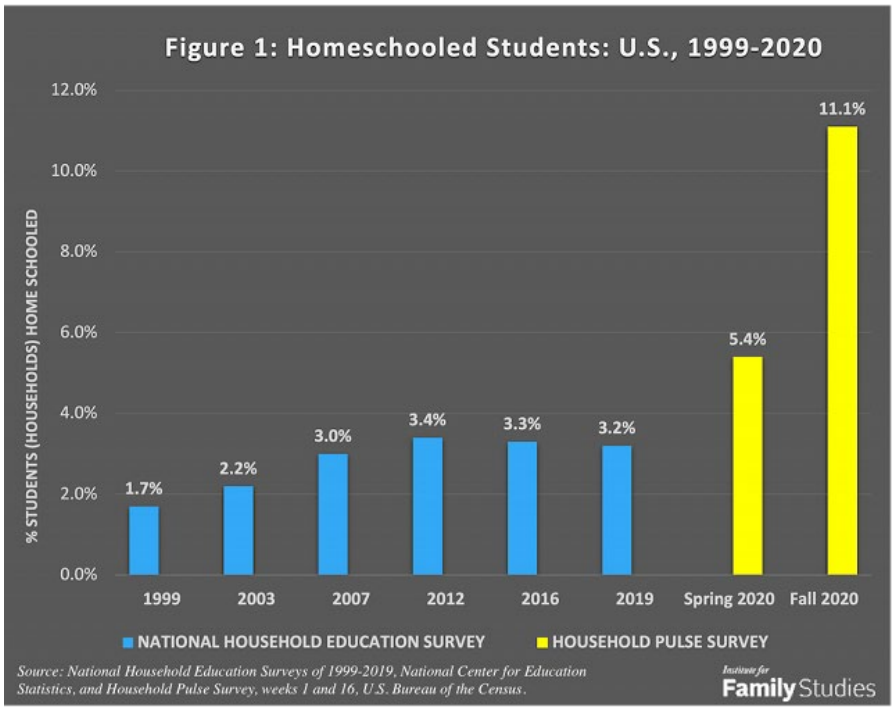


Figure 6 Homeschooled Students : US 1999-2020

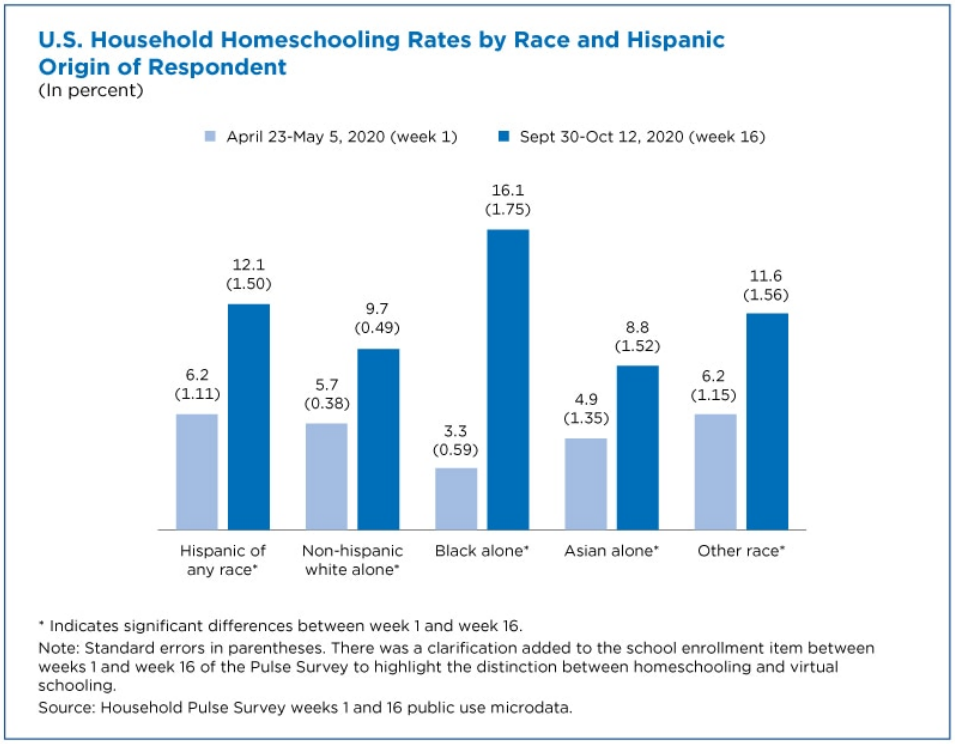


Figure 7 US Household Homeschooling rates

43% had children doing online schooling at home

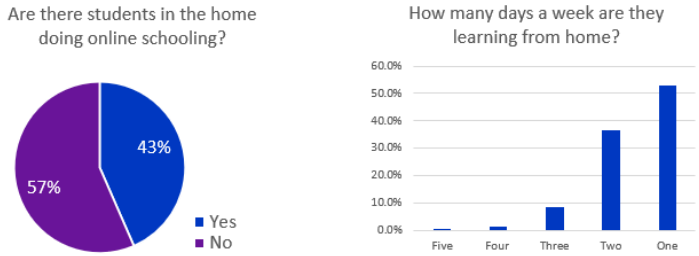


Figure 8 How many days were Children spending schooling at home

During this change in the home environment the importance of being able to work and school effectively raised in performance. There was pressure from several factors

- The ergonomics of a desk environment and replicating our office setup at home.
- Connectivity in particular Wi-Fi performance improvements
    - Bedrooms and every room being used for concurrent family member use exposing issues with Wi-Fi performance at range
- Audio performance – multiple people working in single room and generally trying not to allow the normal noises of home life interrupt work and school times.

- Introduction of Video Conferencing as the new pervasive tool to try and fill the gap for in person meetings and its performance affect on the upstream bandwidth from the home. Video conferencing applications were not as efficient in adaptive bit rate technologies and typically wanted to grab 1.5Mbps to 2.5Mbps to stream video upstream.

Looking at the way the survey respondents spent their money on to improve their new home environment, we see that it covered a number of areas in particular

- 55% of respondents bought desks.
- 40% of respondents improved their home monitors to augment their laptop screens
- Almost 30% of respondents bought separate microphone solutions to improve their audio performance
- Almost 20% of respondents bought new webcams to improve their Video Conference performance
- Over 40% of respondents bought Wi-Fi AP or Wi-Fi mesh solutions

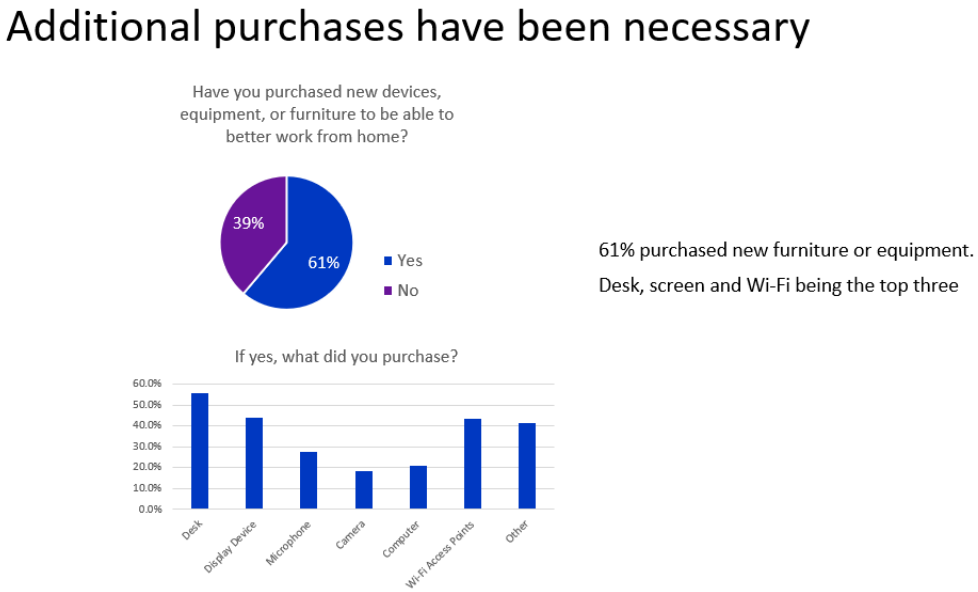


Figure 9 What items did WFH/SFH families purchase during the pandemic

With connectivity the main requirement of a stay at home solution the survey looked to see how respondent's connectivity performance catered to their work and school needs.

For Broadband speed generally 83% of respondents felt that the were on the right and acceptable broadband tier/speed for their requirements. The survey checked what access network the respondents were using with 30% of them on Cable, Fiber and FWA respectively. The ISP's the respondents were using was recorded and showed that 8 ISP's registered above 5% for our survey sample. Comcast, Charter, AT&T and Verizon served about 37% of our respondents. Internationally, China Telcom, Vodafone, Virgin Media, and Airtel also registered covering 5% of our respondents.

The survey also used the opportunity to see what percentage of homes had an MMP (Massive Multiplayer Gamer) in the house and 20% of respondents were gamer homes. As 10%+ of all downstream traffic now is gaming this is a growing segment of traffic that is important to understand for performance implications on home connectivity.

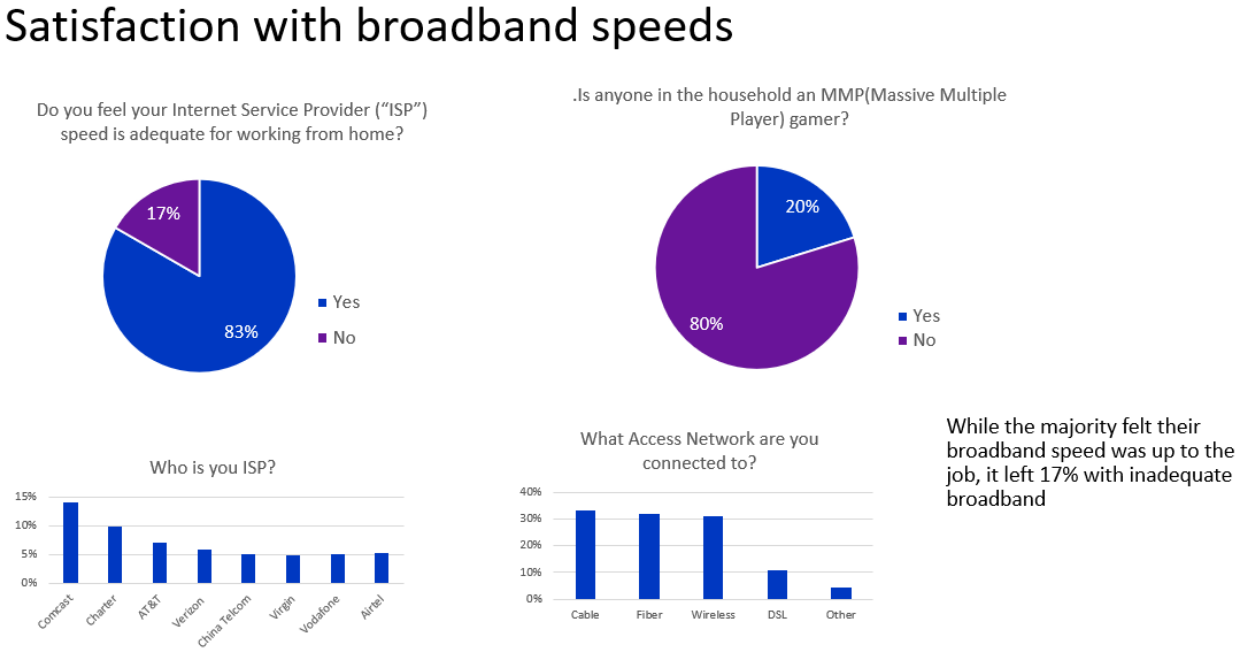


Figure 10 Consumer satisfaction with Broadband speed during pandemic

The survey further asked respondents to check their broadband speed both up and down and showed the following results.

Download speeds

- 25% of respondents were over 200Mbps
  - 7% of these said it was inadequate
- 17.5% of respondents were 100-200Mbps
  - 11% of these said it was inadequate
- 20% of respondents were 51-100Mbps
  - 19% of these said it was inadequate
- 21% of respondents were 21-50Mbps
  - 23% of these said it was inadequate
- And the remaining respondents 16.5% < 20Mbps
  - 40% of these said this was inadequate

Upload speed performance and robustness was the most important metric during the pandemic. The survey respondents answered for their homes that

- 4% of respondents had > 200Mbs Upstream performance (100% of these on Fiber networks)
  - 6% of these said this was inadequate
- 7% of respondents had > 101-200Mbps Upstream performance (100% of these on Fiber networks)
  - 4% of these said this was inadequate
- 10% of respondents had > 51-100Mbps Upstream performance
  - 10% of these said this was inadequate
- 30% of respondents had > 21-50Mbps Upstream performance
  - 10% of these said this was inadequate

- 16% of respondents had < 20Mbps Upstream performance
    - 70% of these said this was inadequate

Generally, it looks like respondents were happy with their broadband speed but room for improvement. It is the cost and hassle factors of upgrading or changing service provider for the return which will drive them to change. As the reader will see later in this paper the survey revisited this a few times as well as introduced reliability and latency questions to the overall view of Broadband performance.

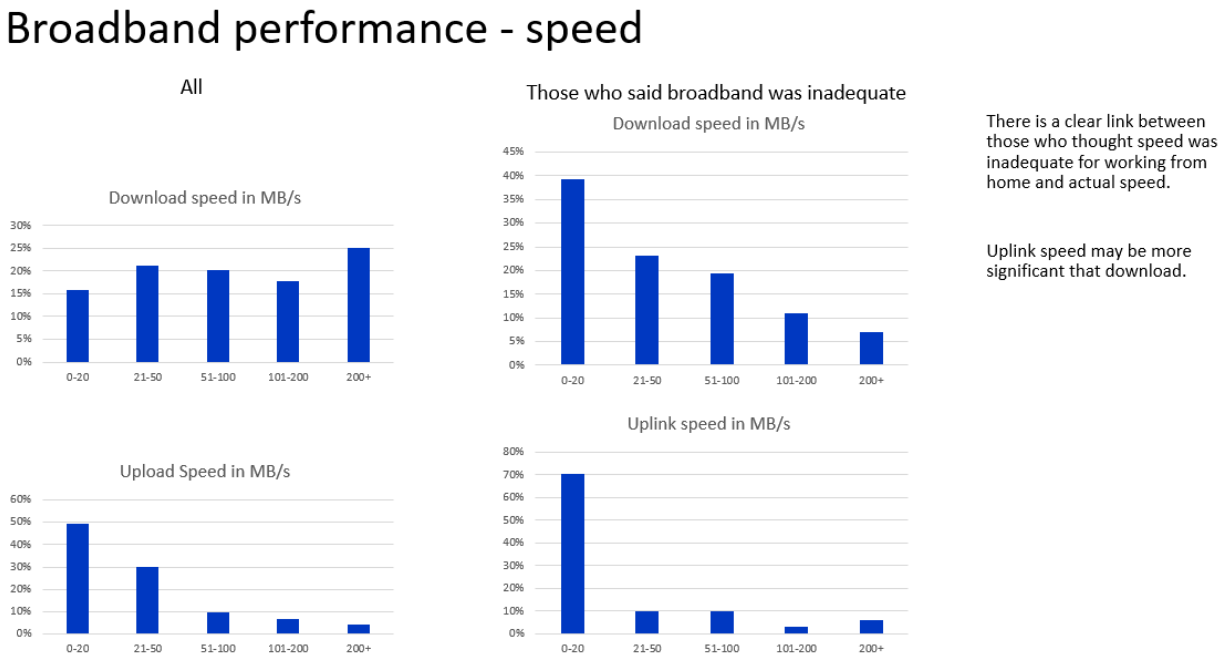


Figure 11 Consumers satisfaction with their Downstream and Upstream performance

As Latency is becoming more important as a performance metric for Audio and Video conferencing, Gaming, and more immersive video experiences like AR/VR – the survey also checked with respondents on their latency performance. Consumers are still not aware of what latency really is or what jitter is and still really view their Broadband quality as their overall speed. They are getting more cognizant of the difference between ISP WAN speed and their Wi-Fi speeds to devices, but latency is still something they don't fully understand. Homes that have gamers are more educated on the value of latency as there is a very specific correlation between latency (or probably more relevantly jitter) in the performance of gaming. With the pandemic consumers were introduced to more latency sensitive applications like Audio and Video Conferencing with echo and problematic video conferencing often related to issues with latency. The survey asked the respondents to use the Netflix fast.com web based speed test which ran loaded and unloaded latency tests (loaded have a traffic mix defined here https://netflixtechblog.com/building-fast-com-4857fe0f8adb). The difference between the 2 numbers is defined as buffer bloat and can certainly affect very latency sensitive applications.

A quick note on the importance of consistent jitter to a gamer versus latency. Gamers seems to be happy with any latency that is less than 55ms (55ping as it is referred to by gamers). What they really value then is having not more than 5ms of jitter within their latency value. This means that it is not good for a gamer

to have 50ms of latency with 10ms of jitter or reducing latency to 30ms but jitter inconsistent across packet arrivals.

The results were interesting with a diverse spread of latency

- 42.5% of respondents had unloaded latency < 20ms
    - 50% of these felt their broadband was inadequate
- 28% of respondents had unloaded latency between 21ms-50ms
    - 30% of these felt their broadband was inadequate
- 15% of respondents had unloaded latency between 51ms-100ms
    - 9.5% of these felt their broadband was inadequate
- 6% of respondents had unloaded latency between 101ms-200ms
    - 7% of these felt their broadband was inadequate
- 7.5% of respondents had unloaded latency > 200ms
    - 4% of these felt their broadband was inadequate

For loaded latency result

- 28% of respondents had loaded latency < 20ms
    - 28% of these felt their broadband was inadequate
- 18% of respondents had loaded latency between 21ms-50ms
    - 23% of these felt their broadband was inadequate
- 23% of respondents had loaded latency between 51ms-100ms
    - 20% of these felt their broadband was inadequate
- 12% of respondents had loaded latency between 101ms-200ms
    - 14% of these felt their broadband was inadequate
- 17% of respondents had loaded latency > 200ms
    - 12% of these felt their broadband was inadequate

These latency results and especially the correlation to broadband inadequacy is has some ambiguity as one would expect the higher latency issues to have increases in dissatisfaction. The comments on the respondent's survey seems to suggest that latency is not well understood by people generally. Those that are gamers or more technical understood latency well enough to in many cases still not be happy with 15ms latency results as they strove for higher performance. Those with > 100ms of latency seemed to also not really be interested in spending money on improving things.

There is also the reliability of the fast.com test and its repeatability and temporal performance levels based on how congested the networks were at time of test as well as what device the test was done.

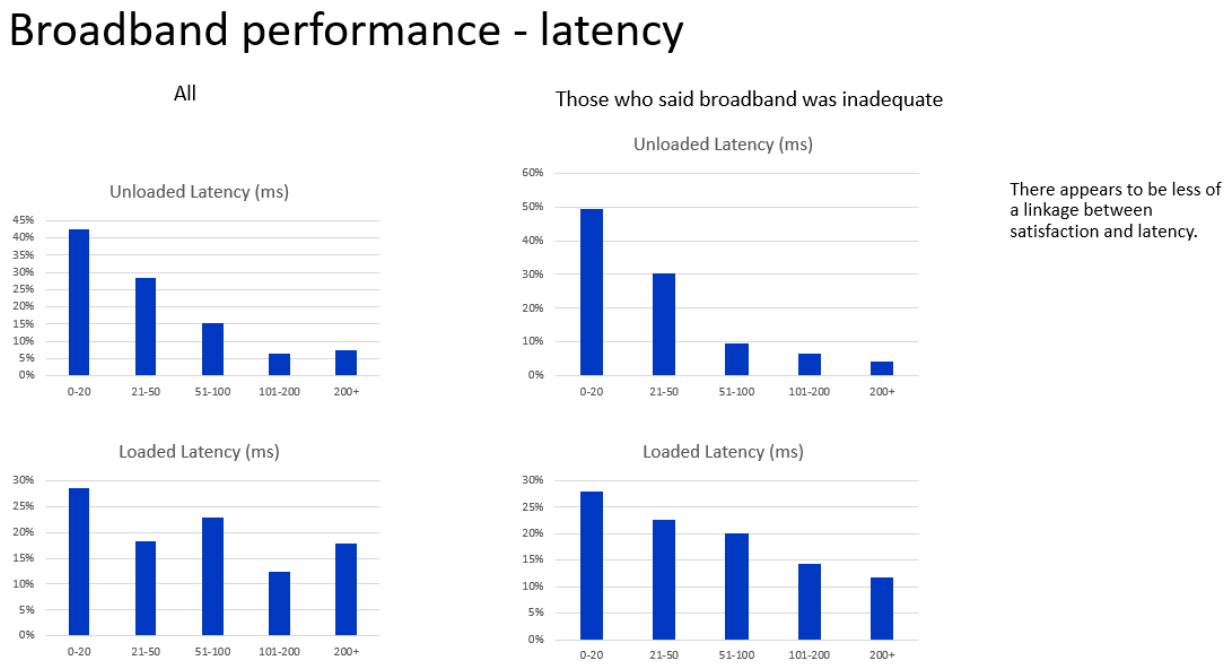


Figure 12 Latency performance of our respondents.

The respondents to rate their ISP and then this was correlated to each Network access type

- 85% of Cable ISP respondents were happy with performance
- 77% of DSL ISP respondents were happy with performance
- 90% of Fiber ISP respondents were happy with performance
- 76% of FWA ISP respondents were happy with performance

Network types

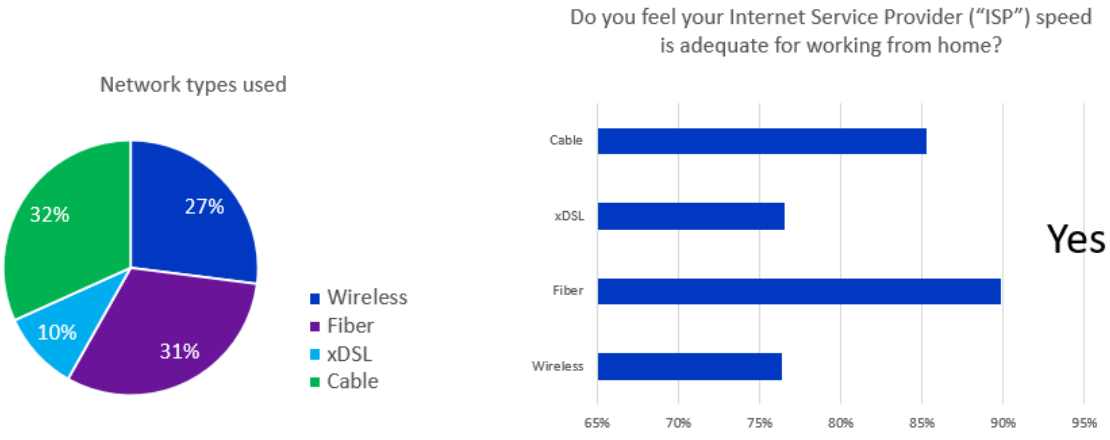


Figure 13 Respondent satisfaction with their ISP by Access Type

As Wi-Fi performance was so critical to everyone at home during 2020 the respondents were asked to explain their Wi-Fi solution and if they had added extenders.

It was interesting that

- 40% of respondents had an AP that was integrated into the SP's GW device
- 49% of respondents had bought their own retail Access Points
- 11% of respondents had an AP issued by their SP

Delving further into the Wi-Fi solution architecture of the respondents

- 31% of our respondents had a Wi-Fi extender/mesh solution with 69% relying in a single AP

Of the 31% of respondents with a Wi-Fi extender solution

- 50% of respondents had 1 extender
- 31% of respondents had 2 extenders
- 14% of respondents had 3 extenders
- 5% of respondents had > 3 extenders

This seemed to show clearly that respondents had to augment their Wi-Fi solution to cater with the increased demand of working and schooling from home at range. The survey may have some bias to more consumer owned devices versus using the supplied SP devices as it was based on employees from CommScope who do skew to more technically aware for a larger base of the sample. However, the majority of respondents were not engineers by job function.

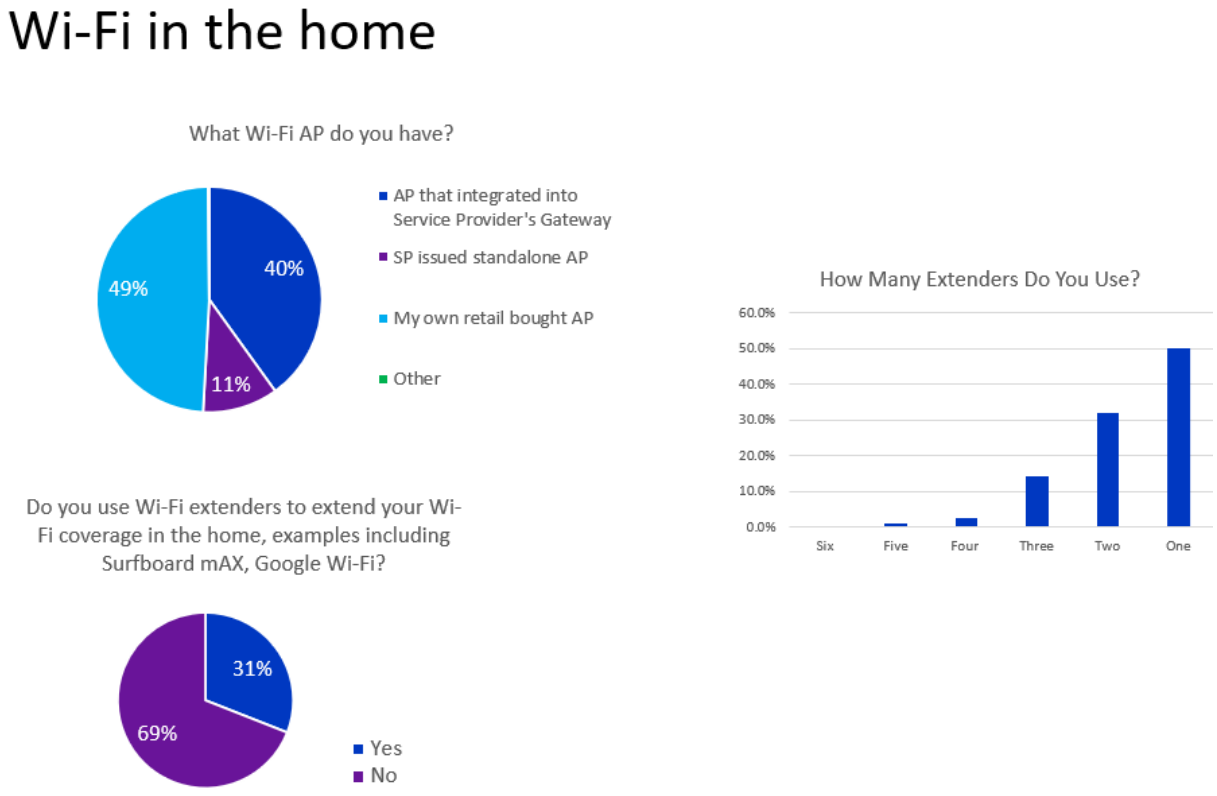


Figure 14 Wi-Fi device use and extenders in our survey

The survey respondents were asked what was their primary device for Working from home

- 95% of respondents said it was company issued laptop, tablet or phone
    - The remaining 5% used Home PC/Laptop or Apple Product

Of the additional devices that respondents used to work from home

- 66% of respondents used a second screen
- 42% of respondents used headset for audio/video conferencing
- 20% of respondents had a dedicated work printer
- 27% of respondents set up a separate webcam for video conferencing

This was to be expected as many consumers during pandemic tried to ensure their home was as comfortable as possible to work and school. Because of security requirements and that almost all of the respondents have office jobs that have company supplied laptop or desktop device – the company issued laptop was the dominant WFH device for the respondents surveyed.

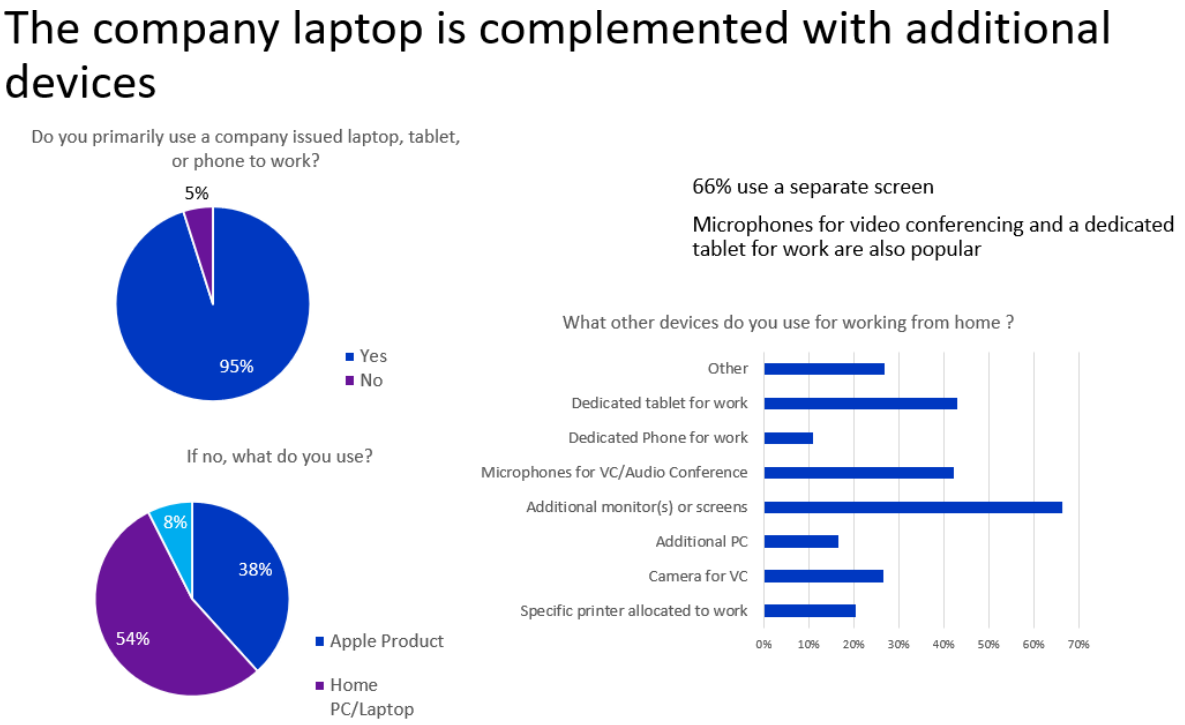


Figure 15 The devices people used at home to Work from Home

The survey asked the respondents if they used VPN to access work resources.

- 97% of respondents were using VPN for work related traffic. (There are some CommScope services that require VPN)

One of the issues at the outset of the global WFH change was the lack of VPN resources scaling and the overhead in speed/performance of VPN. In the survey we checked with the respondents the reliability as well as performance of the VPN

- 30% of the respondents answered they kept the VPN on all the time
    - 15% of respondents said they never turned it off
- 24% of the respondents frequently turned it off

- 30% of the respondents said they occasionally turned off the VPN

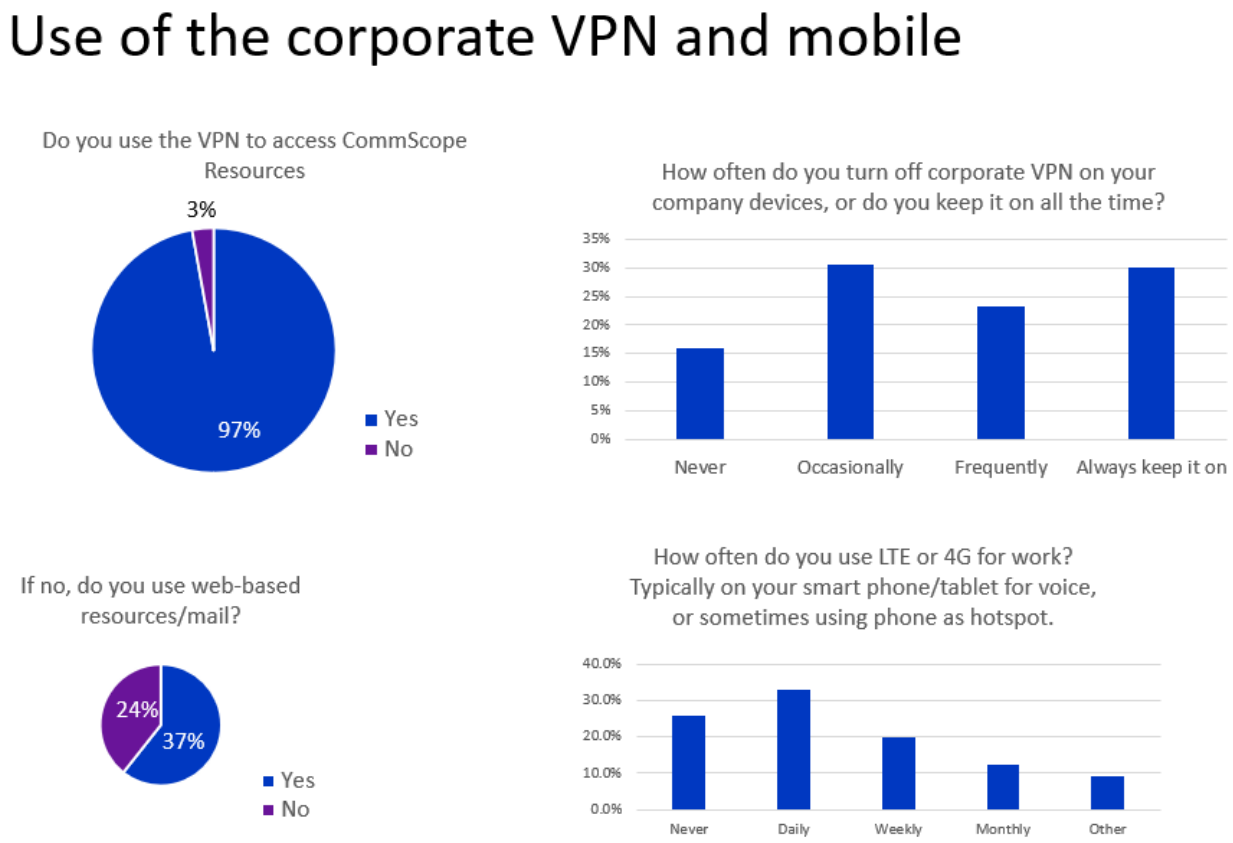


Figure 16 Use of corporate VPN and Mobile device

The survey asked the respondents the direct question of their view of their connectivity issues while working from home

- 31% of respondents said they had connectivity issues
- 28% of respondents identified issues related to using VPN services

The respondents identified the source of the connectivity issues they had

- 59% of respondents said their connectivity issues were the fault of the SP or speeds
- 48% of respondents said the issues were related to Wi-Fi performance
- 50% of respondents said issues were related to VPN performance and overheads
- 28% of respondents said issues were related to Video Conferencing application

Additionally the survey checked with the respondents how often the chose to call IT for support when they had issues

- 56% of respondents never called for help
- 42% of respondents called IT 1-4 times during first 7 months of working from home

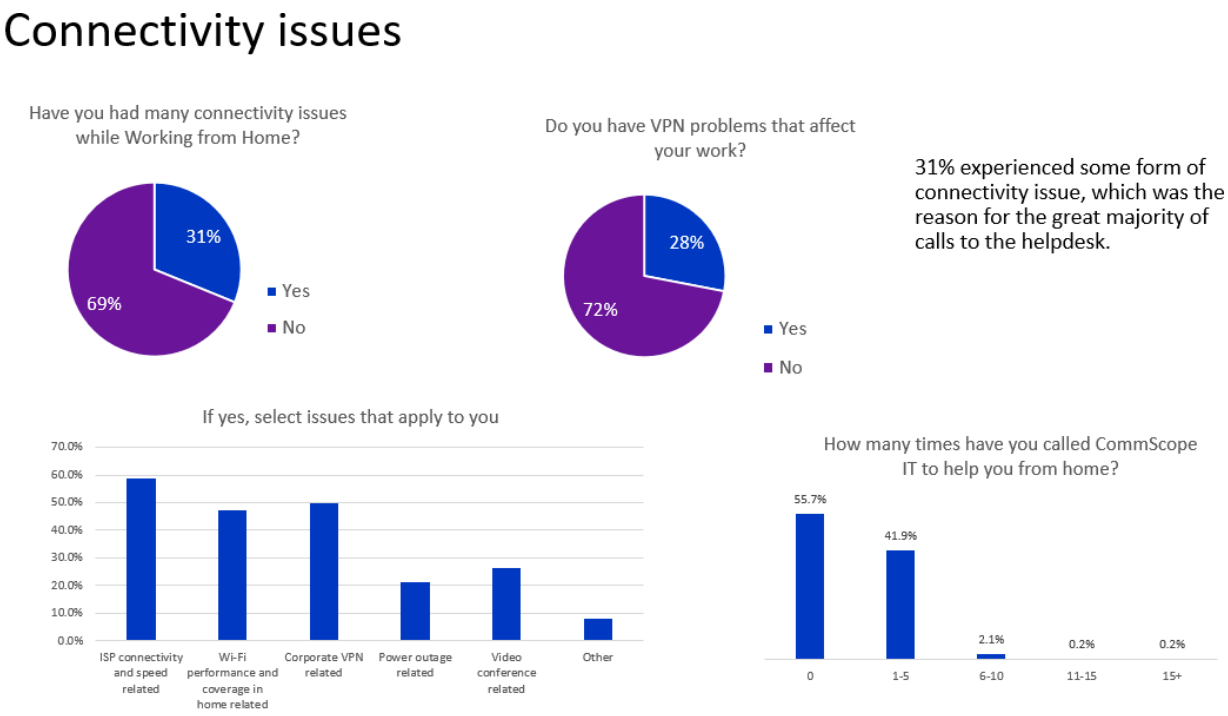


Figure 17 Frequency of Connectivity issues WFH

Being specific again on Low Latency the survey explicitly asked the respondents how they valued low latency

- 42% of respondents valued low latency working from home and for services like gaming

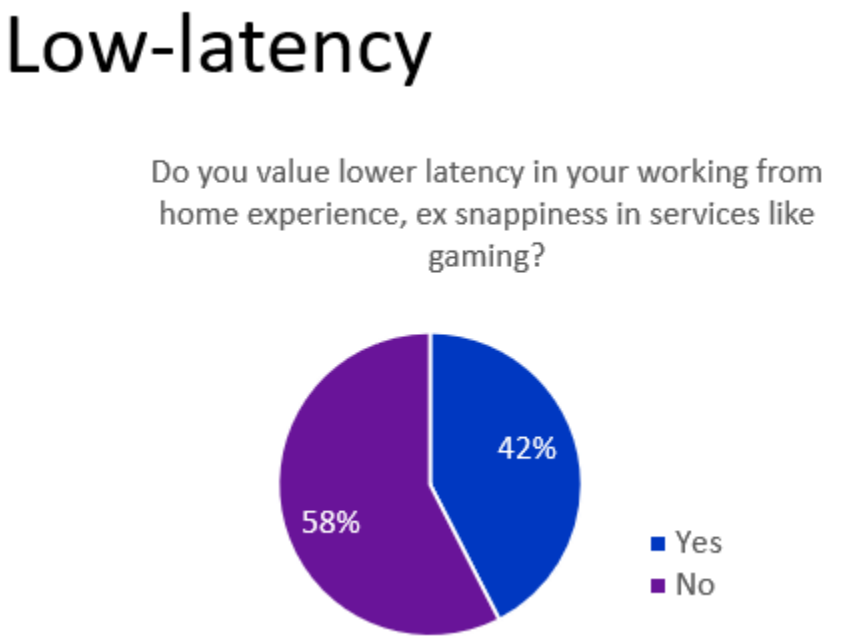


Figure 18 How do we value Latency working from home

Going a bit deeper on the Audio and Video Conferencing side of Working from Home the survey asked several questions

- 69% of respondents used a headset when doing Audio or Video conferencing
    - 40% were wireless; 60% were wired headphones
- 86% of respondents primarily used the camera in their laptop or PC for Video conferencing

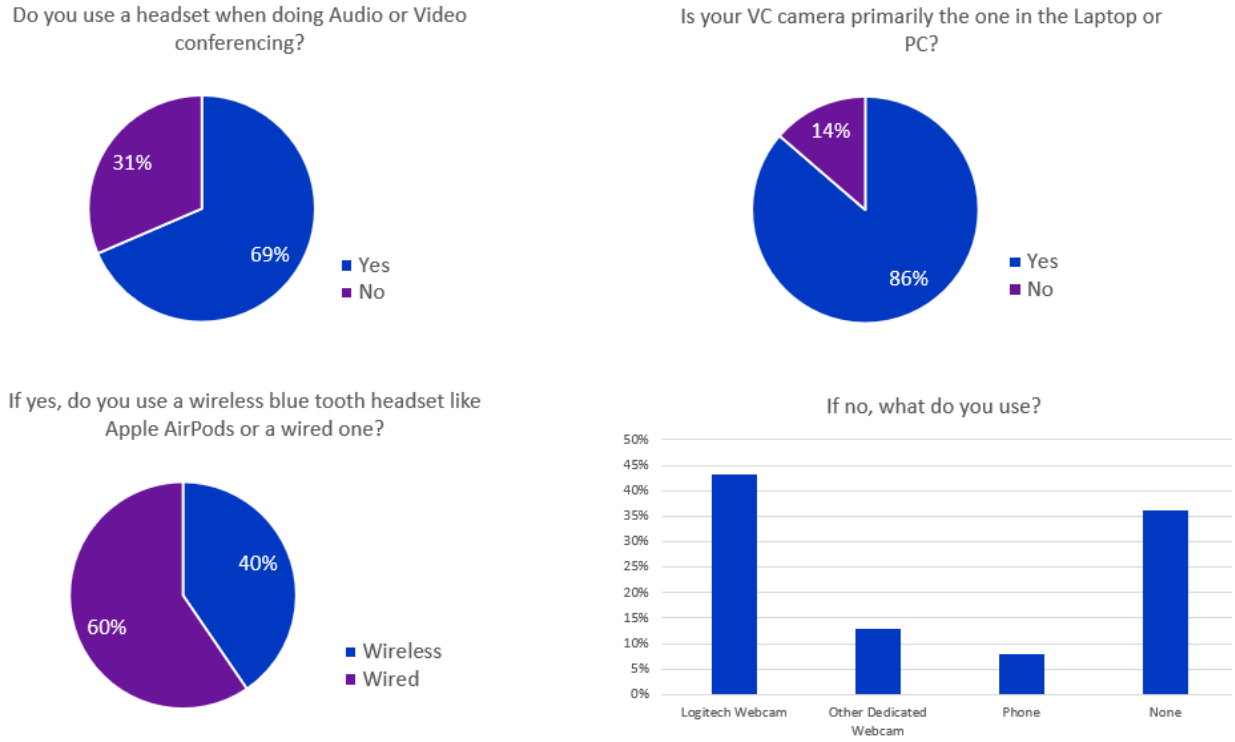


Figure 19 Audio and Video Conferencing for WFH

With the potential of using existing home screens like TV's to help with Working from home the survey asked respondents if they would value using the TV for work video conferencing – only 34% of people felt this was something they would use.

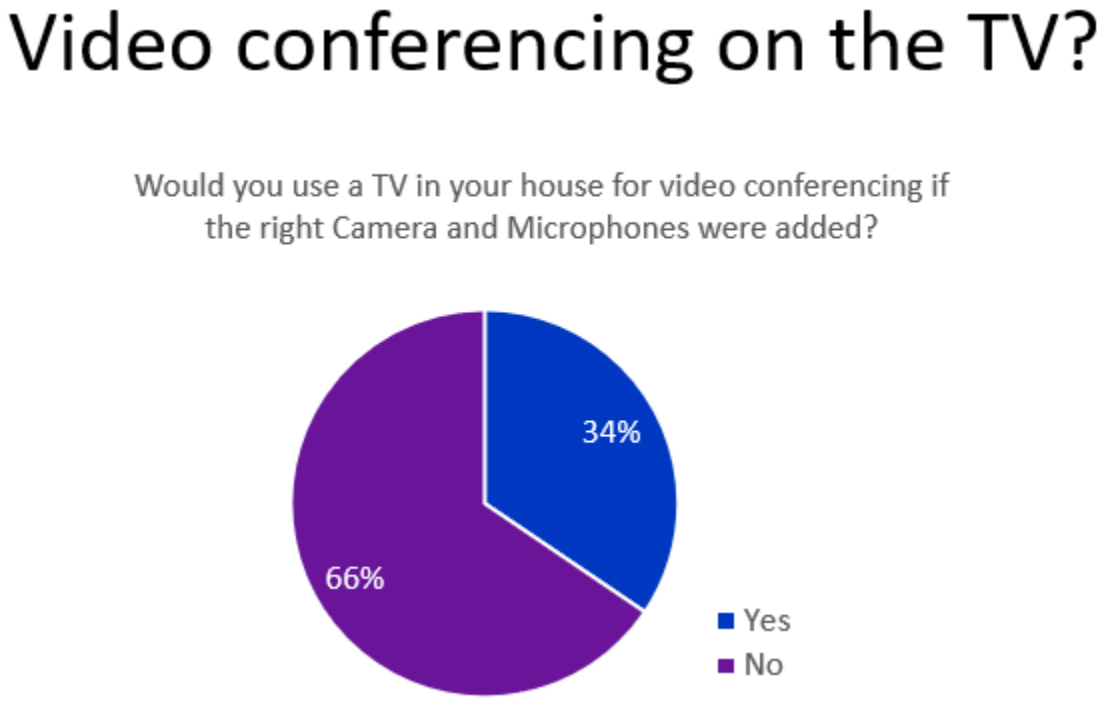


Figure 20 Would you use Video Conferencing on the TV ?

Trying to understand the impact of others Working and Schooling from home on each other the survey got the following responses

- 27% of respondents felt that other home members doing Video Conferencing or Video streaming was affecting their own connectivity
- 44% of respondents said that the had experienced Video Conferencing dropouts that had made doing their job more difficult

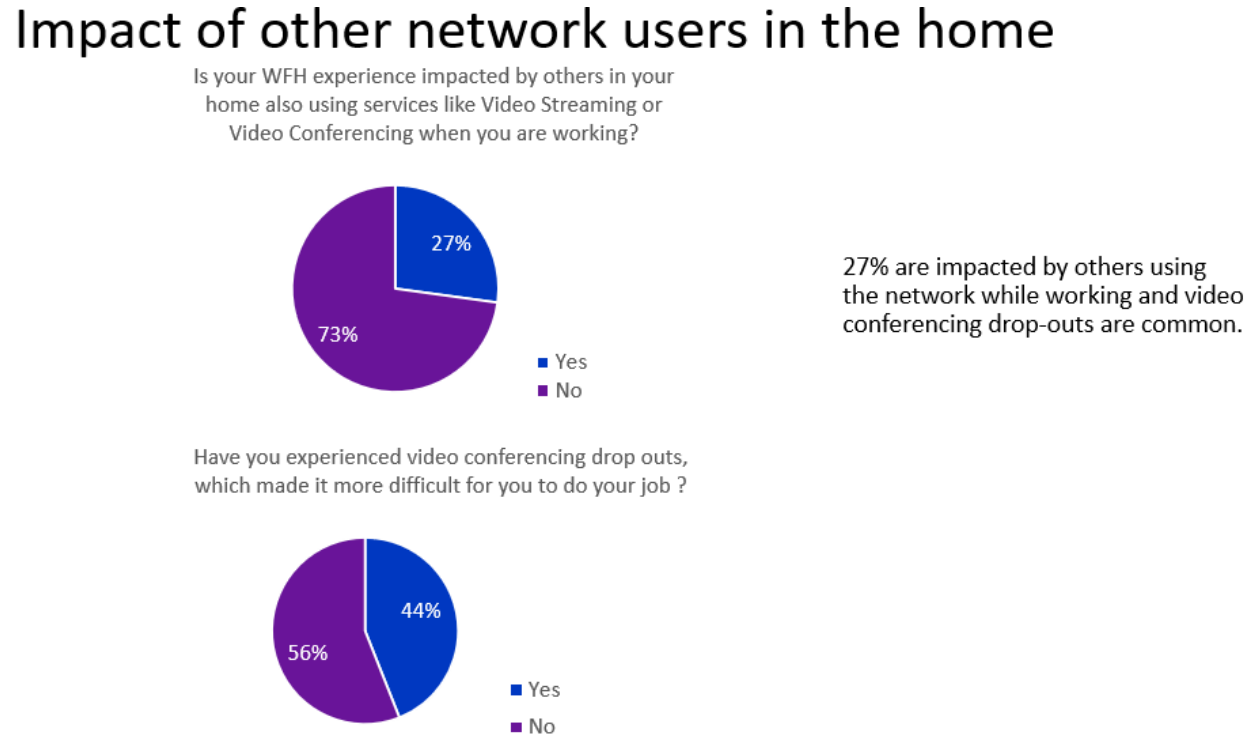


Figure 21 Impact of other home members on your Work connectivity

Going deeper and asking directly if respondents felt that their home working environment was adequate the survey got the following responses

- 45% of respondents felt they needed more reliability
- 37% of respondents felt they needed a better home office setup
- 52% of respondents felt they could benefit from lower latency broadband
- 28% of respondents felt they had Wi-Fi coverage and performance issues
- 22% of respondents felt they had inadequate broadband speeds

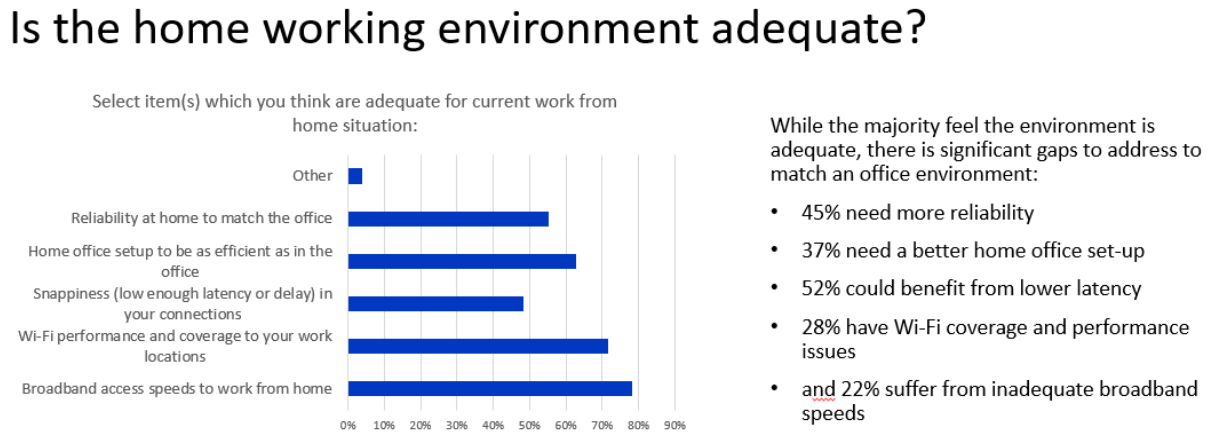


Figure 22 How adequate is your Working from Home environment

Now that we have learned the key issues and things that our respondents' value in their Work from Home lives the survey wanted to check their interest and value in improving the issues and adding new value services

The survey asked them to rank High, Medium and Low several items

- Use of additional screens at home
    - 57% of respondents ranked it high; 29% of respondents ranked it medium and 17% low
- Security on connection to the office
    - 46% of respondents ranked it high; 34% of respondents ranked it medium and 20% low
- Power Backup solution
    - 25% of respondents ranked it high; 36% of respondents ranked it medium and 39% low
- Separate Wi-Fi SSIDs for company work vs personal work
    - 24% of respondents ranked it high; 36% of respondents ranked it medium and 40% low
- Higher Reliability from ISP with backup WAN connection
    - 51% of respondents ranked it high; 33% of respondents ranked it medium and 16% low
- Better Hands-Free microphone solution
    - 42% of respondents ranked it high; 33% of respondents ranked it medium and 25% low
- Separate Corporate Network and Service to residence
    - 25% of respondents ranked it high; 38% of respondents ranked it medium and 37% low
- Better Cameras for Video Conferencing
    - 21% of respondents ranked it high; 36% of respondents ranked it medium and 43% low

This was somewhat different than some of the similar other responses asked in the survey but if we are looking for responses that show over 50% pull from the respondents from the cumulative High+Medium results, there seems to be some merit in number of areas below – in particular high reliability , security. Other teaser questions like the need for corporate SSID and network separation also seem to have some merit but may need to be driven more from the company side vs a pull from the consumer to really make them emerge as potential services a Cable Operator can sell back to corporations.

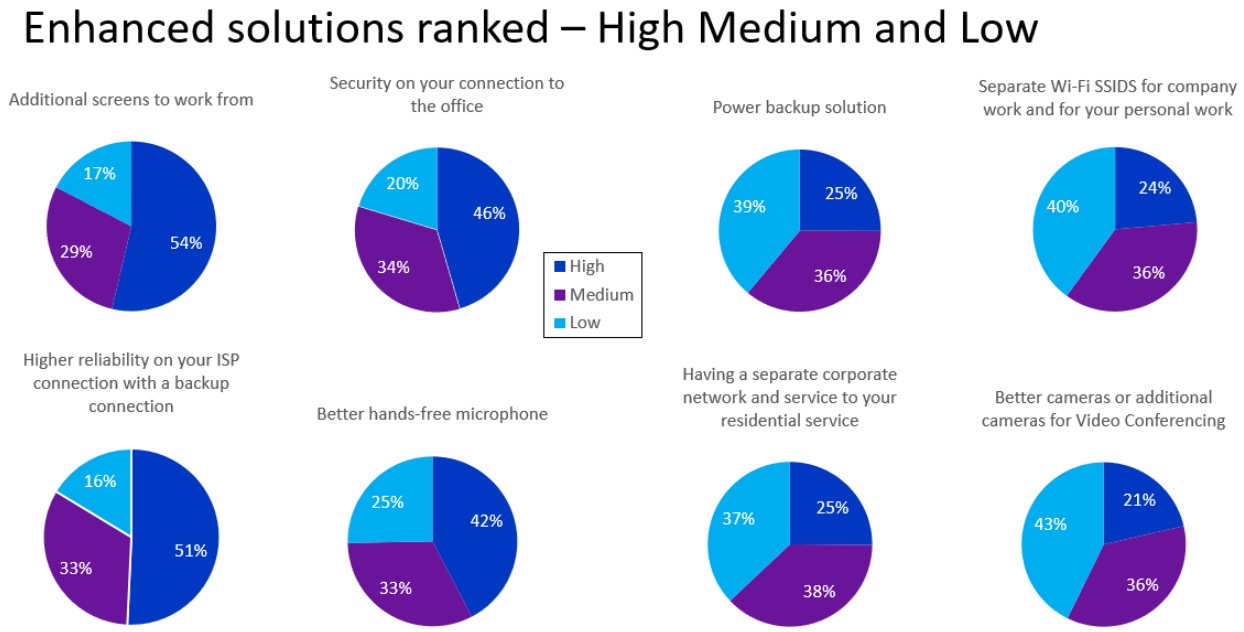


Figure 23 What additional Services and Features would WFH respondents value

Moving to the more social side of working from home the survey asked respondents how they would value more separation of Work and Home time through technology features

- 87% of respondents stayed in the same room while working
- 40% of respondents would like a technology or connectivity solution that separated Work time from Personal time
    - As you will see below when the question is posed as a summary set of questions with the specific use case of turning off corporate resources at certain times there was a higher interest level

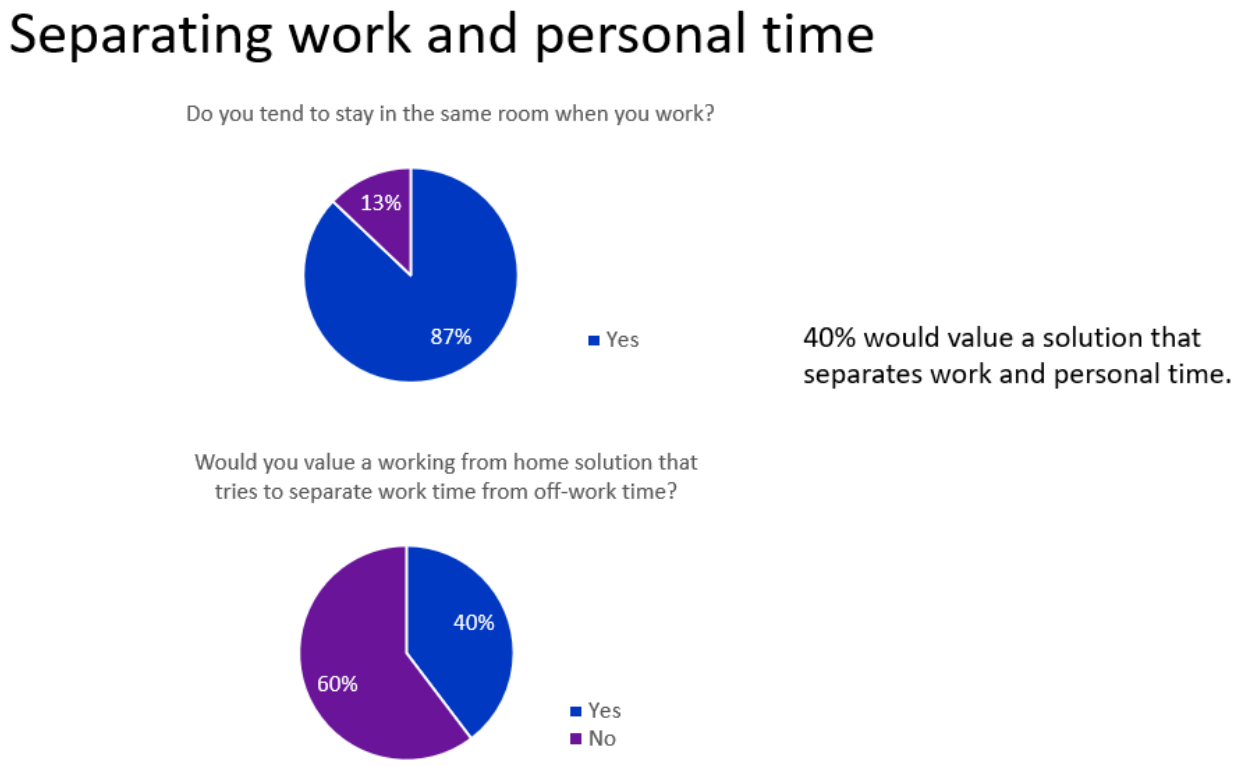


Figure 24 Value to our respondents of solutions to separate work and home life

Trying to analyze further some ideas that may benefit working from home the survey asked the respondents to give their view on some work life balance ideas

- 48% of respondents thought a timer that shut off WFH activities at defined time would be a worthwhile idea
- 73% of respondents thought that there were value in tools that analyzed your working habits to help you stand up from your chair, get some exercise and take breaks from balancing working and screen time
- 51% of respondents thought a dedicated work Wi-Fi, VPN solution may be something they valued
- 6% of respondents interestingly felt they wanted their company to pay for broadband service to the home. This question when followed up with respondents was interpreted as the company would pay for and own all broadband services to the home. There was some fear of privacy that came with this interpretation. When explained that it could be a completely separate service that was independent from residential broadband there was a much higher acceptance of company sponsored broadband and connectivity in the home.
- 68% of respondents felt that they valued physical separation of Work from home location with dedicated devices to home working.

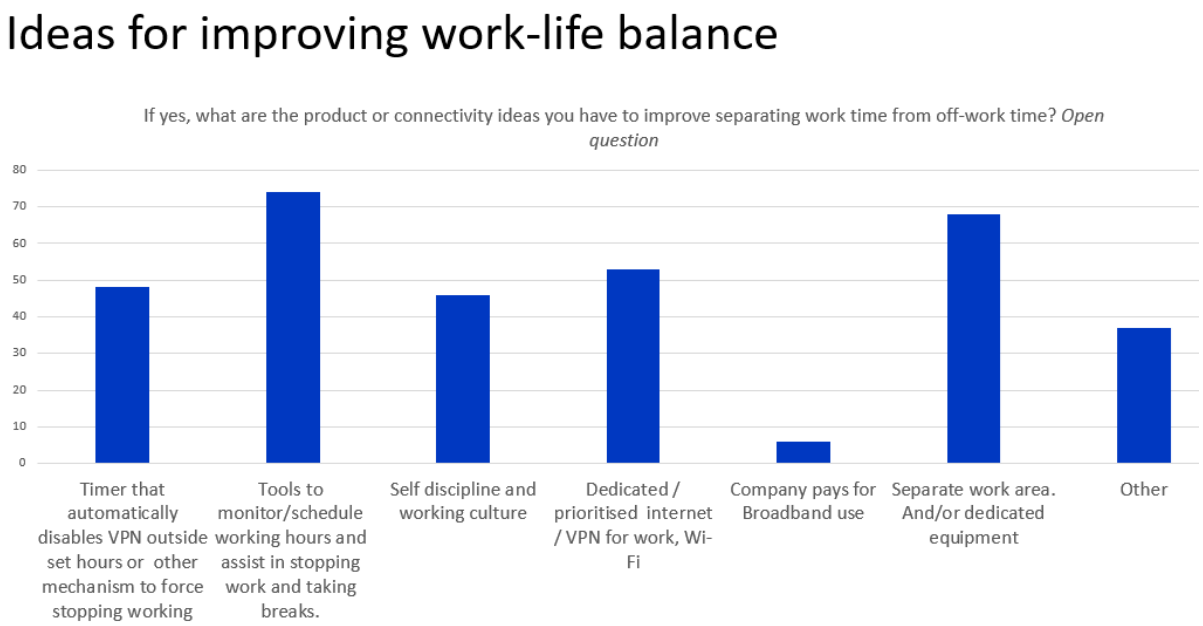


Figure 25 How do you value tools and features for work from home and life balance

As we now try and ascertain if there is a product, service, or market for a better working from home solution we delve a little deeper with the survey to ask some additional questions.

The survey directly asked the following

- Would you be interested if your SP offered a better WFH solution with higher downstream and upstream speed, better home Wi-Fi and more reliability – 75% of respondents said Yes. (For this initial analysis we did not push on how much the respondents would pay for this)
    - To further clarify some of the elements we asked
    - Are you likely to upgrade Wi-Fi because of issues – 43% of respondents said Yes
    - Are you likely to upgrade your WAN connection – 39% of respondents said Yes

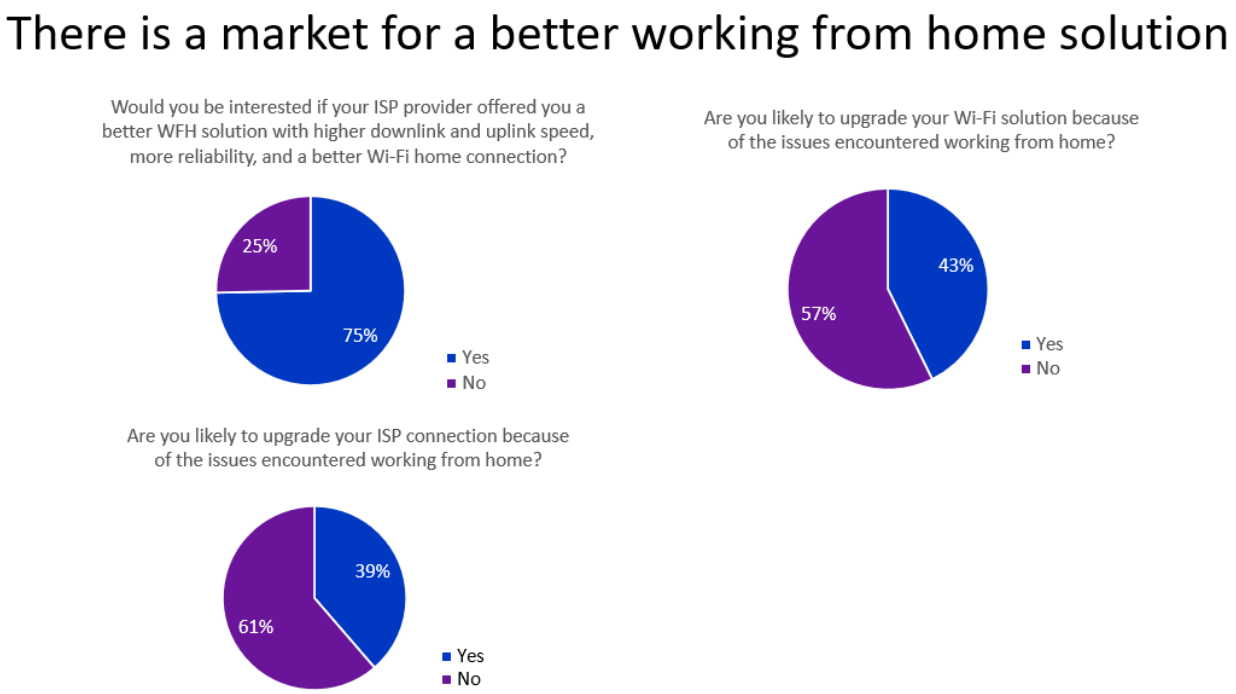


Figure 26 Would Homeowners want to upgrade their connectivity because of WFH

This survey result has some clear responses and leaves a number that are still ambiguous when trying to figure out the value of

- Reliability and backup – the results seem to show clear potential value here
- Security – the results seem to show clear potential but also highlight (with privacy) the importance for clear separation of any WFH sponsored services that they do not bleed into residential use.
- Separation of work and home life – the respondents were a little ambiguous in a couple of shaped questions but there seems merit in exploring this area. There are a number of separation areas from data, time and even Wi-Fi channels and frequencies. Additionally, even separate devices are also potentially something the respondents would value.
- Performance – its clear that consumers and the respondents in this survey vary on their understanding of the performance issues affecting their video conferencing and ability to work from home. Some just rely on their top line speed and maybe aware of their upstream speeds, others very cognizant and aware of Wi-Fi performance in the last 30ft and others who are also aware of latency derived issues.

So, let's move to the next section and explore the potential for the current stock issue of Cable Operator devices to change to reflect some of the new drivers for the home beyond supporting entertainment primarily.

3. Is there a better Home solution now for supporting Working and Schooling from home

Let us now explore the potential opportunity for the new home architectures that cater for

- Residential and Corporate services to a home
- Reliability of Service
- Extension of Corporate Office to the home
- Security and anti-theft of corporate information
- Separation of residential, schooling and working from home services
- Separation of work life and residential life from home

But before we do that lets baseline the simple service offered by Cable Operators today where there are predominantly

- 4 devices – Gateway (or EMTA and AP) , Extenders/Wi-Fi Mesh , STB and for a growing number of Cable Operators a Smart Phone MVNO play.

From a service perspective though there are the following assumptions and residential class of services that are offered

- For the majority of Cable Operators – a single Coax cable to the home with no backup solution for Coax side service failure.
- An ever-decreasing number of Gateways that provide battery backup of the VoIP service
- No solution for battery backup service for data services
- Typically, only 2 SSID in the home – the primary residential service and guest Id. Additional SSIDs can be Community Hotspot and for a few Cable Operators a Smart Home Wi-Fi network. Current networks on 2.4GHz and 5GHz frequencies.
- VoIP services using EMTA function in Gateway as part of DOCSIS network – which has been diminishing in use – in favor of smart phone calls at home and increasing use of FaceTime, WhatsApp and other popular video conferencing applications.

Figure 27 Current Cable Operator devices and services

Is there an opportunity to develop the following logical solution to create separate networks to the home and split the residential and corporate/schooling services for different SLA and policy management or possibly just increase all residential SLA to be able to support the same levels as Enterprise. There are several factors in consumer want and adoption of a solution like this

- Who pays for this uplift to Enterprise or Corporate level reliability and SLA's
- Privacy and separation of home and work life
- Can an Operator provide an SMB/Business like SLA and solution en-masse to consumers with or without subsidization of corporations?
- Will there be enough employees stay at home for work to have enough of a market and need for this solution
- Does the Cable Operator have the economics to be able to provide wireless backup solutions to wired services and without control of the wireless network in MVNO solutions can it work
- Will a consumer value having corporate IT control some of the elements of a home network or see it as invading their home privacy?
- Does the DOCSIS GW technology or indeed even standalone Access Points have low enough power draws to support meaningful holdup times for home data services off of backup wireless solutions. Indeed, can a HFC solution also work long enough for data in the event of a power failure in the home but HFC network still capable of passing data.
- Is it less complex when a consumer wants a WFH solution with higher reliability and performance SLA's to just offer a separate set of devices to separate the residential service/devices?

- Giving Corporations the Security, Performance and Reliability that they offer their employees in the Office
- Providing differential QoS and Separation for Residential and Corporate Svcs
- In HFC plant – even a separate eMTA

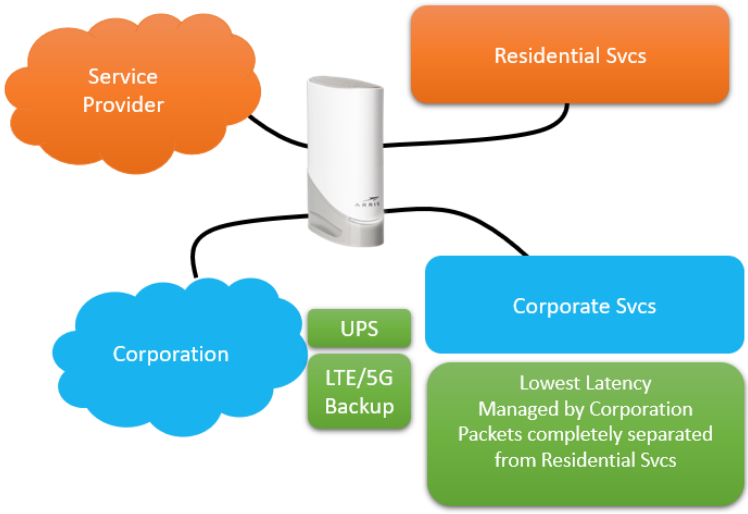


Figure 28 New Potential split Residential and Corporate services model

Expanding on this Gateway/Broadband/Wi-Fi architecture more could the following be a set of potential requirements and features for the better Gateway

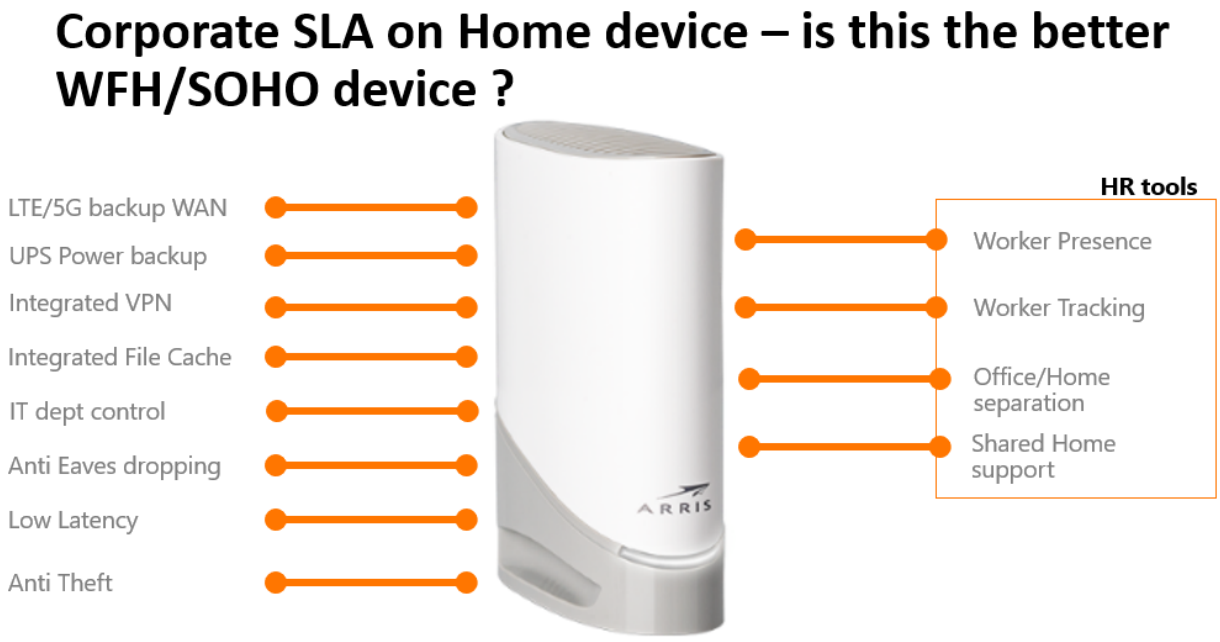


Figure 29 Potential additional features for the better WFH/SOHO Gateway

Some of the features we are defining are not new and have been always in the mix for inclusion in residential services but with the pandemic now really highlighting the stress points of reliable services and devices will these features like LTE backup and UPS/Battery backup now have a place for Operators to sell an improved service to consumers. These services cannot be free to the consumer as they have

additional capital and operational costs to provide them, so consumers must see value to their home requirements to increase their monthly payment to their service provider.

We could also look at the SMB and Enterprise market and apply some of the technology directions there to the home. For example with SD-WAN services and architectures squarely focused on Enterprise and SMB – would the solution above be better defined as an SD-WAN play and provide a Universal CPE (uCPE) component to the connectivity gateway (integrated or separate compute/storage device) and use SD-WAN software solutions for orchestration and management of WFH services independent of the residential services. Who would provide the uCPE device and is there a collaboration with Corporation and Operator to provide the SD-WAN services ? Corporations for example providing their own authorized VPNs but Operator providing other value add services.

Let's go through the services mentioned above to see their relevance to the consumer and if each one itself makes a difference or the combination of all these services is the 'killer application' for the consumer.

3.1. LTE/5G Backup WAN

Probably one of the most often discussed residential features over the years has been the potential of WAN backup to DOCSIS or xPON with LTE or now emerging 5G networks. This has been something that Telco/MNO's have offered more in the past given their ownership of LTE spectrum. Generally, because of the high availability of the wireline network it has not really been deployed at large scale. This paper does not contain any refence information on the number of outages or average yearly outage level on the HFC elements of the network. Residential services had not tended to be enterprise critical but now with increased WFH residential use WAN backup has now also become a potential service that consumers may now pay for the additional capex and cellular packet costs to have a convenient and ready to go backup service.

The increase in higher speed LTE services and more importantly the additional 5G capacity being developed for mobile and FWA applications makes a high function WAN backup feasible.

Some Cable Operators also offer their own MNO services and more and more Operators are moving to offer MVNO cellular services. In North America some Cable Operators have also purchased CBRS band 3.5GHz spectrum so have the potential to use that spectrum for WAN backup if they develop the network and CPE infrastructure to support

As well as a backup solution the addition of an LTE/5G module with a DOCSIS or Fiber WAN port also offers the potential for

- Carrier/Hybrid Aggregation – potentially combining at the cellular spectrum levels across different channels or potentially across wireless spectrum and wired service. This solution has been employed in some cases to augment low bandwidth xDSL networks with LTE and is less likely to be of high value to aggregate high bitrate DOCSIS 3.1 or xPON networks.
- Policy/Service based routing – the potential to be able to send packets over either of the WAN links based on achievement of SLA levels (latency) or congestion issues in either network.

The key issue in this solution particularly at the time of writing this paper is the overall cost of the addition of the wireless backup solution

- LTE UE modules of higher category performance and 5G UE modules are almost the same cost as a Wired Gateway so can the cost of supporting 2 high speed Access networks be covered in the service charge to consumer.
- MVNO costs for data make it less desirable to have wireless backup as the goal of the Operator is to minimize data traffic over LTE/5G or pass the cost of backup services to the consumer.
    - In an MVNO relationship packet priority for MVNO packets may also be constrained to create longer latency in congested areas of cell use.
- The speeds and latencies offered by the wireless backup solution also must be of sufficient enough speed and low enough latency to provide a WFH solution. This may be a lot lower than the Wired service but sufficiently high enough to support WFH service. A reliable 25Mbps/25Mbps could potentially support the majority of WFH workers for quality Video conferencing and low latency enough connectivity for the workers in the home. To make the lowest bitrate possible, software policies can be added to the WFH SLA to prioritize all WFH packets over any other residential devices to at least minimize latency for WFH traffic.
- Probably, most importantly consumer view on the reliability of Wired Broadband vs the investment in a backup solution for Wireless. Additionally, their views on power outage vs wired broadband outages and the downtime exclusively for wired outages. Its also a tougher sell from a wired Service Provider to provide a Wireless backup to their inability to maintain reliability on their wired network. However, consumers are becoming more aware of the environmental issues of fallen trees and storms affecting infrastructure outside the control of the operator so there is some sensitivities now in consumers to want to pay for this additional reliability. As we saw from our survey 51% of people felt it was a High value feature to them with a further 33% saying it was of medium value. Only 16% of respondents felt it was low in their home broadband requirements.

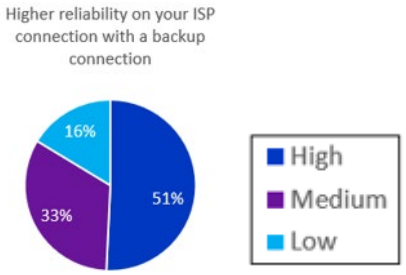


Figure 30 Respondents view on needing a backup higher reliable WAN service

So, for a multi wan backup solution there does seem to be a more heightened view by residential subscribers that it is a worthwhile option to have for their homes to keep them reliable for working and in the rare time outages occur.

3.2. UPS/Battery Backup

It's always an interesting question whether the HFC plant is inherently more reliable than the power grid. Certainly, in parts of the country with overhead power lines and HFC cables and with lots of trees, there tends to be correlated power and HFC outages when a tree takes down both power lines and HFC cables. The US Energy Information Agency reported https://www.eia.gov/todayinenergy/detail.php?id=45796 in November 2020 on the 2019 grid metrics. The average US power interruption was 4.7 hours or (284 minutes) per year (about 99.95% availability or 3 and a half 9's).

Including major events, the average interruption of 4.7 hours (284 minutes) in the United States during 2019 was nearly half the average interruptions experienced in 2017, a year with more hurricanes, wildfires, and severe storms. Excluding major events, the average duration of interruptions customers experienced was 1.5 hours (92 minutes), relatively consistent with previous years.

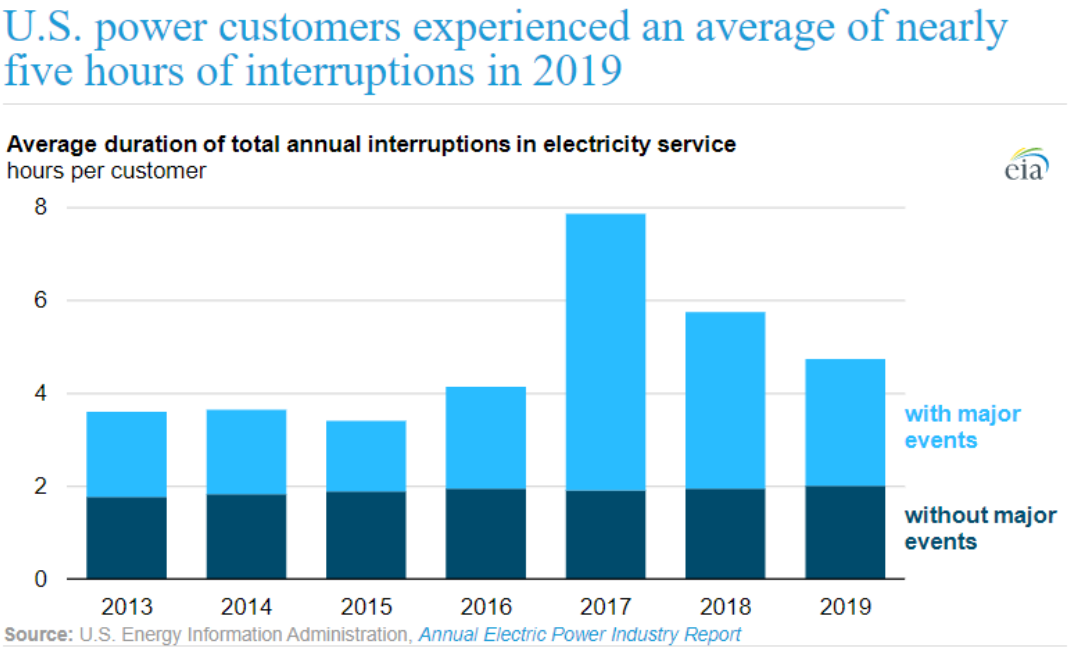


Figure 31 Average Duration of total annual Electricity interruptions in the US

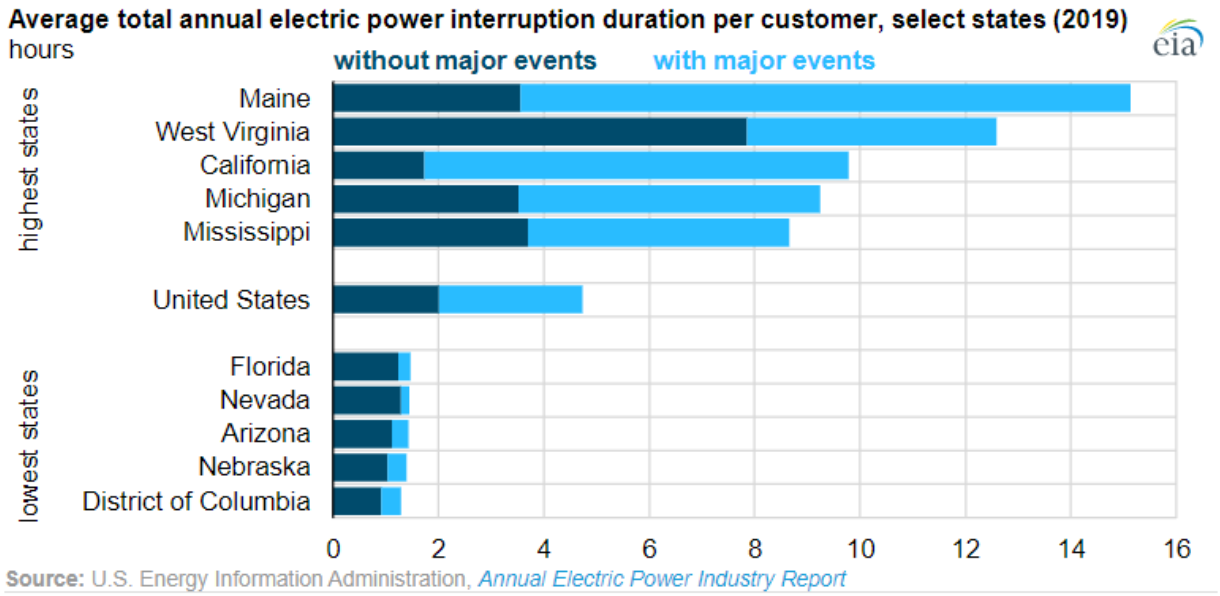


Figure 32 US Power interruption average duration per consumer - select States

So, how often does the power drop in the home really cause the consumer and the SP to want to invest in backup solutions. Certainly requirements for E911 calls still remain but smart phones and power backed up local base stations tend to be the main fall back for consumers now for critical communications and we

have seen a big drop in the use of battery back up in eTMA for VoIP calls on DOCSIS/Cable as consumers and Operators see less of a reliance on the main landline phone for critical or even comfort level services in a power outage.

The correlation of both a power and internet outage are also another interesting data point. As of writing this paper – we did not have correlation information of power vs internet outage. Certainly for major events like Hurricanes, flooding there is often a 1:1 correlation between both power and HFC infrastructure outages (Trees bringing down cables) that also diminish the value that a Power outage solution can bring to somebodies home.

All that said – given the right marketing and cost points WFH consumers who are in areas with higher rates of power outage may opt for a UPS system. Marketing departments in Operators could test the market for UPS to support up to 65W for a single GW device (not including Mesh Wi-Fi) from 600VA/330W at under an hour of hold up time and about $60 to 3000VA/850W solutions closer to $1,000 exceeding the total average power outage time in the US and far exceeding the typical spike fallout or non-major event outage. This UPS would be targeted to support the Wi-Fi GW only with scope for USB charging of phone devices and potential to keep a WFH laptop trickle charging as well.

Of course, homes that generally feel that they experience multiple outages also opt for home generators and typically place their refrigerators and lighting circuits on the generator but with more sensitivities around the importance of Wi-Fi connections in the home also tend to add the Residential Gateway to the Generator.

3.3. Integrated VPN

Many corporations still use VPN for their WFH security requirements. More and more corporations do provide non-VPN support for Applications like Office365 and some companies require specific USB keys like Yubikey to authenticate the use of laptops and devices outside the corporate network. Generally, there is continual work to improve the performance and security of VPN's – supporting context aware connections, zero trust networks and what is defined as Software Defined Perimeters (SDP) to improve security.

When a corporate VPN is used – it is typically run as a software application or extension on the Laptop, Tablet or Phone and in most cases is not application aware. Certainly, corporate VPN's may route all traffic from the corporate device through their VPN servers. In some cases, the VPN client can offer split VPN capability where corporate domain applications or servers/data are sent over VPN and other IP addresses are sent directly through the Gateway to their internet destinations. The Corporation itself will be the ultimate driver of this feature and will determine its viability from their IT policies

- Does the corporation want ALL traffic from a WFH device to go over the VPN for its security and employee application use policies?
- Putting a specific corporate VPN application into an ISP provided Gateway is something not supported currently. With the advent of new containerized Services Delivery Protocols, the feasibility of allowing a corporation to add their own VPN client to the GW runtime.
- Does the corporation favor a single VPN client in the GW supporting multiple WFH devices and having clear channel connections to the client only encrypted by Wi-Fi?
- For regular mobility use of VPN the end device like the laptop must support VPN client anyway so is a WFH solution worth it.

This last point is potentially the most interesting one from the WFH employee's perspective. Often while at home the employee may use their WFH device for non-work applications and typically drop their VPN

for these applications. Potentially, listening to music or watching videos during their lunch break or after hours. Bringing up and down the VPN client manually from the laptop device is an employee inconvenience that could be better accommodated with a GW based context/application aware split VPN. It depends totally on the IT policies on laptop and WFH device use and what is monitored from the company's perspective. Sometimes IT policies for workers while in the office will apply tools like WebSense to prohibit access to certain websites even general use of YouTube or Filesharing tools like Dropbox. If the device itself is not the restricted element but the work hours usage is – then some IT policies can reflect this with the ability to use laptop for personal uses when not on VPN or outside working hours. Ultimately if an IT department has a policy that a corporate issued device can only be used for company activity and all internet access is subject to the company's policies then split VPN tunnel at the GW itself may not be a useful feature.

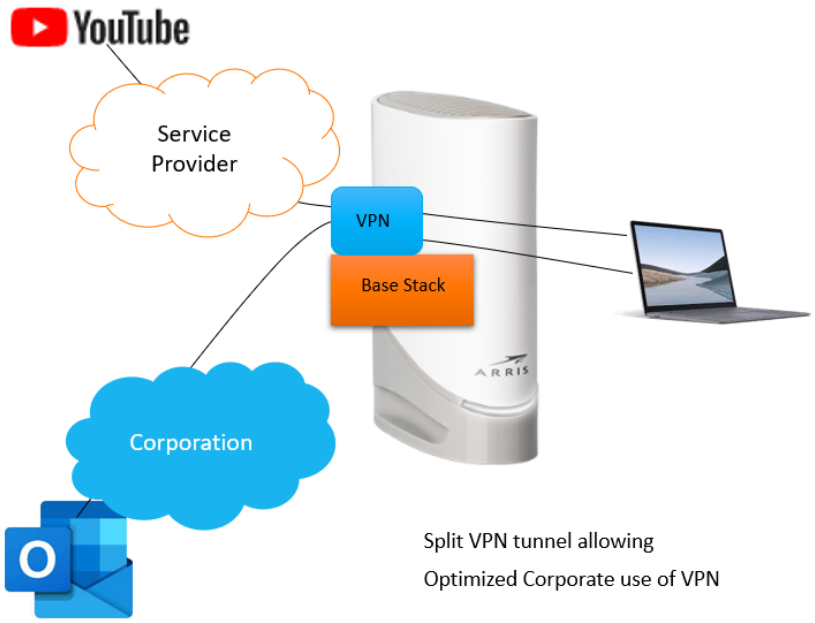


Figure 33 Integrated VPN and Split VPN solution

One additional comment on VPN from a Laptop vs VPN from a Gateway is the ability to provide quality of service to individual services like Corporate Video conferencing vs treating all VPN traffic as higher priority. When every service from a corporate laptop is encrypted in a VPN tunnel then differentiated service policies cannot be applied. While there are additional security implications with WPA3 encrypted traffic to a VPN in a Gateway – it does allow the GW to apply split VPN solutions easier and treat different services with different QoS policies better. A VPN client on the laptop applying split tunnels can also be used.

3.4. Integrated or Home File Cache

With much of the increase in upstream traffic in a WFH environment now starting to be file syncing to solutions like Microsoft onedrive – is there the potential to move that onedrive location to a local file cache in the home and then replicate it also in the corporate or Microsoft cloud? Depending on sync and user policies this could also help with Operator upstream bandwidth crunches where the syncing could be dependent on available bandwidth on the network vs currently more realtime and during office hours. It is

unlikely today that file syncing or local home storage solutions would be integrated in the Gateway (SSD technology making it at least easier from a size perspective) but could be something that is done in a Universal CPE device that offers local WFH processing and storage functions for SD-WAN like services to the residence

3.5. Corporate IT department control

Looking at the potential for the GW to be an extension of the corporate office and maintained as a virtual office end point by the IT department, what are the basic elements of a solution that may extend the office to the home.

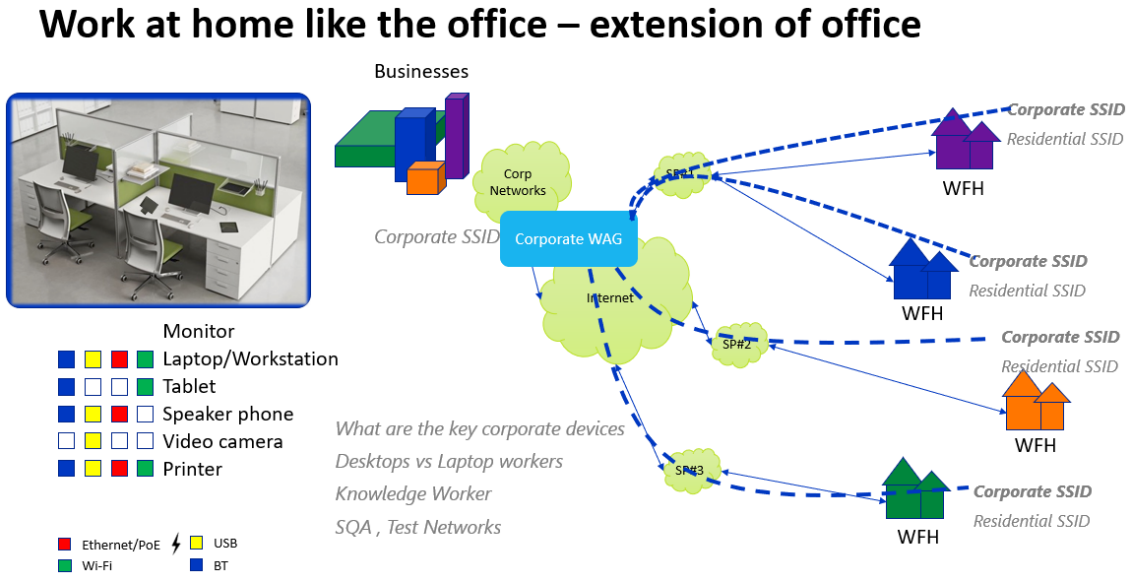
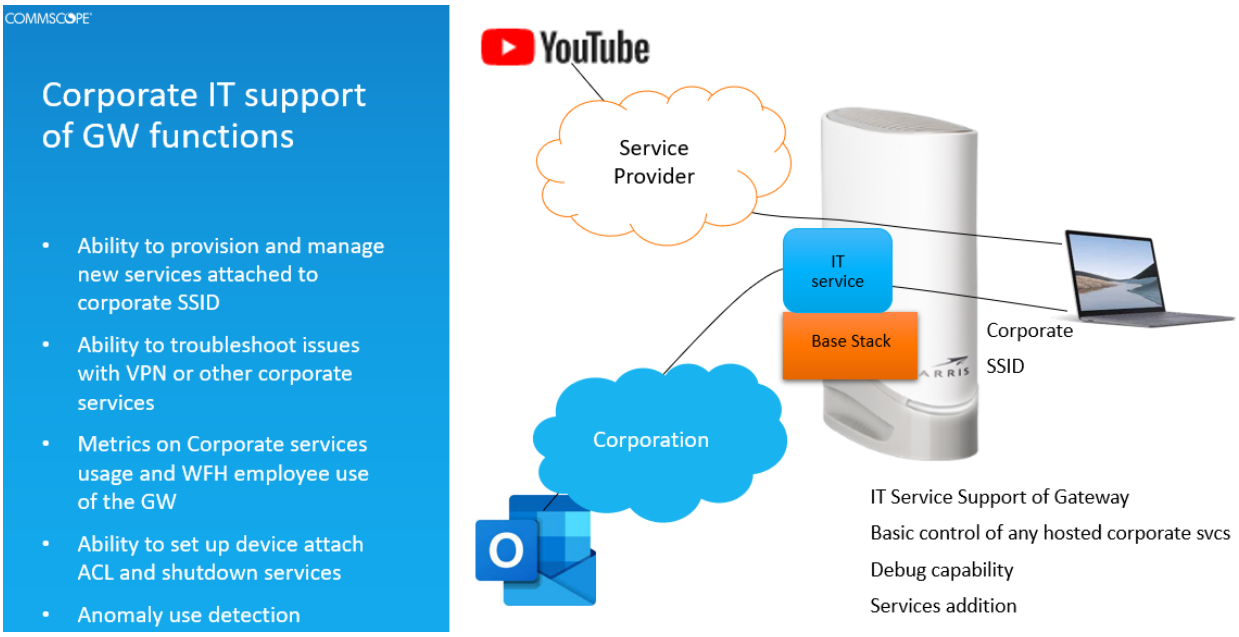


Figure 34 Extending Corporate IP Scope and SSID to the WFH GW

Consider the above approach of providing a dedicated Corporate SSID extending the typical office setup. Using both WPA3 encryption for the Wi-Fi SSID and VPN for all corporate bound traffic and services there is scope to have these residential GW's offer this service by downloading a specific containerized s/w solution on a per company basis. Allowing an Operator to charge to host this service and SSID for the separation of residential and WFH services. Creating containerized edge solution that can allow Corporations to apply their own security policies and ability to reset the service to the SSID when required. There could also be some options where the Operator provides telemetry information normally picked up by its own TR-181 pull and sends a subset of this to the IT department in the case of issues with the Broadband or Wi-Fi services in the home. Simple elements like – the broadband service speeds, wired broadband service is down and 5G/LTE backup is on and its usage statistics, Wi-Fi performance and any anomalies, VPN status, Security Status, CPU status, Memory status etc. An application and telemetry demarcation that (i) allows the Operator to pass the issues for any WFH service to the IT department to resolve (with special access to reset the Wi-Fi or even the Broadband Router device) and (ii) sells a service to Corporations that ties in with Employee presence and use on the Corporate connection could have a double benefit to the Operator.

Something more to consider as its realization is much easier to implement now as Gateways move towards containerized services models and away from monolithic images. Being able to add these services as well as protect the Operator/Consumer in residential mode from the Containerized memory

space and runtime and services that are exposed – enables these applications more to be able to be realized.



3.6. Anti-Eaves Dropping Solutions

With working from home now exposing corporate information outside the company office at huge scale and now having lots of endpoints distributed like leaves on a tree over a huge geography – the ability to steal corporate assets seems to have increased. Consider the following scenarios

- Shared GW and Wi-Fi networks by people working for different companies in the same house or apartment. More frequent and easier access to both Wi-Fi data from specific devices and even open laptops in the apartment. Will corporations pay to improve security in these more often and more frequent situations.
- Increased number of hours that someone is connecting to their corporate network. Everyone usually does work at home, but this can be intermittent vs WFH had day long corporate devices connected to corporate network.
- Higher probability of someone trying to get access to company information targeting a worker at home to steal passwords or create an intrusion to the corporate network. Still requires sitting outside someone's home or being in a neighboring apartment where Wi-Fi connection of the target is available.

One of the reasons why corporations favor running VPNs on the Laptop directly is that all passwords and traffic is VPN encrypted additional to the Wi-Fi encryption on the link to the corporate device. To provide VPN from the GW vs from the Laptop does leave a potential security hole between for example a corporate laptop and the GW it connects over Wi-Fi. Typically, this link will now be WPA3 encrypted but can be exposed with Man in the middle attacks and typically what is know as Evil Twin Access Points that masquerade as the users own Access Point and SSID. Solutions such as a **wireless intrusion prevention system** that monitors the radio spectrum within a wireless network's airspace for unauthorized or unexpected activity and frequencies could also be deployed and be part of a new improved WFH security solution.

Shared homes with different company employees in the same market are also now exposed more for theft of IPR/Assets with the increased frequency of WFH and the opportunity to target employees and corporations by sharing homes and apartments. Solutions that

- Separate traffic, networks, SSID's, encryption schemes as much as possible between co-habitants in a SFU or MDU are potentially desirable. Again, typically corporations require on VPN from the corporate issued laptops to provide security for their IPR. Would SSID separation and WPA3 encryption separation also be desirable?
- Minimize exposure to passwords employing two-factor authentication to a second issued corporate device like a smartphone/SIM is another solution that should be employed more for WFH as the location of the WFH employee and device could be easily spoofed.
- Provide Operator created schemes that are like 2FA using other associated home devices to help a corporation ensure their employee is located in a specific WFH location while working? For example, using other devices such as extenders and Wi-Fi STB to be present in Wi-Fi scans to correlate to an employee being in their home while working and using the specific corporate issued VPN or other connection.
- Track and manage large downloads – potentially requiring 2FA schemes to continue and or limiting the number of bytes that can be accessed in a day for WFH. While the corporation can track the capacity through the VPN client and concentrator the operator can potentially also provide some correlation numbers on this endpoint as well. This also highlights potential applications too of work and home separation (discussed later) to be able to provide billing information for WFH traffic vs residential traffic. This is particularly important given that many service providers operate traffic caps where overages cost the consumer. With WFH increasing both Downstream and Upstream traffic usage, there could very well be an offered solution from a Cable Operator to offer corporations a separate bill for any corporate traffic, identifying it from per instance VPN bandwidth and/or destination addresses, application use etc.

If WFH is set to be a high percentage consistent activity these eavesdropping security issues will only increase so there may very well be value to expand some of these ideas to provide the most secure WFH Wi-Fi or Ethernet connection with some of these ideas below – offered by Cable Operator to corporations as a potential chargeable service.

3.7. Low Latency Service

Latency is very topical now as people have experienced video conferencing issues, audio breakup while working from home as well as the rise of online gaming and general gaming during the pandemic. Is there an opportunity to apply low latency features like Low Latency DOCSIS and Low Latency Wi-Fi into the WFH application space? Is the following thesis potentially valid enough to create these new services and architectures?

- WFH traffic during working hours should be prioritized higher than residential traffic?
- Certain WFH applications like the audio from Video conferencing should be prioritized ahead of other traffic when those flows are detected. This can be impossible to do when Video Conferencing applications are always put into the VPN tunnel and hard to identify their service or destination address to apply QoS and priority policies.
- Being able to apply DSCP and TOS marking to WFH services with or without affecting residential priorities.

There is potentially other latency issues that occur in MDU environments with congestion on Wi-Fi networks in the 5GHz and 2.4GHz. Jitter and latency in congested Wi-Fi environments can affect Video conferencing and potentially in the future other work from home low latency sensitive solutions like

running visualization application on cloud servers but rendered in home devices – for example mechanical engineers running immersive video-based visualization tools and using hand gesture to move around objects in real time.

Enter the new 6GHz spectrum that has just been released in US and is rapidly being approved across the globe. Is 6GHz spectrum destined to be used as a WFH private network to optimize performance, separate service, improve security and create an extension of the office. One simple representation of a Wi-Fi 6E WFH application is illustrated below. Using additional corporate sponsored or supplied devices that connect to the Cable Operator Gateway using a private corporate supplied 2 box Wi-Fi 6E network could create the best blend of home Wi-Fi performance for WFH, total separation of residential and WFH traffic using not only VPN and IP separation but now channel and frequency and device separation. There could be also potential improvements for Wi-Fi RF eavesdropping as well using a solution where 25mW in room VLP is used to connect corporate devices to the first 6GHz AP and then VPN tunneled from their through 6GHz backhaul double protected with WPA3. VLP 6GHz propagation is highly unlikely to be sensed outside SFU and even in the majority of MDU environments so could provide a range limited but effective WFH Wi-Fi spectrum to use for both lowest latency and security.

From a latency perspective using Wi-Fi 6E and 6GHz spectrum provides access to Wi-Fi 6 scheduling of device transmission, clear spectrum for the most part, 160MHz (and 320MHz channels with Wi-Fi 7) all contributing to be able to provide deterministic latency and high probabilities of retaining a 1ms latency per AP hop. This would provide a WFH solution independent of congestion, contention, and latency to be able to power all WFH applications.

Figure 35 Is 6GHz the opportunity to create a private WFH low latency network

3.8. Anti-Theft of Service/Device solutions

Assuming that there are corporate software solutions added to Cable Operator Gateway or Corporate supplied home solutions like those defined above for a 6GHz VLP based in home office WFH solution –

then is it desirable by corporations or their employees alike to also have a solution that supports anti-theft of the device. Having the device only work in WFH environment for the specific corporation and in specific locations and with specific people and even specific applications.

Providing a series of checks to allow operation of the device could be done by several means

- Two or more factor authentication – using Cable Operator other Wi-Fi based devices that are defined to identify location and person requesting use of the WFH gateway
    - One or more STB need to be present
    - Bluetooth beacons need to be present from designated devices like Smartphones and STB
- Specific association of a corporate supplied device with a specific Cable Operator GW or Extender to allow WFH operations
- Nonoperation and boot of a GW if removed from its regular client environment. Regular client environment could be defined as simple as some mandatory presence of the devices in a persistent ARP table. (Note MAC randomization makes these simple associations more difficult and other schemes may have to be used).

There seems to be some merit in adding a feature like this to review with corporations if this level of anti-theft of service and device makes them feel more comfortable allowing employees to work from home.

3.9. HR Tools : Worker Presence & Tracking

A more sensitive or controversial set of features that may be relevant to employers and corporations as they move to completely virtual work solutions – is the application of presence and tracking technology for their workers. This all has to be done within HR guidelines as well as GDPR, CCPA, CCPR privacy regulations. Employers have been concerned that employees working from home, may not retain the same time and work diligence at home than in the office and can take opportunity to skip out from work to do home activity or other local activities during the day. There are simple technology ways to validate worker presence (presence not totally being indicative of the person working to any quantitative level) and a potential Realtime clocking in solution. Potential ways to do this are

- Traffic activity from the specific WFH device or on specific WFH VPN, SSID etc
- Connection of the WFH laptop or corporate issued phone on the Wi-Fi AP and correlating its frequency and potential type of traffic

Additionally, there could very well be some policies applied that an employee must have Bluetooth on in a corporate issued phone and this BLE beacon presence potentially also combined with Wi-Fi presence on the Operator or Corporate supplied GW/AP/IoT hub defines worker presence. This can obviously be spoofed by leaving home without the corporate issued phone but a somewhat restrictive solution to define trust that you are at home and working.

Other more intrusive but potentially allowable schemes are

- Using camera on issued laptop (Exam taking schemes now require the use of camera on the device exam is being taken on. Employers could also mandate something like this for their employees to work virtually).
- Radar and Wi-Fi motion solutions. Trying to signature a specific employee vs another home member is the key algorithmic requirement to make this solution offering viable.
- Remote clock in and clock out solutions – that are also aided by correlating to traffic and network usage

- Keyboard clicks, mouse usage are more invasive presence detection solutions but may not be qualitative of work or who is using computer or mouse.
- Turning on location services on corporate issued devices and company having agreements with employee to track their location during working hours or as a condition of being supplied corporate phone or even working from home.

And of course, a combination of all of the above may be able to reliably track the presence of an employee.

Of course, this may not be an appropriate or acceptable way to create a trust culture with employees and it may be better to trust the time-honored tradition of managers being responsible for the qualitative performance of their subordinates – so does presence or the metric pursuit of presence make any difference if any employee constantly excels at their tasks and role.

One security related use of presence that would be good to improve a WFH environment where multiple employees of companies in similar sectors are sharing an apartment. Having a local only presence solution between a smartphone and a GW/AP where the AP/GW or the WFH virtualized services running the GW are paused/blocked/not accessible unless the GW detects via BLE presence of the corporate phone and/or Wi-Fi attachment of same. A similar approach could be applied directly to the laptop or other corporate devices where they are logged out in the absence of another 2FA solution like BLE presence of smartphone.

A very sensitive and controversial area but there could be some happy medium of presence detection that is both beneficial to employee and employer.

3.10. HR Tools : Office/Home Separation

One of the more employee centric and mental health elements of working from home is the issue of separating Work and Life, This includes work and life from time and location perspective. Working in the place you also live can have issues on many levels. One of the simpler ones is to try and ensure that employees have the same natural triggers and events to start and end their day as well as take breaks for health purposes. For sure in a virtual world online meetings and meetings in general have now become scheduled rather than impromptu with a check of someone in their office. This has caused many people in WFH environment to say that they walk downstairs directly to their computer and surface late in the evening without the regular rhythm of commute to work, beating the traffic home and the odd call by to see can you do lunch.

Also, in certain countries there are very specific rules for time worked per week and the times that you can work. A virtual work environment makes it harder to enforce and check on this and there could be in the future WFH employees who have stress and mental issues that they blame on this WFH virtual working environment.

So, is there a technology way to separate Office and Home?

Separation can be defined in several ways

- For those fortunate enough to have a home office – it does provide some logical and physical separation of work and lifetime. When you enter the office – you are in the company domain even in your own home. When you step outside you are back into your residential environment. As we discussed above something like 6GHz VLP for example could actually separate Wi-Fi

connectivity from work and home and effectively only allow work devices to connect in a specific room of the house and not follow you to the kitchen.

- For time separation, policies can be applied to corporate issued devices or even in the case of a Cable Operator GW providing a WFH SSID for example – it be shut down at specific times to conform to companies working policies. The employee themselves may have to unlock things to complete certain critical tasks they deem necessary, but the company has complied with its or country working directives.

And of course, presence and tracking mechanisms coupled with solutions that have to comply to specific hours worked may be complete solution.

This may be something that could be applied to contractors in particular who understand their requirement to offer value typically billable on hourly basis.

Another potentially controversial topic but again in this new world of increasing virtual work and contracts based on virtual work stipulations giving HR tools to understand working patterns or even to shut down employees or throttle their work based on work directives that companies have to show compliance.

4. Conclusion

As we all wait to see what the permanent effect of the 2020/2021 pandemic is on the number of employees who work fulltime from the office, full time from their home or a flexible mix of both it does seem that there is opportunity to develop a better home connectivity and services solution that caters to the WFH elements and its new effect on the typical residential connectivity solutions. This paper tried to create a discussion on the development of this new mousetrap and see does it have merit for the employer, employee, the employee family or all of the above. Is there a business plan that can show who would pay for these improved services for productivity, security, anti-theft of IPR and even remote worker wellbeing? It does seem that most of the funding for these new services would have to come from corporations paying the Operator to allow these services to be dropped into their devices as containerized home services or easy onboarding of corporate supplied devices into a primary residential service provider solution. With the emergence of containerized service delivery and the arrival now of 6GHz spectrum this does also provide at least new technology elements that can offer path to be able to create some of the solution ideas discussed above. Is the better path to follow the SMB SD-WAN architecture and move this to manage a single WFH employee? Does it scale down to support that and does the new WFH solution also need a new rethink on the Universal CPE solution and role for split residential and WFH services. The authors conclusion from both the almost 2,500 WFH pandemic time employees and the wire brush of some of the ideas above is that there is an opportunity to extend the current residential only connectivity, security, and services model to add WFH as a service with feature additions like those above BUT this is dependent on the indelible change in the ratio of hours spent working in an office moving to the home environment. As of today in August of 2021 , the prediction is that an hours work ratio in 2019 pre pandemic of Office:Home ratio something like 90:10 is now going to be something like 70:30 or 60:40 is probably a big enough shift to make the better WFH GW and Solution mousetrap.

Humanoids Optional: Deploying vCMTS at Scale with Automation

A Technical Paper prepared for SCTE by

Bhanu Krishnamurthy
VP, Software Development and Test
Comcast
1800 Arch Street, Philadelphia PA
217-437-2811
Bhanushree_krishnamurthy@comcast.com

Gregory Medders
Principal Engineer
Comcast
1800 Arch Street, Philadelphia PA
Gregory_Medders@comcast.com

1. Introduction

For more than two decades, cable internet has provided the means to access the internet for hundreds of millions of people. The traditional cable modem termination system (CMTS) has played a key role in facilitating this access. The advent of cloud computing and containerization has caused a shift in access network solutions, with proposals to replace the traditional "integrated" CMTS (iCMTS) with a virtual CMTS (vCMTS), where the hardware-based iCMTS is replaced with an cloud native architecture that delivers the same functionality. By decoupling the CMTS implementation from the underlying hardware, vCMTS was proposed as a way of providing a modern, more flexible alternative to traditional models. For example, when deployed in a distributed access architecture, where the digital nodes can be placed closer to the homes they serve instead of in the headend, vCMTS can provide improved service with significantly increased density while also reducing headend infrastructure costs.

Broadly, vCMTS brings cloud-native approaches to bear on a problem that was traditionally solved through iCMTS hardware/appliances (Cloud Native Computing Foundation, 2020). At Comcast, vCMTS is comprised of a collection of servers running Kubernetes to orchestrate the containerized vCMTS microservices (The Kubernetes Authors, 2021). These micro services communicate with a remote physical device (RPD) through a networking layer composed of various switches in a leaf-spine architecture (see Figure 1). While cloud providers such as Amazon Web Services (AWS) offer services that can stand up Kubernetes clusters with a click of a button, latency and timing between vCMTS containers and the RPDs is a critical consideration for DOCSIS protocol application, making it impossible to use a public cloud for compute resources. As a result, the vCMTS cluster are typically deployed at the edge in the hub sites where their iCMTS counterparts were previously deployed.

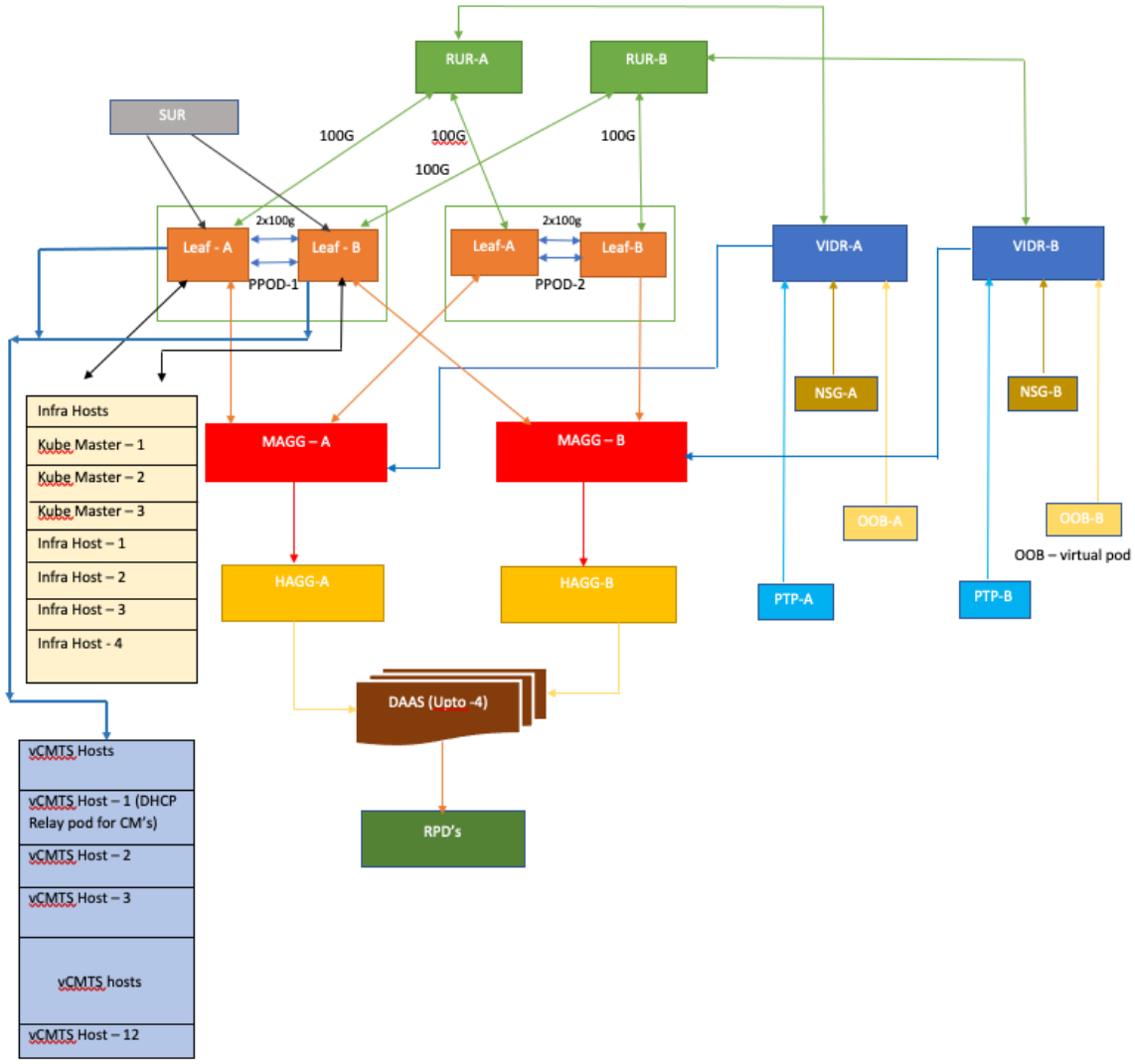


Figure 1- VCMTS Network In A Leaf/Spine Architecture, Where VCMTS Microservices Are Hosted In Servers On The Leafs And The Spine Facilitate Connectivity To The Rpds.

In 2018, Comcast deployed its first production vCMTS cluster. At that time, bringing the first customers online in vCMTS architecture was a major success, proving the feasibility and performance of the distributed access architecture (DAA) using vCMTS (Cable Television Laboratories, 2014). However, as Bob Gaydos, a Fellow at Comcast, remarked at the subsequent SCTE meeting, "Realistically, because Intel came in and helped us … make the vCMTS work on their platform, that's not the hard part. The hard part is actually the operations" (Robuck, 2018). Though the vCMTS architecture provided a huge improvement over iCMTS in terms of visibility through real-time telemetry and logging, operationalizing vCMTS proved challenging in several different ways: how the clusters were built; how changes were deployed to the clusters; how operations teams monitored the systems for issues and responded to incidents.

After our initial success, we started scaling out vCMTS across the Comcast footprint and quickly realized that our operational model would not scale. While each subsequent cluster was identical to the original in terms of hardware and software, each required unique configuration (e.g., IP addressing), making each new cluster build a significant effort. It was not uncommon for a cluster to take months to build from the time when the hardware arrived in the data center until we were ready to cut the first customers onto the new cluster. Because the clusters were painstakingly built by hand by different people, the vCMTS clusters were inevitably considered "pets" — each with their own slightly different personalities. While effort was made to keep the cluster definition in code, because the clusters were hand-configured from inception, it was a constant battle to prevent the clusters from "drifting" from their initial configuration over time.

The culture of customization proved particularly problematic when scheduled maintenances were performed. Since the clusters were built manually by humans, changes to the clusters generally also occurred by humans, often involving a long and complicated standard method of procedure (SMOPs) to both implement and validate the change. Though the SMOPs were scripted as much as possible, because the running configuration of a cluster could differ from the configuration stored in source code, the same change may go perfectly on one cluster but face an unexpected issue in another cluster. These issues could compound; when an unexpected issue was encountered, changes that should have been trivial to revert could snowball into outages that were difficult to resolve. Consequently, changes to software were approached with wariness and generally avoided, resulting in a delay or outright cancelation of feature deployment in this first generation of clusters.

Having proved the validity of the vCMTS model, it became clear that something needed to change if we were to deploy on a national scale. Staffing estimates predicted that we would require between 10-100 times the number of current operational resources to support the vCMTS clusters across the Comcast footprint; more importantly, root cause analyses of incidents indicated that a significant number of issues encountered were caused by human error. While DAA improves the fault tolerance of vCMTS, this comes at the cost of complexity. Fortunately, we were able to use much of the learnings from the DevOps movement to identify three key areas where we needed to improve: clusters must be identical, changes must become "boring" (i.e., frequent and predictable), and incident detection and mitigation must be automated (Kim, Humble, Debois, & Willis, 2016). Practically what this meant was that everything in the cluster lifecycle, from creation to maintenance to incident detection and mitigation, needed to be automated.

2. Infrastructure as Code: GitOps and a Beginning to the End of Customization

Even before the commitment to "automate everything", it was very clear that we were contributing to our own problems by requiring humans to deploy changes. After the first successful vCMTS customer trial, significant effort was invested in codifying the vCMTS deployment into an "infrastructure as code" (IaC). This idea, which has been central to the DevOps movement, is that the desired state of a system should be expressed as a versioned (e.g., Git) code that completely describes how to build that state. Therefore, any desired change to the running state should be achieved by altering the desired state.
A common variant of IaC is the "GitOps" model, where the infrastructure code is stored in Git and, upon commit to Git, the change conveyed in that commit is applied to the running system via a continuous integration/continuous deployment (CI/CD) pipeline. In our initial vCMTS systems (aka "Gen1") human operators were still required to implement changes; nonetheless, investing in IaC enabled us to rapidly rebuild systems, bringing us a step closer toward eliminating the culture of customization that made changes unpredictable.

3. Automating the vCMTS cluster stand up

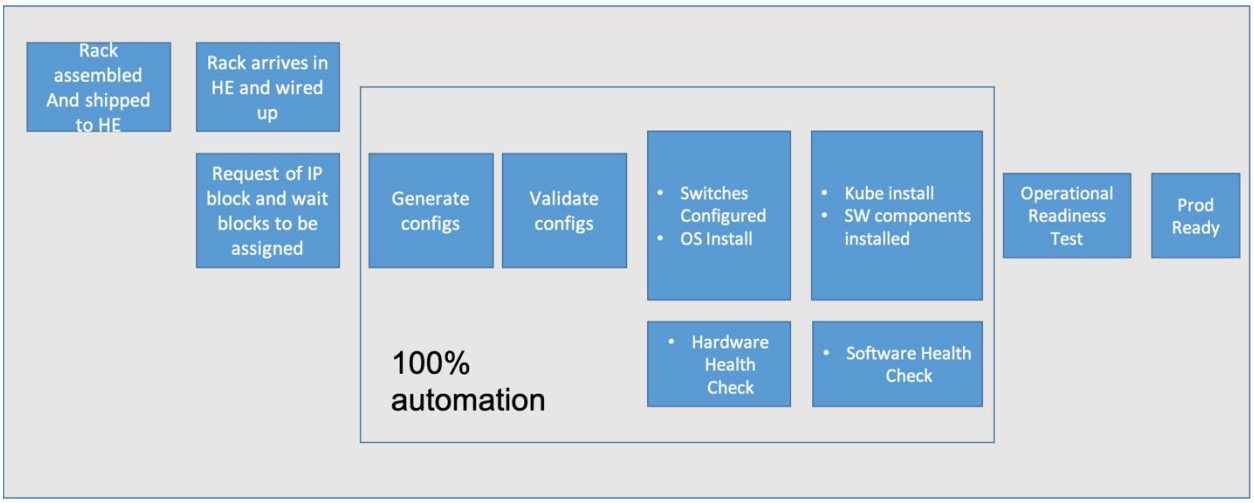


Figure 2 - Automated vCMTS cluster build process

As we were beginning the process of automating the vCMTS cluster build, an obvious but important step was identifying exactly what was required to build a cluster. A cluster build would typically occur over the course of many weeks. Eventually, we found that nearly a dozen teams or specific individuals were responsible for the various steps in a cluster build; this makes sense given that the original build was an "all hands" effort, different individuals or teams naturally became the subject matter experts on different aspects of the build. Since it evolved organically, the resulting build process resembled a web of dependencies as opposed to an assembly line. In the same way that the physical cluster assembly benefitted by consolidating within a team that "owned" the assembly process, as the dependencies were untangled and codified into a CI/CD process (Figure 2), the network and cluster build process was also recentralized around a dedicated team.

Recall that vCMTS consists of microservices (performing the DOCSIS scheduling and control functionality) and that communicate with RPDs via a networking layer. Having streamlined the issues in cluster builds, the remaining challenges can be grouped into two categories: network (switch OS and configuration) and vCMTS (cluster operating system, Kubernetes, configuration, and microservices). In both cases, the configuration (primarily related to IP assignments) is automatically generated based on the planned topology. Running a vendor-supplied network operating systems (NOS), the network primarily implements IS-IS to provide dynamic routing between the microservices and RPDs. When the cluster racks arrive in the local data center and are powered on, the first part to provision are the network devices. The switches automatically begin zero-touch provisioning (ZTP), download the latest OS (IaC, versioned in Git), and load the generated configuration (again, versioned in Git). When the switch becomes reachable, health checks are automatically performed to verify that the switch has the correct OS, configuration, and that the network is healthy.

Once the networking layer is functioning, a CI/CD pipeline brings the vCMTS online. This begins with the servers loading their base operating systems via ZTP. Subsequently, firmware is updated, Kubernetes is installed, and the vCMTS and supporting microservices are installed. A constant source of friction

during the original vCMTS deployments was integrating new clusters with external applications within Comcast (i.e., services that were not part of vCMTS itself, but that were required for a vCMTS cluster to function normally); this process was initially informal (e.g., an email), but by formalizing these dependencies as APIs, a constant stumbling point with external teams became a seamless integration. The CI/CD pipeline concludes by performing health checks for the application level, verifying that the vCMTS system is functionally working, and concludes by registering itself as a built cluster. By structing the build as a CI/CD process, we prevent an unhealthy cluster from accidentally being used to serve customers.

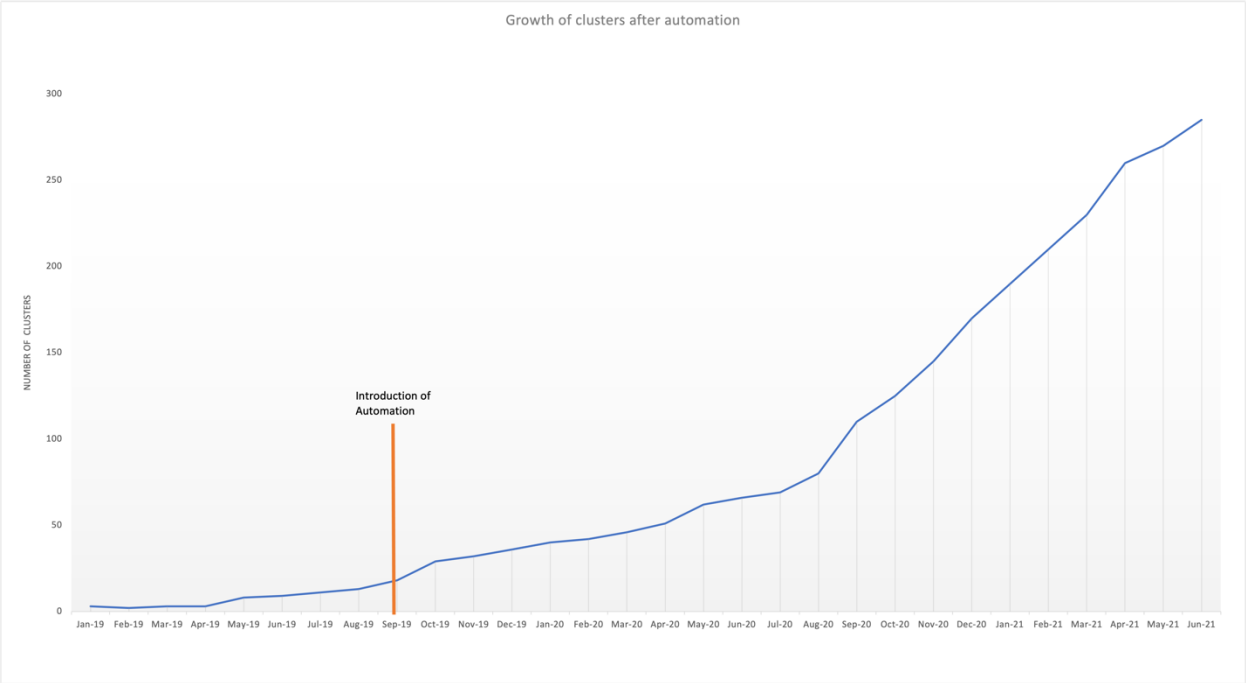


Figure 3 - Number Of VCMTS Clusters Versus Time. The Automated Build Process Began In September 2019, Significantly Accelerating The Cluster Build Process.

While the cluster could theoretically cut-over customers upon conclusion of the CI/CD pipeline, we have maintained a final step of human-driven manual testing (e.g., a cable modem is brought online, voice calls are performed, speed check, connectivity to backend tools, video channel lineup tests). This is consistent with our overall goal of automation—humans should: be removed from tedious and error-prone work that can be readily performed by machines, function as supervisors of machines by providing high-level instruction (e.g., "build a cluster with these parameters"), monitor for patterns and issues that are currently unknown and difficult to detect through automation.

Revisiting the overall build process, human involvement has been limited to 1) physically wiring the cluster/network, 2) some upstream network configuration into the core network (which currently cannot be automated), 3) final manual validation testing. This automated build process was released in September 2019 and as shown in Figure 3, has enabled us to grow the number of vCMTS clusters deployed across the Comcast footprint exponentially without a corresponding exponential growth in human resources as previously predicted. Concretely, provided that the clusters have been wired correctly, the automation has decreased the cluster build time from weeks to a few hours. The cluster

build process has become so predictable that all clusters which had been hand-built prior to automation were rebuilt in automation with essentially no customer impact. Now that the clusters were all identical and consistent with their IaC definitions, we were able to focus on the automation of changes.

4. Automating software and network changes

A critical issue we encountered in Gen1 was that, over time, the clusters drifted from their configuration defined in Git. This was primarily due to humans needing to deploy changes and address incidents, each which inevitably introduced small (unintentional) changes that compounded into significant drift. Thus, having spent considerable effort to ensure that the vCMTS clusters were built with no human involvement from a CI/CD pipeline using (automatically generated) configuration stored in Git, it was critical that any subsequent change to that git also be deployed to the cluster automatically.

While DevOps is often framed in terms of using tooling to achieve CI/CD, a mantra of the DevOps community is that a DevOps transformation is a cultural shift, not just a change in tooling. One of the first significant steps toward this was realized by Eric S, Principal Engineer at Comcast. A frequently encountered issue in Gen1 was that people tended to inadvertently work outside of version control. Whether it be by writing a SMOP and storing it in the change ticket (only to have the underlying cluster definition change in Git, rendering the SMOP "stale"), applying a script that was valid in a previous software version but invalid in the currently deployed one, or contaminating one site by accidentally using configuration from another, we constantly encountered issues with incorrect configuration being accidentally applied to clusters. By 1) breaking apart the deployment scripts for the various microservices into isolated and independently versioned bundles and 2) changing the workflow to require users to always import the current IaC into a local, containerized environment any time they wished to interact with the cluster, Eric was able push the organization toward that DevOps transformation by making it significantly more convenient to practice IaC.

A second step toward DevOps maturity was achieved through the CICD pipeline (Figure 4). Capitalizing on Eric's work to isolate cluster components into visioned collections of deployment scripts (in our case, Ansible roles), deployment pipelines were developed using the same IaC. Operating at the vCMTS cluster level, the pipelines were automatically triggered when a software component version was upgraded for that specific cluster. In Gen1, each change was required to have a series of pre-deployment and post-deployment checks; these were generally included in the change ticket so that if a deployment did not go as planned, the incident could be better triaged. As part of the Gen2 deployment pipeline, these pre-/post-check deployment snapshots were automatically captured and linked with the associated change ticket. By automating this tedious but important part of deployment process, we were able to further incentivize adoption of the CICD process.

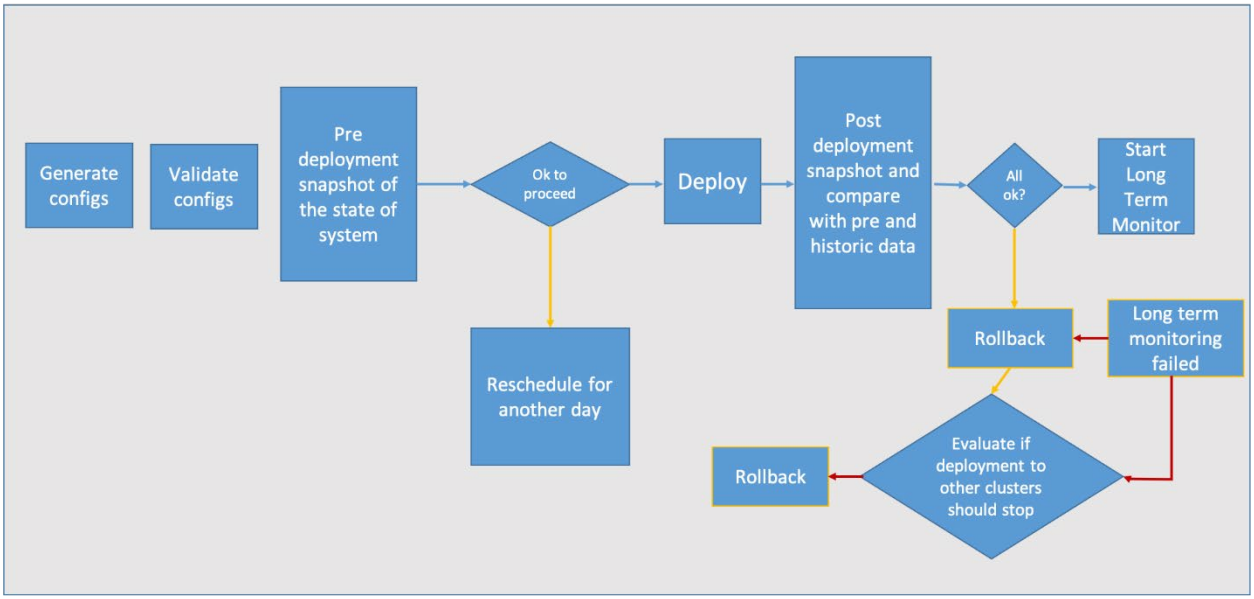


Figure 4 - Software And Network Change Deployment Pipeline

However, as both the number of clusters (due to stand up automation) and the number of changes (due to the cluster deployment automation) began to exponentially increase, we had an increasingly urgent need to orchestrate the deployment of changes across clusters. Since each cluster was identical, we were able to track changes as they were deployed and be confident that any issues encountered were due to the change, as opposed to unexpected differences in the clusters. Here, there are two competing interests. An important theme in DevOps is that developers should have the ability to easily introduce changes themselves and rapidly receive feedback about the success/failure of that change; however, the reality is that ISP operations are critical services, limiting the use of practices like A/B testing of potentially service impacting changes. Here, a compromise is achieved using canary testing, where each change is introduced gradually across the footprint. Changes are categorized by risk and deployed to minimize potential customer impact—first in QA, then staging environments (those without subscribers), and ultimately staggered into production environments.

In terms of changes to the network layer, the switches require both upgrades to their NOS as well as configuration changes. There is redundancy in the network so that if any given switch goes down, the redundant path is used with no impact to service. Our automation has taken advantage of this redundancy to turn maintenance into a zero-downtime event. Maintenance actions are first performed on one device (“A”) in the redundant pair (“A”, ”B”); traffic is then shifted to “A” side and only if everything behaves normally does the maintenance continue to patch the “B” side; if unexpected behavior occurs, the change is reverted. Through this entire process, the subscriber served by that network is unaware of any change being performed—their services continue uninterrupted.

Due to the large number of switches in the vCMTS access network (currently in the thousands), it was imperative that we fully automate the network operations and maintenance. To this end, configurations are generated purely through automation. Network operators are not generally allowed to make changes to the configuration on any given switch. Instead, network engineers edit the switch’s template file and commit it to version control where, consistent with the GitOps model, changes are detected and automatically scheduled for deployment across the entire network. Similar to the software

deployment automation, a network change will ripple across the Comcast footprint, with the engineer who performed the change having direct visibility into the impact and success/failure of their change through health checks. Consistent with the DevOps principles, providing this immediate feedback to engineers has significantly increase the overall quality of network changes and stability. Rather than the classic pattern of engineers "throwing a change over the wall" to operations to implement, the person who initiated the change is empowered to safely deploy a change across multitude of devices with a click of a button.

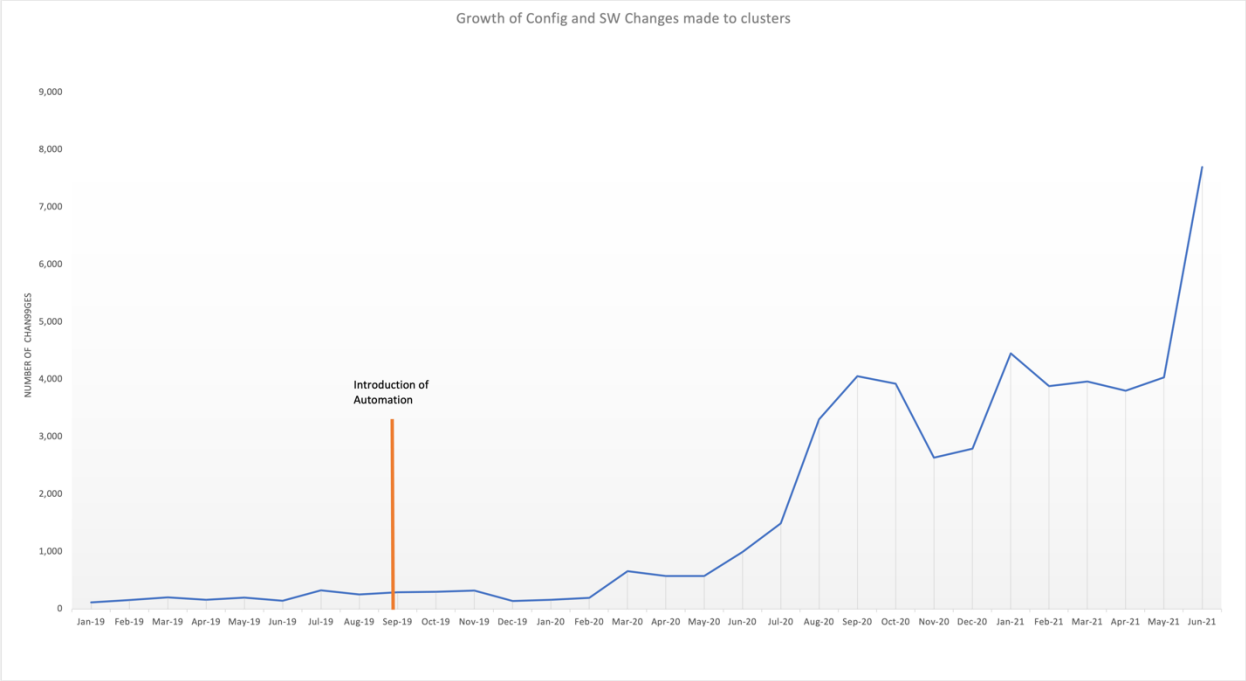


Figure 5 - Number Of Changes Versus Time. Automated Change Deployment Pipelines Were Introduced In September 2019

As is shown in Figure 5, automation was crucial in enabling our ability to deploy changes across the ever-expanding footprint of vCMTS devices. In Gen1, software and network changes were high-stress and high-risk events. Even though we had fewer than 25 Gen1 clusters, any given change would take several weeks to roll out since each change was manually performed by a human; furthermore, due to the high incidents of human error causing accidental outages, all changes (even to supporting applications that were not required for customer service) were typically performed in a scheduled maintenance window in the middle of the night out of a general fear of an unexpected issue escalating into a customer outage. As will be shown in the subsequent section, automation was crucial to our ability to scale out and rapidly deploy changes. Furthermore, the DevOps adage that "infrequent changes decrease service availability, frequent changes (e.g., continuous deployment) increases availability" was borne out; frequent deploys, including empowering the engineer who is making the code change to deploy the code, has significantly improved reliability and service availability, making changes "boring".

5. Automating incident detection and mitigation

After automating the VCMTS cluster build and change processes, our next focus was to automate the incident detection and mitigation. Because each cluster is identical and because no humans interact with the clusters directly during normal operations, our overall incidence of human induced errors dropped to zero essentially overnight. Correspondingly, the availability of the clusters across the

footprint has increased significantly (Figure 6). While issues certainly arise day-to-day (e.g., hardware failures), due to the built-in redundancy of the vCMTS clusters, these issues are typically not service impacting. Furthermore, because these issues are often routine, the response to them is known and can be codified in a script. For example, when a server in the Kubernetes cluster becomes NotReady, we automatically determine the cause of the Kube node being in NotReady state; if the problem is fixable and non-impacting (e.g., by rebooting the host or even reinstalling the host OS), it is done automatically and noted for subsequent analysis rather escalating to on-call operational support. If the issue does require human intervention (e.g., arranging an RMA with a vendor, re-seating an optical device in a maintenance window) steps like scheduling the maintenance are automated. By using machines for these routine and repetitive tasks, we significantly decrease the load on our ops team, enabling them to act as supervisors of the automation by codifying their responses into scripts and reserving them to troubleshoot non-trivial issues.

Figure 6 - System Availability Vs Time. Availability Is Plotted On The Right Side (Excluding Scheduled Maintenances) And The Number Of Incidents Per Month Caused By Human Error Is Plotted On The Left.

A hallmark of cloud native architectures is the improved visibility into the internal workings of the systems. Using a combination of open-source solutions, our vCMTS operators have visibility into everything ranging from network health to the performance of individual microservices. When an incident arises, this makes troubleshooting that cluster a much easier process because nearly everything that could go wrong is monitored. However, considering that we now operate more than hundreds of clusters, we cannot rely on humans having "eyes on glass" monitoring the clusters for potential issues. In a very real sense, we are simply inundated with data. Fortunately, in the age of data science and machine learning, being inundated with data is a great problem to have. A customer going offline could be caused by many sources: a fiber cut, power outage, network issues, backend unavailability, cluster software errors, etc. Troubleshooting the root cause can therefore be a lengthy process, resulting in a long mean time to repair (MTTR). Instead of immediately escalating to a human to investigate, our automated triage system can quickly probe various failure domains and rank likely root causes. This allows us to rapidly deploy the correct fix-agent.

During a deployment, variable usage patterns in the network can make immediately identifying issues difficult. While many changes are known to be non-service affecting in advance, some changes are intrinsically higher risk. Hence, long-term monitoring for anomalies is required. Using machine learning, external systems monitor telemetry for deviations from historical patterns. Since higher-risk changes are staggered across the footprint, this enables the automation system to detect anomalies and recommend that deployment be halted and already-deployed changes be reverted. This is integrated in our deployment orchestration system, allowing for the safe and hand-off deployment of changes.

6. Conclusion

In the past few years, vCMTS has evolved from a prototype to full-fledged production infrastructure providing cable services to a considerable portion of the Comcast footprint. While our initial Gen1 deployments proved the technology of vCMTS, the last three years have been devoted to scaling vCMTS operations. Often when we discuss our work with colleagues in other departments, a common response is that their work is automated too—via scripts. We want to make the distinction that, while Gen1 was fully scripted, the automation in Gen2 is neither a collection of scripts nor a simple CI/CD pipeline. Indeed, while one can write a CI/CD pipeline for a full stack application and deploy it in the public cloud in an hour, we have (somewhat painfully) learned that is not trivial to automate the operations of hundreds of bare-metal Kubernetes edge clusters running in hundreds of data centers across the US, in total comprising of thousands of switches and tens of thousands of servers. After our early work defining our infrastructure in code, we naively thought we had solved the difficult part of the problem. In fact, someone in our organization famously predicted it should take “a few weeks” to whip up the automation to allow us to manage vCMTS at scale. Instead, what began with a few engineers has evolved over the past three years into multiple teams comprised of 15 people. Building capabilities such as self-recovery, complex decision-making logic, risk-based scheduling to name a few took months of hard work by some of our highest performing engineers.

In a traditional ISP operations model, even our current footprint was projected to require hundreds of operational support engineers. Our goal ultimately is to reduce that to tens of engineers. While we released our initial automation almost two years ago, this work has grown over that time into a full-fledged product involving the collaboration of numerous teams at Comcast. Critically, this work has enabled us to achieve the exponential growth necessary to expand vCMTS across the Comcast footprint while also improving on the reliability and maintainability of iCMTS. As this work continues, we are further integrating our operational automation into the monitoring and core network infrastructure at Comcast, with the goal of achieving a zero-downtime and high-speed experience for our customers.

Abbreviations

vCMTS	Virtual Cable Modem Termination System
iCMTS	Integrated Cable Modem Termination System
IaC	Infrastructure as Code
DAA	Distributed Access Architecture
SCTE	Society of Cable Telecommunications Engineers

Bibliography & References

Cable Television Laboratories. (2014). *Remote PHY Specification.*

Cloud Native Computing Foundation. (2020). *CNCF Annual Report.*

Kim, G., Humble, J., Debois, P., & Willis, J. (2016). *The DevOps Handbook.* Portland: IT Revolution Press, LLC.

Robuck, M. (2018, October 29). Retrieved from Fierce Telecom: https://www.fiercetelecom.com/telecom/comcast-execs-lessons-learned-from-deploying-vcmts

The Kubernetes Authors. (2021). Retrieved from Kubernetes: https://kubernetes.io

Implementing Multi-layer Infrastructure Management for Multi-Access Edge Computing (MEC) Services Using Kubernetes

A Technical Paper prepared for SCTE by

David K. Bainbridge
Senior Director, Software Engineering
Ciena Corporation
7035 Ridge Road
Hanover MD, 21076
dbainbri@ciena.com

Stephane Barbarie
Software Engineer
Ciena Corporation
7035 Ridge Road
Hanover MD, 21076
sbarbari@ciena.com

Dmitri Fedorov
Embedded Software Engineer
Ciena Corporation
7035 Ridge Road
Hanover MD, 21076
dfedorov@ciena.com

Marco Naveda
Senior Director, Network Architecture
Ciena Corporation
7035 Ridge Road
Hanover MD, 21076
mnaveda@ciena.com

Raghu Ranganathan
Principal & Distinguished Engineer, Advanced Architecture
Ciena Corporation
7035 Ridge Road
Hanover MD, 21076
rraghu@ciena.com

1 Introduction

Communications networks are at the heart of advancing society and bringing people and places closer together. The evolution of communications services will be central in transforming how we work, play, collaborate, and interact with the environment around us. Emerging collaboration technologies such as augmented and mixed reality (AR/MR) promise to offer highly immersive, multi-user, real-time and content rich experiences that will simplify business operations, improve productivity, and unlock new services and revenue sources across a wide range of verticals. This type of application relies on large amounts of bandwidth and extremely low network delay to do real-time processing of very large data sets and tracking user and virtual object movement, while enabling fine-grained interactions between remote users, the physical world, and holographic objects. This will be possible as network application intelligence and cloud platforms converge at the network edge in Multi-Access Edge Computing (MEC) locations.

Over the last decade, communication service providers (CSP) have invested in significant network modernization to keep up with a growing demand for bandwidth hungry applications and increasingly distributed service consumption patterns. The adoption of Telco Cloud architectures for virtualizing network services has improved the operational responsiveness of the network. However, despite advances in network automation, the traditional top-down BSS/OSS operating model has not adapted to the realities of delivering dynamic, cloud-native network services to meet the needs of distributed MEC applications. This new application delivery paradigm requires new operational tools that enable CSPs to maintain carrier-grade operations for virtual machine-based virtual network functions (VNF), while evolving to the on-demand and intent-based deployment of cloud-native containerized workloads for the next generation of network services and MEC applications.

MEC infrastructure and connectivity services are expected to be a growing revenue source for service providers who build a distributed edge compute network platform for application delivery from cloud to edge to the customer premise [11]. However, no single provider will be able to address this massive opportunity, thus, there will be a need to coordinate resources across multiple layers of the network infrastructure as well as the federation of services from different providers in the wide area network (WAN). This type of inter-provider coordination requires the flexibility to define a specific network topology for a given application and user endpoints as well as exposing Telco Edge Operator Platform capabilities [9]. Such a system must include a tighter coupling with application networking for optimal placement of MEC workloads given the specialized requirements for compute resources and proximity to endpoints.

As enterprises adopt hybrid, multi-cloud strategies to support their digital transformation initiatives, pressure is mounting on the traditional telco cloud environment to align with the same level of service agility and developer experience offered by the hyper-scaler cloud providers. This is driving the need to support multiple providers for different components of the application infrastructure. For example, a cloud provider could be responsible for portions of the application infrastructure while a network provider could be responsible for portions of the network infrastructure and latency-constrained application components. Additionally, the use of the same MEC environment to support components from multiple application providers will require different hard or soft isolation techniques at MEC locations. A service provider must also find

ways to align to a cloud provider's edge deployment and operations model with suitable hooks for tighter visibility and control of the service provider's network. Service providers will need tools to do this at scale with the likely need to manage 100s-1000s of highly distributed sites.

In this paper, we propose a Kubernetes-based control plane with built-in intent-driven automation to address these operational challenges and facilitate deployment of both virtual machine-based and container-based network functions (VNFs/CNFs) alongside MEC applications in a hybrid, multi-cloud architecture with a multi-layer connectivity network underlay.

This paper first describes the use of Kubernetes technologies and extensions to those technologies introduced as a solution to address the operational challenges implementing a MEC architecture (section 2). These technologies are then applied as the paper describes deploying a MEC architecture based on the Kubernetes system (section 3). Finally, the findings and recommendations based on the work completed are summarized (section 4).

2 Background and Technologies

The European Telecommunications Standards Institute (ETSI) [1] provides a reference architecture for the deployment of MEC hosts and applications. With the shift to containerized workloads in a cloud-native environment, ETSI [7, 8] work supports the use of container management systems such as Kubernetes for providing platform-as-a-service (PaaS) services. Additionally, the mapping of the ETSI management and orchestration (MANO) information model to the container workload deployment model enables an approach for implementation using a Kubernetes model. The MEC architecture can be deployed using a Kubernetes model and, with extensions, a more complete MEC model can be achieved with virtual machines (VM), VM to container service chaining, and constraint policy-based connectivity.

ETSI [Section 8 of reference 1] defines MEC service as a service provided and consumed either by the MEC platform or a MEC application. In the context of this paper, the term is inter-changeably used for both user application services, e.g., a MEC application like AR/VR rendering, as well as host or platform services, e.g., traffic management service or domain network service. As an example, some host level MEC services could be offered by a network provider while some application level MEC services could be offered by a cloud provider. Some of these MEC services may be part of the Kubernetes implementation.

2.1 Kubernetes In Brief

Kubernetes is a portable, extensible, open-source platform for managing containerized workloads and services, that facilitates declarative configuration and automation. It has a large, rapidly growing ecosystem. It provides a framework to run distributed systems resiliently, providing standard patterns for application deployment, scaling, failover, security, and load balancing. A full description of Kubernetes and its capabilities can be found at [10].

Since Kubernetes primarily operates at the container level rather than at the hardware level, it provides some general features common to PaaS offerings, such as deployment, scaling, load balancing, and lets users integrate their logging, monitoring, and alerting solutions. There are

also extensions to Kubernetes that allow it to manage VM based workloads which will be discussed later in this section of the document.

Kubernetes is not monolithic, and its default solutions are optional and pluggable. It provides the building blocks for building platforms but preserves user choice and flexibility where it is important.

A single Kubernetes deployment is known as a cluster and consists of a set of machines (physical or virtual) called nodes, which are utilized to host containerized applications. The nodes within a cluster can be classified as either a control-plane node, where workloads that implement the Kubernetes system are executed, or a worker node, where primarily application workloads deployed to the cluster are executed.

The Kubernetes control plane manages the worker nodes and the pods in the cluster, and it makes global decisions about the use of cluster resources. A Kubernetes pod is a schedulable entity that is comprised of one or more containers. The control plane components can be run on any machine(s) in the cluster, however, the usual practice is to run all control plane components on one or more machines and avoid running user containers on the same machines as the control plane components.

The following illustrates Kubernetes cluster elements described above [12].

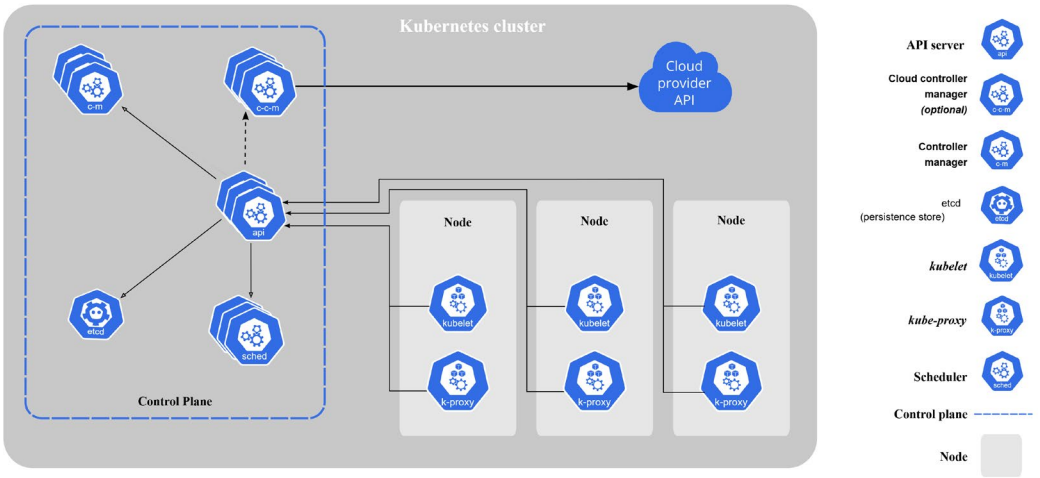


Figure 1 - Kubernetes cluster with all its components

As mentioned, Kubernetes is an extensible system and a key mechanism to extending Kubernetes is implementing a custom resource definition (CRD). A resource created through this feature can be used to store and manipulate information in the Kubernetes system. These custom resources are normally used in combination with a custom Kubernetes controller that interprets the data for the custom resource type contained in the Kubernetes store and then reacts to changes in the data (adds, deletes, modifications). This extensibility via CRDs highlights the fact that at its core, Kubernetes is a declarative based resource management system, and this can be leveraged when implementing a MEC architecture with Kubernetes.

2.2 Multi-cluster Strategies

Enterprises are adopting Kubernetes as a platform to enable application portability and agile deployment across public clouds, private environments, and more importantly on the network edge to optimize local service performance. This is critical for enterprises running retail, hospitality, and manufacturing operations with 100's if not 1000's of locations where application infrastructure is needed to support business to consumer (B2C) and business to business (B2B) applications, as discussed in section 2.8.

Kubernetes supports mechanisms such as pods and name spaces to isolate application components, and ensure resources are allocated optimally within a multi-tenant edge cluster. However, as Enterprise MEC applications proliferate at the network edge, industry trends are starting to emerge to define mechanisms to spread workloads across multiple clusters in different geographic areas. The chief technical reasons for multi-cluster deployments are:

- Lower latency by deploying applications closer to end users
- Service availability with fail-over support and geo-redundancy
- Workload scalability across distinct physical clusters with specialized resources
- Workload isolation & security with physical separation

The main two dimensions of these multi-cluster trends are the *distribution* of an application's resources and the *delegation* of lifecycle control of the distributed application resources, (see Figure 2).

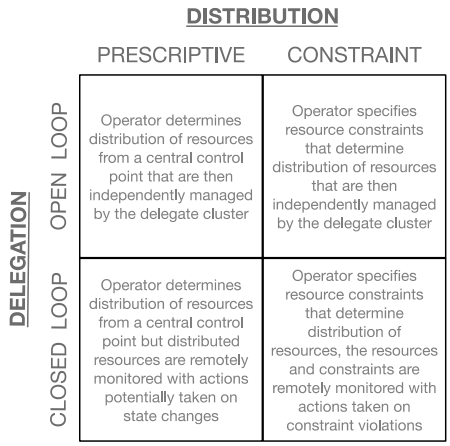


Figure 2 - Multi-Cluster Scheduling Strategies

Distribution of an application's resources refers to how an operator specifies the initial distribution of the resources across the available clusters. Distribution may also reference how an application's resources are redistributed based on a failure or other event. In a *prescriptive* system, the operator specifies the cardinality and location (Kubernetes cluster) for each

application resource. In a *constraint-based* system, the operator specifies the constraints for application resources, such as CPU, memory, network bandwidth, and network latency to an internal application resource or another external MEC application. These constraints are used by a scheduler to determine the optimum placement of the application resources.

Delegation of application resource lifecycle control refers to how the lifecycle of a resource is managed, including initial assignment to a member cluster, and any reassignment to a different cluster based on manual intervention or an event. In an *open-loop* system, once a resource is delegated to a participating cluster, the resource's lifecycle is completely managed by that cluster and will never be removed from that cluster except via an explicit action. In a *closed-loop* system, once the resource is delegated to a participating cluster, a feedback loop is used to monitor the resource and decisions about moving a resource would be based on a defined policy.

Both the distribution and delegation dimensions reflect the level of automation in a multi-cluster system. Systems that fall to the lower right quadrant (see Figure 2) tend to be more autonomous where systems that fall to the upper left quadrant tend to be configuration systems that strictly enact actions in the exact way the operator specifies without any remediation based on failures or resources violations.

2.3 Node/Cluster Capability Discovery

Applications are increasingly looking to leverage available hardware accelerators (GPUs, TPUs, etc.) and software data plane technologies (DPDK, VPP, etc.) to meet their performance requirements. This information is useful to a MEC control plane when placing MEC service components on inter-connected compute nodes. Leveraging Kubernetes and CNCF projects enables the deployment to self-discover a node's capabilities and report or expose those capabilities to a control-plane to be used during scheduling of workloads. Specifically, the CNCF node feature discovery project [55] provides this capability by discovering node features and labeling nodes in a standard format to allow features to be used as part of the standard Kubernetes scheduling capability. This is of critical importance in space & power constrained MEC environments, where full visibility of resource capabilities and programmability of the network infrastructure enable optimal allocation of premium resources.

2.4 Constraint Policy

Kubernetes provides a mechanism for a workload to specify resource constraints that can be used by the control-plane to influence the node selected when scheduling a workload. The existing mechanism is simplistic and predefines only the CPU and memory resource. Kubernetes does allow for other resources to be defined but limits the requested value of those resources to be an integer without a unit specification.

The system we have developed defines a constraint policy model that allows for arbitrary resource constraints to be specified and then leveraged by a custom scheduler as well as de-scheduler [6] extension.

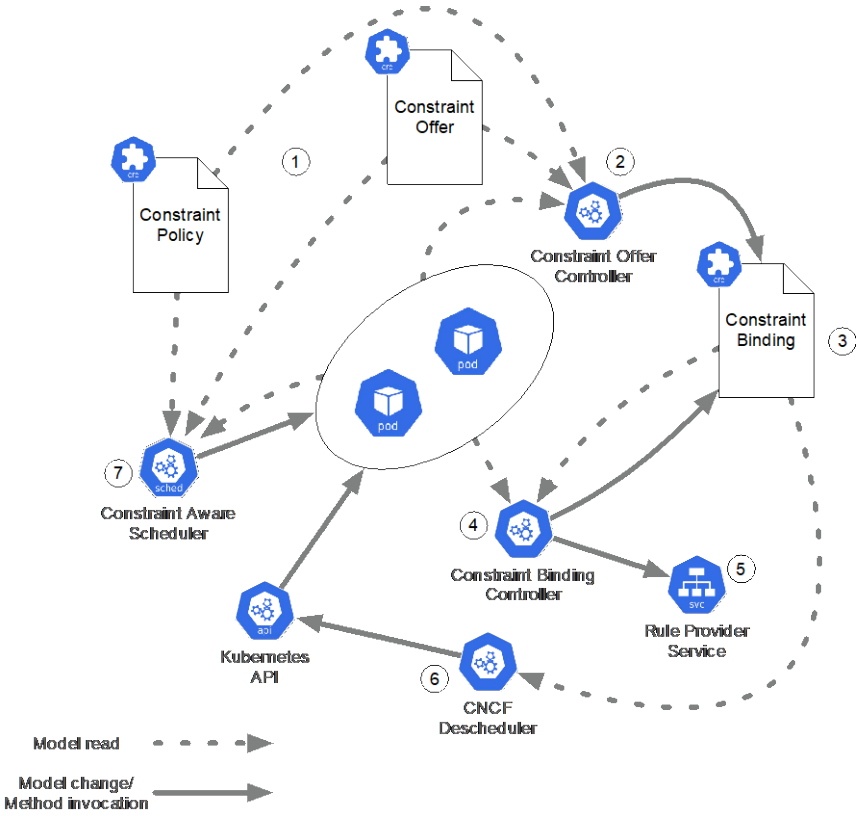


Figure 3 - Constraint Policy Overview

Figure 3 depicts the model and interaction of the constraint policy extension. A constraint policy is a set of constraint rules, where each rule is a tuple of constraint name, constraint request, and constraint limit. The constraint subsystem is designed to be dynamically extensible, and as such the constraint name is a moniker that is used to locate a constraint provider implementation at runtime. The implication is that the system does not pre-define any set of constraints. The constraint request and limit mirror the semantics of the existing resource constraints in Kubernetes in that a request is the preferred value, and the limit is the "worst" allowed value.

A Constraint offer ① is used to associate a constraint policy to one or more workloads (Pods, Services, or NSM network chains in the current implementation). The association is discovered by the Constraint Offer Controller ②, using a selector based on the tuple of Kubernetes ApiVersion, Kind, and Name. For each association discovered, a Constraint binding ③ is created to track the specific policy-workload association including its compliance status. Offers are periodically evaluated and the set of bindings is updated accordingly, deleting bindings that are no longer valid and creating those that now exist.

The Constraint Binding Controller ④ periodically evaluates the policies against the list of bindings leveraging the various provider services ⑤. A provider service is identified via a well-known label based on the constraint name. This allows providers, and thus constraint types, to be

dynamically managed. The compliance of a binding is updated as part of the status value for the binding.

This implementation extends the CNCF de-scheduler project ⑥ to monitor the status of the bindings and when a binding is found to be non-compliant, the de-scheduler may, based on a policy setting, evict the pod via the Kubernetes API ⑦.

2.4.1 Network Connectivity Constraints

Utilizing the constraint policy capability, this implementation defines a set of network connectivity-based constraint providers: bandwidth, latency, and jitter. Using these constraints, an operator can specify the requirements for connectivity between two or more workloads.

By implementing a declarative model for connectivity within Kubernetes, the network becomes part of the overall resource model within the environment, opening new automation use cases, including the ability to declaratively specify constraints that affect the underlay and overlay networks to meet the operator specified application requirements.

2.5 Scheduling Optimization

By default, the scheduling context for Kubernetes is a single pod. During scheduling, Kubernetes selects a node and assigns the pod to that node. Once a pod is assigned to a node, the containers defined within the pod are created and invoked. This can lead to sub-optimal scheduling when constraint policies (see Section 2.4.1) represent a binding between two or more pods as is the case with a connectivity constraint. For optimum scheduling, the entire set of connected pods to be scheduled should be known, and a "plan" should be created such that the pods can be scheduled according to the optimized plan.

To provide this capability, a scheduler extension was developed that operates as an optimized schedule plan builder, as well as a gating function to prevent pods from being scheduled until a trigger is detected. For the initial implementation, we are using a "quiet" timer, but this is easily extendable to support additional trigger types. The quiet timer simply fires when no new pods are defined over a specified period, thus the assumption being that all required pods have been defined and an optimal schedule can be produced.

Before the trigger fires, the scheduler is called repeatedly for each pod. When the planner is invoked, it queries the list of all unscheduled pods and creates a candidate plan, utilizing any specified constraints via the constraint policy resources. If the candidate plan is preferred over the existing plan, the existing plan is replaced by the candidate plan. In either case, an empty node list is returned indicating to Kubernetes that the pod cannot be placed at this time and the pod will remain in a "Pending" (non-assigned) state. After the trigger fires and the scheduler is called, the scheduler uses the plan to determine the node to which to assign the pod. As pods are assigned to nodes, they will be instantiated, and their containers will be created. At this time the connectivity-based constraint policy bindings will be created based on the scheduled pods.

2.6 Network Controller

As described in previous section, the scheduler extension developed as part of this work can leverage the connectivity-based constraints when scheduling workloads.

During development of the solution, while simulating network latency between two nodes, it was noticed that the connectivity-based constraints could not be met by the existing network configuration, causing workloads never to be scheduled, even though the underlying network had the capacity to meet those constraints. As these conditions could exist in a production deployment, especially a multi-cluster deployment, the concept of a network controller was introduced into the solution.

A network controller was represented by a defined interface and could be used by the scheduler to request network resources when the current network configuration could not meet the specified constraints. From the perspective of the scheduler, the network controller is an external entity located by a label on the Kubernetes service resource.

When invoked, the network controller has the flexibility to modify the underlay, overlay network, and/or network slicing to meet the resource request, returning to the scheduler enough information so that the pods that will be created can leverage the modified resources. This information can be used with the network service mesh project to ensure the connectivity to containers by creating the proper network interfaces and configuration on the containers. When the existing network does not meet the constraints and the network controller is not able to modify the network to meet the constraints, the pods will remain in the pending state until the constraints are modified, or the network comes into compliance.

The network controller was then integrated into the de-scheduler capability. When a connectivity-based constraint was found to be out of compliance, rather than immediately evicting the pod for rescheduling, a capability was added to the network controller to mediate this situation via network configuration. If the network controller is not able to bring the network back into compliance, then based on policy, pods may be evicted to be rescheduled or the violation can be ignored to prevent service disruption.

2.7 Common Application Function Model

Telco cloud implementations based on the ETSI network function virtualization (NFV) model have been in production for several years, delivering data and control plane network functions in a much more flexible and software-based format. These virtual machine-based VNFs evolved from the software applications delivered via dedicated hardware appliances for traditional switching, routing, firewall, and signaling services, among others. From a compute environment perspective, these network applications are no different than enterprise or consumer type applications that are delivered from cloud-native environments today, except for requiring specialized hardware assist for packet processing functions. These functions include protocol encapsulation and decapsulation, packet header classification, inspection and manipulation, wire-speed forwarding, encryption, and traffic protection, to name a few. These functions typically require traffic to be steered through multiple functional blocks that make up a network application service chain. The implication is that it is possible to describe a common model for

service function chaining of network application components independently of the specific software logic running within these components. This service function chaining specifies the communication patterns and processing policy for function chaining blocks.

2.7.1 Service Function Chaining

A network function is composed of one or more deployable components. These components can be inter-connected workloads to provide the overall features intended by the network service. The workloads form a chain with traffic flow sequence dependencies, and therefore, should be deployed as a unit on a single compute cluster node for optimal performance. In a hybrid virtualization configuration, a network function could potentially inter-connect hypervisor and container-based workloads on the same cluster node. For this reason, the MEC compliant platform should expose a common network function model to onboard and chain different workload formats.

Although sub-optimal, there may be cases where service function chains span multiple nodes within the same cluster or even multiple clusters separated by an edge network. This may result from constraints imposed on the service chain and the availability of specialized resources such as GPU or smart NICs required to support hardware accelerated functions. In such cases, constraints such as network latency, bandwidth, packet delivery guarantees, and traffic balancing must be taken into consideration when composing the end-to-end service chain through cluster federation mechanisms and network connectivity constraint policy.

2.7.2 Kubernetes Controllers for Function Chaining

When a purpose-built device and associated objects that consume compute resources are not directly modeled by Kubernetes, they can be represented through CRDs and the Kubernetes API can be extended to expose the configuration and capabilities of that new device. A custom controller can then be implemented to manage the lifecycle and translate the resource data into instructions that the target device may understand. Once CRDs and controllers are installed in a Kubernetes cluster, the orchestration of the device can be done through the Kubernetes control plane by abstracting the interactions through the custom device controller.

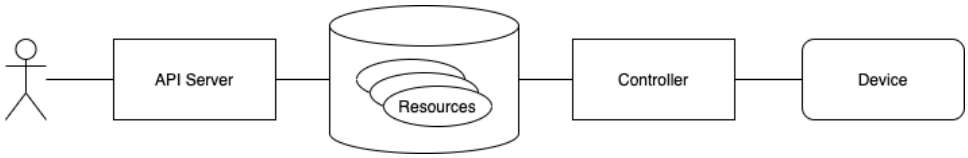


Figure 4 - Orchestrating Device Configuration via Kubernetes CRDs

We used this approach to support the orchestration of VM-based and container-based network functions (NF) on a common operational platform and service chained on a common network layer. This common orchestration framework was achieved by defining models to abstract the network function and chaining complexity. The models are then converted to Kubernetes custom resources with their corresponding controllers. The NF controllers are responsible for orchestrating a VM or container, based on the specified NF type. The chaining controller can

instruct the cluster to establish connectivity between the NFs on the underlying data plane, which may consist of software-based switching and hardware-based traffic processing.

This orchestration walkthrough for VMs and containers is a simplified and high-level view of what really needs to occur. Two scenarios can be considered to configure a system that supports virtualization for both VMs and containers.

1. System with hypervisor engine only with a VM instance running container native constructs, such as Kubernetes, to host containers.
2. System with both hypervisor and container engines to host VMs.

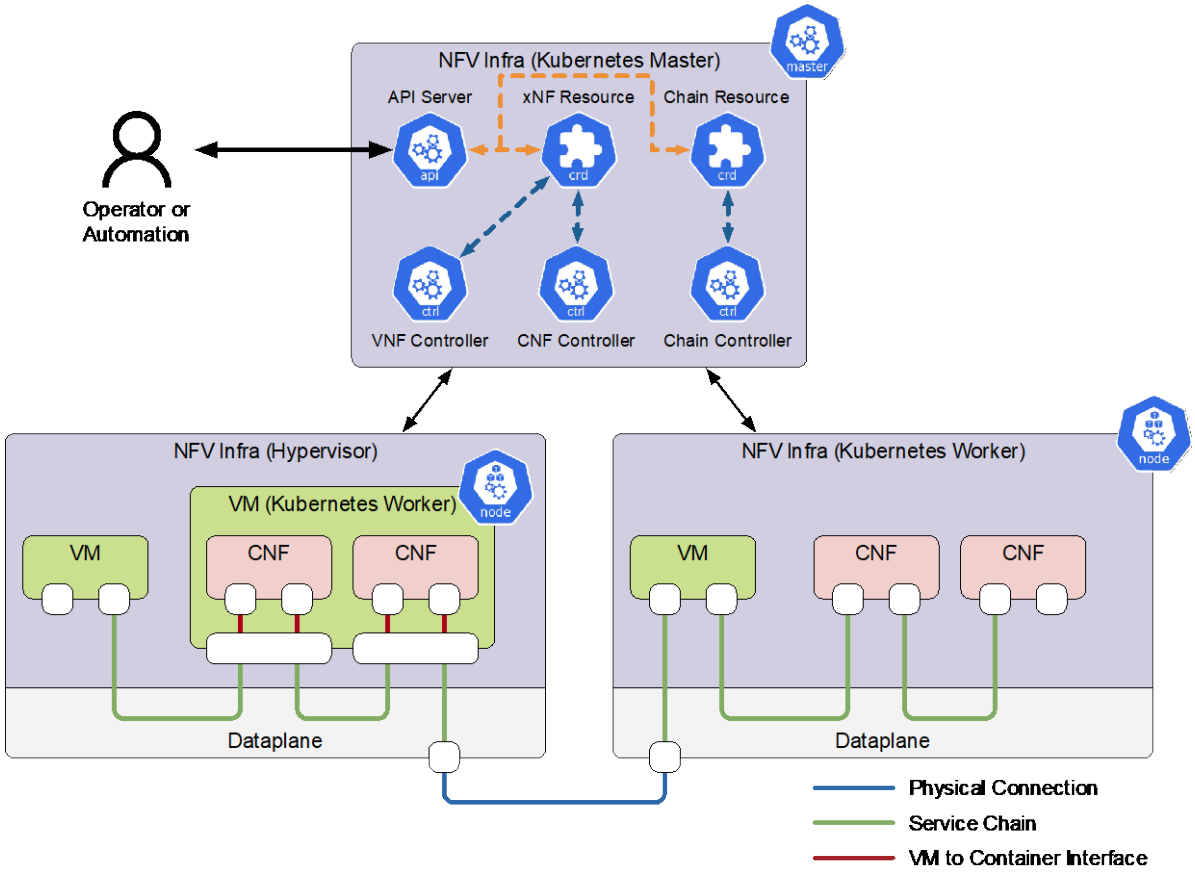


Figure 5 – CRD Driven Chaining

In the first scenario (Figure 5, left side), VMs are orchestrated through the system's hypervisor. A VM running Kubernetes is also used to deploy the container-based network functions. The connectivity between network functions is handled by service chaining the interfaces allocated to the VMs and then linking the containers with the Kubernetes VM interfaces using container networking interface (CNI) plug-ins (for example, Multus). In the second scenario, VMs and containers are orchestrated by their respective virtualization engines. The connectivity between the network functions is handled by service chaining the interfaces allocated to the VM and containers. In both cases, the interface allocation is provided by the data plane embedded in the NFV infrastructure.

In the second scenario (Figure 5, right side), the VM/Kubernetes capable infrastructure is used to create both the VMs and the containers. The connectivity between the network functions handled by using the infrastructure interfaces to chain the VM interfaces across the data-plane and then into the container-based NFs.

2.8 Public/Private Cloud Integration

When architecting a multi-cluster Kubernetes deployment, it is important to understand common industry deployment models. Currently, it is common that network operators and enterprises leverage both their private cloud resources and resources available from cloud hyper-scalers such as Google, Amazon, and Microsoft. Consider a scenario where an Enterprise customer with a large chain of retail stores is planning the introduction of new cloud native applications for inventory management, store security, advertising, and in-store customer engagement. This customer uses one of the major cloud providers to run their own DevOps environment and a national network operator to inter-connect all their stores to MEC locations, private data centers and cloud. A key requirement for the Enterprise IT operations team is to unify the management and delivery of containerized applications to 100's of locations (premise and edge) while maintaining a common network and security policy nationally. As Figure 6 illustrates, this leads to designing a fabric of Kubernetes clusters deployed at many locations, managed through a cloud provider's control-plane and interconnected by a network operator that hosts some of the clusters within the MEC locations.

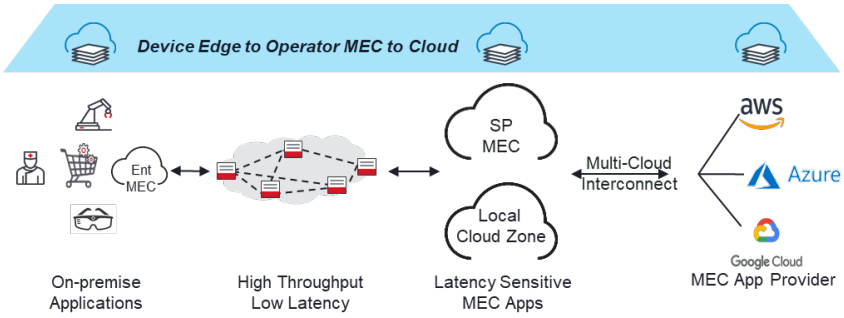


Figure 6 - Enterprise Multi-MEC Applications

While the capabilities described in the sections above can be deployed into Kubernetes clusters under the administrative control of the network operator, it is not always possible to deploy these capabilities on the Kubernetes clusters provided by the hyper-scalers. This is more obvious when it comes to the aggregation or federated level as each hyper-scaler typically provides a custom federation solution as one of multiple tightly integrated cloud-based services, i.e., Google Anthos, Amazon EKS, Microsoft Arc, Rancher, etc.

How these capabilities can be integrated with the various hyper-scalers' offering depends on the amount of customization each allows. In the case where a custom scheduler extension cannot be deployed, this can be "worked-around" by creatively assigning node affinity to resources before

allowing the hyper-scalers scheduler to be activated. Node affinity is a standard Kubernetes capability available on all distributions.

This can be implemented by shifting the capabilities described above, particularly scheduling and de-scheduling, from the Kubernetes domain to a DevOps domain, as depicted in Figure 7. In this situation, the DevOps pipeline would be leveraged such that when an application is pushed to storage ①, the pipeline would evaluate the scheduling needs of the workload and augment the resources with the node assignment encoded as a node affinity configuration ②. Additionally, the constraint aware scheduler may, depending on availability, contact a network controller provided by the network provider ③ to request network capabilities compliant with the constraints specified in the constraint policy. This in turn might trigger the network provider to reconfigure the underlay network ④. If a network controller is provided by the hyper-scaler, then the scheduler may also make requests via that interface which could affect both the overlay and underlay ⑤. After the scheduler updates the manifests and commits those back to storage, the DevOps pipeline receives the augmented manifests ⑥ and pushes the manifests to the hyper-scaler managers ⑦. The hyper-scaler managers process the manifests using their standard schedulers ⑧, adhering to the standard affinity rules, and enact the set node assignment ⑨.

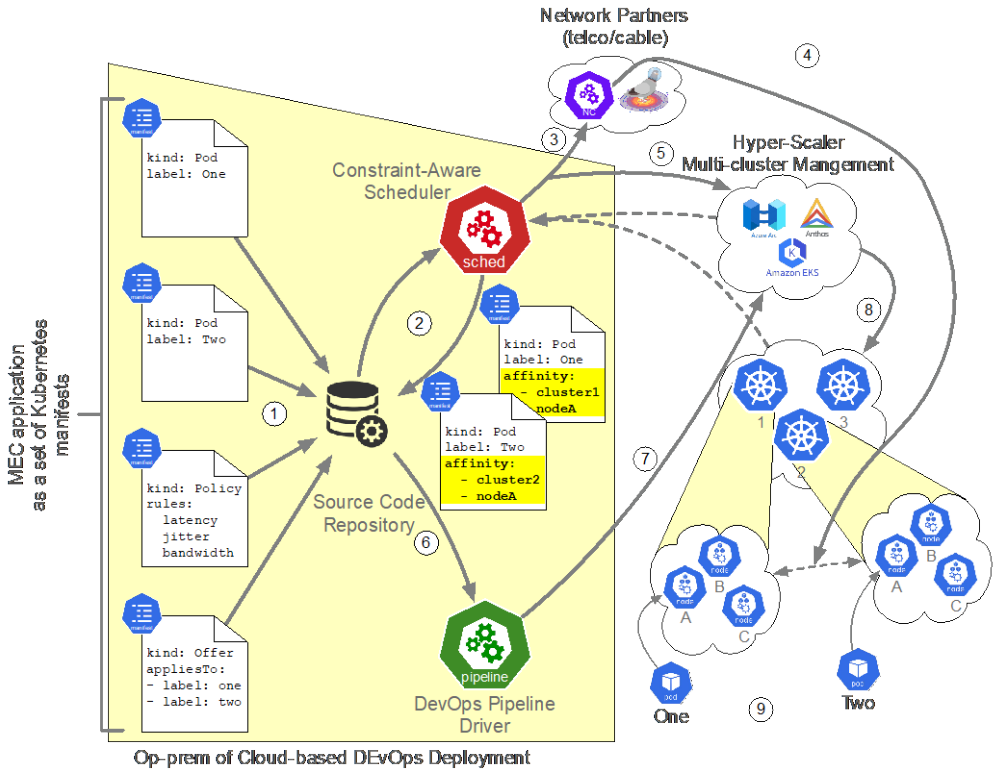


Figure 7 - Using Constraint-Base Scheduling with Public Clouds

Other than the scheduler extension, the described technologies should be able to be leveraged within a hyper-scaler's environment as the other technologies either are common user-based extensions (CRDs + controllers) or components that run outside the core Kubernetes control-

plane (de-scheduler). The one exception is the network controller, which may require support from the hyper-scalers to support underlay control, although it is possible to implement a network controller that only affects the overlay.

3 MEC Architecture on Kubernetes

In this section, we describe how the MEC architecture can be implemented using the de facto industry standard container orchestration system originally developed by Google, i.e., Kubernetes. Figure 8 depicts the standard MEC architecture on the left and on the right depicts that same architecture with an overlay that indicates the cloud native technologies that can be leveraged to implement the MEC architecture. The following sections will detail how each component of the MEC architecture can be implemented using specific cloud native technologies.

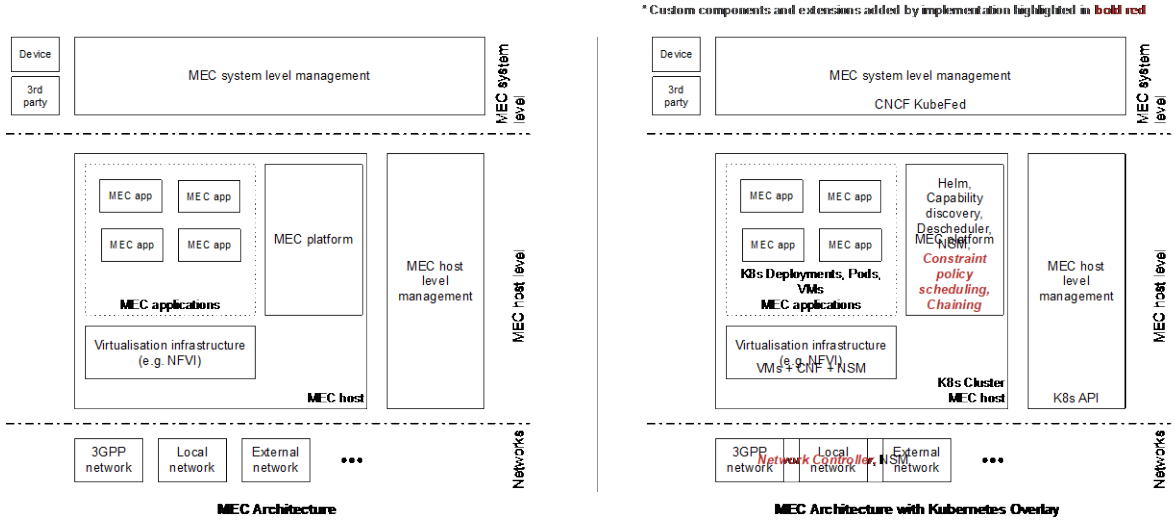


Figure 8 - MEC Architecture with Kubernetes Overlay

3.1 MEC Host

A MEC host is defined as “an entity that contains the MEC platform and a Virtualisation infrastructure which provides compute, storage and network resources for the MEC applications. The Virtualisation infrastructure includes a data plane that executes the traffic rules received by the MEC platform and routes the traffic among applications, services, DNS server/proxy, 3GPP network, other access networks, local networks and external networks.” [1]. By this definition the MEC Host is functionally equivalent to a Kubernetes cluster, which is defined as “A set of worker machines, called nodes, that run containerized applications” [2] The Kubernetes cluster provides the virtualisation infrastructure and data plane as required by the MEC definition. While Kubernetes’ original focus was orchestration of containers, several virtualizations extentions have been added to Kubernetes to provide a run-time for virtual machines and networking.

3.1.1 MEC Virtualization Infrastructure

The virtualization infrastructure of a MEC compliant system should have the ability to orchestrate hybrid service deployments where VMs and containers can coexist on the same platform. To achieve this hybrid configuration, there is a need to accommodate VM based network functions within Kubernetes. A system with such capabilities would have to share resources (compute, memory, storage, networking) to offer a seamless integration with the hosting platform. KubeVirt is one of several projects of the Kubernetes ecosystem that can manage the lifecycle of virtual machines within a Kubernetes cluster while also supporting container workloads.

Additional solutions and platforms exist that provide VM/container capability through the Kubernetes declarative model solution. Further, some of these solutions provide a tight integration with the network interfaces such that they are purpose built to support VM and container-based network functions.

Today the declarative models used to define VM based resources vary across the available solutions and most focus primarily on the detailed attributes for creating a VM and less on the concept of chaining NFs. The solution proposed in section 2.7 bridges this gap by allowing the specification of network services that can contain both VMs and container-based NFs, abstracting away the specific virtualization choice and focusing on the connectivity between those functions. Further, this approach can be extended in the future to support additional virtualization techniques and/or new infrastructure as it is released by vendors.

3.1.2 MEC Applications

A MEC application "runs a virtualized application … on the infrastructure provided by the MEC host" [1]. Within Kubernetes, the typical executable workload is known as a pod, which is a set of containers run on a single Kubernetes node that share storage and networking. As shown above, with the use of CRDs, Kubernetes can be extended such that a workload may be either a pod (container) or a VM.

Kubernetes provides several "higher" level resources constructs that help the operator group and deploy the basic building blocks of an application. These include a Deployment, which is a set of distinct pod definitions and the cardinality for each of the pod types, as well as a ReplicaSet, which maintains a stable set of pod instances for a single pod definition. In addition to pods and other deployment constructs, Kubernetes also provides mechanism to enable load-balancing and high availability for applications.

These basic building blocks provided by Kubernetes provides the basis on which MEC compliant applications can be built. Because an application typically requires more than a single Kubernetes resource, a higher-level application construct can be created using the CNCF Helm [3] tool. This abstraction allows a MEC application developer to specify any number of Kubernetes resources as a set and then deploys that set of resources under a single name, thus allowing a complete application to be deployed instead of dealing with the applications piecemeal.

At its core, Helm is a template engine that create instances of resources based on defined templates, parameterized Kubernetes resource definitions, substituting configurable values for the required parameters. Templates can be core Kubernetes resources as well as CRD defined resources, thus providing access to the full resource model.

3.1.3 MEC Host level management

As described above, the MEC virtualization infrastructure capability can be provided by Kubernetes with extensions to support VMs. Through the defined NF CRDs, both VM and container-based capability can be specified and deploy via a common abstraction. Once deployed, the Kubernetes control-plane will monitor the lifecycle of the resources accounting for scalability and high availability. Additionally, Kubernetes provides a security and network infrastructure to support application deployment.

With the additional of the connectivity-based constraints, scheduler extensions, de-scheduler, and network controller, Kubernetes provides the base capabilities required of MEC host level management.

3.1.4 MEC Host Level Scheduling

MEC host level scheduling is the equivalent of scheduling on a single Kubernetes cluster. The previously described technologies (constraint-based scheduler, optimized scheduler, de-scheduler, and network controller) work in concert to provide the scheduling of workloads to nodes.

3.1.5 MEC Host Level Networking

While a single Kubernetes host provides basic MEC host networking capabilities, through the use of add-on capabilities such as the network service mesh (NSM), additional MEC host (or Kubernetes intra-cluster) networking can be leveraged. A key consideration when deploying a NSM into a Kubernetes cluster is the ability to declaratively define the network connectivity such that there is a separation of concerns between the development of the application and the deployment of the application., i.e., the expected connectivity should not be “baked” into the application code and instead be left to deployment (declarative) configuration. This can be achieved with the NSM implementation.

A network function deployed within a MEC host must focus on serving its intended purpose and should remain unaware of any chaining requirements with other network functions. The Network Service Mesh framework (NSM) in the Kubernetes eco-system fills the role of creating chains and managing the assignment of a network function within a chain. It does so by augmenting orchestrated network functions with a sidecar container and controlling the interactions between the sidecars. The NSM manager can then implement the desired topology by establishing links between sidecars through the NSM data plane.

3.2 MEC System Level Management

The Cloud Native Computing Foundation (CNCF) has defined a special interest group, the Multi-cluster Special Interest Group (SIG), whose charter specifies that this SIG focuses on "solving common challenges related to the management of multiple Kubernetes clusters". [4] As indicated above, if a MEC Host represents a single Kubernetes cluster, then the MEC system level management is meant to manage multiple MEC Hosts and thus multiple Kubernetes clusters.

The CNCF Multi-cluster SIG facilitates the development of a solution for deploying workloads across multiple Kubernetes clusters known as "KubeFed". KubeFed allows an operator to specify the cardinality and location of workloads that are part of an application. Thus, an operator can deploy an application and prescriptively control which cluster a pod is deployed on and how many instances of that Pod are deployed to that cluster. While this is required, it is not sufficient for an autonomous MEC system that can deploy MEC services based on their compute, storage, network, and other resource constraints. Sections 2.4, 2.5, and 2.6 describe a constraint policy extension to Kubernetes that can be applied to a multicluster Kubernetes deployment to provide the capability to deploy MEC services across multiple MEC hosts based on operator specified constraints.

3.2.1 Why Multi-MEC Host is needed

In modern deployment architectures, a single MEC host is not always sufficient to meet the MEC service requirements for latency and/or performance. MEC services will be designed around network and performance bottlenecks, but these designs cannot always compensate for the limitations imposed by the constraints of a single MEC host.

To truly meet the requirements of modern and near future MEC services, deployments must take advantage of multiple MEC host deployments where some of the MEC hosts may be "network close" to the end client with lessor compute power, commonly called edge, and other hosts may be "network distant" with greater computer power.

It is important when deploying a MEC application across multiple MEC hosts that the MEC application is not "topology aware" in that it is not aware of the network location of the compute nor the network on which it is deployed. Instead the MEC applicaiton must specify the constraints it requires and allow the "MEC control-plane" to allocate resources to meet the specified constraints. Providing this separation of concerns between the MEC application and the MEC control-plane allows operators to better align their resources and provide the expected quality of service (QOS) to their clients.

3.2.2 MEC and Edge Computing

Edge computing is the delivery of computing capabilities to the logical extremes of a network to improve the performance, operating cost and reliability of applications and services to the user of the services.

As ETSI GS MEC 003 document [1] states, “Multi-access Edge Computing enables the implementation of MEC applications as software-only entities that run on top of a Virtualization infrastructure, which is located in or close to the network edge.”

While this means that MEC can define a network edge capability, it is also true that MEC can define a metro or central data center. In the context of MEC, any Kubernetes cluster is considered a MEC host regardless of the “nearness” to any given client. As such, at the system level, MEC hosts create a mesh of connectivity that can be leveraged by users that deploy MEC applications.

A MEC *location capability* is simply the MEC host that is “network near” the client of a given MEC application. Thus, any given MEC host may be *at the edge* to some client regardless of its actual location.

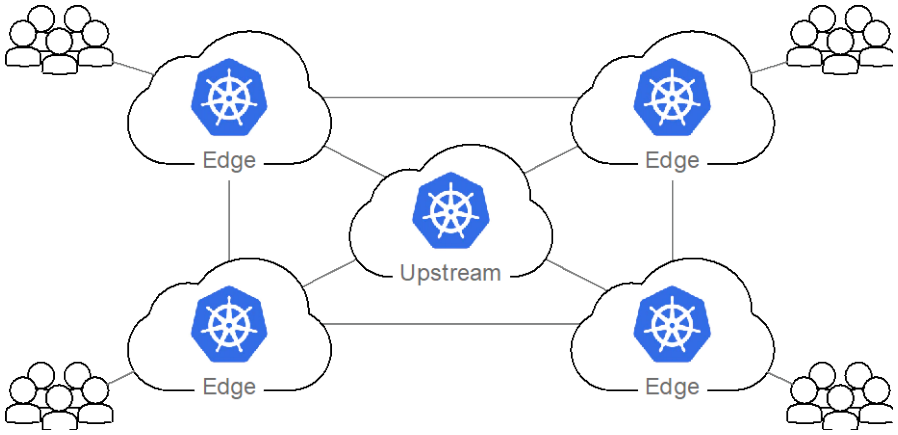


Figure 9 - Depiction of a Kubernetes Cluster Mesh

The illustration above shows an example of how Kubernetes clusters can be inter-connected to represent a MEC host mesh. While some of the clusters are labeled “Edge” or “Upstream” it is important to note that all clusters are functionally equivalent. Where the cluster may differ is in resource capacity or nearness to a given client, but these are operator deployment choices, and an “Edge” could have just as much or more capacity as an “Upstream” cluster.

Based on the above, it is possible to qualify existing or purpose-built MEC hosts as edge clouds for placement of services required by applications that use them.

This edge computing requirement driven optimization of network and compute resources also can be achieved by re-configuration of the underlay connecting MEC hosts.

3.2.3 MEC System Level Scheduling

At the Kubernetes multi-cluster (multi-MEC host level), scheduling is provided via the KubeFed project. At the federation level, the scheduling process changes from scheduling a single pod to a node to delegating or replicating Kubernetes resources to a cluster. The constraint-based

scheduling described above in the context of a single cluster can be applied at the system level with minor additions to the capabilities.

In KubeFed's existing implementation of scheduling, an operator specifies how a resource is federated across the set of member clusters. This includes the specification of the cluster as well as the cardinality of a resource assigned to that cluster. There is a capability to allow the federation to be ratio-based as opposed to completely explicit, but this still equates to a prescriptive federation.

By augmenting KubeFed's scheduling algorithm, as it does not provide the same extension mechanism that base Kubernetes provides, constraint-based scheduling, including connectivity-based constraints, can be achieved. Instead of specifying a cardinality and a cluster, the cardinality and connectivity constraints can be specified allowing the scheduler to place the workloads across the multiple clusters. After the workloads are placed, the constraints can be monitored and upon violation, the resources can be rescheduled within the currently assigned cluster or to another cluster.

3.2.4 MEC System Level Networking

Kubernetes does not provide inter-cluster networking capability natively nor as part of KubeFed project. CNCF provides multi-cluster DNS capability that can be used in a multi-cluster deployment.

Inter-cluster connectivity can be facilitated via the exposing of cluster services via a standard Kubernetes ingress controller or the NSM. Additionally, as part of the scheduling process, a network controller can be used to establish new network paths or modify existing paths.

When using an ingress controller, the services provided through a given cluster are exposed on a public IP address and port. This allows services from other clusters to access these services. The downside of this approach is that it only supports layer 3 (L3), and in some implementations only HTTP connections.

With an NSM implementation, inter-cluster networking can be established through peer to peer connections between NSM managers in each cluster. This allows the establishment of layer 2 (L2) and L3 connections. Further, using sidecars, this connectivity can be declarative, maintaining the SOC between application development and application deployment.

Where a network controller can be integrated, either through a scheduler extension or a DevOps pipeline, new network connections can be established that meet the connectivity-base constraints specified via the constraint policy system. Between the use of the network controller and the NSM complex, inter-cluster networking scenarios can be supported.

4 Summary

This document described extensions and additions to the standard Kubernetes deployment that provide constraint-base, specifically connectivity-based constraints, scheduling of Kubernetes workloads. Additionally, support for VM as well as container-based workloads was introduced,

including chaining of those workloads. How these capabilities can be applied to a single Kubernetes cluster and a federation of Kubernetes clusters was described. Additionally, how these technologies could be applied to non-operator cloud capabilities (i.e., hyper-scaler Kubernetes clusters) was described.

This document then showed how the Kubernetes-based technologies could be deployed to provide an architecture that is compliant with the MEC architecture and how the components of the Kubernetes deployment map to the MEC architecture.

In summary, this document has shown how a MEC compliant multi-host system can be deployed using existing CNCF projects with a few key extensions providing a declarative based, autonomous system for MEC service deployments.

Abbreviations and Definitions

API	application programming interface
AR/MR	augmented and mixed reality
BSS/OSS	business support system / operations support system
B2B	business to business
B2C	business to consumer
CNCF	Cloud Native Computing Foundation
CNF	cloud-native network function
CNI	container networking interface
CRD	custom resource definition
CSP	communication service providers
DNS	domain name service
ETSI	European Telecommunication Standards Institute
HTTP	hypertext transfer protocol
K8s	Kubernetes
L2	layer 2 networking
L3	layer 3 networking
MEC	multi-access edge computing
NF	network function
NFV	network function virtualization
NSM	network service mesh
NFVO	network function virtualization orchestration
MANO	management and orchestration
PAAS	platform as a service
QOS	quality of service
SCTE	Society of Cable Telecommunications Engineers
SIG	special interest group
SOC	separation of concern
VM	virtual machine
VNF	virtual network function

WAN	wide area network

Mixed Reality	Merging of physical and virtual worlds to produce new environments and visualizations, where physical and digital objects co-exist and interact in real time.
Augmented reality	Related to Mixed Reality term, and it takes place in the physical world, with information or objects added virtually.
Edge Computing	The delivery of computing capabilities to the logical extremes of a network in order to improve the performance, operating cost and reliability of applications and services. By shortening the distance between devices and the cloud resources that serve them, by reducing network hops, edge computing mitigates the latency and bandwidth constraints of today's Internet, ushering in new classes of applications. In practical terms, this means distributing new resources and software stacks along the path between today's centralized data centers and the increasingly large number of devices in the field, concentrated, in particular, but not exclusively, in close proximity to the last mile network, on both the infrastructure and device sides.
Edge Cloud	Cloud-like capabilities located at the infrastructure edge, including from the user perspective access to elastically-allocated compute, data storage and network resources. Often operated as a seamless extension of a centralized public or private cloud, constructed from micro data centers deployed at the infrastructure edge. Sometimes referred to as distributed edge cloud. Implementation of these capabilities with Kubernetes clusters is this paper's focus.

References

[1] ETSI GS MEC 003 v2.2.1 (2020-12): *Multi-access Edge Computing (MEC); Framework and Reference Architecture*; European Telecommunications Standards Institute; https://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/02.02.01_60/gs_MEC003v020201p.pdf

[2] Kubernetes Documentation: Glossary (10-AUG-2021); https://kubernetes.io/docs/reference/glossary/?all=true#term-cluster

[3] Helm: project home page (10-AUG-2021); https://helm.sh/

[4] CNCF Multicluster Special Interest group (10-AUG-2021); https://github.com/kubernetes/community/tree/master/sig-multicluster

[5] CNCF Node Feature Discovery (10-AUG-2021); https://github.com/kubernetes-sigs/node-feature-discovery

[6] CNCF Descheduler (10-AUG-2021); https://github.com/kubernetes-sigs/descheduler

[7] ETSI GR NFV-IFA 029 V3.3.1 (2019-11): Network Functions Virtualisation (NFV) Release 3; Architecture; Report on the Enhancements of the NFV architecture towards "Cloud-native" and "PaaS"; European Telecommunications Standards Institute; https://www.etsi.org/deliver/etsi_gr/NFV-IFA/001_099/029/03.03.01_60/gr_NFV-IFA029v030301p.pdf

[8] ETSI GS NFV-IFA 040 V4.2.1 (2021-05): Network Functions Virtualisation (NFV) Release 4; Management and Orchestration; Requirements for service interfaces and object model for OS container management and orchestration specification; European Telecommunications Standards Institute; https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/040/04.02.01_60/gs_NFV-IFA040v040201p.pdf

[9] Global System for Mobile Communications (13-AUG-2021): 5G Operator Platform; https://www.gsma.com/futurenetworks/5g-operator-platform/

[10] Kubernetes Documentation (13-AUG-2021): https://kubernetes.io/docs/home/

[11] Mobile Experts, “Edge Computing for Enterprises” (July 2019); https://mobile-experts.net/Home/Report/1152

[12] Kubernetes Components (13-AUG-2021), https://kubernetes.io/docs/concepts/overview/components/

Improving Pedestrian Safety using Computer Vision, Machine Learning and Data Analytics

A Technical Paper prepared for SCTE by

Parmjit Dhillon
Director Wireless R&D
Charter Communications Inc.
6360 S. Fiddlers Green Circle, Greenwood Village, CO 80111
(303) 793-4465
Parmjit.Dhillon@charter.com

Mohamed Daoud
Director Wireless R&D
Charter Communications Inc.
6360 S. Fiddlers Green Circle, Greenwood Village, CO 80111
(720) 699-5077
Mohamed.Daoud@charter.com

Hossam Hmimy
Senior Director Wireless R&D
Charter Communications Inc.
6360 S. Fiddlers Green Circle, Greenwood Village, CO 80111
(720) 536-9396
Hossam.Hmimy@charter.com

1. Abstract

Pedestrian fatalities are on the rise, with more than 6,000 pedestrians killed each year in the United States.[1] There are several technologies and use cases that can help cities make the roads and highways safer. The Smart Intersection proof of concept (POC) deployed by Spectrum is one such example that demonstrates how cities can use technology for protecting pedestrians.

The Smart Intersection proof of concept uses computer vision, edge and near-edge computing to detect and monitor the pedestrian and vehicle movement at the intersection. The anonymous pedestrian and vehicular traffic data is stored in the cloud for further analysis, planning and design of the components of the intersection, including stop signs, traffic/pedestrian light timing, crosswalks and sidewalks to improve pedestrian safety.

This paper discusses the Smart Intersection architecture, machine-learning (ML) model and computer vision technology. The paper also explains the collection, storage, visualization and analysis of metadata on pedestrian and vehicle movement at the intersection.

2. Introduction

As cities around the world grow at an unprecedented rate, there has been an uptick in traffic accidents. Pedestrians are the most vulnerable users of the roads. As per the National Highway Transport Safety Administration, “pedestrian fatalities in crashes increased 44 percent in the last decade (2009 to 2018), with the pedestrians’ share of traffic fatalities increasing 32 percent, from 13 to17 percent.” [2]

One initiative cities are working on is Vision Zero. The goal is to bring the number of fatalities or collisions in an intersection to zero. Data collected from road sensors and cameras can help cities better understand everything from crosswalk signal timing to traffic patterns that may lead to these collisions.

Data-driven decisions are essential for improving public safety. This paper discusses smart intersection use cases, technology, and the methodology used to collect and analyze a broad array of traffic data to draw actionable insights to make the intersection safer for pedestrians.

In the Smart Intersection POC, we are using several IoT sensors, edge computing, computer vision, wired and wireless connectivity, cloud services for data storage, analysis and visualization to help a city improve pedestrian safety.

3. Computer Vision

Computer Vision is a field of study that trains and enables computers to process images and extract information similar to the human visual system. Recent advancements in algorithms, computational power, and the availability of large datasets of digital images have helped improve computer vision capability.

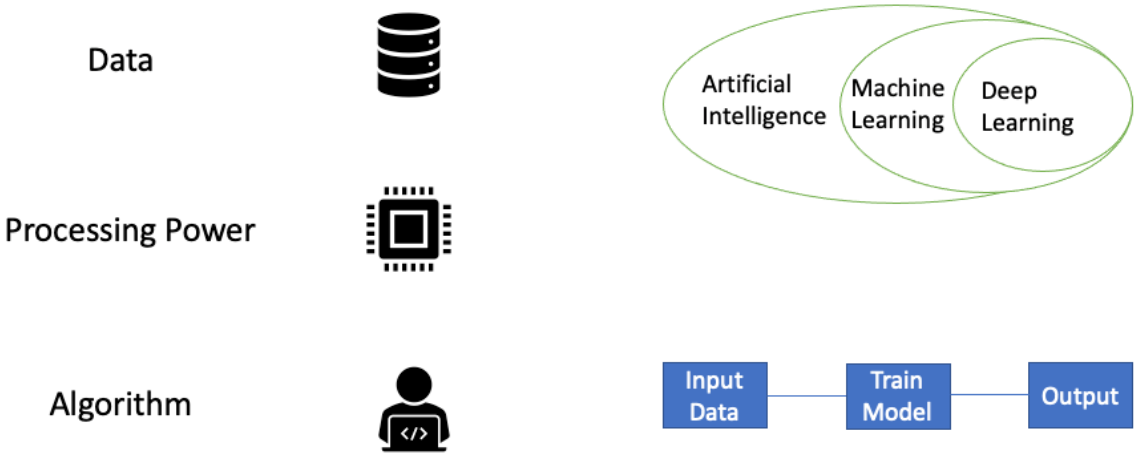


Figure 1 - Computer Vision Components

4. Architecture

The Smart Intersection POC uses cameras attached to the light poles and has the option to use fiber or wireless connectivity to backhaul the data. The switch inside the light pole has multiple Power over Ethernet (PoE) ports providing power to the cameras.

The cameras use edge computing technology with high computing power offered by the graphical processing unit (GPU) to process computer vision. A near-edge server with additional computational power is used for more advanced image processing and machine learning.

The traffic metadata is sent to the cloud for storage, analysis and visualization and can be accessed via a dashboard using a web browser.

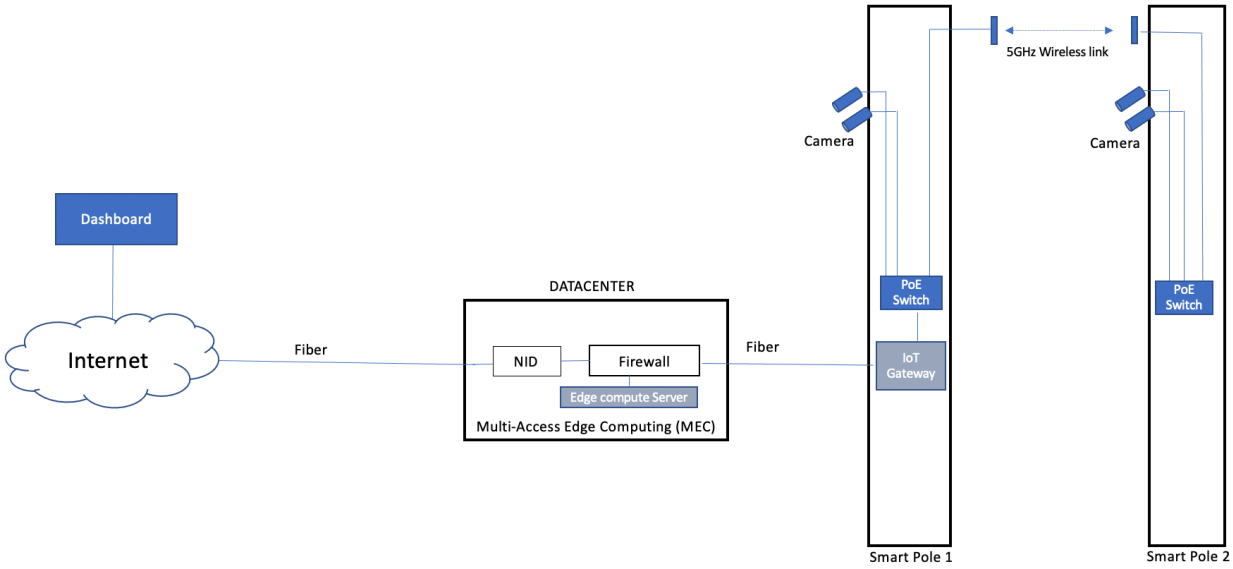


Figure 2 - Architecture

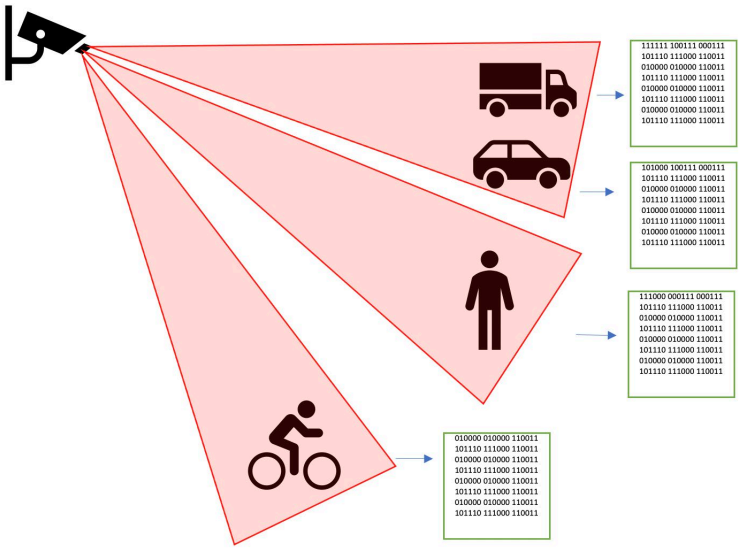


Figure 3 - Digital Image

The Smart Intersection proof of concept uses computer vision and edge computing to detect the presence of a pedestrian. The camera's technology can be used to set the zone within the view of the camera that needs to be monitored for pedestrian activity. The images are anonymized by deleting the image after extracting the metadata and only the metrics are leveraged for analysis. The anonymous pedestrian and vehicular traffic data is also stored in the cloud for further analysis, planning and design of components of the intersection, including stop signs, traffic/pedestrian light timing, crosswalks and sidewalks to improve pedestrian safety.

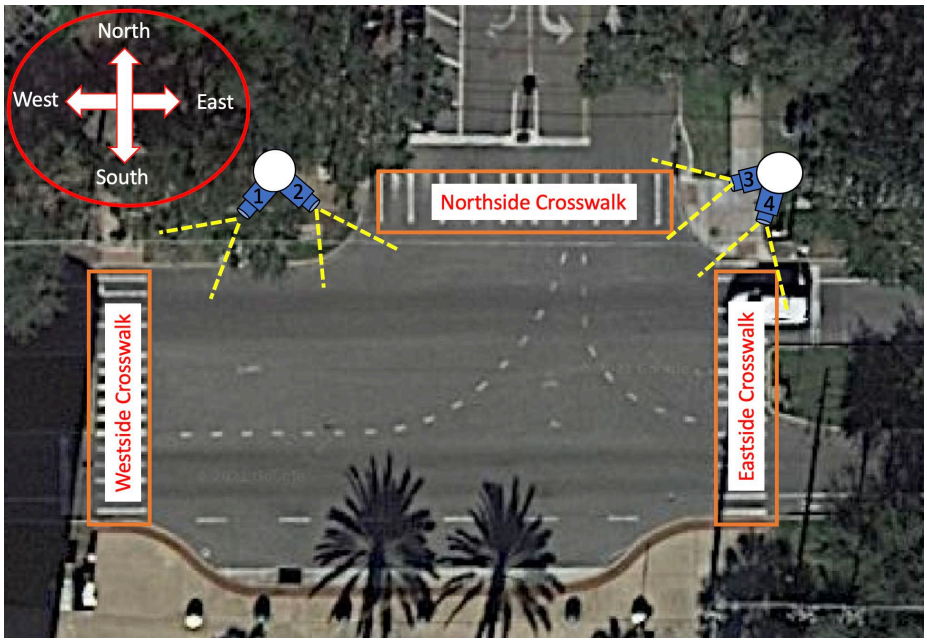


Figure 4 - Smart Intersection

Using computer vision, the cameras are able to detect the presence of pedestrians, bicycles and vehicles at the intersection and also determine the direction of travel.

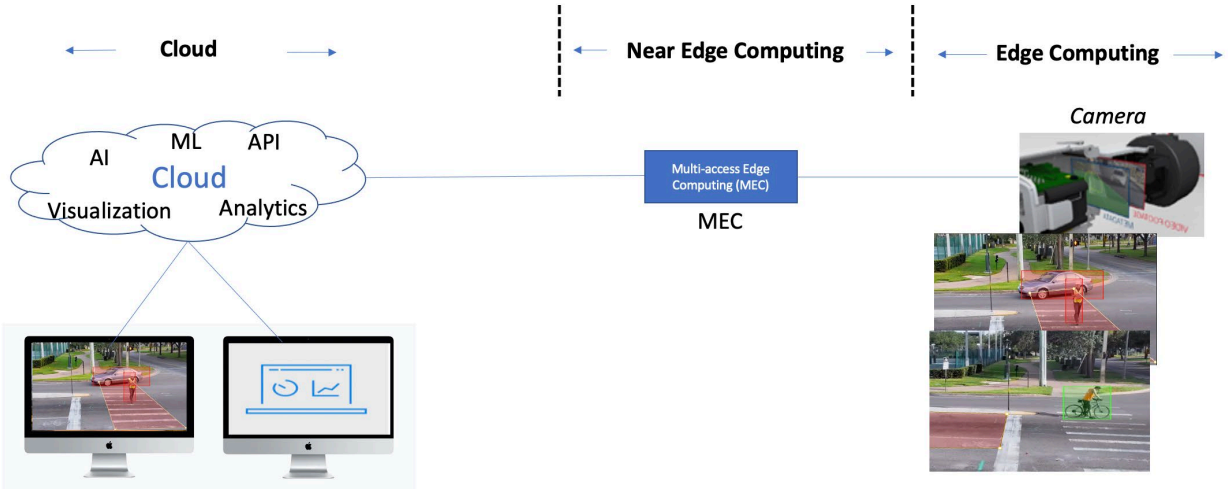


Figure 5 - Computer Vision

The Smart Intersection POC uses a computer vision machine-learning model to train the camera to process complex images such as pedestrians using a walking aid, pedestrians walking with pets or identify different types of vehicles.

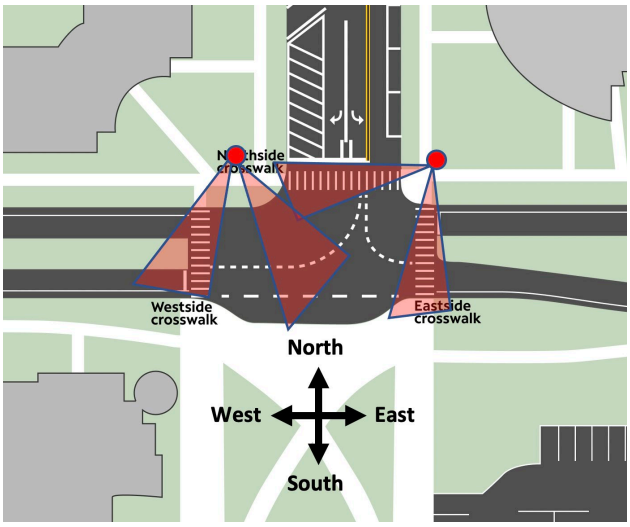


Figure 6 - Camera position

Computer Vision Model Training and Prediction

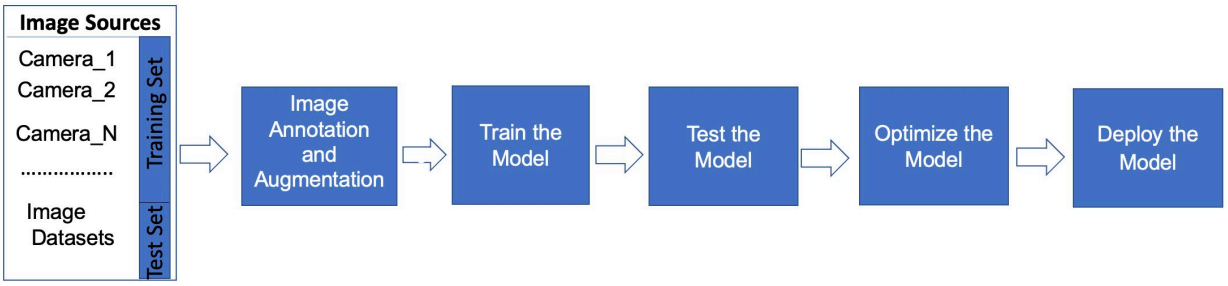


Figure 7 - Computer Vision Training Model

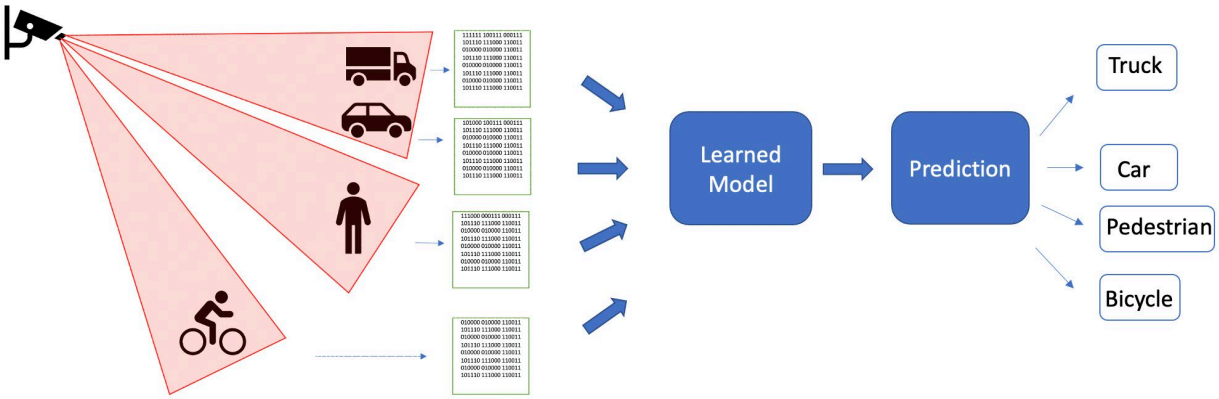


Figure 8 - Learned Model Prediction

5. Data collection and analysis

It is not just the collection of data through the deployment of sensor technology that makes a city smarter; it's also the ability to analyze, to draw insights and make informed decisions. In the Smart Intersection POC, only the metadata is sent to the cloud for storage and analysis. As discussed earlier, for privacy reasons, the image is deleted after extracting the metadata.

When it comes to data, privacy is, of course, very important, especially if the data collected has personally identifiable information (PII). Such data should be handled as per the agreement with the city, and it must also meet regulatory compliance.

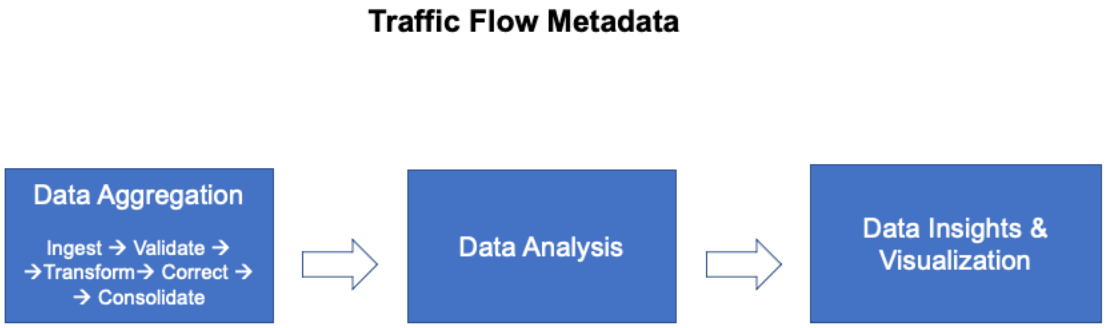


Figure 9 - Data Model

5.1. Pedestrian count data and direction of travel

The camera collects the following metadata on pedestrian movement:

- Number of pedestrians crossing the crosswalk
- Direction of pedestrian movement (e.g., north, south, east, west)
- Jaywalking

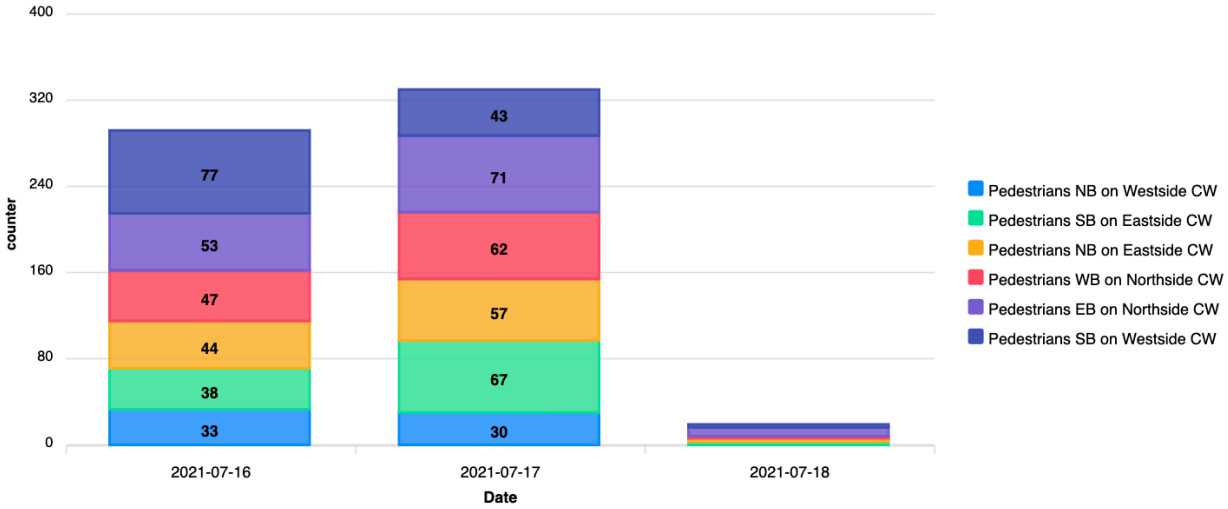


Figure 10 - Pedestrian Traffic Per Day

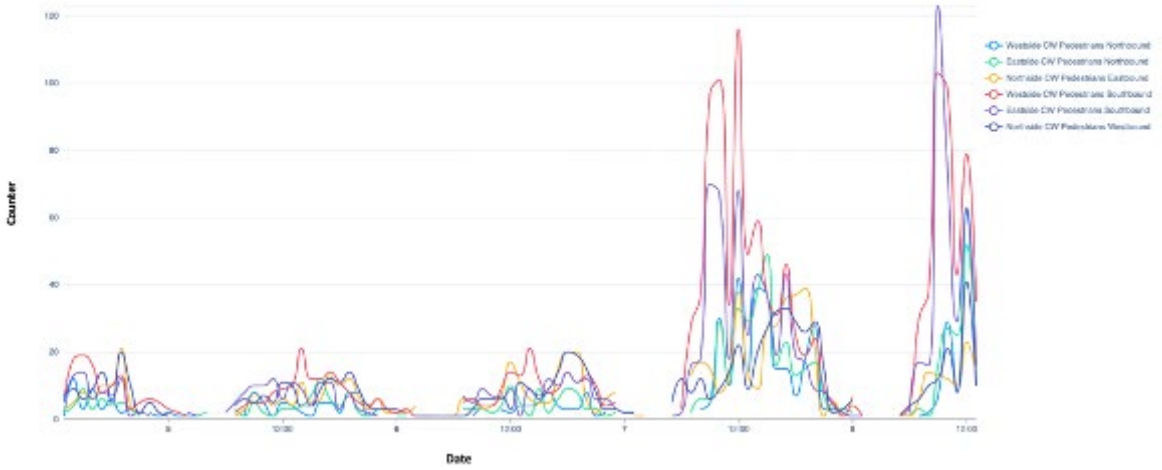


Figure 11 - Pedestrian Traffic Per Hour

The Smart Intersection POC picked up a high number of pedestrians avoiding the crosswalk and jaywalking at the intersection.

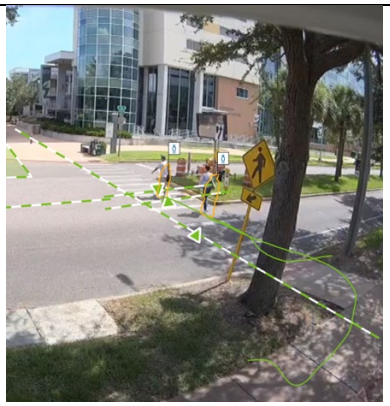
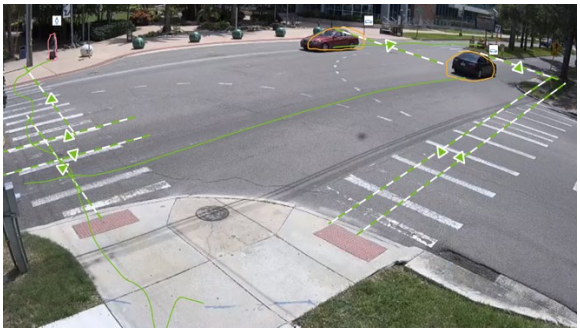


Figure 13 - Pedestrian Image

Figure 12 - Jaywalking

5.2. Vehicle count and direction of travel



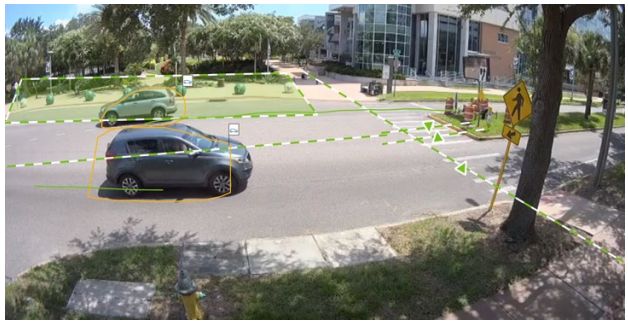


Figure 14 - Vehicle Image

The camera collects the following metadata on vehicle movement:

- Number of vehicles crossing the intersection

- Vehicle making turn at the intersection

- Vehicle making an illegal U-turn

The following data shows more westbound data.

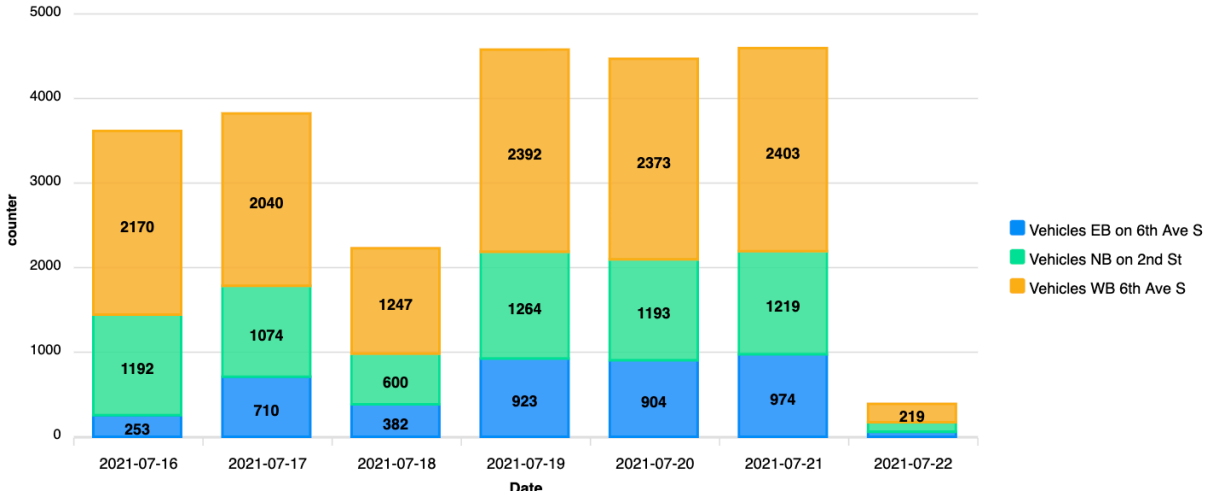


Figure 15 - Vehicle Traffic Per Day

The traffic flow at the intersection shows a consistent and predictable pattern, as shown in the image below. The traffic peaks during a specific time window during the day and shows minimum activity during the night.

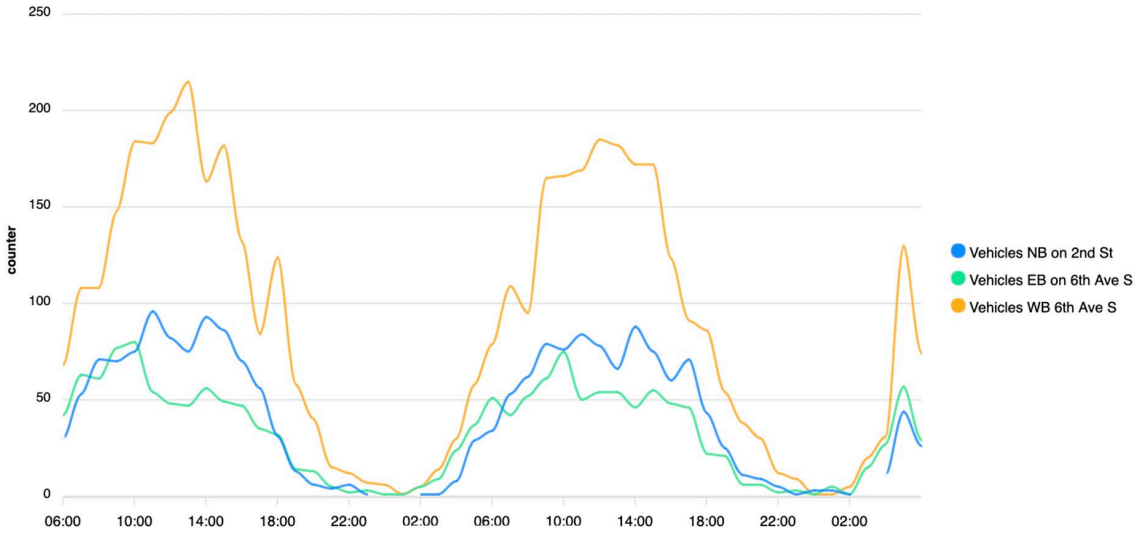


Figure 16 - Vehicle Traffic Per Hour

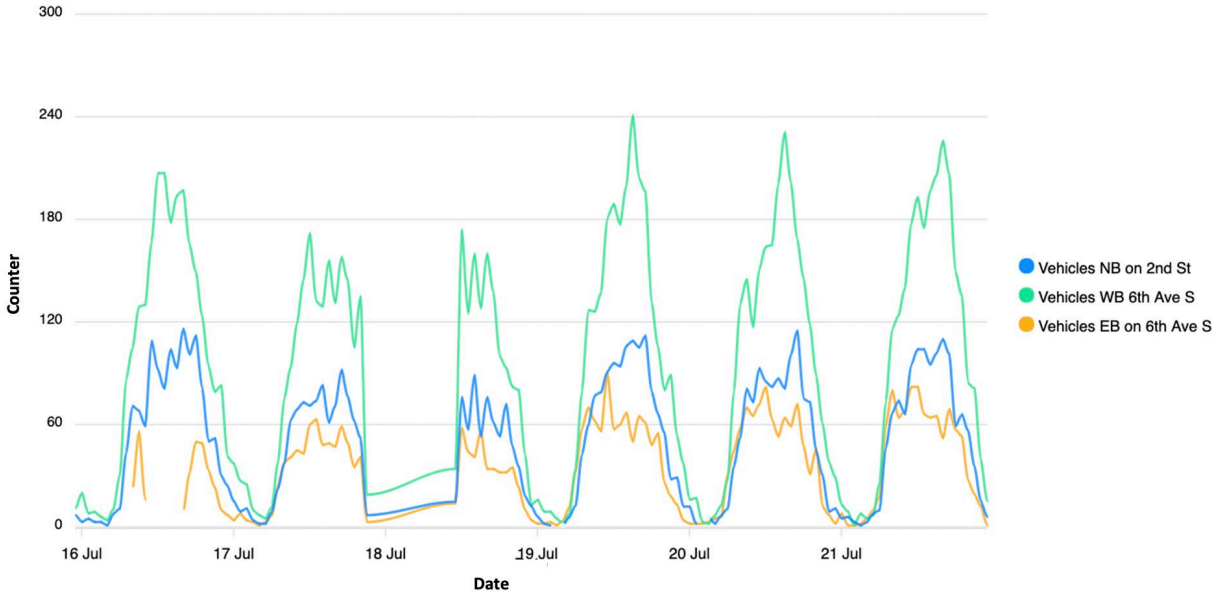


Figure 17 - Vehicle Traffic Per Direction

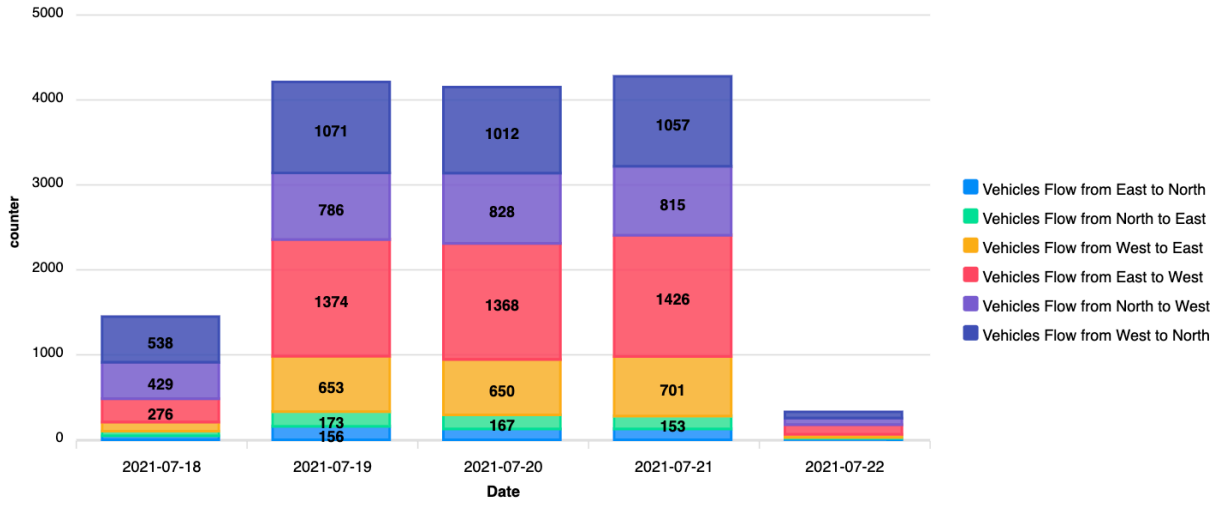


Figure 18 - Traffic Direction of Flow

The Smart Intersection POC even picked up illegal U-turns, putting pedestrians at risk.

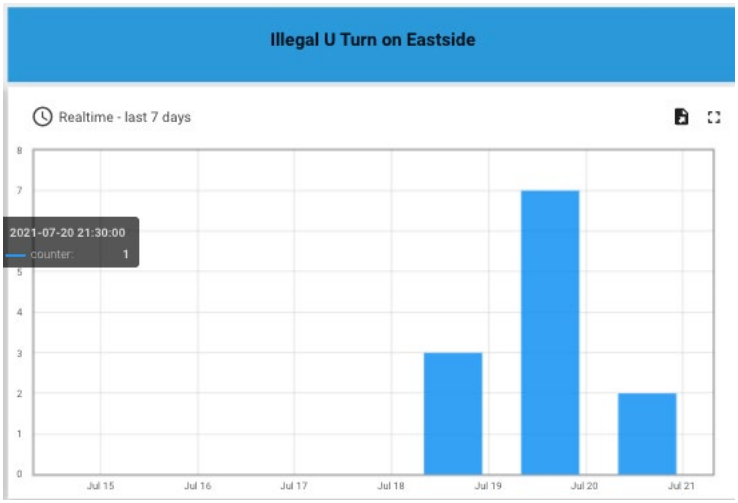


Figure 19 - Illegal U-Turn

5.3. Bicycle count and direction of travel

Figure 20 - Bicycle Image

The camera collects the following metadata on bicycle movement:

- Number of bicycles on the road and crosswalk
- Direction of bicycle movement on the road and crosswalk (e.g., north, south, east, west)

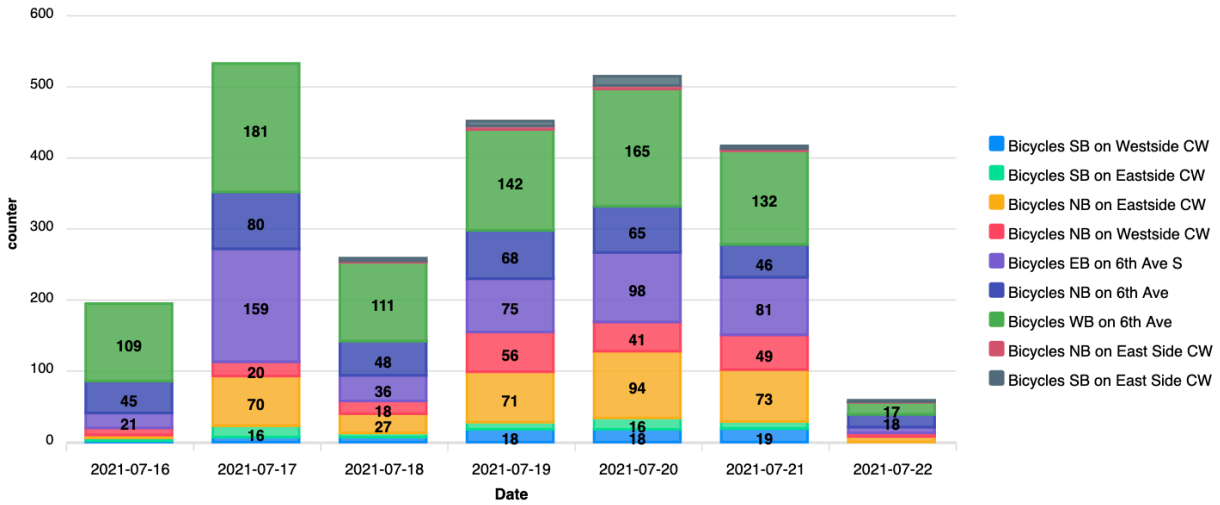


Figure 21 - Bicycle Traffic Per Day and Direction of Flow

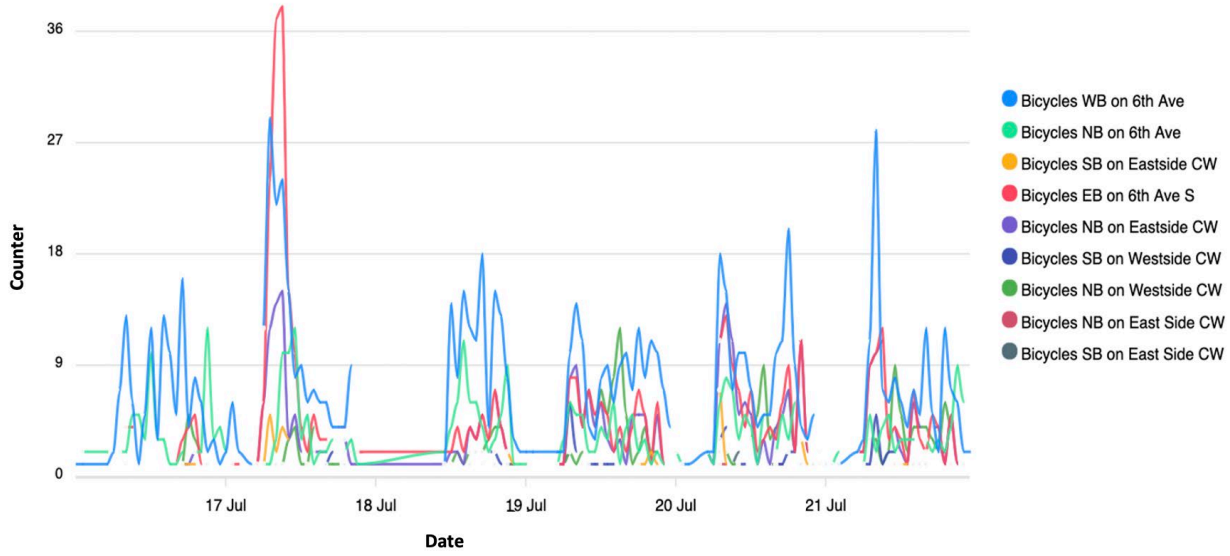


Figure 22 - Bicycle Traffic Per Direction

5.4. Data summary and forecast

This data is currently being used to train and develop a data model to forecast the traffic at the intersection.

P90: The true value is expected to be lower than the predicted value 90% of the time.

P50: The true value is expected to be lower than the predicted value 50% of the time.

P10: The true value is expected to be lower than the predicted value 10% of the time.

Figure 23 - Traffic Forecast

The city is using this anonymized pedestrian, bicyclist and vehicular traffic data to study all traffic activity at the intersection. Analyzing this information is helping the city to draw insights and help city planners redesign intersections to improve pedestrian safety.

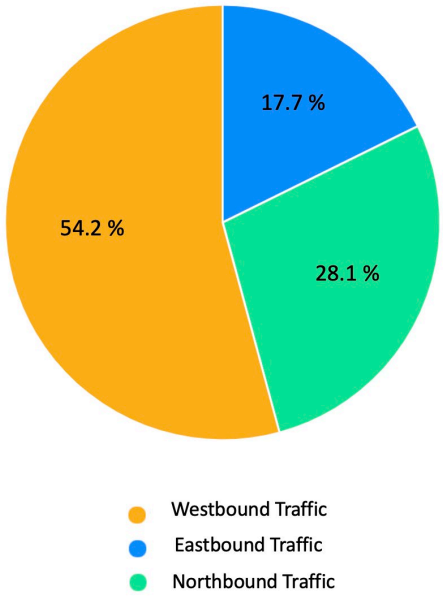


Figure 24 - Overall Traffic Flow by Direction

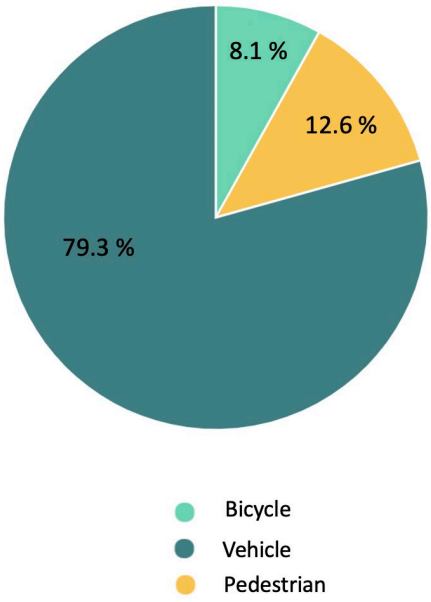


Figure 25 - Percent of Pedestrian, Bicycle and Vehicle Traffic

6. Conclusion

Designing an effective intersection requires an understanding of what activity actually happens within the intersection. The Spectrum Smart Intersection POC exponentially increases that level of understanding over manual methodologies that simply produce a count of pedestrians or vehicles. The POC is collecting, anonymizing, analyzing and visualizing data sets covering pedestrians, bicyclists, vehicles, direction of travel, time of day, counts of accidents and near-accidents and more.

The traffic activity data is helping the city to design and optimize intersections that focus on optimizing pedestrian safety while also streamlining traffic flow through intersections. These types of implementations allow a city to truly emerge and leverage technology on their path to becoming a smart city.

As we collect more data and datasets get richer over time, we expect to draw more insights and forecast traffic. Using the training data sets, we can train the computer vision model and capture additional metadata on the type of traffic at the intersection so that cities can have richer insights and move them closer to the goal of Vision Zero.

Abbreviations

CW	crosswalk
EB	eastbound
IoT	Internet of Things
MEC	Multi-access Edge Computing
NB	northbound
POC	proof of concept
EB	eastbound
PoE	Power over Ethernet
WB	westbound

Bibliography & References

1 “Traffic Safety Facts”, NHTSA’s National Center for Statistics and Analysis, National Highway Transport Safety Administration.

https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813079

2 “Geographic Summary of Pedestrian Traffic Fatalities”, NHTSA’s National Center for Statistics and Analysis, National Highway Transport Safety Administration.

https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813089

Improving Upstream Efficiency

A Technical Paper prepared for SCTE by

Karthik Sundaresan
Distinguished Technologist
CableLabs
k.sundaresan@cablelabs.com

Tom Williams
Distinguished Technologist
CableLabs
t.williams@cablelabs.com

Sheldon Webster
Lead Engineer
CableLabs
s.webster@cablelabs.com

Alberto Campos
Fellow
CableLabs
a.campos@cablelabs.com

Doug Jones
Principal Architect
CableLabs
d.jones@cablelabs.com

CableLabs
858 Coal Creek Circle, Louisville, CO,80027
3036619100

1. Introduction

Cable operators are facing an unprecedented increase in upstream traffic usage because of the shift to working/schooling from home, and the reliability, capacity, and efficiency of the upstream is top of mind for all operators. This paper is focused on improving the efficiency of the DOCSIS upstream (DOCSIS 3.0 SC-QAM and DOCSIS 3.1 OFDMA) channels. The paper will document some of the upstream engineering problems seen by operators and make robust recommendations on how improve those situations. There are thousands of upstream parameter settings and control knobs available to an operator and much of the time operators leave those settings at their default values. This paper will investigate areas (for SC-QAM & OFDMA) such as minislot size, FEC, Cyclic Prefix, FFT size, frame size, pilot patterns, channel size, guard bands, profile definition, CMTS settings for changing IUCs, etc., and make recommendations on each of those parameters and settings. This paper will also explore configuration file service parameters for a cable modem (CM) and make recommendations for better performance on the upstream. Ultimately, this paper will help operators understand optimized upstream configurations and settings for production deployments.

This paper assumes some background understanding of both DOCSIS 3.0 and DOCSIS 3.1 upstream technologies. Here we try to provide an understanding of different upstream channel parameters (which an operator can tweak) and then the pros and cons of choosing different values for that parameter. Based on these tradeoffs we make recommendations on optimal values for each of these channel parameters or network configurations.

2. DOCSIS 3.0 SC-QAM Upstream Recommendations

This section of the paper is a “how-to” on improving SC-QAM upstream efficiency. Access network, outside plant and system engineers would like to understand the benefits and how to realize those benefits as they continue to maintain the DOCSIS SC-QAM upstream technology.

2.1. Minislot Sizes

The DOCSIS bandwidth allocation MAP uses time units of "minislots." On SC-QAM channels, the size (duration) of the minislot is a multiple of the DOCSIS SYNC time ticks (6.25 us). A minislot is the unit of granularity for upstream transmission opportunities and represents the time allowed for CM transmission of a fixed number of symbols. Figure 1 shows the range of values of minislot sizes for an upstream SC-QAM channel, assuming 64 QAM as the modulation order.

To determine the best minislot sizes, one needs to understand how the network upstream traffic behaves and model it accurately. Based on the traffic patterns seen on the upstream one can optimize the DOCSIS minislot size for an upstream channel.

There are overheads involved with different settings of the minislot size. Setting minislot sizes to one of smaller settings can increase the scheduling granularity for the CMTS. Also, for larger grants, the overhead of bytes wasted in the grant will reduce. On the other hand, larger minislot sizes increases the largest grant that can be made to DOCSIS 2.0 modems (which support only one upstream SC-QAM channel). Prior to DOCSIS 3.0 technology modems requested bandwidth in minislots, whereas starting with DOCSIS 3.0 technology, modems (with support for channel bonding) request bandwidth in bytes using the queue-depth based requests. Also, a larger minislot size, may make the computations a bit simpler on the CMTS upstream scheduler and receiver though the benefits may vary by implementation.

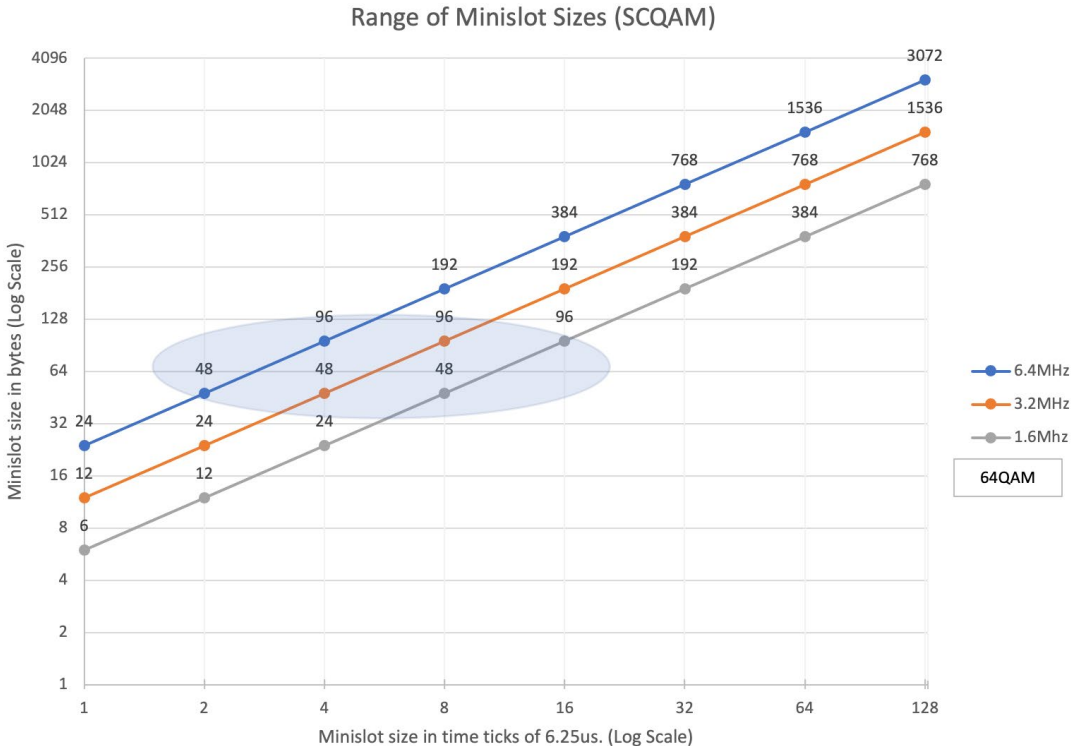


Figure 1 – SC-QAM Minislot Sizes Range

As seen in the graph above, for a 6.4 MHz channel (5.12 MSyms/sec) using 64 QAM modulation, a minislot size of 1 has a raw capacity of 24 bytes. The maximum grant size (a 255-minislot grant) will be 6120 bytes. A minislot size of 4 will have a raw capacity of 96 bytes and allows a maximum 255-minislot grant of 24480 bytes. Based on current patterns of upstream packet sizes, and the size of an Ethernet frame carrying a TCP ACK (64 bytes), the consensus is that minislot sizes in the range of 48 bytes to 96 bytes will enable an appropriate variance in grant capacity and efficiency for upstream scheduling. This translated to a minislot size configuration of 2 or 4 for a 6.4 MHz channel or a size of 4 or 8 for a 3.2 MHz channel.

2.2. FEC Settings

DOCSIS specifications allow for a range of FEC settings for an upstream SC-QAM channel. This is enabled by two settings in the upstream channel which have the following range restrictions:

- FEC codeword information bytes (k), 16 <= k <= 253
- FEC codeword correction setting (T), 0 <= T <= 16.
    - The number of parity bytes is 2T and ranges from 0 bytes to 32 bytes
    - Thus, the total codeword length (k+2T) would be in the range of 16 to 285 bytes

A short data grant may use FEC parameters that are appropriate to shorter packets while a long data grant may be able to take advantage of greater FEC coding efficiency. FEC codeword lengths for Request, Request/Data, and Ranging IUCs can be shorter while the codeword lengths should be longer for Short Data grant (IUC 5) and Long Data grant (IUC 6) and UGS grants (IUC 11).

The FEC parameter selection is a trade-off between channel utilization and robustness to noise. In practice a codeword efficiency of 75%-to-90% (i.e., 10%-25% FEC overhead) looks to be the optimum

for a DOCSIS upstream SC-QAM channel. An operator can use the lower end of that range for noisier channels (e.g., lower in the cable upstream spectrum) and the higher end of the FEC efficiency range for cleaner channels. Of all the possible values of k from 16 to 253, many k & 2T values lead to a codeword efficiency which fall outside this optimal range. The below graph (Figure 2) shows the zone (gray) of these optimal values.

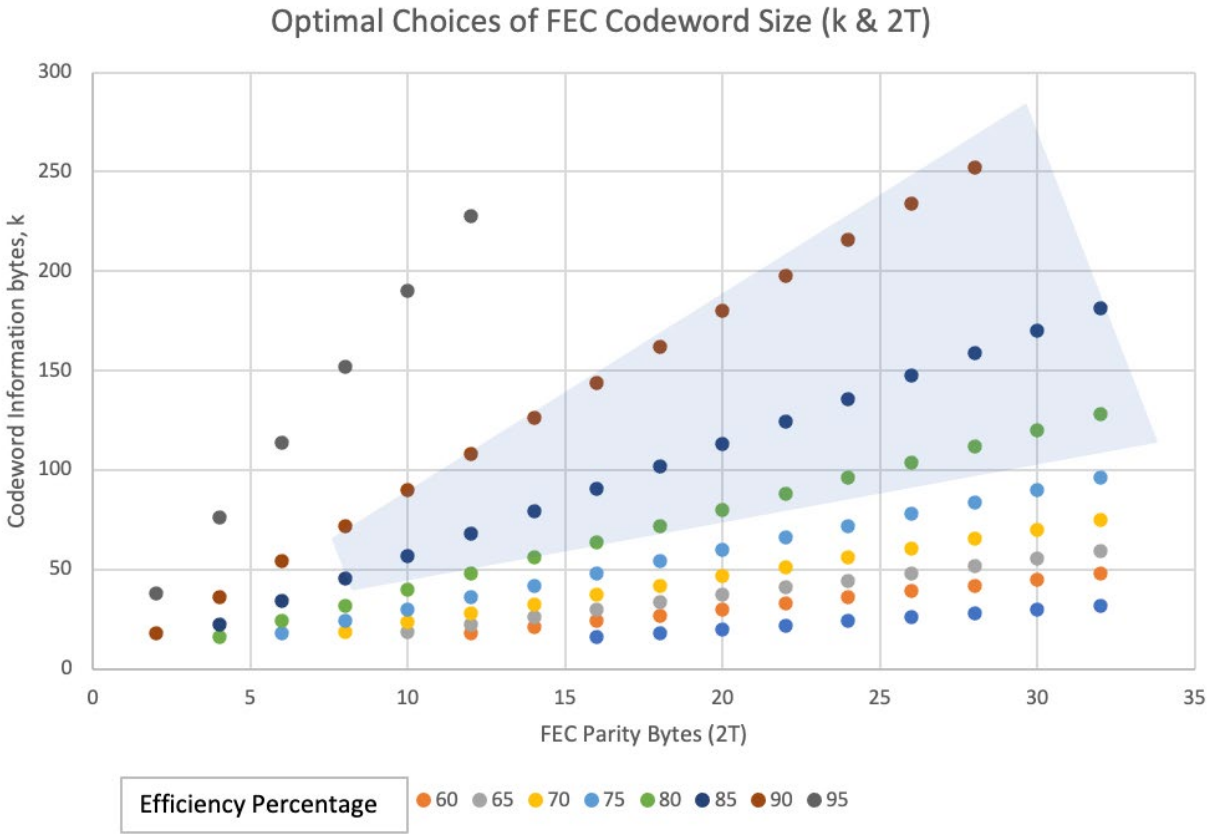


Figure 2 – SC-QAM FEC Sizes Range

Table 1 shows the range of k values (information bytes) for various choices of 2T (FEC Parity bytes). Combinations of k & 2T which give below 50% codeword efficiency are ignored, as they are not useful in the normal cable plant. Shown in green are the range of k values which fall in the 75-90% efficiency range and are the meaningful codeword sizes to choose for Data IUCs.

Ideally an operator would read the RxMER numbers for each of the CMs using the upstream channel and based on those would choose an appropriate FEC setting (k & 2T) and vary the modulation order for that IUC.

Table 1 – Choosing values of k and 2T for SCQAM FEC

Information bytes (k)	FEC bytes, 2T															
Codeword Efficiency	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
50								16	18	20	22	24	26	28	30	32
60						18	21	24	27	30	33	36	39	42	45	48
65					19	22	26	30	33	37	41	45	48	52	56	59
70				19	23	28	33	37	42	47	51	56	61	65	70	75
75			18	24	30	36	42	48	54	60	66	**72**	**78**	**84**	**90**	**96**
80		16	24	32	40	48	56	64	**72**	**80**	**88**	**96**	**104**	**112**	**120**	**128**
85		23	34	45	57	68	**79**	**91**	**102**	**113**	**125**	**136**	**147**	**159**	**170**	**181**
90	18	36	54	**72**	**90**	**108**	**126**	**144**	**162**	**180**	**198**	**216**	**234**	**252**		
95	38	76	114	152	190	228										

2.3. Interleaver settings

The DOCSIS SC-QAM upstream interleaver supports two modes, one operating mode in which the block size is fixed, and a second dynamic mode in which the interleaver depth is determined based on the burst size. The FEC encoded data bytes of the packet are first divided into interleaver blocks. In the fixed mode, the interleaving depth of the last interleaving block of a packet can be small, resulting in low burst noise robustness for this block. In dynamic mode, the depths of the interleaver blocks are chosen such that all blocks have approximately the same depth to achieve nearly optimal burst noise robustness (for the given block size). The current point of view is to choose "dynamic mode" and let the system decide the appropriate interleaver settings. An MSO who uses data analytics on the channel performance metrics could potentially outperform the CMTS but until operators have such software systems built, the dynamic options look to be the best.

2.4. Upstream performance

This section provides guidance on the operational conditions and recommends which modulation orders are recommended for various levels of SNR to have an FEC error-free operation. The ability of the system to support a given QAM level depends on the RxMER values and the mappings to an appropriate QAM level when creating a profile. These mappings are used for DOCSIS 3.0 SC-QAM channels and are summarized in Table 2 below.

Table 2 – Upstream RxMER to QAM Level mapping

Upstream Constellation/ Bit Loading	Upstream MER (dB)
QPSK	10
8 QAM	14
16 QAM	16
32 QAM	19
64 QAM	22+

3. DOCSIS 3.1 OFDMA Upstream Recommendations

This section of the paper is a "how-to" on improving OFDMA efficiency. Access networks, outside plant and system engineers would like to understand the benefits and how to realize those benefits as they plan to deploy DOCSIS 3.1 OFDMA technology.

3.1. OFDMA Channel Location

Many operators have shared data indicating the spectrum below 20 MHz is burdened by the presence of ingress noise as shown in Figure 3. In some plants which are well maintained and cleaner, operators have successfully reported running SC-QAM channels down to 9 MHz. In some cases, if the plant is clean, the spectrum may be amenable to the presence of an OFDMA channel in those lower frequencies. In many cases if the spectrum contains bursty or impulsive noise ingress, then we do not recommend placing OFDMA channels in that region. A few operators have reported OFDMA codeword errors leading to system stability issues with the OFDMA channel in spectrum below 20 MHz, whereas these issues become non-existent when the channel is moved to higher portion of spectrum.

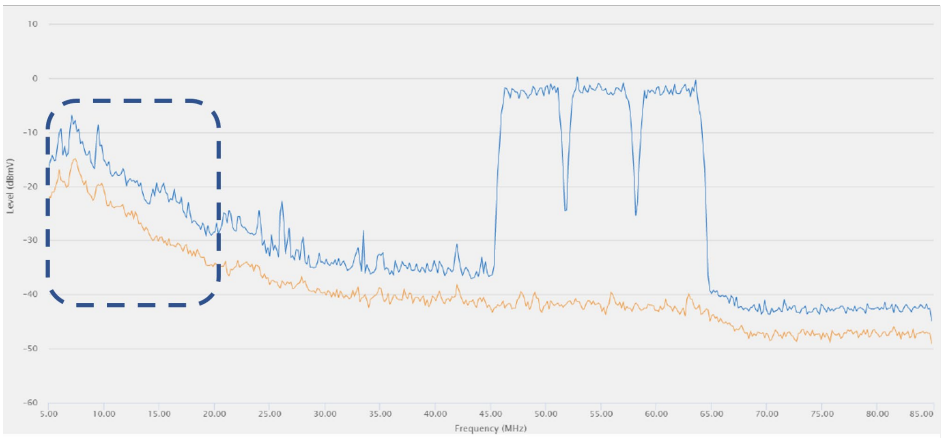


Figure 3 – Noisy Spectrum Below 20 MHz

When initially deploying OFDMA channels it would be prudent to initiate the OFDMA trial in a cleaner (higher) part of the spectrum. The idea is to work out the inevitable system issues and bugs and get the OFDMA channel working reliably first in a cleaner part of the spectrum and then evaluate OFDMA below 20 MHz.

3.2. Number of symbols in a frame (K)

The number of symbols in an OFDMA frame, K, is configurable between 6 (minimal value) and one of the following values:

- With 20 µs FFT duration (2K FFT)
  - Kmax = 18 for BW≥ 72 MHz
  - Kmax = 24 for 48 MHz ≤BW < 72 MHz
  - Kmax = 36 for BW < 48 MHz
- With 40 µs FFT duration (4K FFT)
  - Kmax = 9 for BW ≥ 72 MHz
  - Kmax = 12 for 48 MHz≤ BW < 72 MHz
  - Kmax = 18 for BW < 48 MHz

Figure 4 shows the layout of minislots within an OFDMA channel.

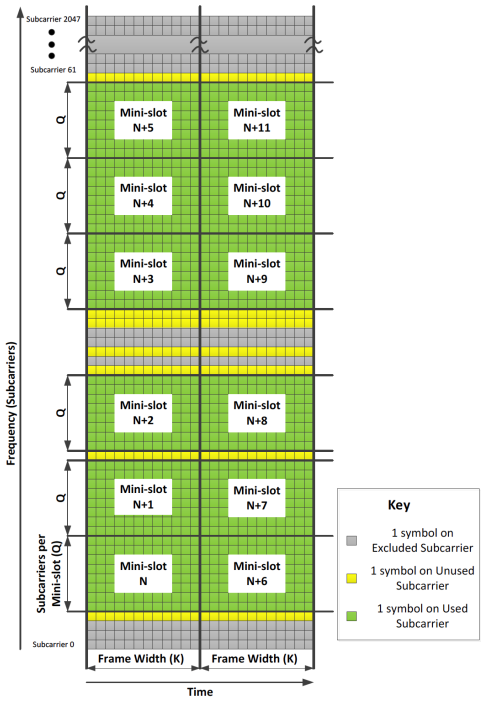


Figure 4 – Minislots in OFDMA

We performed various throughput tests in the lab with 32 MHz and 75 MHz channels, at 2K FFT and 4K FFT with different values of K. We consistently found a 3% to 5 % increase in data throughput across all CMTS platforms when the value of K is increased from 6 to 9/18 (for the larger 75MHz channels) or 6 to 18/36 (for smaller 32 MHz channels).

There are two minislot types, edge minislots and body minislots. Edge minislots have a bit more overhead and carry less data than body minislots. An edge minislot is the first minislot in a transmission burst and also used after an exclusion band or a set of skipped subcarriers. Body minislots are used for all other minislots in a burst. Hence it is advisable for an operator to avoid gaps in the OFDMA channel if possible.

3.3. FEC – Unreliable/Uncorrected Codewords

We tested various DOCSIS systems to verify the FEC performance under different noise level/ type (flat AWGN) to build expectations under different plant conditions. Figure 5, 6, and 7 show the FEC error rate (both correctable and uncorrectable) over different RxMER levels. The lab testing was run with traffic of 200- and 1500-byte packets and a combination of the two. The noise (AWGN) was introduced to reduce the RxMER levels across the OFDMA channel, and the levels of FEC statistics were captured at every level of increasing noise. We also tested with two different pilot patterns, the most dense pattern 4 and the least dense pattern 1.

Note that in most CMTS implementations when unreliable codewords are detected on an OFDMA channel, the upstream receiver discards the unreliable codewords considering them as uncorrectable. In other implementations they are sent to the MAC layer for additional processing using the MAC HCS and Ethernet CRC.

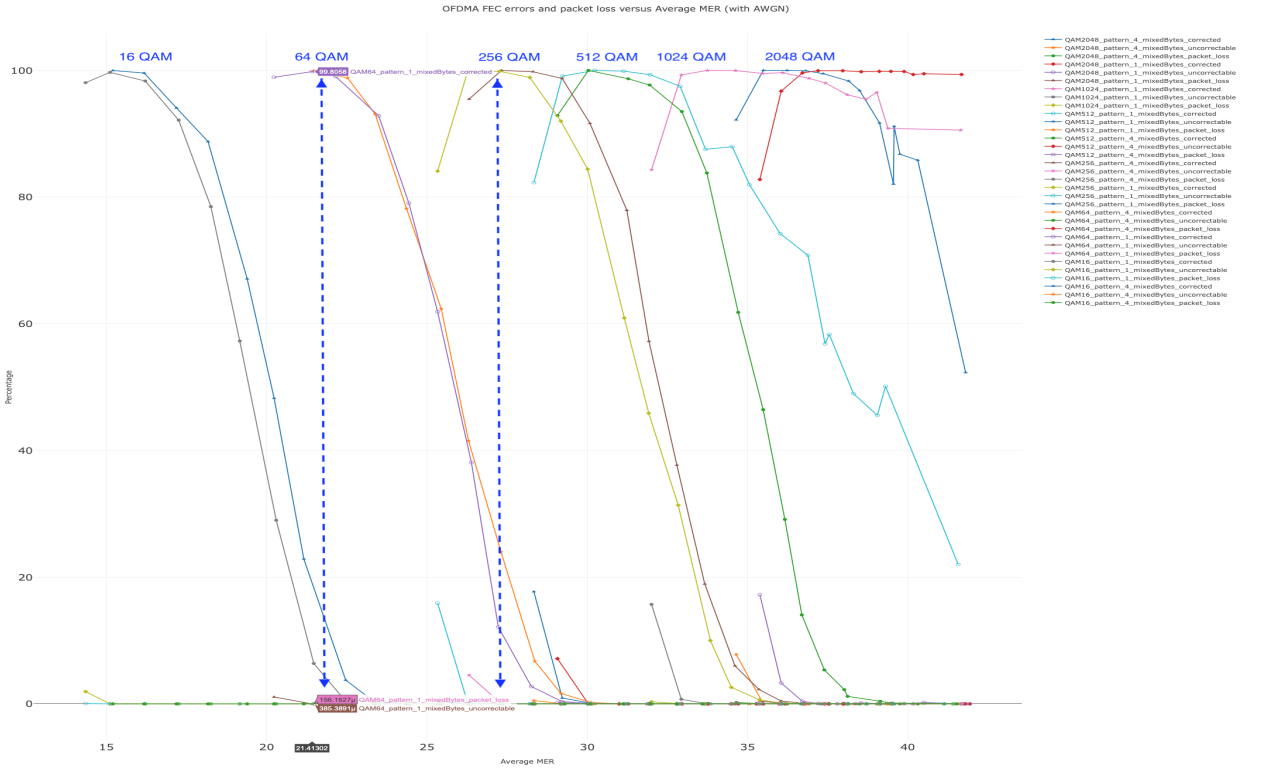


Figure 5 – OFDMA FEC Performance (mixed size packets, Pilot Pattern 1,4)

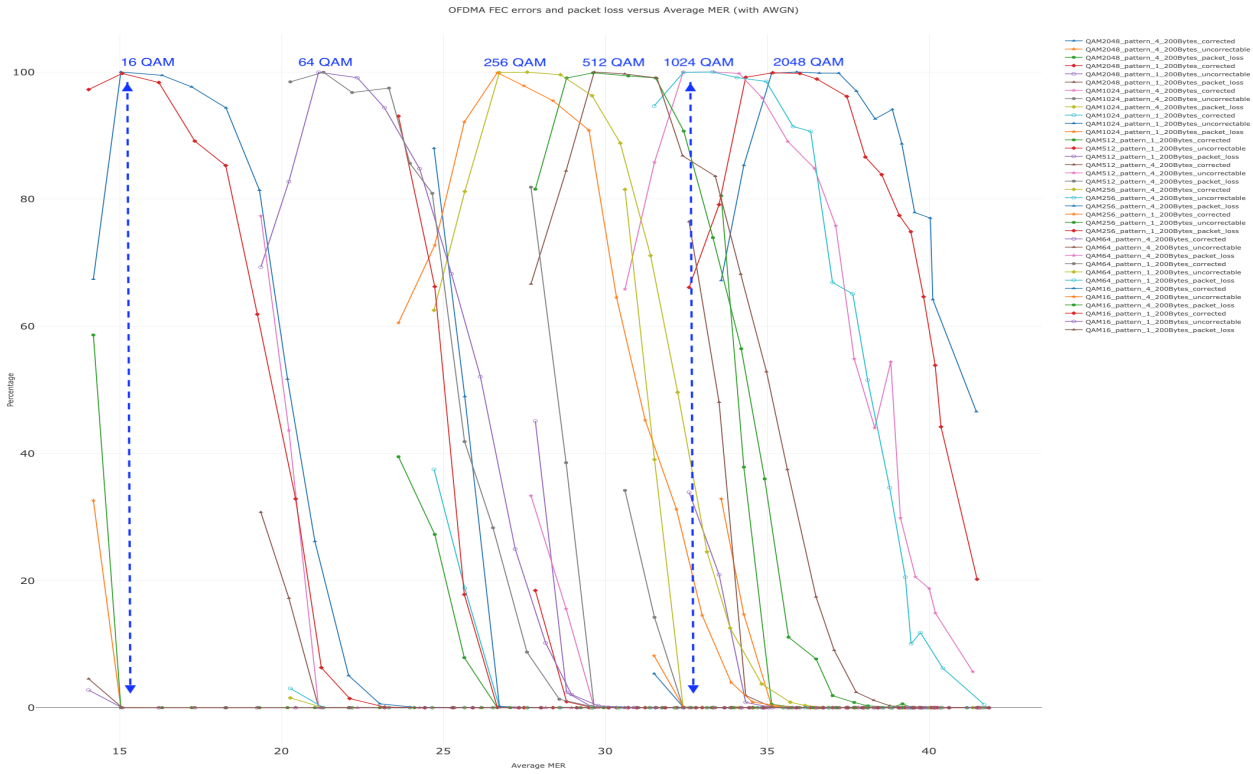


Figure 6 – OFDMA FEC Performance (200 byte packets)

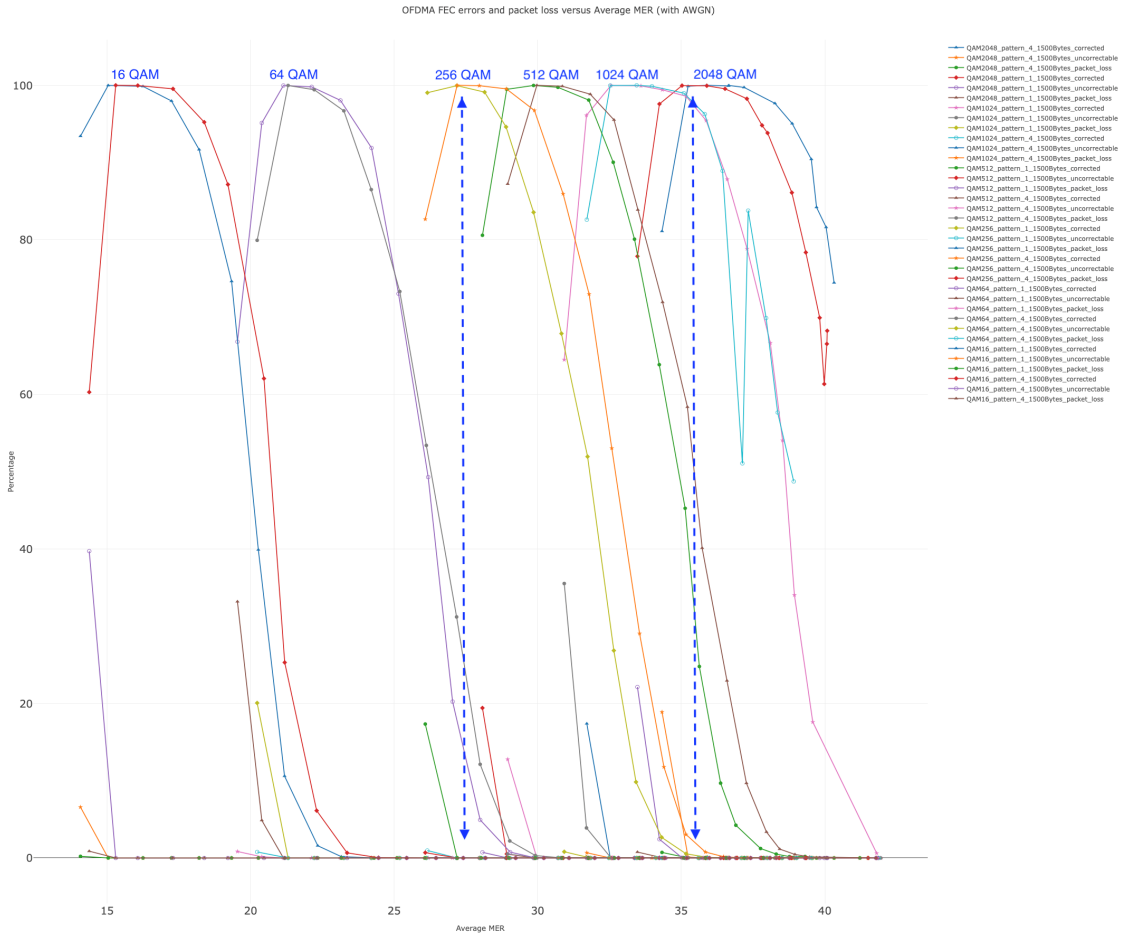


Figure 7 – OFDMA FEC Performance (1500 byte packets)

3.4. Cyclic Prefix (CP) and Roll-off Period (RP)

Two options exist for upstream OFDMA FFT size, either 2048 (2k FFT) or 4096 (4k FFT). A cyclic prefix (CP), which precedes an OFDMA symbol, should be slightly longer than a longest significant echo to be encountered in the channel, which normally is in the range of a few microseconds for coaxial cable. (See Figure below from [DOCSIS PHYv3.1]. The addition of a cyclic prefix enables the receiver to overcome the effects of inter-symbol-interference caused by micro-reflections in the channel.

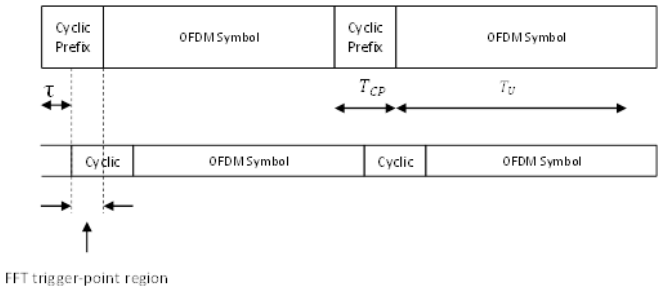


Figure 8 – CP and microreflections

The duration of the FFT useful symbol duration is 20 µsec for the 2k mode, or 40 µsec for the 4k mode. Therefore the percentage of overhead for a 2.5µsec CP can be either 12.5% (2k FFT) or 6.25% (4k FFT). CP sizes vary from 0.9375 µsec to 6.25 µsec, for which the overhead will vary from 4.6% to 31.25% for the 2k FFT, or 2.3% to 15.6%for the 4k FFT. Figure 1 shows the change in capacity for a 40 MHz channel, 2k FFT size, for the range of possible CP values.

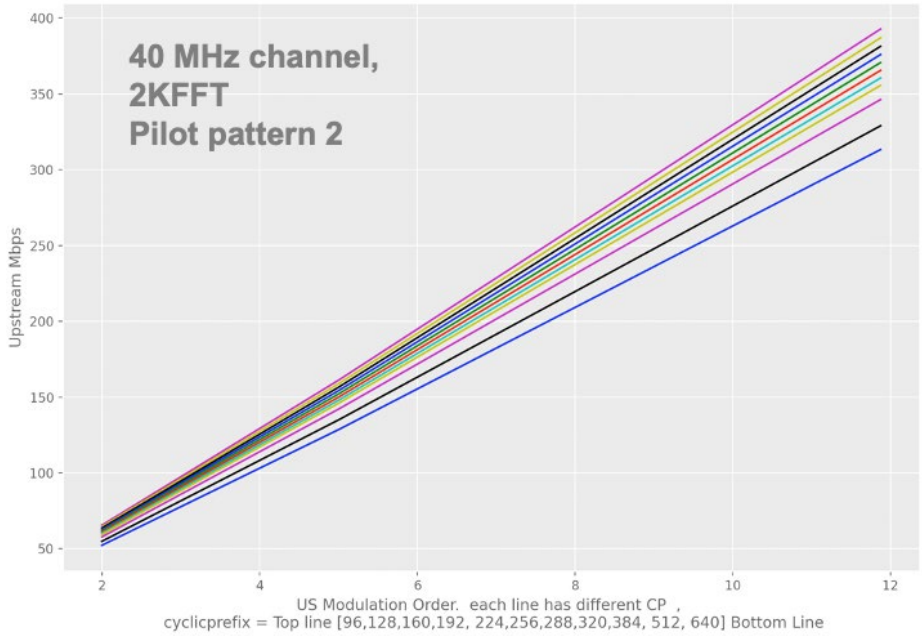


Figure 9 – OFDMA capacity change for different CP

Windowing maximizes channel capacity by sharpening the edges of the spectrum of the OFDMA signal (this applies to an OFDM downstream signal as well). Spectral edges occur at the two ends of the spectrum of the OFDMA symbol, as well as at the ends of internal exclusion band. The roll-off period (RP) setting is another OFDMA overhead. An RP is defined by a number of samples and essentially is a gradual rise or fall in energy and precedes or trails an OFDMA transmission. The purpose of the RP is to limit the spread of interference of an OFDMA transmission to adjacent single frequency carriers. See Figure 14. A longer duration RP produces less interference, and like the CP, decreases efficiency. Note that the effect of a RP is to decrease effective CP time, so if a long RP is chosen, the CP duration should be increased. The RP does nothing to protect an OFDMA channel from energy from adjacent SC-QAM channels.

3.5. 2K vs 4K FFT size or (subcarrier spacing)

The OFDMA upstream multicarrier system is composed of either 25 kHz or 50 kHz wide subcarriers. In the upstream, the subcarriers are grouped into independently configurable OFDMA channels each encompassing up to 95 MHz of spectrum, totaling 3800 25 kHz spaced subcarriers or 1900 50 kHz spaced active subcarriers. Figure 10 shows the possible size of the minislots with 4k FFT (25 kHz subcarriers) and with 2k FFT (50 kHz subcarriers)

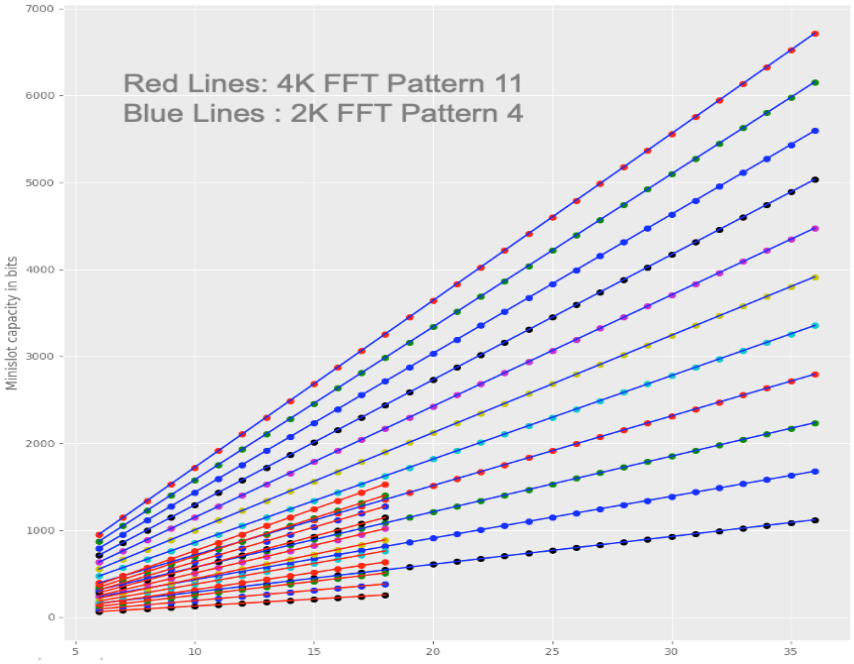


Figure 10 – OFDMA Minislot capacity change 2K vs 4K FFT

Figure 11 shows the capacity difference in OFDMA channels using either the 2K FFT or the 4K FFT, with different pilot patterns.

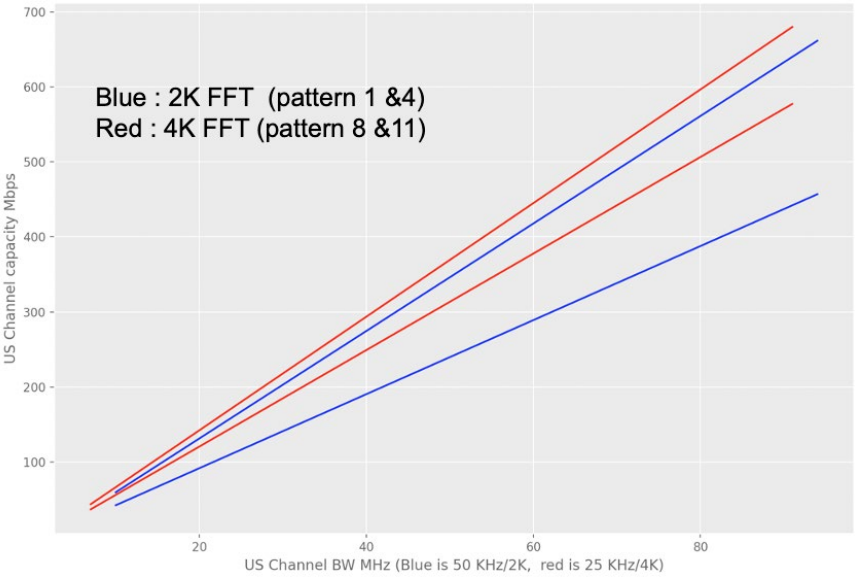


Figure 11 – OFDMA capacity change 2K vs 4K FFT

When processing an OFDM or OFDMA frame with CP, an exact integer number of symbols are chosen for processing by a FFT, either 2k or 4k. The abrupt truncation of symbols in the time domain creates a window in the frequency domain which has frequency domain sidelobes. See Figure 12. If spurious energy, such as a single carrier or a continuous wave (CW), lands on a sidelobe it will cause inter-symbol interference, generally affecting the subcarriers on the edges of the OFDM(A) band the most. One

solution is to provide vacant bandwidth adjacent to the OFDM(A) transmission to prevent sidelobe vulnerability. Another solution that has been used is to use a lower modulation order for subcarriers that are on the lower and upper portions of an OFDM block. Using lower order modulation, such as 256-QAM instead of 1024-QAM reduces the chance of symbol errors.

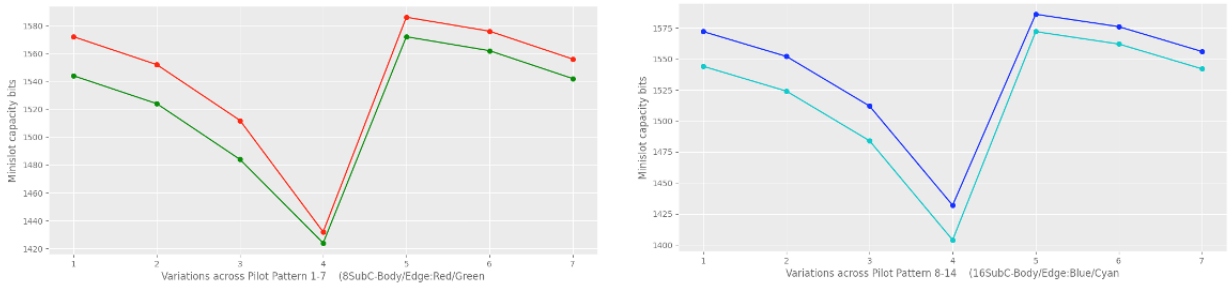
On sidelobe vulnerability, 2k is twice as susceptible because each sidelobe is twice as wide, relative to 4k. Note that this vulnerability affects upstream OFDMA as well as downstream OFDM.

An adjacent OFDMA channel will not cause harm to an OFDMA channel if the interfering frame timing is the same. This implies that both channels are using the same timing including FFT size, CP, and RP.

4k does have a vulnerability relative to 2k. It is more susceptible to phase noise. That is because any wander or movement of phase while an OFDM(A) frame is being captured is detrimental. Phase noise can be caused by a transmitter symbol clock, a receiver symbol clock, or any frequency conversion used in the OFDM(A) signal processing, such as block up and down conversion. Lab tests indicate this effect is a couple of dBs of MER worsening for 4k around the 35-40 dB level.

3.6. Upstream Pilot Patterns

Various pilot patterns are available for OFDMA transmissions. Not all CMTS vendors have implemented all pilot patterns, and different implementations support different default pilot patterns. Generally boosted pilots give a better channel characterization in the presence of noise, and denser pilot patterns improve estimates for signal time, frequency, phase, and amplitude offsets. But more pilots come at the expense of efficiency. The pilot pattern chosen should be sufficiently dense to characterize the ripples in the frequency response caused by echoes. If transmit pre-equalization is not used, less efficient pilot patterns should be used.



K= 20 symbols. 2K FFT
Patterns 1-7

K= 10 symbols. 4K FFT
Patterns 8-14

Figure 12 – OFDMA Minislot capacity change w Pilot Patterns

It is recommended that pre-equalization be turned on for both OFDMA and SC-QAM channels because not using pre-equalization requires a use of more dense pilot patterns, which are less efficient. Operators need to watch for this as in many cases pre-equalization, for OFDMA and SC-QAM channels, is off by default.

3.7. Efficiency of OFDMA channel

The throughput and efficiency (e.g., bits/Hz) of the OFDMA channel changes with respect to channel width and other configuration parameters, e.g., CP, RP, frame size, etc. The expected range has been detailed in a previous paper [D31 Capacity].

4. Upstream Interference Analysis

Generally, these technical discussions apply to both upstream and downstream single carrier (SC, such as SC-QAM) and multicarrier signals (MC signals, such as OFDMA and OFDM), although upstream efficiency is the general topic of this paper.

A problem arises when a single carrier signal and a multicarrier signal are located next to each other in adjacent frequencies. For this reason, when doing spectrum planning, for best efficiency it is generally best not to "mix up" single carrier and multi carrier signals but keep them separated as much as possible

On the upstream, as previously discussed, due to upstream noise funneling, the lower part of the spectrum, especially below 20 MHz, is hostile to all signals, and particularly hostile to OFDMA signals. This is due to impulsive energy spreading by an OFDMA receiver's FFT. This is particularly detrimental when there is switching regulator impulsive noise ingress in the upstream. This example of impulsive energy is created by poor switching power supply filtering combined with a loss of coaxial shield integrity (e.g., shield break). This allows both radiated and conducted energy from the power mains in homes to get onto the coax center conductors and travel upstream. Many switching regulated power supplies operate at frequencies close to the OFDM/A frame rates, such as 50 kHz, so there may be an impulse in every OFMDA FFT, with impulsive energy being transformed (spread) to each OFDMA symbol.

In the downstream both SC-QAM and OFDM signals are continuous, hence if there is adjacent channel interference, it is also continuous. In contrast, the upstream uses bursts of energy, so if a SC-QAM burst occurs when there is no adjacent OFDMA burst, no interference occurs. And if a OFDMA burst occurs when there are no adjacent SC-QAM bursts, again there is no interference. An interference problem only arises when the SC-QAM and OFDMA bursts occur at the same time(s) in adjacent frequency bands.

Interference needs to be evaluated in both directions, OFDMA into SC-QAM, and SC-QAM into OFDMA. Generally, an adjacent SC-QAM signal is more likely to damage an OFDMA signal than vice-versa.

4.1. OFDMA interference into SC-QAM

When an OFDM or OFDMA signal is created by a D-A converter and transmitted, out of band sidelobes are created. This is because, when transmitting data, the OFDMA signal abruptly transitions to different amplitude (voltage) levels. The sidelobes can be seen and measured on a conventional scalar spectrum analyzer and can cause interference with adjacent single carrier signals. Usually, time domain symbol damage is limited to just the one or two symbols that occur at the start and stop of the OFDM frame. The sidelobes interference can be reduced by using a roll-off period specified in the DOCSIS specification, along with providing a guard band between an OFDM(A) block and a single carrier signal. This causes inefficiency both in time for the roll-off period, and lost bandwidth.

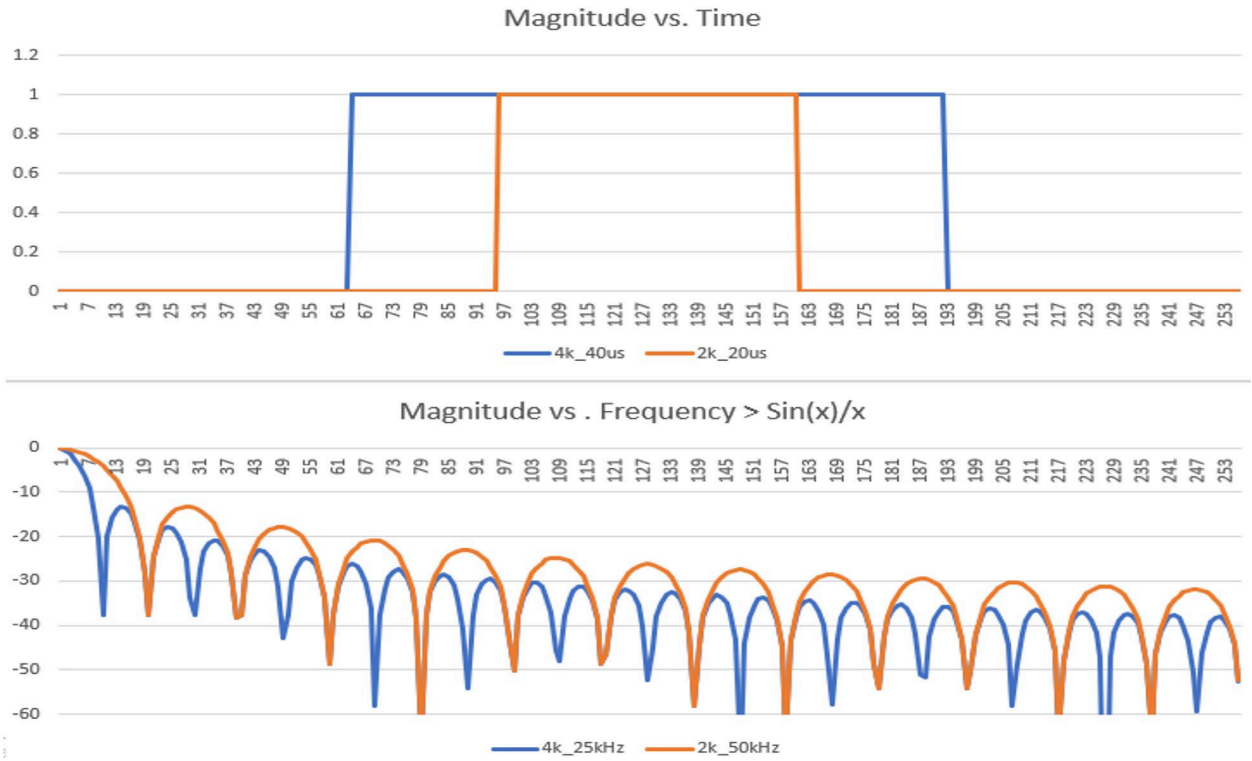


Figure 13 – OFDM/A Sidelobes created by rectangular time domain energy

In Figure 13, the upper plot is a rectangular pulse of magnitude voltage vs. time. The vertical axis is linear voltage, and the horizontal axis is time index. The rectangular pulses have a duration of either 20 µsec (2k mode) or 40 µsec (4k mode). If this waveform is transformed into the frequency domain with an FFT, the lower plots are obtained. The lower plot's vertical axis is dB, and the horizontal axis is frequency. The spectral shape consists of a DC term at 0 dB on the left, and then decreasing sidelobes over the rest of the plot. The sidelobes decrease as an absolute value of a sine(x)/x function. The separation between sidelobes is 50 kHz for 2k mode or 25 kHz for 4k mode. As you place more vacant spectrum between a SC-QAM signal and an OFDMA signal, the OFDMA interference to the SC-QAM signal is less.

The DOCSIS specifications provides another way to reduce interference besides vacant spectrum. The method is to shape OFDMA interference sidelobes and reduce their spectral reach over into the SC-QAM signal. That is done by employing a roll-off period (RP), which is tapering of the time domain envelope so that it gradually rises and falls, not with an abrupt transition, as illustrated in Figure 14. This is also referred to as a Tukey filter. Tapering is illustrated in Figure below (from the [DOCSIS PHY v3.1] specification). The roll-off period is subtracted from the cyclic prefix (CP) length, so CP length may need to be increased if a roll-off period is used. Unfortunately, employing a roll-off period and a CP is a time overhead penalty which decreases efficiency.

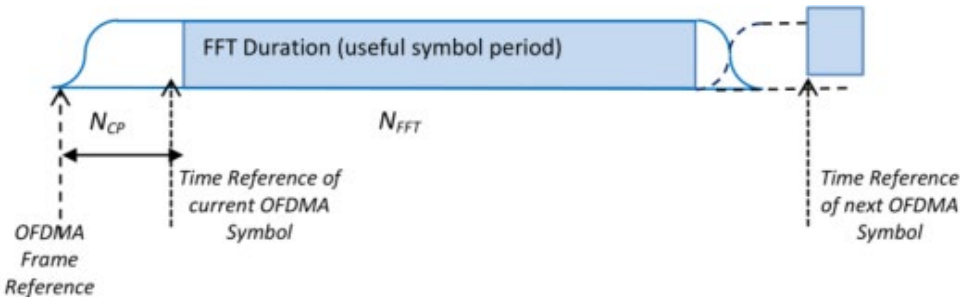


Figure 14 – OFDMA Signal, Useful symbol period and roll off

The above figure shows a time diagram of an OFDMA signal showing roll-off period and useful symbol period.

As mentioned above, for OFDM modulation, in general, there are spectral sidelobes beyond the edges of the modulated spectrum. This is a physical property of OFDM modulation and comes in the form of non-zero modulation energy beyond the edges of the configured modulated spectrum (i.e., below the lowest active subcarrier and above the highest active subcarrier). Figure 15 below, illustrates this modulated energy (sidelobes) and how it varies over different settings of the roll-off period (Nrp).

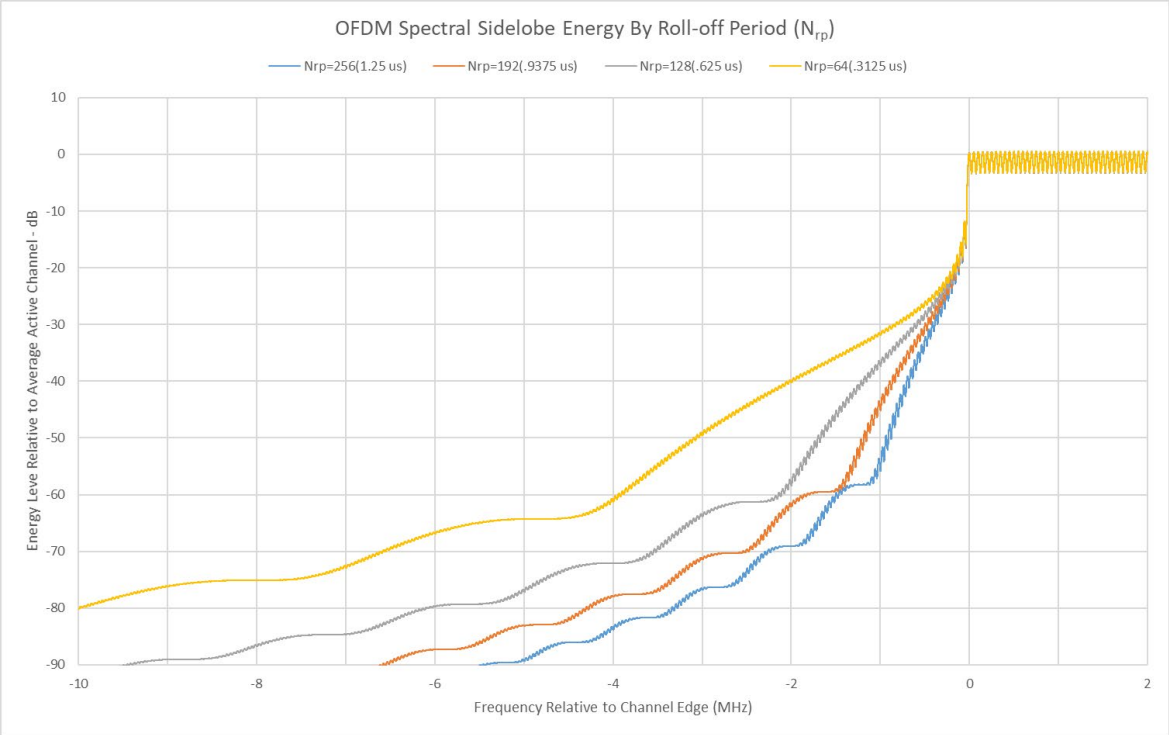


Figure 15 – OFDM/A Spectral Sidelobes (impact on Adj. SC-QAM)

** Figure 15: MATLAB simulation data provided courtesy of Roger Fish & Thomas Kolze, Broadcom*

Note that there is a trade-off between different Nrp settings. Smaller settings may offer more efficiency in the time-domain but have the cost of additional spectrum used by the OFDM channel modulation (larger spectral sidelobes/guard bands).

If you consider the top time plot of Figure 13 carefully, the interference to adjacent SC-QAM channels only occurs in time at the instant when the voltage is abruptly changing to or from zero, occurring every 20 or 40 us. So, if the SC-QAM bursts could be timed to "miss" those abrupt steps, no interference would occur. That could be an opportunity for a CMTS scheduler. Using 40 us (4k mode) generates half as many abrupt steps as 2k mode, so is half as harmful to an adjacent SC-QAM.

There is also a possibility that the damage to adjacent SC-QAM symbols can be eliminated by using good FEC settings on the SC-QAM bursts. The SC-QAM symbol rate is 5.12 MSymbols per second, and symbols occur every 195 ns. The abrupt transitions will damage one or two SC-QAM symbols out of every 256 symbols. Upstream FEC uses a Reed-Solomon block code, so setting the value of T to a large number (e.g., T= 16) can undo symbol damage. Furthermore, the increased FEC code overhead increases the robustness of the SC-QAM signal in general.

If two OFDM(A) signals are using the same exact timing, start time, stop time, cyclic prefix, and roll-off periods, they do not interfere with each other. So, no guard band or roll-off period needs to be used when signals are originating from the same slot/port on a CMTS and are using the same synchronous clock and the same OFDM & OFDMA frame timing. (The same CP etc. on the OFDMA frame)

4.2. SC-QAM interference into OFDMA

When receiving an OFDM(A) channel a comparable timing process occurs, where a first sample is taken continuously to the final sample. Unfortunately, time sampling is abrupt, not tapered. As a result, the same leakage phenomena that occurs with transmission occurs with reception, but this shows up as a susceptibility to interference from foreign energy at frequencies outside of the OFDM(A) channel. This energy can be a nearby SC-QAM and shows up as poor MER on subcarriers near the band edge as seen in the experimental results in Chapter 5.

Back a few years ago, it was observed in the field that the RxMER (receive modulation error rate) per subcarrier for lowest and highest subcarriers appeared worse than MER for subcarriers in the center of the OFDM or OFDMA band. It was determined that the presence of a SC-QAM above or below the OFDM or OFDMA spectrum was the culprit. Figure 17 and Figure 18 are plots of MER per subcarrier with an adjacent SC-QAM signal. Observe that subcarriers nearest the OFMDA band edges were automatically eliminated (not used) by the CMTS to avoid damage. Figure 19 is a MER per subcarrier plot with no adjacent SC-QAM signals. MER is uniformly good.

Why was this happening? An OFMDA receiver uses a rectangular sampling window in the time domain. The rectangular sampling window is illustrated in Figure 14 as the time interval with a blue shading. A resulting receive spectral window is not rectangular, but also extends into adjacent frequencies with sidelobes.

Figure 16 illustrates two receive spectral windows, one in red for 2k FFT and one in blue for 4k FFT. Because the sidelobes associated with the 4k FFT sidelobes decay more rapidly than the sidelobes associated with the 2k FFT, the 4k FFT has a narrower receive spectral window, and thus is more efficient than the 2k FFT. That is, the 4k FFT pulls in less energy from an adjacent SC-QAM signal relative to a 2k FFT.

Can using a roll off period on a transmitted signal improve efficiency by reducing interference? The answer is no, because the tapering is not inside the blue sampled region and is discarded (not used in the transform). Using a long cyclic prefix or a long roll-off period do not improve the problem of poor receive MER on OFMDA band edges. Choosing a 25 kHz (4k) subcarrier spacing over a 50kHz (2k) subcarrier spacing does make an improvement.

This advantage is illustrated in Figure 16. This narrower receive spectral window advantage for a larger transform size applies for both OFDMA and OFDM.

Figure 16 – Receive Windows 2k vs 4k

Receive spectral windows extend into the sidelobes and are wider for 2k (red) relative to 4k (blue) FFT.

4.3. Conclusions on Upstream Interference

The following are some conclusions reached based on the above discussion.

1. Generally OFDMA is harmed more by an adjacent SC-QAM than the SC-QAM is harmed by an adjacent OFDMA.

2. 4k is more efficient than 2k for reasons of narrower receive spectral window and required CP being half the symbol period. 4k also interferes with adjacent SC-QAM less, because of half as many abrupt voltage transitions.

3. Putting on more CP or RP on an OFDMA does not protect it from an adjacent SC-QAM signal. That is because its CP or RP do not modify the receive spectral window.

4. Adding a RP to an OFDMA channel does protect an adjacent SC-QAM channel from interference.

5. This discussion applies to upstream or downstream channel adjacencies.

5. OFDMA and SC-QAM Upstream Recommendations

5.1. Guard Bands

To examine the treatment of spectral guard bands in the DOCSIS specifications for OFDM/OFDMA channels and within different implementations, it is instructive to look at fundamental properties of OFDM modulation, in general, as well as a comparison between DS/OFDM and US/OFDMA specification requirements. As described in chapter 4, it is important to point out that the additional spectral sidelobe energy beyond the edges of the modulated spectrum is part of the channel and carries useful information. Therefore, another signal encroaching on this spectrum, such as a non-synchronous adjacent channel, has the potential to interfere with the OFDM channel and, vice-versa, the spectral sidelobe energy of the OFDM channel can also cause some interference on the adjacent channel.

5.1.1. DS/OFDM Requirements

We think of a full-width DS/OFDM channel as 192 MHz, although the maximum modulated spectrum width is only 190 MHz. This is convenient since, for downstream OFDM channels, the specifications require a minimum 1 MHz spectral guard band at the lower and upper channel edges. This guard band is required to accommodate the spectral sidelobe energy at the edges of the OFDM channel.

The nominal 192 MHz OFDM channel width includes 1 MHz (minimum) unmodulated spectrum at each edge as guard bands. It should be noted that more guard band could be needed, depending upon the setting for the roll-off period (N_{rp}). [DOCSIS PHYv3.1] spec (Appendix V) provides suggested additional guard band sizes (referred to as Taper Regions) for different N_{rp} settings. The MATLAB model used for the plot in Figure 15 was used to obtain the recommended guard bands (Taper Regions). Note that there is a trade-off between different N_{rp} settings. Smaller settings may offer more efficiency in the time-domain but have the cost of additional spectrum used by the OFDM channel modulation (larger spectral sidelobes/guard bands).

5.1.2. US/OFDMA Requirements

The physical properties of the OFDM modulation process still apply for bursted OFDMA upstream signals. Spectral sidelobe energy still occurs during bursts. Since the US/OFDMA channels are scheduled bursts of minislots, adjacent channel interference due to the spectral sidelobes of the OFDMA channel would only occur when these adjacent channels are scheduled and granted to burst simultaneously. Interference can be avoided through the scheduling and granting of these adjacent channels.

Based on our lab experiments, we found that CMTS implementations enforce some amount of guard band at the edges of OFDMA channels. Different CMTS implementations enforce OFDMA guard bands in different ways. In some cases, a minimum 0.5 MHz guard band is enforced by the CMTS at each edge of the OFDMA modulated spectrum. In other cases, 1 MHz or more guard bands are enforced by the CMTS implementation, either fixed or variable, depending upon the configuration parameters of the adjacent channels.

The important take-away here is that different CMTS implementations enforce different guard bands in different ways, even though the specs do not require enforcement of guard bands for US/OFDMA channels. It's important for the operator to examine and understand whether and/or how their specific CMTS implementation(s) enforce(s) guard bands at the edges of US/OFDMA channels.

This automatic enforcement of guard bands by the CMTS, which is not required by the spec, could result in part of the OFDMA channel not being used. For example, the CMTS might not grant a number of minislots at the edges of the channel so that its own minimum guard band requirements are met.

For the OFDMA channel, as an example, the interference comes in the form of reduced RxMER on the OFDMA subcarriers at the edge close to the adjacent channel whenever both channels are bursting simultaneously. Figure 17 shows the degraded RxMER per subcarrier at the edge of a 32 MHz OFDMA channel when an adjacent SC-QAM channel is bursting simultaneously (SC-QAM channel is to the left of the OFDMA channel). Note that, although the SC-QAM is configured immediately adjacent (guard band = 0 MHz), the CMTS in this case enforces a minimum 0.5 MHz guard band by leaving 2 minislots ungranted/unused at the edge of the channel.

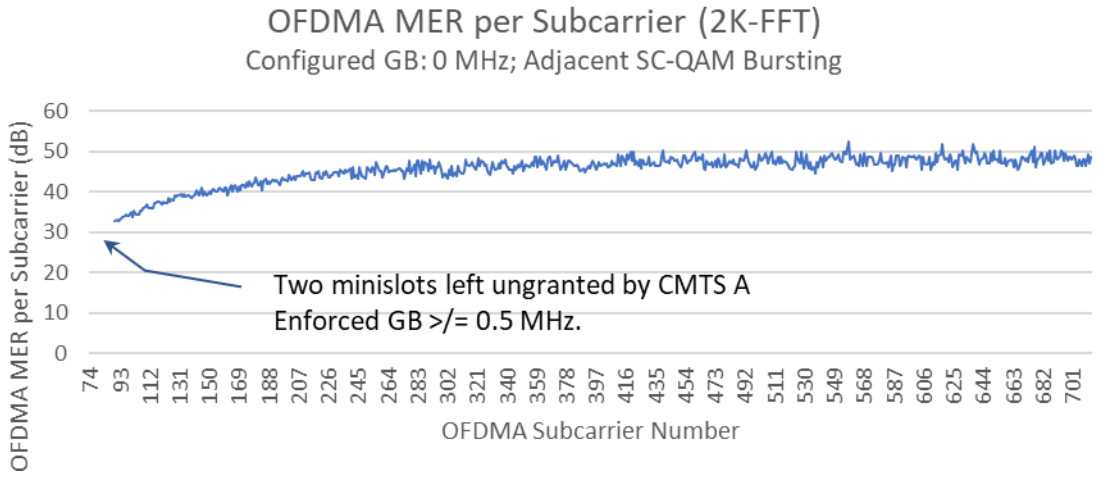


Figure 17 – RxMER : Non-synchronous Adj Channel Bursting Simultaneously (CMTS A)

In Figure 18, after further lab experiments using a different CMTS but with identical channel configuration, we observed that the first 10 minislots are left unused by this CMTS to automatically enforce its own guard band rules.

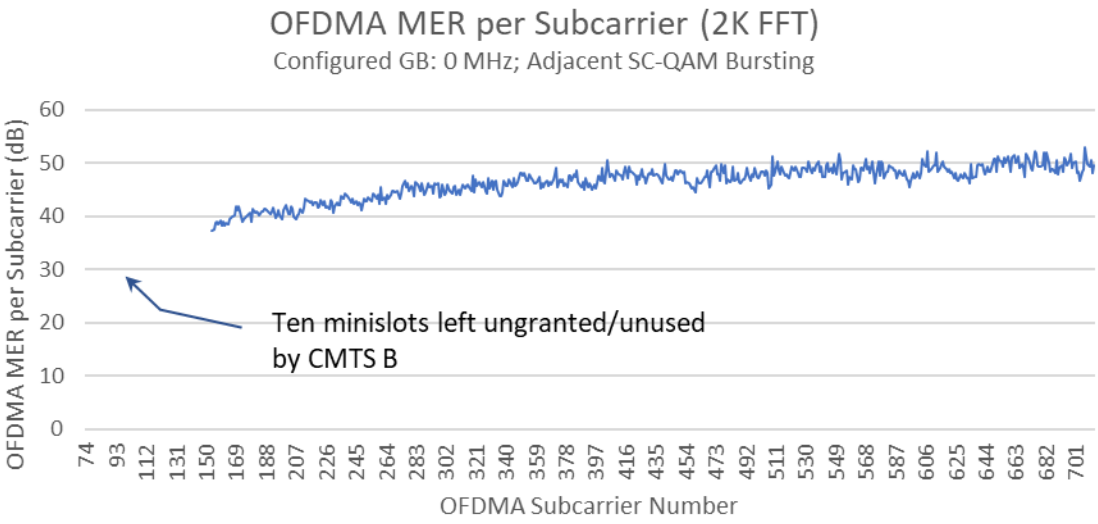


Figure 18 – RxMER : Non-synchronous Adj Channel Bursting Simultaneously (CMTS B)

It is important to remember that the interference affected RxMER and automatically enforced guard bands only occur when both adjacent channels are bursting simultaneously. When bursting alone as shown in Figure 19, there is no degraded RxMER at the edge and there are not automatically enforced guard bands.

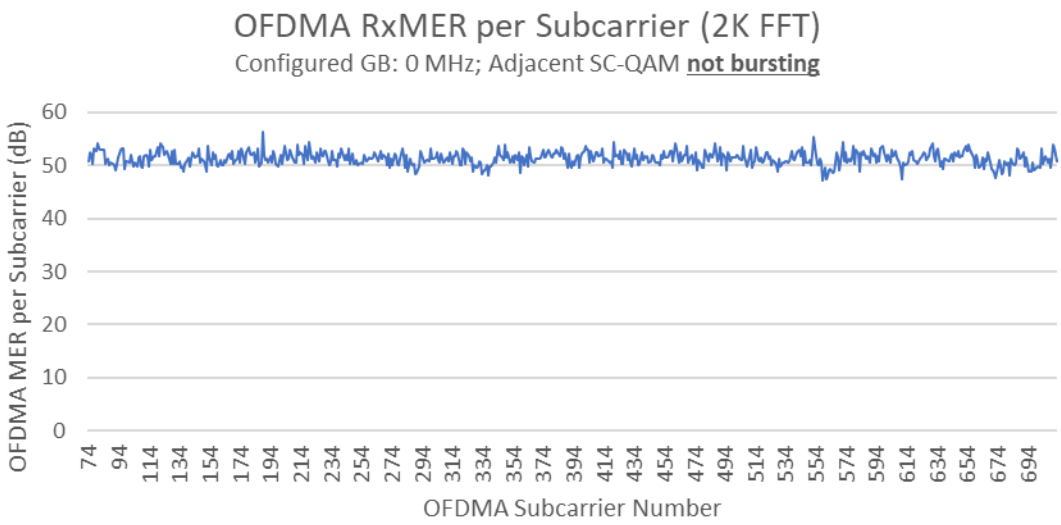


Figure 19 – OFDMA RxMER (OFDMA bursting alone)

5.1.3. Recommendations

It's important for the operator to know and understand how their specific CMTS systems treat adjacent channels with respect to their own enforcement of guard bands. If their own minimum guard band limits are not met by the spectral location of adjacent channels, parts of the OFDMA channel edges might be unused (minislots ungranted) by the CMTS scheduler.

An adjacent SC-QAM channels' interference on the OFDMA channel largely affects the subcarriers on the edge of the OFDMA channel, the impact is that the subcarriers on the edge of the OFDMA channel have worse RxMER than the rest of the channel. The OFDMA interference on SC-QAM can be remedied with a large guard band or moderate guard band with taper (Roll off period). Tapering/Roll-off period does not protect the OFDMA channel from an adjacent SC-QAM channel.

5.2. TaFDM (Time and Frequency Division Multiplexing)

It's useful to examine the use of TaFDM in the context of the discussion of guard bands (above). As previously noted, DOCSIS specifications do not mandate minimum guard bands for US/OFDMA channels. In fact, it follows that different non-synchronous channels may even occupy the same frequency spectrum. This is allowed by the specification and can be handled by the CMTS scheduling the use of overlapping spectrum by the included channels at separate times.

5.2.1. TaFDM and CMTS Guard Band "Rules"

In practice, we find that CMTS implementations tend to adhere to their own guard band enforcement processes, either partially or in whole, to implement TaFDM. In fact, we have seen that some CMTS implementations might automatically trigger their own TaFDM mode when adjacent channels are configured closer than their own expected minimum guard band. Even though the modulated spectrum may not overlap, it could be considered overlapping by the CMTS if a portion of their expected minimum guard band is overlapped.

Other CMTS implementations may extend their own guard band rules to zero out larger portions of the OFDMA channel that is also occupied by, for example, an overlapping SC-QAM channel. This could result in half or even more of the OFDMA channel being zeroed out (not granted for OFDMA bursts).

Like the examples above, Figure 20 shows the MER per subcarrier for the same 32 MHz OFDMA channel, but with a 6.4 MHz SC-QAM channel configured in TaFDM mode just above the center of the OFDMA channel.

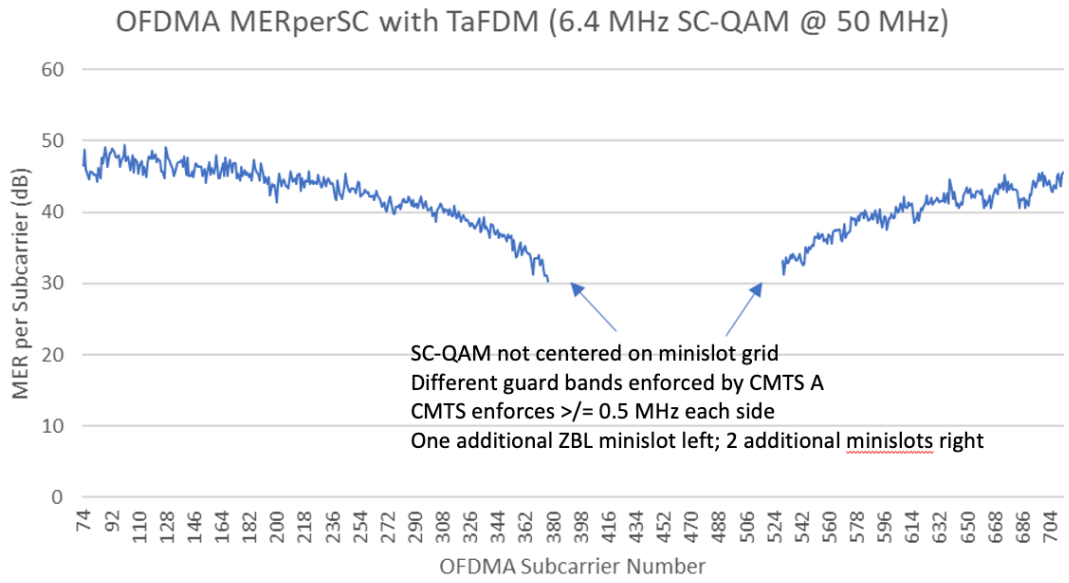


Figure 20 – OFDMA RxMER w TaFDM (SC-QAM bursting in center of OFDMA) (CMTS A)

As shown in Figure 21 when using a different CMTS, nearly half of the OFDMA channel is unused in order to enforce a combination of its own guard band and TaFDM rules.

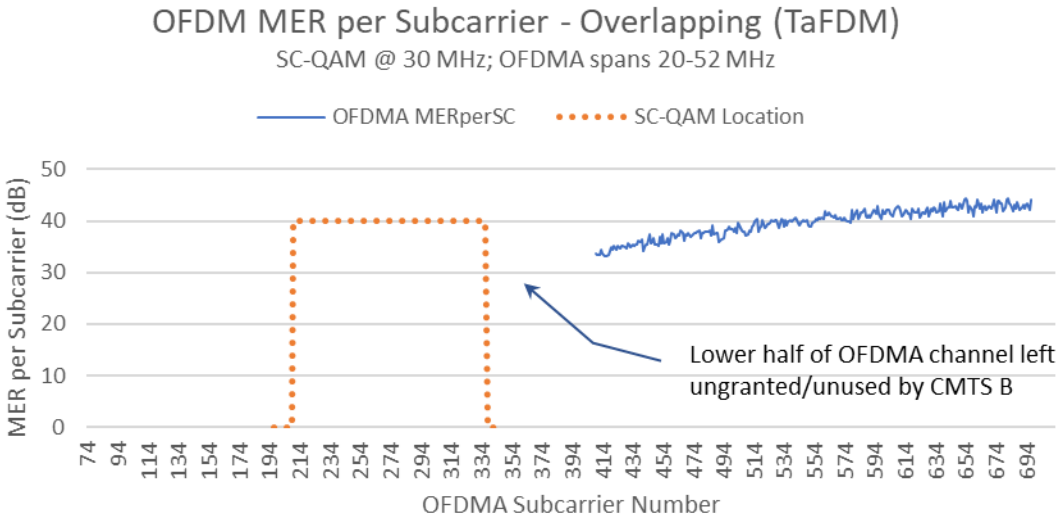


Figure 21 – OFDMA RxMER w TaFDM (SC-QAM bursting in center of OFDMA) (CMTS B)

It is worth mentioning again that when the OFDMA is bursting alone, the full width of the channel is used without MER degradation, as shown in Figure 19.

5.2.2. Recommendations

It is important to understand how a CMTS implementation handles TaFDM to judge whether the associated trade-offs are worthwhile. If a particular CMTS only enforces a small guard band at each end of a fully overlapping spectrum and, therefore, allows simultaneous bursting of overlapping channels using TaFDM, then there may be situations where fully overlapping channels are useful. Of course, it's possible that the use of TaFDM could zero out half or more of the useful modulated spectrum of an OFDMA channel, depending upon the location of the overlapping channel and the specific CMTS "rules" for handling guard bands and TaFDM, and in these cases may not be an efficient use of the spectrum.

6. DOCSIS Protocol and Configuration Efficiency Considerations

DOCSIS protocol relies on several processes for proper operation. These processes include initialization, request for data, the multiple aspects of ranging and the processes to assign CMs to the right channels and their configuration to the right profile. Most of these processes will have some impact on efficiency that are worth further analysis.

6.1. Initialization

The initialization of a CM consists of downstream scanning, upstream channel acquisition, ranging, IP connectivity, time of day acquisition, config file download in addition to registration verification and the establishment of connection privacy. These processes by themselves are part of the expected behavior, however efficiency problems arise when these initializing CMs take a longer time or struggle to complete any of these processes. There may be conditions in the plant, sub-optimal configuration or problems with the implementation that forces a CM to repeat one or more of these processes to recover normal operational state.

6.2. Initial Maintenance

DOCSIS 3.1 CMs use an initial maintenance region that is of significant size in frequency and time. This region is used to allow CMs to perform initial ranging which takes place in a broadcast mode and is used with a very coarse time adjustment as initially the CMs do not know what timing offset to apply to their transmissions. The design of this time and frequency resource area (ranging zone) must be done very carefully to minimize collisions that occur when that initial maintenance region is not big enough or to avoid wasting resources if it is too big. The frequency of occurrence of this initial maintenance region, which is called the insertion interval, must be carefully defined. In conjunction with the insertion interval, the back-off process is configured by adjusting the random range back-off parameter that is used after a collision occurred.

Above all the location of the ranging zone needs to be chosen in a higher part of the spectrum to minimize interference from the noise sources on the upstream. This has been a single point of failure in some early deployments of OFDMA, where the default ranging zone created by the CMTS was not in a robust region to allow all the CMs to range successfully.

Figure 22 depicts the initial ranging zone in a DOCSIS 3.1 OFDM channel, time and frequency map.

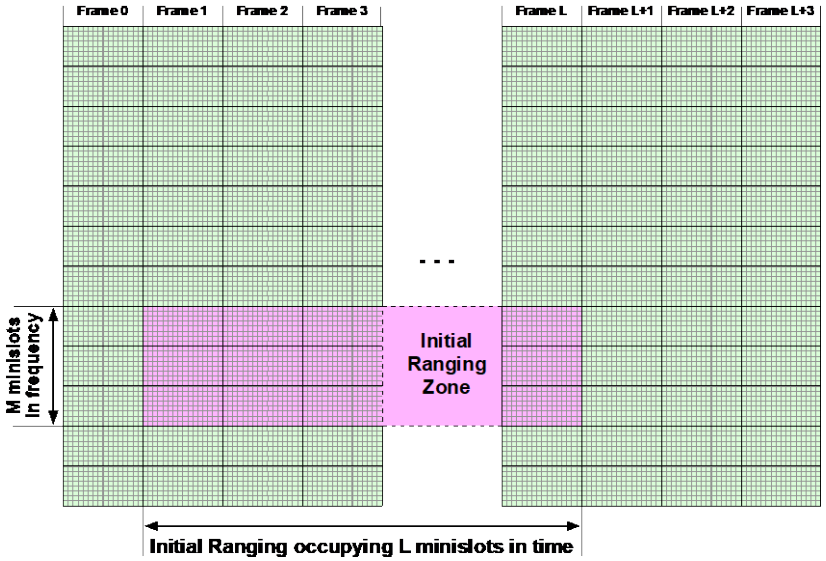


Figure 22 – Initial ranging zone in DOCSIS 3.1 in symbol versus frequency view

When ranging power levels, since initially the CM has no knowledge of what the right power level should be, it must go through many power levels before converging on a final value. The time this process lasts can be significant and any improvement mechanisms are worth implementing. One popular approach is to try the last good known state first. So, if a CM reboots it uses the last good known power and frequency configuration parameters that worked the last time the CM successfully initialized. Once the initial maintenance is successfully achieved, the CMTS and CM can communicate through unicast messages. Nevertheless, impaired channel conditions can force CMs into repeating this process.

6.3. Station Maintenance

The ranging response message is used to adjust transmit timing information, transmit power level information, transmit equalization information among other things. A CM that requires such an adjustment must wait until the CMTS acknowledges that the CM is ready to transmit by changing the ranging response messages from “RNG-RSP continue” to “RNG-RSP success”. The CMs cannot transmit data while still in “RNG-RSP continue”, a state which results in upstream inefficiency. These processes can extend for a while. Every time that a RNG-RSP message is not acknowledged, the T3 timeout counter is increased. After 16 successive T3 timeouts a T4 timeout occurs. The US channel is then deemed unusable, and the US channel enters in partial service mode or if it is the only channel, it forces re-initialization which represents an even greater impact on efficiency. In a good stable environment periodic maintenance is only used to conduct minor adjustment. Monitoring and quantifying the extend a CM spends beyond an uneventful ranging, helps you assess the impact on CM performance and efficiency.

6.4. Upstream Partial Service

The CMTS puts a CM in partial service mode when a CM is not able to range on a particular channel. This CM in partial service stops using the degraded channel. While station maintenance is the default mechanism to put CMs in partial service mode, alternative metrics can also be used to place a CM in partial service which include thresholds in SNR, as well as thresholds in uncorrectable as well as correctable codewords. This provides the operator with different tools to manage which CMs should have channels in partial service mode. These alternatives mechanisms also provide a hysteresis function so that

a CM does not waste time going in and out of partial mode. The CMTS periodically will have the CM attempt to range on the degraded channel and check if it is still impacted. Going into partial service not only has the impact of not having US channel available to a CM, but also these transitions to and from partial service impact efficiency. CMTS algorithms on resolving partial service is a big factor in how quickly a CM returns to full service.

If a system struggles to achieve operational stability due to constant adjustment of power levels, constant adjustment of timing offset and a lack of convergence of equalization coefficients, efficiency is significantly impacted. Proper maintenance of a healthy US and DS channel prevents such occurrences.

6.5. Contention and Piggybacking Bandwidth Requests

The process of requesting upstream bandwidth in DOCSIS systems could also introduce inefficiencies. To request bandwidth, (minislots is the upstream resource currency), the CM takes advantage of a contention period designed for that purpose. This contention period allows CMs to send a request message containing how much data is desired. The size and periodicity of this contention period must be configured carefully by the CMTS to limit the number of collisions. Like contention in initial maintenance, when request messages collide, a random back-off is used to reduce the probability of further collisions

In addition to the contention request messages, DOCSIS systems also leverages a piggybacking process to reduce the number of contention requests needed thereby having a side benefit of reduce the likelihood of collisions. In piggybacking a bandwidth request is appended behind the packet just transmitted thereby not requiring a contention request.

Monitoring the number of regular request messages sent and number of piggyback requests sent in addition to monitoring the number of bytes requested and bytes granted are ways of determining how efficiently bandwidth is being allocated.

6.6. Traffic Characteristics and Signatures

The amount of traffic as well as its characteristics such as packet rate, packet rate variability and packet size distribution, stream duration as well as resources available are useful in determining how many mini-slots and how often to allocate request message opportunities. Since the traffic characteristics are dynamic, the CMTS will likely react and adapt according to these changing traffic conditions. The intelligence and efficiency of the CMTS scheduling algorithms will result in lower or higher transport efficiencies.

In the upstream, in DOCSIS 3.1 technology there are short, medium, and long codewords, each with different efficiencies. Very short packets will have lower efficiency and longer packets will have higher efficiencies. Also, in SC-QAM upstream operation each packet is preceded by a preamble and a guard time. This represents a time overhead that is fixed and for short payloads the payload size could be comparable to the preamble and guard-time, while with longer payloads such as is the case of 1500 Bytes, the preamble and the guard-time is only a small fraction of the packet duration. This inefficiency is aggravated with higher order modulations as the payload is further reduced in time. Efficient concatenation reduces the percentage of shorter packets thereby resulting in higher efficiency.

In DOCSIS systems the CMTS measures consumption and allocates resources in minislots, not bytes. One can have an accurate estimate of the resources being used when we leverage the MIBs that provide minislot utilization. An operator needs to make sure that the system is not starving for minislots, simply tracking the bytes can be deceiving.

Figure 23 shows an example of packet size distribution characteristics for some applications and aggregate packet size distribution of CMs on a service tier to highlight the specific impact that packet size distribution could have on performance. Some other packet size studies have found that 80% of upstream packets have a size below 80 bytes, and 90% of upstream packets are below 160 bytes.

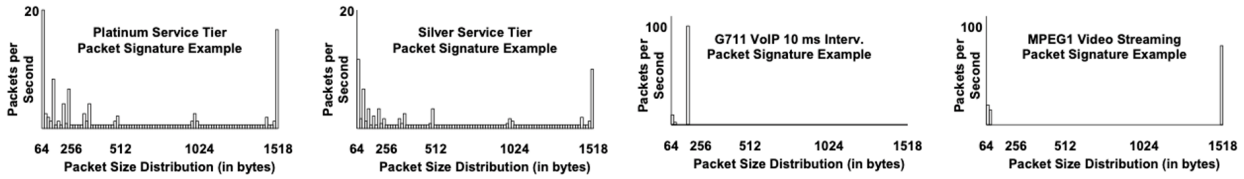


Figure 23 – Traffic Signatures for specific applications and for sample service tiers

In today's very high service tiers, bursty traffic is more prevalent. Tweaking the token bucket parameters is a tool that has been used in the past to optimize performance and can play well under bursty conditions, this could get large bursts of traffic out of the way to free up the channel for other transmissions.

6.7. Service Flow Configuration

The service flow configuration for the upstream directly affects the user experience. There are many Service flow parameters which an operator configures for an upstream service flow. These include CM Service Flow Parameters such as the Max sustained rate, Peak traffic rate, Max traffic burst, SF Priority, R/T Policy, Buffer control, AQM etc. Using the Buffer control setting (50 ms recommended) and enabling AQM will reduce the latencies observed by the user. The Max Sustained traffic rate for the service flow is set by the Service tiers offered by the operator.

One recommendation that will improve the user experience across all the service tiers is setting the Maximum Traffic Burst to something beyond the default of 3044 bytes, which is too small. The recommendation is to bump up the Max Traffic Burst (B) setting for a service flow to something in the range of 50 kBytes to 65 kBytes. Higher values of B can give users a boost in the initial data rate seen by the user, until the CM uses up the tokens allowed by the B value.

6.8. Profile Management / Upstream IUC management

DOCSIS 3.1 specifications introduced the concept of profiles so that performance of CMs could be optimized based on the SNR that they have available on the plant. DOCSIS systems have thereby evolved from the lowest common denominator conditions operation to a more dynamic opportunistic resource usage approach.

The CMTS assigns OFDMA IUCs (profiles) to the CMs based on the measured plant conditions. It is intended that the Data Profile IUC 13 is configured to be a robust OFDMA profile usable by any DOCSIS 3.1 CM served by that upstream channel and is used for all OFDMA data grants to modems which have not completed registration and for transport of MAC management messages and data grants after registration.

During or after modem registration, the CMTS has the option of assigning the CM to use any other configured data profile. Typically, the data profiles other than IUC 13 will be configured with higher modulation orders than IUC 13, although not all these profiles will be usable by all modems. The CMTS typically assigns IUC 13 and an additional data profile (one of IUC 5/6/9/10/11/12) to the CM for an OFDMA channel. The CMTS grants bandwidth on the OFDMA channel for data transmissions to a CM using this data profile (e.g., IUC 5). Now at some point the plant may have a noise ingress and interfere

with the upstream transmissions. If the CMTS detects codeword errors on the IUC/profile used for a CM, it will temporarily use IUC 13 for data transmission, and initiate a dynamic change process to swap out the errored data profile with lower modulation data profile (e.g., swap out IUC 5 with IUC 6). This process continues (and the CMTS may further downgrade profiles) until the CMTS can receive codewords from the CM without errors. At a later point in time when the noise ingress goes away, the CMTS based on US RxMER measurements (using probes) or OFDMA Upstream Data Profile (OUDP) testing bursts, can evaluate the CM performance on a higher modulation profile and then choose to assign those profiles to the CM if appropriate. (e.g., upgrade from IUC 6 to IUC 5)

On noisy upstream spectrum, ingress noise issues may interfere with the signal enough to cause FEC codeword errors, to initiate the whole process described above, where a CM's profile is continuously being upgraded or downgraded, or in the worst case where the CM goes into partial service and is unable to use that channel for periods of time. Intermittent noise will cause this profile flapping behavior in the upstream OFDMA channel, such instability may be triggered by plant conditions and ultimately ends up impacting CM performance. (Similar profile flapping has been observed in noisy downstream channels as well).

DOCSIS CMTS settings need to be configured correctly to minimize the profile flapping. This includes the thresholds for the number of FEC errors, the RxMER dB thresholds for a particular modulation order, the time for which a CMTS downgrades a CM before re-evaluating if a profile is appropriate for a CM, the choice of looking at RxMER vs FEC errors in the upgrade/downgrade decision.

A well-designed Profile Management application (PMA) or system can create a good set of profiles, adjusted to the noise characteristics of that node, and minimize the profile flapping for each CM. A lot of the profile management concepts, algorithms, field experiences and deployment lessons have been described in a previous paper [US PMA].

The ability of the system to support a given QAM level depends on the RxMER values and the mappings to an appropriate QAM level, when creating a profile. These mappings are defined in [PHYv3.1] and are summarized in the Table below.

Table 3 – Upstream RxMER to QAM Level mapping

Upstream Constellation / Bit Loading	Upstream MER (dB)
QPSK	11.0
8 QAM	14.0
16 QAM	17.0
32 QAM	20.0
64 QAM	23.0
128 QAM	26.0
256 QAM	29.0
512 QAM	32.5
1024 QAM	35.5
2048 QAM (optional on CMTS)	39.0
4096 QAM (optional on CMTS)	43.0

6.9. Channel Conditions and PNM

In the previous sections, we have seen how unstable operations of DOCSIS processes directly impacts transport efficiency. This change from a stable to an unstable DOCSIS operation may have been triggered by changes in channel conditions. The DOCSIS system configuration may not be robust enough to overcome the channel impairments. These impairments may be sporadic or time of day sensitive. Even though these conditions may be difficult to detect, there are different parameters that are worth monitoring, including:

RxMER on a per channel and per subcarrier basis, FEC statistics, transmit and receive power levels and the different PNM tools such as equalization coefficients/channel estimate, full band capture, upstream triggered spectrum analysis, quite probes and CM specific probes analysis.

The goal is to have a good understanding of the conditions of the plant, troubleshoot and correct problems that are performance impacting and/or verify performance and profile assignment is as expected.

Of the upstream PNM functions defined in the specifications, we have obtained and looked at the following data sets: upstream RxMER per subcarrier, upstream Equalizer Coefficients (CMTS), upstream Triggered Spectrum Capture Analysis, and upstream FEC Statistics (aggregated over the channel / all CMs). All of these have useful PNM applications in understanding the upstream performance. Upstream and downstream equalization coefficients contain valuable channel information, please refer to [PreEq Analysis] for further details.

So far, we have not been able to get access to data from CMTSs for the following PNM features: upstream Capture for Active and Quiet Probe, upstream Impulse Noise Statistics, upstream Histogram, upstream Channel Power. These upstream PNM functionality needs to be implemented on CMTS platforms and then further investigated and applications developed.

7. Upstream Split Migration Scenarios

There are a few different approaches for increasing the upstream bandwidth within the cable plant. The DOCSIS 3.1 system will have options of several split configurations that can be exercised based on traffic demand, services offered and the capability of the cable plant. These include the classic Low-split, the mid-split, and high-split frequency plans for DOCSIS 3.1 equipment and going to the ultra-high split options with DOCSIS 4.0 technology. To reach the target service goal in the upstream direction, plant changes on the upstream/downstream spectrum split are expected.

7.1. Adding Upstream Spectrum

Information theory says there are two ways to increase channel capacity:

- Add spectrum
- Increase the signal-to-noise ratio of the spectrum in use

In brief, doubling the upstream spectrum will double the upstream capacity; that is, a 100% increase in capacity. Staying in the same spectrum and increasing the Gaussian SNR by 6 dB to go from 256 QAM to 1024 QAM provides a 25% increase in capacity. Below 20MHz the noise is usually bursty, so a Gaussian noise assumption is not quite appropriate. Furthermore, the impulsive noise is spread to all frequency domain symbols by the FFT at the CMTS receiver. At some point, increasing the SNR has diminishing returns and the big gains will come from increasing the amount of spectrum allocated to the upstream.

Hence DOCSIS 3.1 technology includes both mid-split and a high-split options will be discussed in more detail. DOCSIS 4.0 technology includes ultra-high splits which provide additional upstream spectrum. Figure 24 shows the relative amounts of upstream spectrum with sub-split, mid-split, and high-split plant in DOCSIS 3.1 systems.

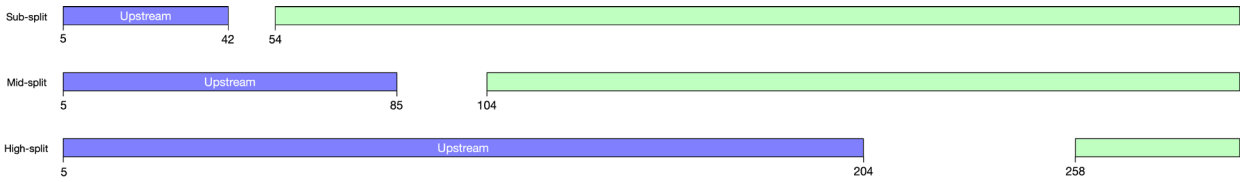


Figure 24 – Split Scenarios

7.1.1. Impact on Legacy Customer Premise Equipment

Existing cable modems deployed in the field may not support different splits, however, they will operate with the DOCSIS channels that reside in their existing upstream spectrum option. Set-top boxes (STBs) are limited to forward data channel (FDC) frequencies between 70 MHz and 130 MHz in the downstream direction as described in the SCTE 55-1 and 55-2 standards. As a result, these set-top boxes can continue to operate on a mid-split network but not a high-split network.

7.1.2. More Upstream spectrum and Downstream Spectrum

As more upstream spectrum is added, the downstream starts higher in the spectrum. For example, with a sub-split the downstream spectrum starts at 54 MHz. With a high-split network, the downstream starts at 258 MHz, and to maintain enough downstream spectrum the top end of the useful spectrum is usually increased, and the typical number is 1000 MHz or 1218 MHz.

As a summary, Table 4 shows the aggregate upstream throughput for the examples given in the following sections.

Table 4 – Upstream Throughput Examples

Case	Channels	Approximate Throughput
Sub-split Use Cases		
1	Four 3.2 MHz SC-QAM at 32 QAM	44 Mbps
2	Two 3.2 MHz SC-QAM at 32 QAM and Two 6.4 MHz SC-QAM at 64 QAM	72 Mbps
3	Four 6.4 MHz SC-QAM at 64 QAM	100 Mbps
4	Four 6.4 MHz SC-QAM at 64 QAM and One 3.2 MHz SC-QAM at 64 QAM and One 1.6 MHz SC-QAM at 64 QAM	115 Mbps
5	Two 6.4 MHz SC-QAM at 64 QAM and One 3.2 MHz SC-QAM at 64 QAM and 13 MHz OFDMA at 1024 QAM	160 Mbps
Mid-split Use Cases		
6	Ten 6.4 MHz SC-QAM at 64 QAM	250 Mbps
7	Four 6.4 MHz SC-QAM at 64 QAM and 40 MHz OFDMA at 1024 QAM	400 Mbps

8	70 MHz OFDMA at 1024 QAM	525 Mbps
High-split Use Cases		
9	Four 6.4 MHz SC-QAM at 64 QAM and One 66 MHz OFDMA at 1024 QAM and One 96 MHz OFDMA at 1024 QAM	1350 Mbps (1.3 Gbps)
10	Two 96 MHz OFDMA at 1024 QAM	1700 Mbps (1.7 Gbps)

7.2. Sub-split HFC network

A sub-split HFC network has a return path up to 42 MHz and has room to carry multiple upstream DOCSIS channels. In North America, sub-split is a hold-over from the days of analog television channels and cable-ready TV sets (both of which are rapidly going extinct, if not already there). Sub-split allowed the old Channel 2 to be carried on the coaxial cable starting at 54 MHz. The old analog Channel 2 has been replaced by digital TV carriers and IPTV, so the reasons for sticking to sub-split are diminishing. The following sections will show examples of how to get the most out of various configurations of an upstream channel.

The configurations in Cases 1-5 shown below will fit within the spectrum of a sub-split HFC network and as can be seen the upstream capacity can more than double by increasing both the amount of spectrum used for broadband and increasing the modulation order of the SC-QAM channels.

7.2.1. Making the most of Upstream SC-QAM Channels

For sub-split operation operators should experiment with various recommendations to get more upstream capacity. Strategies include:

7.2.1.1. Case 1: Remove Guard Bands

Removing large guard-band between upstream SC-QAM carriers; these guard bands are not needed. The DOCSIS specifications are written such that upstream SC-QAM channels can be directly adjacent to each other. Deployments exist where the upstream SC-QAM carriers have large guard-bands between them as shown in Figure 25 below.

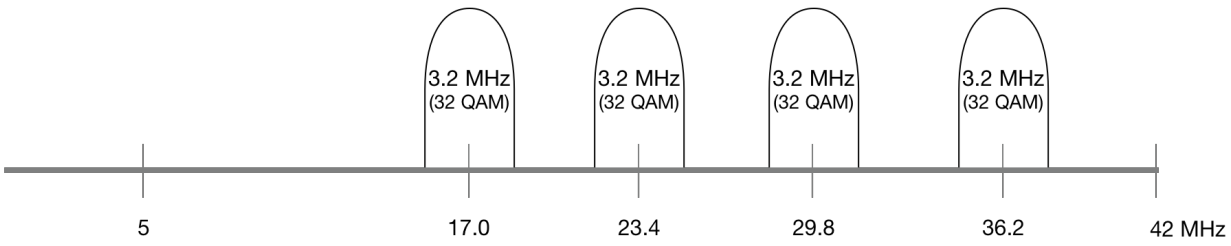


Figure 25 – Large Guard Bands Between Upstream Channels

With four 3.2 MHz wide channels, operating at 32 QAM, the aggregate throughput of this upstream channel is about 44 Mbps which leaves a lot of unused capacity on the table. The guard bands between upstream SC-QAM carriers are not needed; the carriers may be directly adjacent to each other on each side with no guard-band. However, as you place a carrier near the roll-off starting at 42 MHz it is best to

leave around a 500 kHz guard-band from 41.5 MHz to 42 MHz to avoid non-linearities related to amplifier cascades.

7.2.1.2. Case 2: Increase Modulation Orders

The next step for an operator would be to increase the modulation orders to 64 QAM for the upstream SC-QAM carriers. The DOCSIS SC-QAM technology is quite resilient, and the recommendation is to increase carriers to 64 QAM modulation, widen the carriers to use available spectrum, and remove the guard-band between the carriers. As shown in Figure 26 below, the moves should be done stepwise to confirm proper operation.

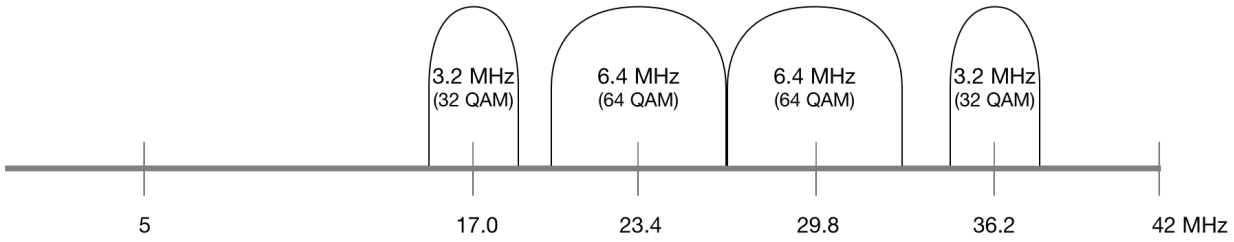


Figure 26 – Increasing Channel Width To Increase Upstream Capacity

With both two 3.2 MHz wide channels operating at 32 QAM and two 6.4 MHz wide channels operating at 64 QAM, the aggregate throughput of this upstream channel is about 72 Mbps.

7.2.1.3. Case 3: Use available spectrum

Another option for an operator would be to use all the available spectrum and using the widest channels. Figure 27 shows using all available spectrum, which could include widening all upstream carriers and abutting those carriers.

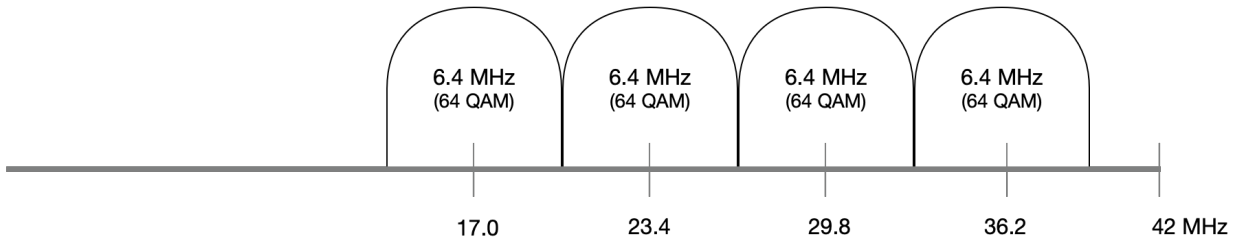


Figure 27 – Maximizing Upstream Capacity With Wide Upstream Channels and Small Guard Bands

With four 6.4 MHz wide channels operating at 64 QAM, the aggregate throughput of this upstream channel is about 100 Mbps.

7.2.1.4. Case 4: Squeeze in an additional carrier

As a final step for a sub-split network, operators could try adding one or more additional upstream SC-QAM carriers, up high or down low as shown in Figure 28 . The paper [Bandwidth Growth] draws the conclusion that even adding 10% additional upstream capacity can help alleviate upstream congestion.

Operators have been successful at adding new carriers, which is a testament to maintaining the plant more diligently over the last decade.

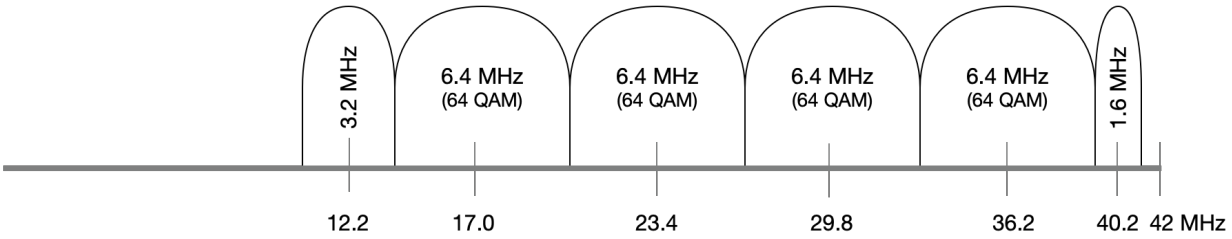


Figure 28 – Adding Additional Upstream Channels

This example adds both one 3.2 MHz wide channels operating at 64 QAM and one 1.6 MHz wide channels operating at 64 QAM, the aggregate throughput of this upstream channel is about 115 Mbps.

The DOCSIS technology has many options for optimizing channel layout and is different from even a decade ago. The digital transceivers have increased in sensitivity and capabilities to enable optionality not thought of even 5 years ago. Homes passed has decreased which lowers the effect of noise funneling at low frequencies. Cascades have shortened which lessens the impact of group delay close to the diplex filter cut-off frequency. Operators have been successfully running narrow carriers (typically with lower order modulation) both down to 10 MHz and closer to the diplex filter.

7.2.2. Sub-Split and OFDMA Channel

If the deployment must stay sub-split and there are enough DOCSIS 3.1 modems on the network, consider removing upstream SC-QAM channels and replacing them with an OFDMA channel. OFDMA technology makes better use of spectrum because it can operate nominally at 1024 QAM whereas an upstream SC-QAM channel is limited to 64 QAM.

An OFDMA channel low in the spectrum will be susceptible to impulse noise. Very strong noise impulses in the time domain might result in the loss of blocks of symbols at the receiver because that impulse noise is spread across multiple symbols.

7.2.2.1. Case 5: Add a small OFDMA Channel

The recommendation here is to put the OFDMA channel high in the spectrum where there is less impulse noise, (at least initially until the operators get comfortable with OFDMA) as shown in Figure 29.

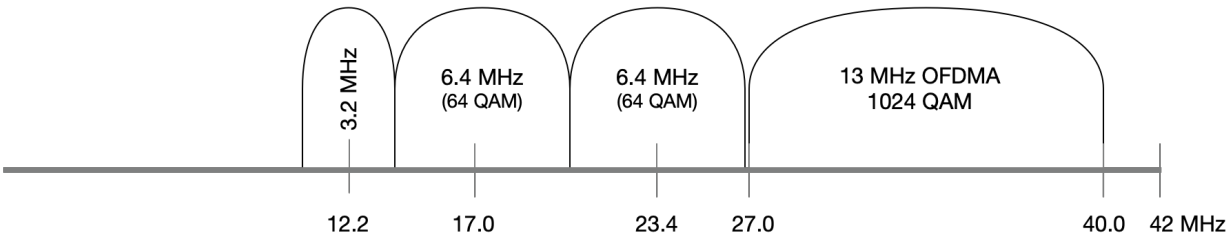


Figure 29 – Sub-split HFC Network with OFDMA Channel

The arrangement of channels shown above can provide an aggregate upstream capacity of 160 Mbps, showing the efficiency of the OFDMA spectrum as compared to SC-QAM spectrum. Note though that only DOCSIS 3.1 modems can take advantage of the OFDMA channel, therefore, there should be a large percentage of DOCSIS 3.1 modems on that plant segment.

Time and Frequency Division Multiplexing (TaFDM) is discussed in Section 5.2.

7.3. Mid-split HFC network

A mid-split network has a return path up to 85 MHz, or two times the spectrum of a sub-split network. With a mid-split, the forward path begins around 108 MHz so set-top box-based video services should be able to be maintained because the forward data channel can be up to as 130 MHz.

The additional upstream spectrum provided by a mid-split network can provide about 500 Mbps of capacity.

7.3.1.1. Case 6: Start with SCQAMs

In terms of additional upstream spectrum, a mid-split can be configured in several ways. As shown in Figure 30, a mid-split can fit ten traditional upstream SC-QAM carriers of 6.4 MHz width (for a total of 64 MHz of upstream spectrum allocated to broadband).

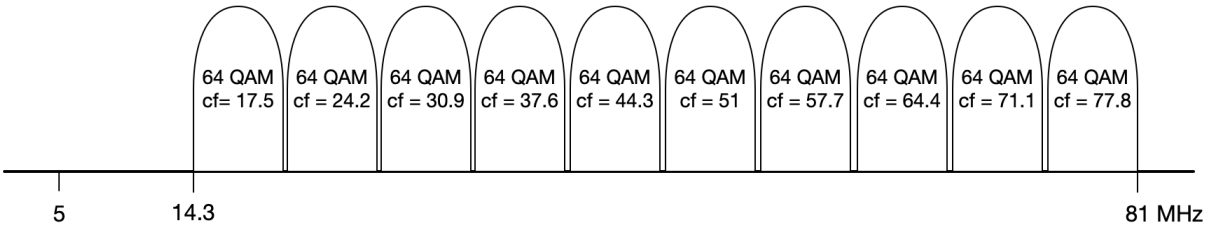


Figure 30 – Mid-split with 10 SC-QAM channels

This configuration can yield up to 250 Mbps of aggregate upstream capacity. Note that DOCSIS technology allows SC-QAM to be modulated up to 256 QAM and no higher. That is, with SC-QAM the options of 512 QAM and 1024 QAM are not available. However, the newer DOCSIS 3.1 technology does allow higher order QAM modulation.

7.3.1.2. Case 7: Add an OFDMA channel

As shown in Figure 31, another example is the traditional 4 upstream SC-QAMs up to 40 MHz, and then 40 MHz of OFDMA running at 1024 QAM from 42 MHz to 82 MHz.

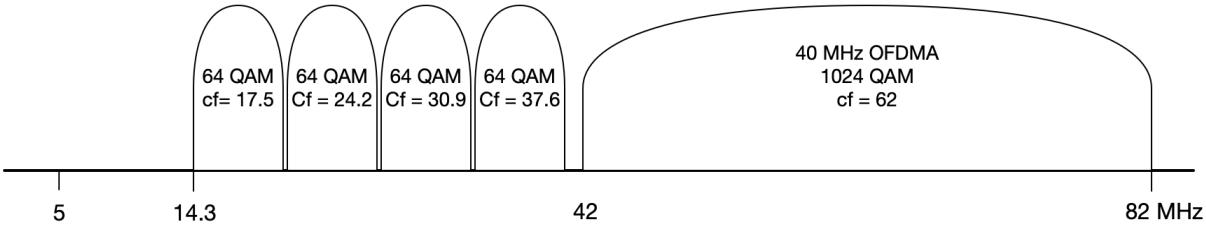


Figure 31 – Mid-split with Both SC-QAM and OFDMA channels

This configuration can yield up to 400 Mbps of aggregate upstream capacity and increase of 150 Mbps using the same spectrum because the OFDMA carrier can operate at a higher order of QAM modulation than the single-carrier QAMs.

7.3.1.3. Case 8: Replace with OFDMA

As shown in Figure 32, another example is shown using 70 MHz of OFDMA at 1024 QAM which can offer 525 Mbps of capacity.

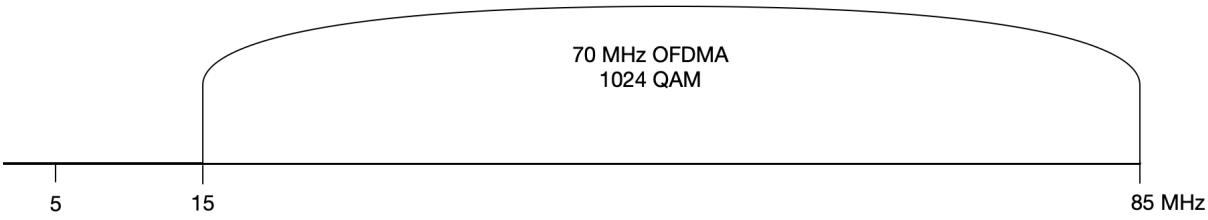


Figure 32 – Mid-split with All OFDMA channel

Though it may be necessary to retain a single SC-QAM upstream channel for DOCSIS 2.0 modems.

7.4. High-Split HFC Network

A high-split network has a return path up to 204 MHz, or four times the spectrum of a sub-split network. With a high-split, the forward path begins around 258 MHz hence set-top box-based video services cannot be maintained because the forward data channel (FDC) can only be moved as high as 130 MHz as described in the SCTE 55 standards.

However, the additional upstream spectrum provided by a high-split network can provide more than 1.5 Gbps of capacity.

7.4.1.1. Case 9: Add a second OFDMA

Figure 33 shows a configuration that retains the four SC-QAMs for DOCSIS 3.0 (and earlier) modems and uses the rest of the spectrum for OFDMA using up to 1024 QAM modulation.

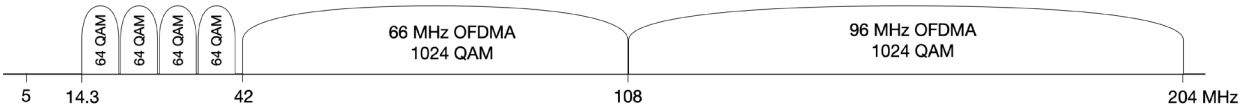


Figure 33 – High-split with Both SC-QAM and OFDMA Channels

This configuration can provide up to 1.3 Gbps of capacity while retaining backward compatibility for DOCSIS 3.0 and earlier modems.

7.4.1.2. Case 10: Full OFDMA

Figure 34 shows all the high-split spectrum using OFDMA and this configuration is capable of over 1.6 Gbps of capacity.

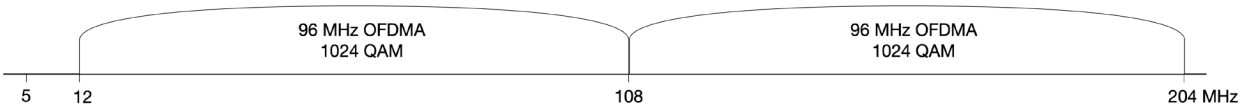


Figure 34 – High-split with All OFDMA Channels

7.5. Ultra-high split Networks

The DOCSIS 4.0 specification builds upon DOCSIS 3.1 OFDM and OFDMA technology with an extended Frequency Division Duplex (FDD) DOCSIS alternative. DOCSIS 4.0 FDD supports both mid-split and high-split and provides extended upstream splits up to 684 MHz in an operational band plan that is referred to as Ultra-high Split (UHS). DOCSIS 4.0 FDD also introduces expansion of usable downstream spectrum up to 1794 MHz to support the higher upstream splits.

8. Conclusion

For the past 20 years in most of the world, the DOCSIS upstream has included up to four SC-QAM channels. With DOCSIS 3.1 technology, OFDMA technology became available for the upstream. As compared to SC-QAM, OFDMA has more configuration options which can be used to optimize upstream efficiency.

As DOCSIS technologies evolve, so do the tools and understanding the operators need to ensure that their networks are running as efficiently as possible, while still maintaining robust service offerings. There is a fundamental trade-off between robustness and throughput. This paper analyzes those trade-offs and discusses methods to increase the efficiency of the DOCSIS upstream. Included are discussions on:

- SC-QAM technology,
- OFDMA technology,
- Interference considerations when combining SC-QAM and OFDMA on the same upstream,
- DOCSIS protocol efficiency,
- Adding more spectrum to the upstream.

A first step to increase upstream efficiency is to use as much of the available spectrum as possible with existing SC-QAM technology. SC-QAM technology has a well-established history and experience, though potentially less flexible when compared to OFDMA technology. There are still multiple parameters like minislot size or FEC size choices which when optimized will improve the upstream efficiency.

A second step includes consolidating SC-QAM technology and allocating spectrum to OFDMA technology. Mixing both SC-QAM and OFDMA technology on the same upstream allows support for older cable modems (DOCSIS 3.0 technology and earlier) as the penetration of DOCSIS 3.1 modems increases. Adding OFDMA allows the DOCSIS 3.1 modems to take advantage of the newer technology to increase the overall throughput of the upstream.

OFDMA technology can provide more flexibility for spectral efficiency but carries more configuration complexity to be considered across all networks as well as individual network conditions. A prerequisite for deploying OFDMA technology is working through the optimal configuration of the upstream parameters which include trade-offs for robustness and capacity. For example, an operator needs to understand how a CMTS enforces guard bands in the upstream and think through the implementations of TaFDM on the CMTS. The overhead of the cyclic prefix can gain robustness but at the cost of capacity. The 2K FFT is more robust and the 4K FFT provides more throughput. 4k is more efficient than 2k for reasons of narrower receive spectral window and required CP being half the symbol period, 4k also interferes with adjacent SC-QAM less. Designing profiles to match the channel conditions on a particular part of the plant, using a profile management application (for OFDMA and SCQAM channels) allows an operator to gain robust operation and extra capacity. OFDMA is worth the effort because it is the new foundation for upstream technology just as OFDM has become the new downstream technology.

Lastly, to get more out of the upstream consider adding new spectrum by changing the upstream split. Just continuing to segment the HFC network does not provide more speed, rather, segmentation replicates the existing spectrum and provides the same speeds for fewer users which does have benefit but does not enable the true capability of the HFC network. To get to both more upstream capacity and really fast speeds, more spectrum must be allocated to the upstream. This new spectrum should be allocated to OFDMA technology to make the best use of it.

Abbreviations

AP	access point
AWGN	average white gaussian noise
bps	bits per second
CM	cable modem
CMTS	cable modem termination system
CP	cyclic prefix
CW	continuous wave
dB	decibel
DOCSIS	data over cable service interface specifications
FDD	frequency division duplex
FDX	full duplex
FEC	forward error correction
FFT	fast Fourier transform
GB	guard band
Gbps	gigabits per second
HD	high definition
Hz	hertz
IUC	interval usage code
kHz	kilohertz
MAC	media access control
Mbps	megabits per second
MER	modulation error ratio
MERperSC	modulation error ratio per subcarrier
MHz	Mega hertz
MSyms/sec	Mega symbols per second
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PMA	profile management application
PNM	proactive network maintenance
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RNG-RSP	ranging response message
RP	roll-off period
RxMER	received modulation error ratio
SC-QAM	single carrier QAM
SCTE	Society of Cable Telecommunications Engineers
TaFDM	time and frequency division multiplexing
UHS	Ultra high split

Bibliography & References

[DOCSIS PHYv3.1] *DOCSIS 3.1 Physical Layer Specification,* CM-SP-PHYv3.1-I18-210125, 2021, Cable Television Laboratories, Inc.

[DOCSIS MULPIv3.1] *DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification*, CM-SP-MULPIv3.1-I21-201020, 2020, Cable Television Laboratories, Inc.

[US PMA] *Field Experiences with US OFDMA and Using US Profile Management,* K. Sundaresan, J. Zhu, and J. P. Fernandes, *SCTE Expo 2020*

[D31 Capacity] *Accurately Estimating D3.1 Channel Capacity,* Karthik Sundaresan, SCTE Expo 2017

[PreEq Analysis] OFDMA predistortion coefficient and OFDM estimation decoding and analysis, Tom Williams, Jason Rupe, Alberto Campos, SCTE Expo 2021

[Bandwidth Growth] *Managing the Coronavirus Bandwidth Surge: How to Cope with the Spikes and Long-term Growth,* John Ulm & Dr. Thomas Cloonan, CommScope, SCTE Expo 2020

It’s 9:00 AM And Your Fiber Is Still Dark

Update to “It’s 10:00 PM Do You Know Where Your Wavelengths Are” (SCTE Expo 2020)

A Technical Paper prepared for SCTE by

Justin Riggert
Principal Engineer II
Comcast
1401 Wynkoop Street, Denver, CO 80202
+1 (303) 378-4227
justin_riggert@cable.comcast.com

Joel Swan, Comcast
1401 Wynkoop Street, Denver, CO 80202
Joel_Swan@comcast.com

Simone Capuano, Comcast
1401 Wynkoop Street, Denver, CO 80202
Simone_Capuano@Comcast.com

Tony Curran, Comcast
1401 Wynkoop Street, Denver, CO 80202
Anthony_Curran@cable.comcast.com

Scott Johnston, Comcast
1401 Wynkoop Street, Denver, CO 80202
Bryan_Johnston@cable.comcast.com

1. Introduction

Last year's paper, titled "It's 10 PM: Do You Know Where Your Wavelengths Are?" provided an overview of Comcast's move into Distributed Access Architectures (DAA) and the need to reinvent how we manage and monitor fiber assets. The DAA architectures call for individual fiber nodes to be smaller and placed farther out into the network which leads to a significant increase in the number of fiber assets that are being installed. All this new fiber needs to be monitored and managed.

"It's 10 PM: Do You Know Where Your Wavelengths Are?" highlighted several innovations in fiber monitoring using a triad of tools. This includes a headend racked Optical Time Domain Reflectometer (OTDR) and Optical Spectrum Analyzer (OSA) monitoring device capable of continuous monitoring of 48 fiber links, a handheld optical meter, and cloud integration software and storage. With these tools it becomes possible to quickly identify fiber cuts and other service impacting fiber degradation events. The tools are used to accurately locate fiber cuts with closest address and cross streets and determine the service impact caused by these events.

But we still find ourselves with a tremendous amount of fiber unmonitored. DAA calls for a great deal of new construction of fiber, and that infrastructure gets much attention and focus. At the same time, we still maintain a huge legacy network. Existing analog fibers, commercial subscribers, and fiber to the premise products also have fiber assets to manage and monitor.

When any unmonitored fiber is cut, a human technician must identify the root cause of the damage. This can sometimes take hours of time -- just to isolate the location of the cut. Yet the goal is that our fibers will not be left alone or dark for long. Whether DAA or legacy infrastructure, digital or analog, residential or commercial services, all services can and will benefit from improved fiber monitoring tools.

"With the CPM (Continuous Pervasive Monitor), we provide infinite attention to individual fiber links with continuous and pervasive monitoring." -- Venk Mutalik, Executive Director / HFC Architecture, Comcast, 2021, in an interview for this paper

In this paper we will explore the steps being taken to expand coverage of our fiber monitoring capabilities. This includes process changes in our operational and development organizations, configuration optimizations and automation to help manage information on hundreds of thousands of fiber links and introduce a new optical passive hardware solution to address physical limitations such as bad optical test ports. We will also cover some of the latest features that have been built into these tools. Enhanced GIS mapping and location of fiber events. Network DVR and time averaging to play back data over time. Refined dispatch tools that combine information about our fiber optics with real time customer equipment status, commercial power outage impacts, and natural disaster events. And we have learned about specific event signatures that allow us to build workflows to automatically route alarms and instruct a fix agent on what to do in easy-to-follow steps, removing a lot of the human guesswork necessary to isolate root cause and required fix. Join us in our ongoing documentary as we improve our fiber monitoring capabilities.

2. DAA Systems

In a DAA system, a CMTS core is located at the Primary Headend (PHE) sending digital signals over fiber optic strands out to a Secondary Headend (SHE). The SHE contains a collection of equipment organized together into a DAA Switch Point of Deployment (DAAS POD).

In the POD, one of the components is the new Continuous Pervasive Monitor (CPM). The CPM is a 48-port rack-mounted device which includes a 1611nm OTDR and an OSA for monitoring individual fiber links. This monitoring unit is capable of continuously tracking fiber length with the OTDR, to isolate physical abnormalities within the fiber strand, while the OSA is full fledged spectrum analyzer on the C-Band and provides monitoring of all wavelengths of light present on the fiber. These monitoring capabilities allow us, in near real-time, to detect when a fiber length shortens, indicating a fiber cut. This monitoring unit also allows us to track when power levels on existing wavelengths go down, when specific wavelengths disappear, or when wavelengths show up that were previously not there.

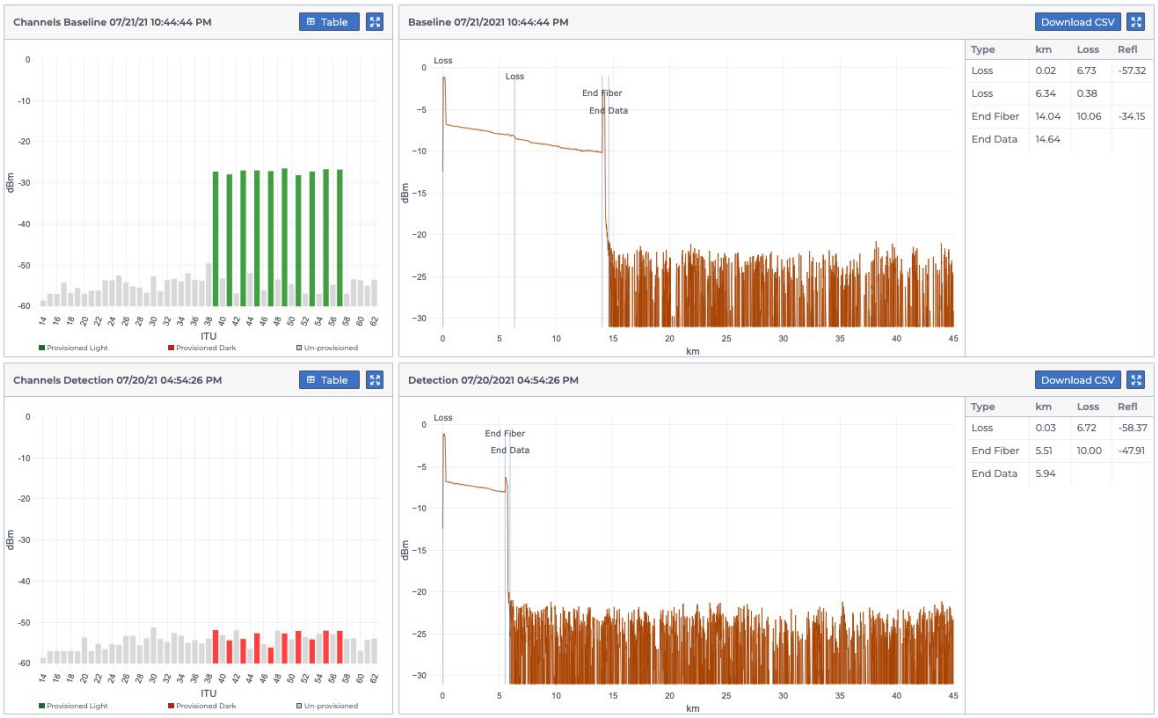


Figure 1 – Event View: Baseline vs Fiber Cut Side by Side

In a DAAS POD, each CPM port is connected to the Inside Plant Multiplexer (ISP MUX) test ports for monitoring of the fiber links and wavelengths served by this DAAS POD. The ISP MUX is a hardware component that takes many different wavelengths of light and combines them together onto a single fiber strand, enabling dozens of wavelengths to share transport on a single fiber. The ISP MUX is paired with an Outside Plant MUX (OSP MUX) in the field to separate those wavelengths back out and send them on to individual fiber nodes.

These new monitoring and troubleshooting tools provide next generation capabilities to pair with our next generation networks, allowing us to pinpoint root cause of a major outage event within minutes. Not only are we alarming on the outage in 90 seconds, on average, but also immediately identifying the root cause

of the outage and where the break in the network is located -- both as a distance from the headend, and with an estimated geo-location, including latitude and longitude, and a list of all impacted services. This detailed data combined with plant design data allows us to know where to drive to find the cut, what kind of fiber is installed there and if it's aerial or underground plant, and how many other fiber strands are in the same bundle as the fiber in alarm. 100% of new DAA installations are required to be installed with these next-generation tools.

3. Legacy Systems

Although DAA currently takes the spotlight and the most focus on fiber monitoring coverage, we still maintain a tremendous amount of legacy architecture with existing fiber links that serve millions of subscribers. They also need to be monitored better than we do today. The good news is the CPM monitoring tool is not specific to DAA architectures and can be deployed as a monitoring tool on legacy systems as well. That said, there are some specific hurdles to overcome for legacy-facing deployments.

For example, one advantage of DAA is that for new construction, we use all the latest technology, including ISP MUXes that have consolidated test ports. These consolidated test ports provide convenient access for monitoring fiber assets with the CPM, because a single test port can connect to a single monitoring port and provide monitoring capabilities for both OTDR and OSA features. Legacy architectures are not as consistent on what is available for test ports on older equipment. In many cases there are no functional test ports available at all, presenting challenges about how to connect the headend monitoring equipment.

For these cases where there are no available test ports for our monitoring equipment, there is a new rack-mounted device with inline optical test points that can be placed in-line between the ISP MUX and Outside Plant (OSP) MUX. This device samples a small amount of light from the fiber optic strands, for the CPM to monitor. Connecting this equipment in-line removes the dependency on available and functional test ports. This optical device only reduces the optical power levels by about 0.5dBm, which in practice has negligible effect on plant performance. The gains that we see in improved fiber monitoring are well worth the small loss of power.

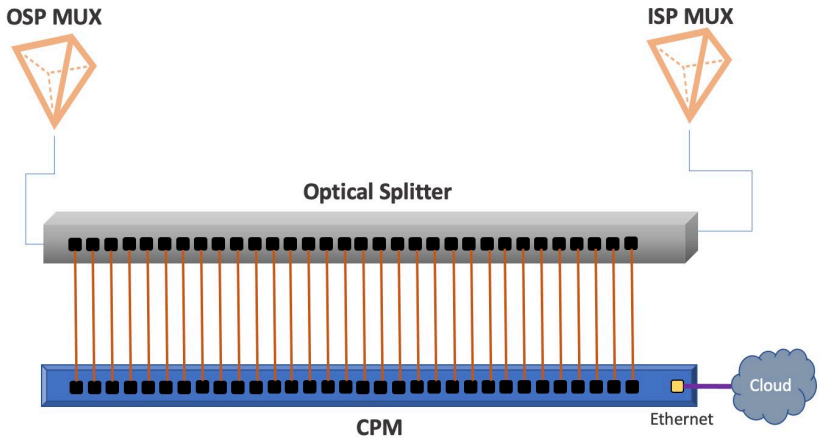


Figure 2 – Inline Optical Test Points Connected to CPM

In other situations, we do have functional test ports available, but they do not support consolidated OTDR and OSA measurements. In these cases, we have modified the CPM software to logically combine 2 physical ports for monitoring of a single fiber link. In this situation, we configure one port to perform the OTDR functions, and another port measures the OSA channels. The application then combines these two ports for monitoring and alarming purposes, so we represent them as a single fiber link to the end user.

Other legacy scenarios include the need to monitor analog optics, or to monitor fibers with wavelengths that are outside the range of the OSA. Both the OTDR and the OSA feature of the CPM are capable of operating independent of each other. The OTDR feature can monitor a fiber for length and physical abnormalities on its own in cases where we do not have visibility into the wavelength power measurements. And the OSA feature can be used independently to monitor wavelengths in places where the 1611nm laser is blocked from giving us monitoring visibility of the physical fiber attributes. We have many reasons to be optimistic about our legacy networks benefiting from these new innovative monitoring technologies.

4. Handheld Optical Multimeter & OTDR

This past year brought many new powerful features to the handheld unit. The home screen has new intuitive control bar and menu selection area for the technician. The control bar provides real-time status on both the handheld unit and the SFP connected to it. There is a real-time indicator for laser status, whether the SFP is programmable, the current ITU wavelength configured on the SFP, and a threshold warning if there is too much power coming into the OSA or OCC ports. There is now an SSO login, for technicians to connect their scans to their account in the cloud. The control bar also provides status on WIFI connection, power status, and battery levels.

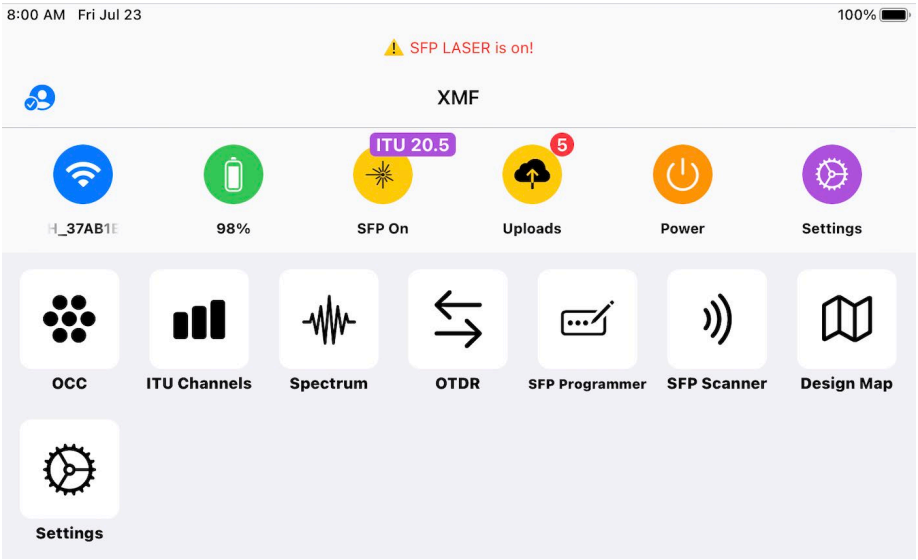


Figure 3 – Handheld Unit – Control Bar and Menu

With the SSO login integration and ability to upload scans to the cloud, we now can search and view those scans by technician, job ID, node, or geographical bounds drawn on a map. This allows the user to compare different scans taken at different times or locations, view progression of a particular job, and hand off information about a particular job from one technician to another.

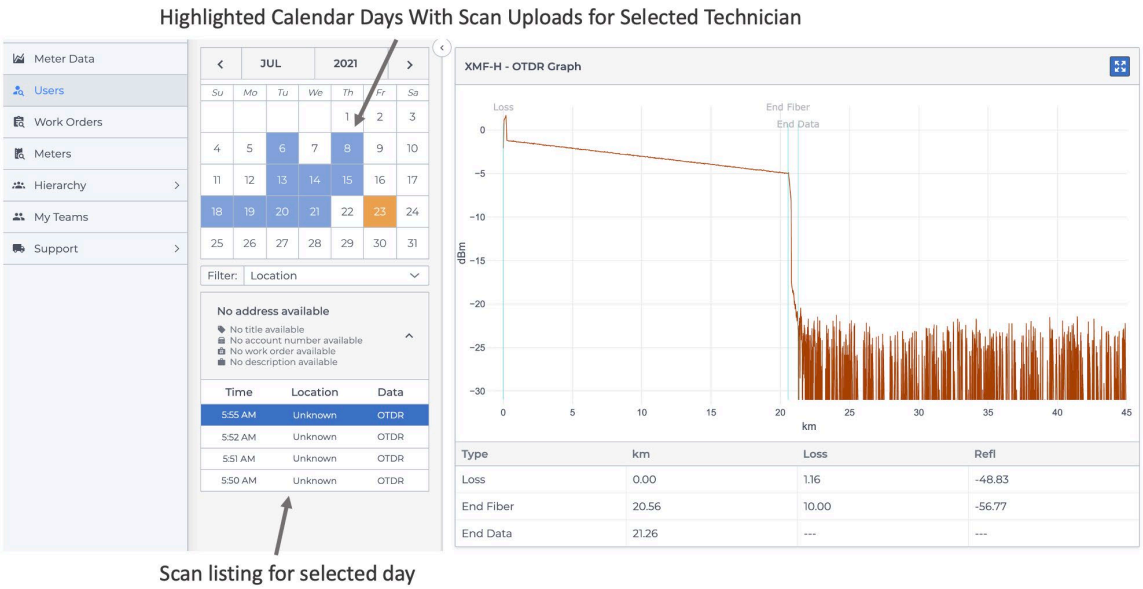


Figure 4 – Cloud Navigation for Uploaded Scans (Technician View)

The handheld device also includes an Optical Channel-band Checker (OCC) which is a quad-band power meter that can sense power levels on the fiber. The OCC provides power measurements across the entire usable spectrum. It measures and displays the detected power levels, grouping them into four bands:

- 10G EPON Downstream (L-Band 1577 mm)
- BAU DWDM (C-Band 1550 nm)
- 10G EPON Upstream (O-Band 1270 nm)
- All Other Bands not specified above

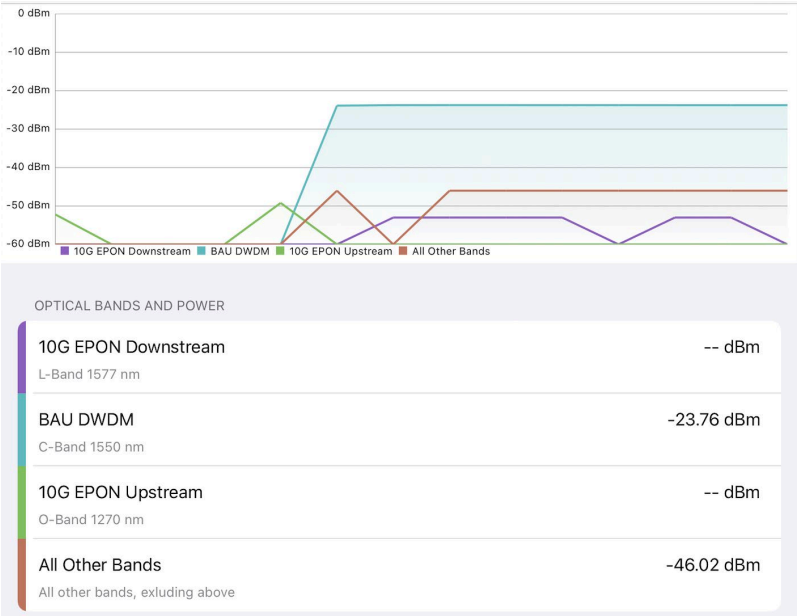


Figure 5 – Optical Channel Checker Application View

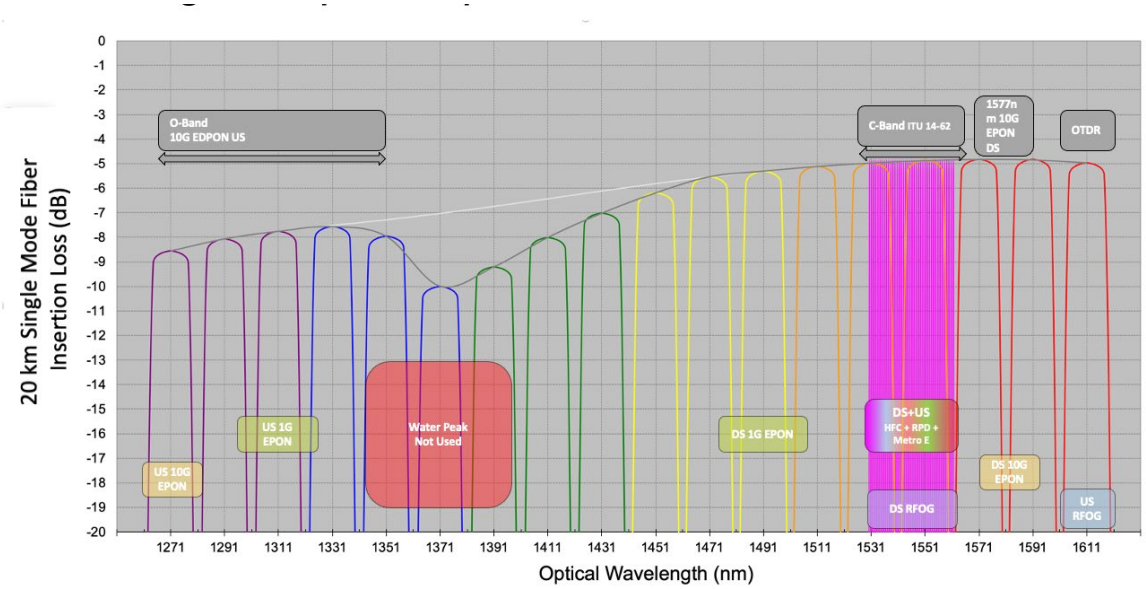


Figure 6 – OCC Wavelength Band Overview

Another powerful new feature on this platform is the integration of the handheld device with the headend rack-mounted CPM. With these technologies integrated with the cloud, a technician can be out in the field troubleshooting an outage. While physically located miles away from the headend, the technician can remotely search the system for the fiber being tested, and connect to the CPM that is monitoring the same fiber link. The technician can trigger an OTDR from the headend, down the fiber link, and within seconds shoot the fiber from their current location back towards the headend with their handheld device.

With cloud storage, all measurements are stored for future playback and analysis, and to provide visibility to other technicians who may need to assist with the same outage. The combination of headend and handheld equipment is designed with a similar look and feel, and with comparable features on both platforms. This helps to foster a common language and a shared platform used by both the headend and field technician, where the two disciplines now overlap in ways that were not previously possible.

While connected to the remote CPM in the headend, another benefit to the technician is the ability to be "virtually hands-on" with a fiber strand located miles out into the network, taking measurements of the fiber from the perspective of the headend CPM. For example, by creating a bend in a fiber strand, the technician can look for loss in the micro bend on the OTDR traces of the CPM, enabling validation of the fiber strand and where it connects. Changes to the network can be made instantly and verified by the changes in the receive power at the headend.

While troubleshooting the network and requesting live traces from the CPM, the software is also configured so that the continuous monitoring feature doesn't miss a beat. The CPM can quickly switch between ports as part of its ability to monitor 48 fibers with a single optical receiver. This switching capability allows a technician to request live data from the CPM, and this request is interleaved with the scans that are required for ongoing monitoring. The CPM creates a queue for any live requests coming in, and any live requests are processed in the order received. But no matter how many live requests are in queue, every other scan is still prioritized for continuous monitoring. This matters because we will never miss detecting the next major outage and creating a dispatch ticket with detailed information about that fiber cut, while in parallel we can still provide real-time troubleshooting capabilities to the technicians working in the field.

5. Installation and Configuration

With a target of 100% of DAA architecture including CPM monitoring on every fiber link, keeping up has proven challenging. The initial installation of the CPM requires power, network, and configuration for every wavelength being served. And BAU updates over time need to be updated on the device configurations.

Manual configuration of each fiber link and wavelength becomes a tedious and time-consuming process as the number of DAA installations increases. The hours required for manual configuration has been one of the roadblocks encountered in our efforts to get 100% adoption from the field. This led to one of the major new features enhancements this year, allowing configuration that scales and better acceptance from the user base managing these configurations. This new feature is bulk import capabilities from the fiber configuration spreadsheets used by each local headend.

Each headend is already required to maintain an up-to-date configuration of its fiber links, and the format of these configuration sheets is standardized across the various regions. By supporting bulk configuration of the data format already in place, this reduced the configuration time from a multi-hour effort to just a few minutes.

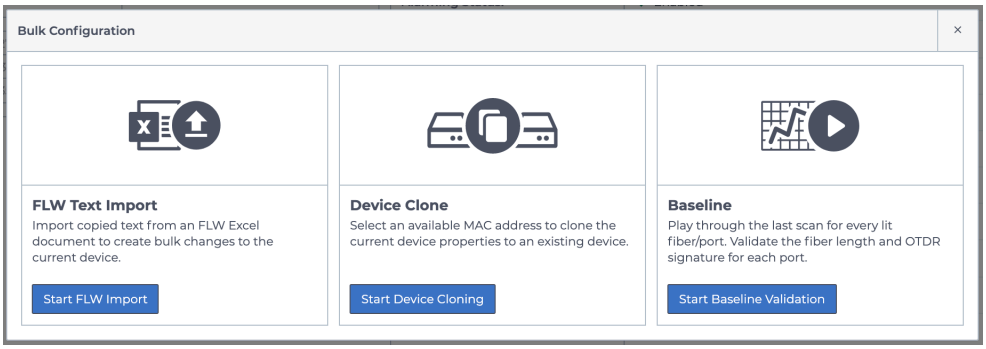


Figure 7 – Bulk Configuration Selection Screen

The next phase of configuration automation is to centralize all the local fiber configuration data into a database, with APIs the CPM can use to fully automate the configuration and relationships between ISP MUX, OSP MUX, CPM, wavelength, and fiber node. By centralizing and providing APIs to access all the fiber configuration data, the additional configuration required that is specific to the CPM becomes minimal, and the installation process for the CPM becomes scalable even for a busy field operations staff. By centralizing fiber configurations into a database, we also enable the CPM to receive real time updates as BAU configuration changes are made over time.

Another critical element of configuration is establishing a baseline of the expected fiber length, as well as the expected power of each wavelength at the receiver in the headend. The baseline is what allows us to determine if a fiber is short or if individual wavelengths are degraded from their normal operational state.

The initial plan was to manually certify each baseline by putting "eyes on glass," and allowing a person to identify flaws in the OTDR trace that may have "slipped through the cracks." This provides an opportunity to clean up bad splices or other physical flaws during the installation process and provides a better physical network that gets handed off to operations after installation and configuration.

Although manual verification was well-intentioned, as we started to deploy monitoring on what quickly became thousands of fiber links, it became a roadblock to field adoption of the tool and willingness to put in the time to configure. So, we pivoted to an automated baseline feature where the software looks for trends over time. When we have specific wavelengths that stay continuously lit for a specified period, this triggers that wavelength to become eligible for ongoing monitoring. And during this time, when we flip wavelengths from unmonitored to monitored, the automation software updates the baseline for that link with the current length of the fiber and the power levels of the wavelength. In addition to the automatic creation of fiber baselines, the automation software also must be aware of trends over time and watch for fibers that may have been decommissioned or changed monitoring ports.

Another installation and configuration feature that has been added is related to how we monitor various fibers with different physical attributes. For DAA, we see fiber lengths ranging from a few hundred meters to 50 km or more. We have variations in the amount of light coming into the headend receivers. And there are various kinds of glass strand and physical anomalies in the fiber link that lead to different attenuation and refraction patterns.

With the OTDR portion of the CPM we continuously monitor the length of fiber looking for when the fiber is cut or damaged. The challenge has been that an OTDR must be tuned to accurately analyze the fiber strand based on distance and other physical properties.

One of the tuning parameters is pulse width. A short pulse width provides a small "dead zone" at the beginning of the fiber, where this dead zone represents a blind spot where it is not possible to accurately measure the end of fiber or find anomalies. A short pulse width also provides the ability to detect distinct events that are stacked close together. However, a short pulse width also reduces the dynamic range of what we can measure, so in cases where there is a high amount of attenuation in the DAAS POD installation, the short pulse width can cause the signal to get lost in noise before being able to measure the end of fiber. Short pulse width also loses its energy over long distances, so the signal gets noisy at longer distances.

Long pulse widths provide better dynamic range, but we lose monitoring capabilities within the blind spot of a longer pulse width/ larger dead zone. And we lose granularity.

Averaging time is another tuning parameter available to us. In general, the greater the averaging time, the better the results on the OTDR trace. But the cost of higher averaging time is longer time to fiber cut detection. As the CPM loops through the 48 ports, it is looking for fiber cuts and other events in real time. If the averaging time is set too short on a particular port, we gain a quick measurement but lose to a noisy signal and inaccurate analysis of what is happening on that fiber. If the averaging time is too long, we gain accuracy but lose on the time it takes to detect a significant issue like a fiber cut.

IOR (Index of Refraction) is another parameter that is important to tune so as to achieve accurate distances on the OTDR analysis. IOR is a representation of how fast light travels through a particular medium. Set IOR too low, and all the distances measured will be too long. Set IOR too high and all the distances will be too short. Inaccurate distances cost the field technician time when trying to find a fiber cut. And inaccurate distances also can cause us to misidentify which sheath the fiber is contained in, which gives us access to what other fibers are bundled together, and then what other services are connected to all those fibers in the sheath -- which may all be down, if the entire sheath is cut.

To solve these different tuning requirements, we continue to look to automated configurations based on fiber design, combined with AI and machine learning using the data collected and stored in the cloud. This allows us to customize the settings for each individual fiber link.

6. Cloud Integration

One of the most empowering new features that these new tools have added is integration with the cloud and other internal tools, to enrich the data we gather and publish. In the handheld meter, one of these features is integration with node inventory and jobs. Field technicians are often troubleshooting fiber nodes, and one of the features in the handheld meter is the ability to search for a node to tag onto existing scans. By tagging scans with a fiber node (or other equipment), when the scan is stored in cloud this data is indexed on node name. This allows anyone to search the cloud for scans that have been taken on this node over time, which provides visibility into patterns and historical context. Another feature for the field technician is "job search," where a technician can tag handheld meter scans with a job currently assigned to them. The job association allows the work the technician is doing to be visible to others, which benefits other technicians if an issue is escalated to them, or it gives more visibility to all technicians working the same job in tandem.

On the CPM headend tool, another focus over the past year has been enhanced DVR functionality. The DVR is a tool that allows the user to view a particular fiber link over time, in a movie-like experience, where every scan taken over the selected period is played back in sequence at a selectable frame rate. As the user watches the playback, they can isolate changes in pattern and visualize how the OTDR trace changes over time.

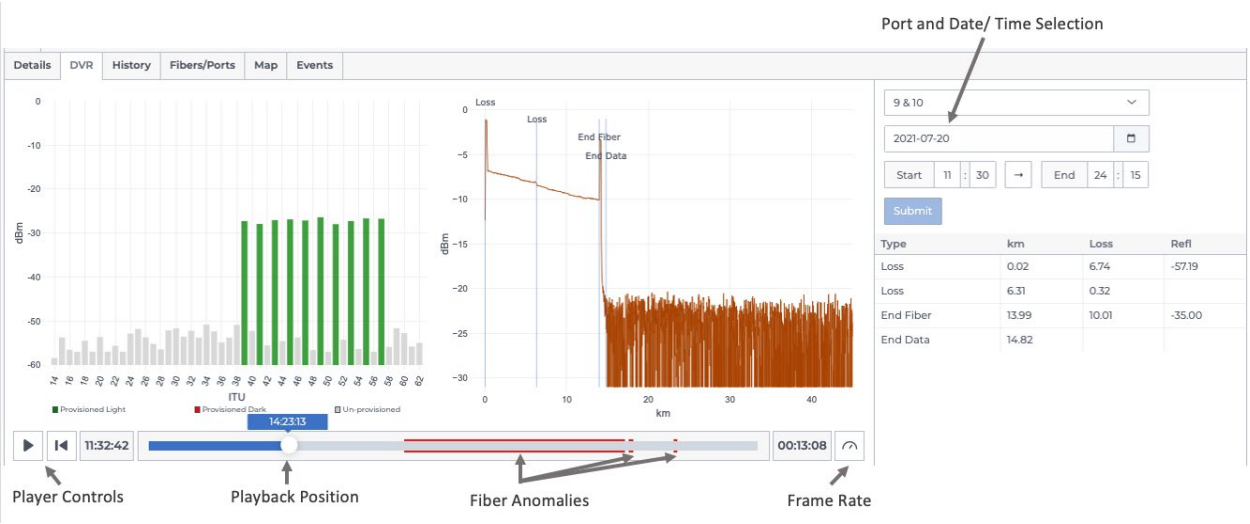


Figure 8 – DVR Playback View

An alternate tool to the DVR is the history tool, where the user can select a time window, but instead of playing frame by frame, they are presented with min, max, and average lines over the selected time.

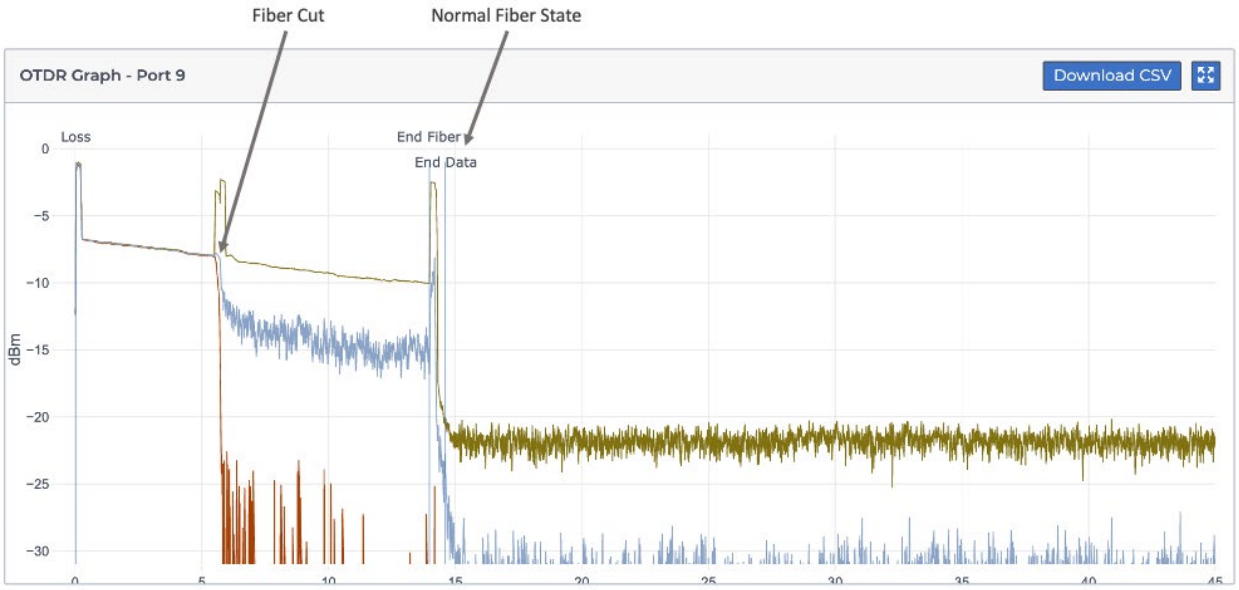


Figure 9 – Historical Averaging View

7. Validating Design Maps

One of the ongoing efforts is to validate the design data for our fiber assets. All our fiber assets are stored in a GIS asset database that includes the physical path where the fiber travels, the length of each fiber segment, and includes the helix factor, amount of sag, and amount of fiber contained in slack loops. By using the OTDR and measuring physical distance from the headend, we can audit the design data to validate distances and correct mistakes that are found. We can also store regional averages for helix, sag, and slack that let us better estimate fiber cut locations by adjusting the distance of the fiber to account for the regional helix, sag, and slack averages. One of the other benefits of our GIS asset database is that the entire fiber run is connected in the database from the fiber node to the headend. With this connectivity, we can trace from the fiber node to each splice, through MUX and various fiber sheaths, to the termination panel, OCEF, and ISP MUX in the headend. With access to all the physical connections being made along the fiber path, we can correlate specific events in the OTDR trace to specific equipment. We can also help audit the GIS asset database to flag fiber lengths that do not match the OTDR measurements and provide reports to the field that can be used to clean up bad or missing data.

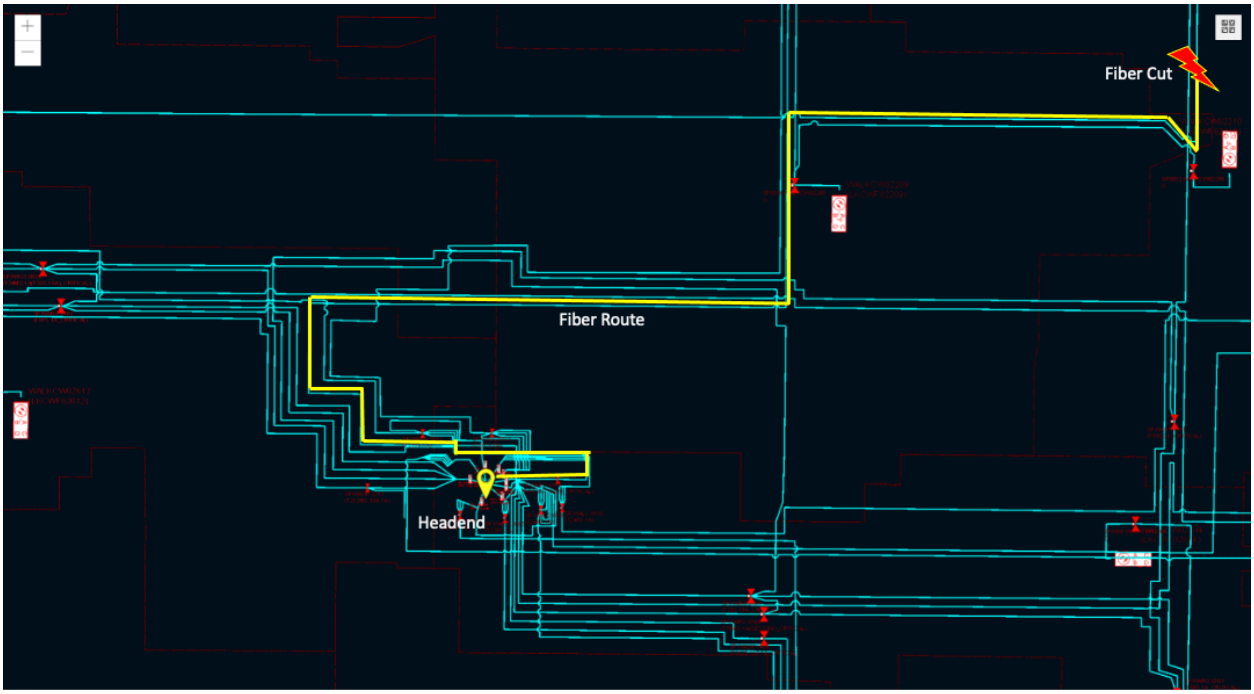


Figure 10 – GIS Asset Maps/ Route to Fiber Cut

8. Natural Disasters

Another opportunity for the CPM to provide unique value is for large scale natural disaster events. During these large events, there are an overwhelming number of alarms that are generated by all the different monitoring tools on the network. Trying to triage the alarms and work them in a prioritized order is not an easy task. The OTDR fiber cut detection capabilities of the CPM becomes an amazing tool to help with this prioritization process. With the OTDR functionality, identifying a fiber cut is highly accurate. Many of the other alarming tools deployed on our network are good at finding outages but require manual research to find the root cause of the outage. The CPM can immediately identify root cause for an outage

as a fiber cut. Combining the fiber cut with configuration data that tells us what is deployed on that fiber lets us identify the scope of impact of each fiber cut that is still active and helps the field better triage which cuts they want to fix, in what order.

Another reality with natural disasters is that during large scale events there are some outages that stay open for many days, or, in rare cases, many weeks. As the CPM tools get better about learning the network and setting baselines, we also need to program the tool with a method to anticipate these large-scale events, so the tool behaves in an appropriate way in context of all open alarms. For example, one of the asks of the CPM is the change the baseline length of a fiber in cases where there has been an intentional change to the fiber length. These scenarios happen in cases like node splits, or when the headend makes physical configuration changes. If the fiber monitored by the CPM does physically change, the baseline needs to change with it. One method to automate this change is to look for configuration changes applied in the fiber configuration data. But we do not always have visibility into those configuration changes. In these cases, we can look for trends on the physical characteristics of the OTDR trace and decide when to update the baseline using those trends. But then in a disaster scenario, we want to suspend those trend-based updates, so the algorithms do not make mistakes like setting a fiber to "inactive" in a case where we have a real and ongoing customer-impacting fiber cut.

9. Conclusion

As we reflect on the past year and look at lessons learned and the new development that has gone into this monitoring platform since we last convened to describe it, the value of this innovation continues to surpass what the development team had initially envisioned. Fiber optics are a critical part of the service delivery platform serving our customers, and our fiber assets continue to expand deeper into the network. Our customers rely on this network around the clock for video, internet, phone, home security, home automation, and emergency services. Our customers depend on our network in their everyday lives; therefore, reliability is of paramount importance. Fiber optic monitoring tools are one of many ways we are improving on that reliability. We began with DAA as the target for these tools, but over time we continue to uncover so many remarkable new opportunities to provide better insight into our fiber optic networks. So many legacy systems, analog optics, commercial services, and fiber to the premise products can benefit from these monitoring innovations. And in parallel to the technology development, working with the field and operational leadership to shape process changes and training around these new products. We look forward to new discoveries of what this platform can do for our company. We also embrace this technology as just a single part of a broad effort at Comcast to build a culture of reliability across all products. We are winning the battle against fiber dark forces, ensuring that we wake up in the morning knowing our fiber is light and watched over.

10. Acknowledgements

OTDR and Optical Spectrum Analyzer technologies have been around for decades and are common tools used for monitoring and managing long haul fiber networks. But it was not until 2019 that our own Venk Mutalik, Executive Director of HFC Architecture was inspired to take these technologies used in long haul environments and apply them to the local fiber optic networks serving our residential and commercial customers. It has been a privilege for the team to follow Mutalik's technical leadership in our journey -- to take his invention and build enterprise scalable applications and tools that provide real operational value to Comcast.

Abbreviations

BAU	Business As Usual
CMTS	Cable Modem Termination System
CPM	Continuous Pervasive Monitor
DAA	Distributed Access Architecture
DAAS	Distributed Access Architecture Switch
DWDM	Dense Wave Division Multiplexing
EPON	Ethernet Passive Optical Network
GIS	Geographic Information System
ISP	Inside Plant
FLW	Fiber Loss Worksheet
OCEF	Optical Cable Entrance Facility
OCC	Optical Channel Checker
OSA	Optical Spectrum Analyzer
OSP	Outside Plant
OTDR	Optical Time Domain Reflectometer
PHE	Primary Headend
POD	Point of Deployment
SHE	Secondary Headend
SSO	Single Sign-On

Bibliography & References

1. *It's 10pm: Do You Know Where Your Wavelengths Are?* Venk Mutalik, Dan Rice, Rick Spanbauer, Simone Capuano, Rob Gonsalves, and Bob Gaydos, SCTE EXPO 2020

Leakage Detection in a High-Split World

Industry Progress Toward a Viable Solution

A Technical Paper prepared for SCTE by

Rex Coldren
Principal Access Architect
Vecima
Phoenix, Arizona
+1 (602) 206-2690
rex.coldren@vecima.com

Michael Cooper
Principal Engineer
Cox Communications
Atlanta, Georgia
+1 (404) 858-4276
michael.cooper4@cox.com

Greg Tresness
President
Arcom Digital
Syracuse, New York
+1 (315) 422-1230
tresness.greg@arcomlabs.com

1. Introduction

System leakage monitoring and detection in a Low-Split and Mid-Split world involves detecting leakage of transmissions originating from a CMTS, CCAP, R-MACPHY node and R-PHY node. Legacy methods for accomplishing this have been in place for many years and are well-understood.

In Low and Mid-Split scenarios, the aeronautical band from 108 MHz to 137 MHz lies within the downstream spectral band. With DOCSIS® 3.1 High-Split and DOCSIS 4.0 Ultra-High Split, the aeronautical band will fall within the upstream spectral band. As a result, system leakage monitoring and detection in High-Split and Ultra-High Split scenarios involves detecting leakage of transmissions that originate from a cable modem (CM). This requires a completely new way of approaching the problem.

The cable industry has made significant recent progress analyzing the alternatives available to solve the High-Split and Ultra-High Split leakage detection problem, specifying the necessary support in industry standards, and validating in laboratory and controlled outdoor environments. This paper discusses the progress that has been made on the most promising of these alternatives which has now become part of the DOCSIS 3.1 specifications.

2. Basic Concepts of the Chosen Solution

A paper from Comcast and Arcom Digital [4] described the problem well and offered four possible solutions. It suggested that the most viable of these alternatives is one where the CM is granted opportunities to transmit OFDMA Upstream Data Profile (OUDP) test bursts under the control of the CMTS, CCAP, or R-MACPHY node (hereafter referred to simply as "CMTS"), in an operator-specified leakage detection test spectral region in or near the aeronautical band. Figure 1 illustrates an example of such a leakage detection test region that is placed between 138.1 MHz and 139.7 MHz in an OFDMA channel that spans 108 MHz to 204 MHz.

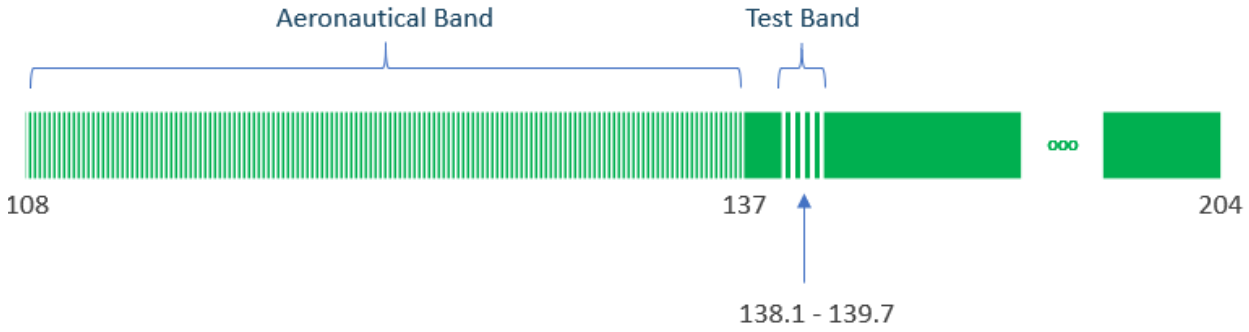


Figure 1 – Example Leakage Detection Test Region

DOCSIS 3.1 CMs are required to support transmission of OUDP test bursts on all OFDMA channels, as specified in [2]. An OUDP test burst has a specific MAC header and payload data format. It is transmitted by the CM using the modulation order and Pilot Pattern configured in a burst profile definition for the minislots granted to the CM for the test burst.

The CMTS is configured to make OUDP test burst grants covering the operator-configured leakage detection test spectral region. Field detectors tune to this spectral region and detect egress of the OUDP burst signal from the cable network by looking for the specific known Pilot Patterns from these test burst transmissions made within the region by CMs. The Pilot Patterns that are easiest to detect and

correspondingly provide the greatest detector sensitivity are those which are most densely populated with pilots. The two densest Pilot Patterns are illustrated in the [1] excerpts in Figure 2 and Figure 3. Pilot Pattern 4 is recommended for 2K FFT leakage detection test bursts. Pilot Pattern 11 is best for 4K FFT leakage detection test bursts.

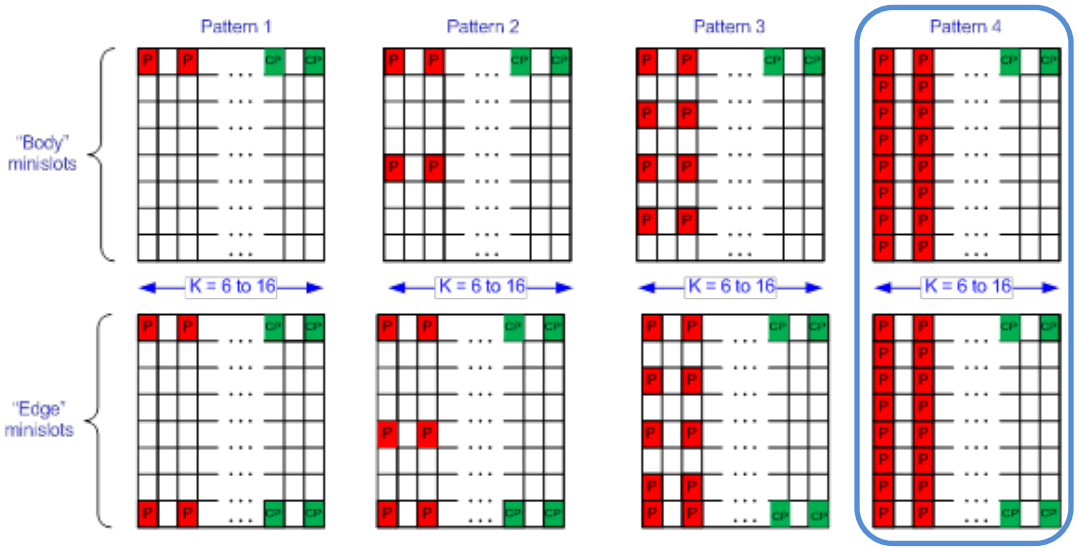


Figure 2 – Recommended 2K FFT Pilot Pattern for Leakage Detection Bursts

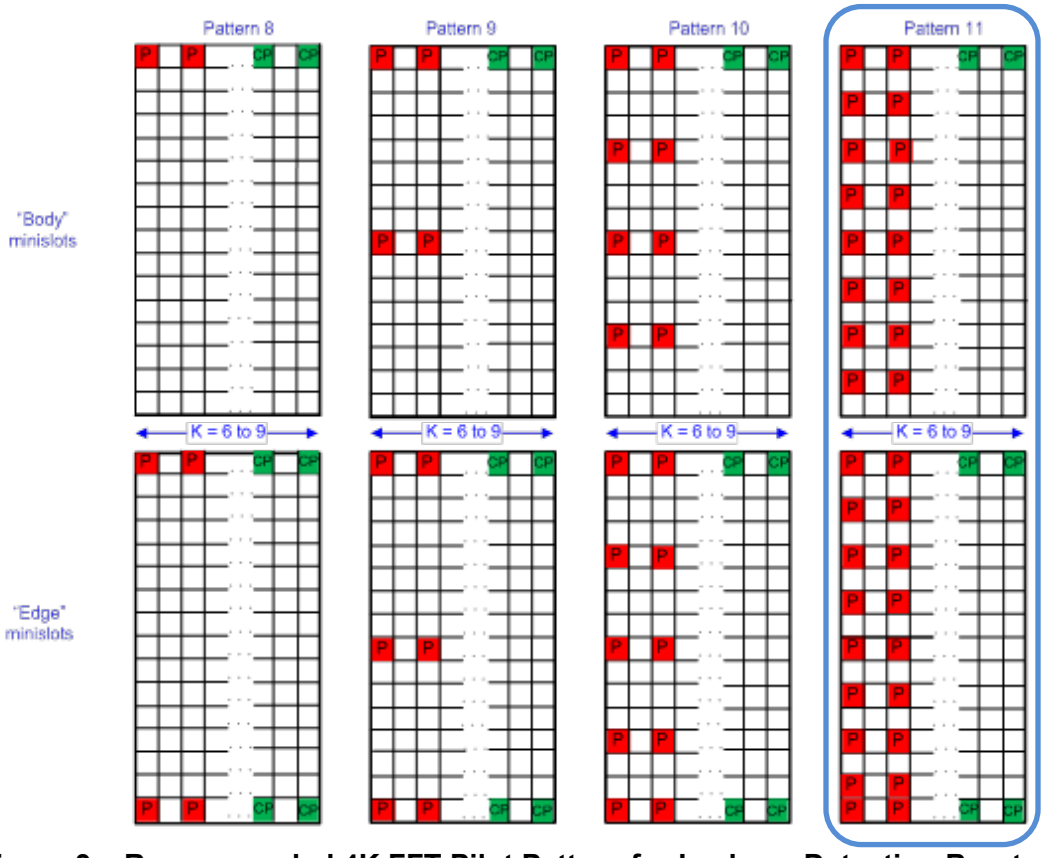


Figure 3 – Recommended 4K FFT Pilot Pattern for Leakage Detection Bursts

A user interface, Command Line Interface (CLI) and/or Application Programming Interface (API), is enabled at the CMTS to define the scope of leakage detection tests and to control scheduling of OUDP test bursts within the specified scope. For example, a CLI would specify which CM(s) to test, the frequency range of test region to use for grants, OUDP test burst duration, etc. An information model representing the configurable attributes for such a user interface has been specified in [3].

The user interface also supports reporting metrics that verify test performance. In a perfect world there will be no leakage detected. Without test metrics, there is no way for the operator to be certain that grants were made to CMs and that granted CMs responded with test burst transmissions. For this reason, a complete solution provides leakage detection test metrics per tested CM.

3. Leveraging the Solution

3.1. Leakage Detection Testing Use Cases

The DOCSIS 3.1 specifications support testing of a single CLI/API-specified CM, a CLI/API-specified ordered list of CMs, a CMTS-determined ordered list of CMs based on a CLI/API-specified upstream scheduling domain (e.g., OFDMA channel ID, MAC Domain), or CMTS-determined ordered lists of CMs created for every upstream scheduling domain in a CLI/API-specified scope, such as an entire CMTS or a specific R-PHY or R-MACPHY node.

The variety of use cases supported allows the operator to perform leakage testing according to their desired implementation. This facilitates lab testing of CMs for support of OUDP test burst capabilities and field detectors and applications for performance in various leakage scenarios. It enables both targeted (e.g., per CM, per upstream scheduling domain, per node) and sweep-based (e.g., per CMTS) testing and provides an opportunity to combine use cases to support maximum automation and intelligence in the leak identification process.

As one possible use case, consider a scenario with a theoretical Flexible MAC Architecture (FMA) MAC Manager which provides management functionality for three R-MACPHY nodes that are deployed in the same geographic area. Each of these R-MACPHY nodes has two Upstream Service Groups (i.e., upstream scheduling domains). A single leakage detection test session is configured at the MAC Manager to cover all three of these R-MACPHY nodes. All of the upstream scheduling domains in all of these R-MACPHY nodes will then simultaneously have leakage detection test sessions running. Each leakage detection test session automatically includes all DOCSIS 3.1 CMs which are currently using the specific OFDMA channel that is undergoing testing. This widespread sweep coverage approach, illustrated in Figure 4, is most suitable for operators that just want to set up leakage testing and leave it. It would also be appropriate at times when aeronautical flyovers are being performed for leakage detection. Leakage would be detected by field detectors encountering leaked signals from OUDP test bursts. Leakage can be isolated manually by monitoring detector strength-of-signal and detector proximity to the leak.

Automation can be introduced via field detectors equipped with GPS and associated leakage applications. Field detectors would transmit their current GPS locations to the application, and the application would communicate with the MAC Manager via an API to schedule testing, for example, only in those nodes which currently have field detectors operating in proximity to the node. Scheduling would need to be refreshed every few minutes to account for vehicle movement, but the process is easily automated. Since vehicles equipped with leakage detection equipment are only in a small percentage of nodes at any one time, this process could decrease the cumulative bandwidth required for leakage detection, allowing the test region to be scheduled for normal data transmission when there is no leakage detection occurring on the nodes where service vehicles are not nearby.

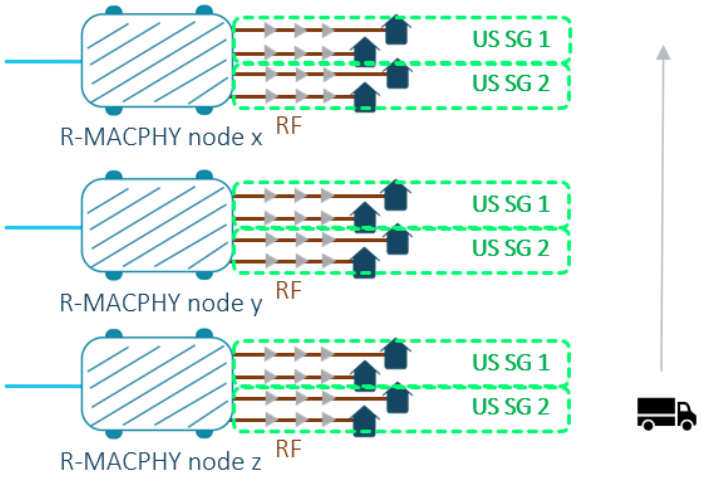


Figure 4 – Example FMA Leakage Detection "Sweep" Test

An alternative use case might be to configure test sessions using an ordered list of CMs. The list of CMs on-line and available to test could be provided by the MAC Manager to the leakage application using an API. The API then would instruct the MAC Manager as to the desired test burst sequence and the desired scheduler control parameters. Since CMs are generating OUDP bursts in a known order and the characteristics of the bursts are known, it is possible using intelligent detection algorithms to determine exactly which CM in the sequence was bursting when leakage was detected. This type of targeted leakage testing is illustrated in Figure 5.

The leakage application could then query a database that associates the MAC address of the CM causing leakage to the physical address of the leak such that repair can be scheduled. This could happen without the field technician leaving the vehicle. As such, utilizing ordered lists can potentially result in significant time savings in the last step leak identification process. The ordered list technique can also be useful in remotely determining whether the leak source is in the hardline (where multiple CMs could be detected as the leakage source) or the drop (where only one CM could be resolved as the leakage source).

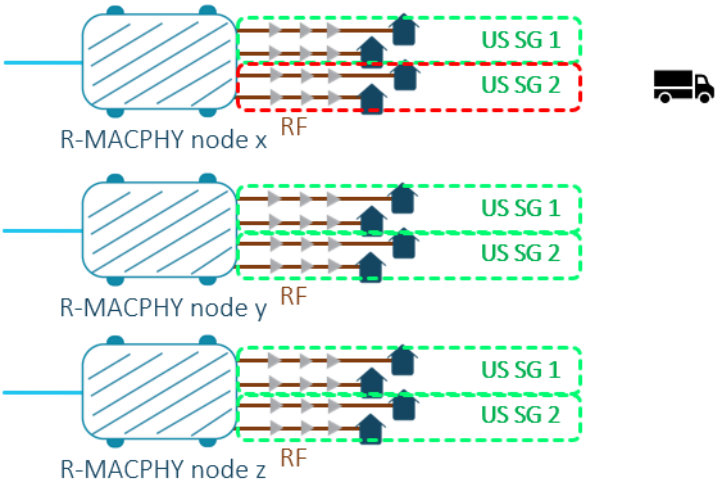


Figure 5 – Example FMA Leakage Detection Targeted Test

A similar process using an ordered list containing only one or a few CMs could also be used as technique to confirm that the CM whose transmissions are suspected to be source of the leak is in fact the exact source. In Figure 6, a single CM in an Upstream Service Group is specifically tested for leakage.

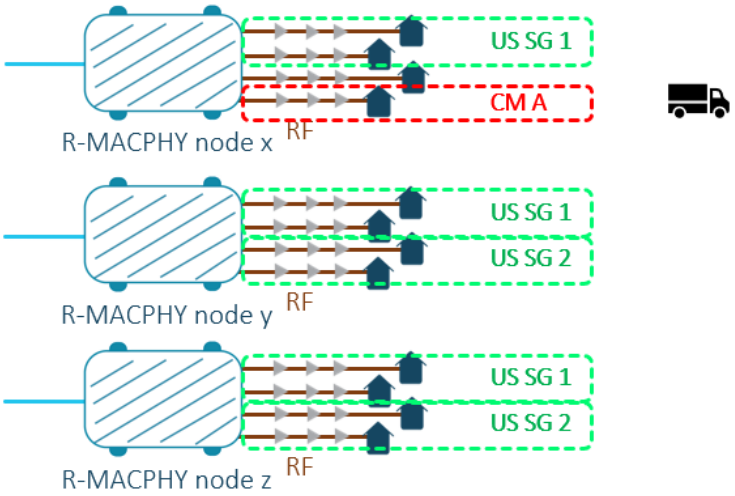


Figure 6 – Example FMA Leakage Detection CM Specific Test

3.2. Configuration of OFDMA Channels for Leakage Detection Testing

Operators are responsible for configuring their OFDMA channels to maximize data throughput. They are also responsible for OFDMA channel configuration aspects to enable leakage detection in High-Split and Ultra- High Split band plans. A leakage detection test region needs to be defined where OUDP test burst grants can be made to CMs by the CMTS. This test region can be in or near the aeronautical band, which lies between 108 MHz and 137 MHz. Ideally the test region is just above the aeronautical band so that leakage testing does not actually result in leakage into the aeronautical band.

A modulation profile, or Interval Usage Code (IUC), needs to be configured for leakage detection OUDP test burst granting into the test region. This profile must cover the minislots which fully span the leakage detection test region frequencies. The profile should also specify the densest Pilot Pattern available on the OFDMA channel for use in the test burst. As previously mentioned, Pilot Pattern 4 is recommended when the channel uses a 2K FFT and Pilot Pattern 11 is recommended when the channel uses a 4K FFT.

Since a DOCSIS 3.1 CM is only required to support two IUCs per OFDMA channel, it is recommended that the OUDP test burst configuration within the test region be assigned to IUC 13, which is generally the most robust profile. Another IUC can be defined to cover optimal data transmissions by the CM when the test region is not being used for leakage detection testing.

3.3. Configuration of Leakage Detection Testing Sessions

[3] specifies the configuration necessary for supporting a variety of leakage detection testing use cases. It defines a generalized information model that can be used for CLI-based or API-based implementations. The information model can be supported with CableLabs® standardized YANG code. A protocol for the API has not yet been specified by CableLabs®. Obvious alternatives are available such as NETCONF, which is a standard OSSI protocol, and gRPC/gNMI, which is becoming commonly used in Streaming Telemetry applications and in several FMA interfaces.

The configurable parameters provided by the information model are:

- Test region start and end frequencies
- Test scope:
    - Single CM MAC address or
    - Ordered list of CM MAC addresses in a single upstream scheduling domain or
    - Interface name of upstream scheduling domain (OFDMA channel ID, MAC Domain) or
    - Specific R-PHY or R-MACPHY node or
    - Full CMTS/CCAP/MAC Manager scope
- Scheduler control parameters:
    - Duration of test burst per CM, in number of frames
    - Duration of gap between CMs in a list, in number of frames
    - Duration of gap between cycles through a list of CMs, in number of frames
    - Scheduled start time of test (i.e., time of day; not precision timing)
    - Scheduled stop time of test (i.e., time of day; not precision timing)
    - Explicit immediate test enable/disable

Leakage detection tests are configured as independently controllable test sessions that focus on lists of CMs within a single upstream scheduling domain. Each session is identified by a session ID. When a leakage detection test session is created for CMTS, CCAP or MAC Manager scope, a master session ID is created as a control envelope for the independently controllable constituent test sessions in scope, each of which receives its own session ID.

3.4. Execution of Leakage Detection Tests

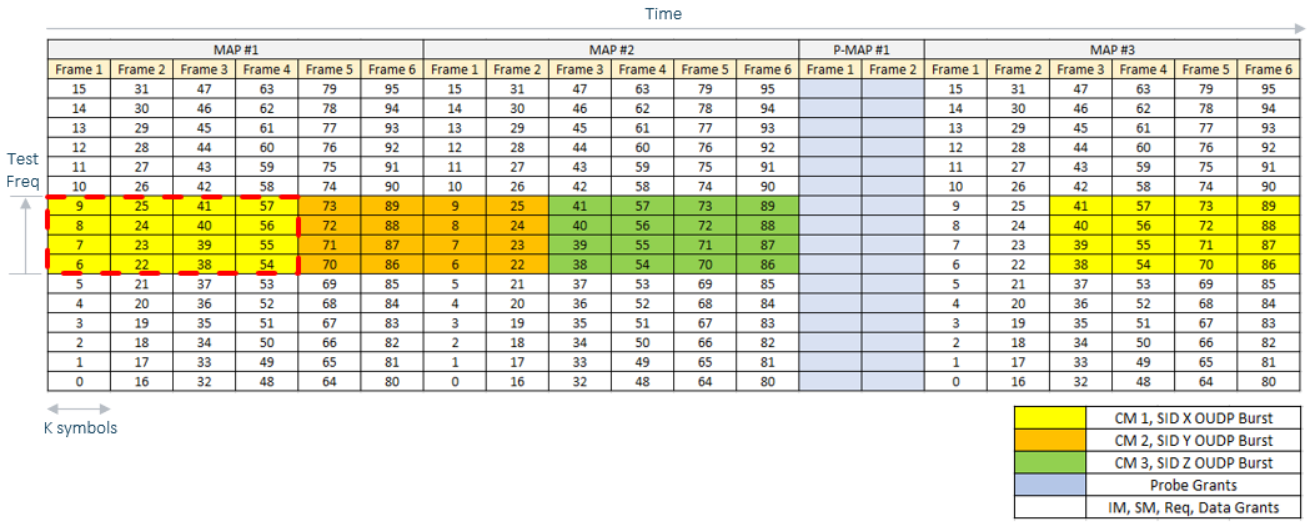
During a leakage detection test session, the CMTS upstream scheduler cycles through each list of CMs repeatedly and independently. Note that the single CM use case is a degenerate case of a list of CMs with only one CM in it. The rules followed by the CMTS upstream scheduler per configured test session are as follows:

- Grant OUDP test burst opportunities to each CM in the list over the entire test region spectrum and for the number of frames configured by burst duration
- Do not grant to multiple CMs in a given list in the same OFDMA frame
- Provide a gap between CM grants if one has been configured (intra-CM gap)
- Provide a gap between cycles through the list if one has been configured (inter-cycle gap)
- When a cycle through a list completes, including the intra-CM and inter-cycle gaps, start a new cycle through the list and continue this process until the test completes.

Intra-CM and inter-cycle gaps provide flexibility to support different scenarios. In one case it may be that the amount of time required to get through a full cycle is quite low and available time remains to transmit data. Gaps can provide the opportunity to do so. In addition, Probes need to be transmitted periodically by CMs as part of regular maintenance activities. Probes cover the entire spectrum of an OFDMA channel and cannot be scheduled if there is constant transmission of OUDP test bursts in the channel. As another example there several Proactive Network Maintenance (PNM) capabilities such as Active and Quiet Probe and Upstream Triggered Spectrum Capture where the entire OFDMA channel may be needed for other purposes for a specific period. Gaps in the leakage detection burst grant cycle can be leveraged for these activities as well.

Figure 7 illustrates a theoretical example of a 16 minislot OFDMA channel with a four minislot leakage test region. In practice the OFDMA channel will be much wider and there would be more data minislots relative to the test region. The test region in this example is being scheduled for four frames each to three CMs in the upstream scheduling domain. There is no intra-CM gap and there is a four frame inter-cycle

gap where the inter-cycle gap is being used for Probe transmissions and for data transmission. The CMTS scheduler inserted a P-MAP for the Probes and left the first two frames of MAP #3 for data transmission.



MAP #1						MAP #2						P-MAP #1		MAP #3					
Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6	Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6	Frame 1	Frame 2	Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6
15	31	47	63	79	95	15	31	47	63	79	95			15	31	47	63	79	95
14	30	46	62	78	94	14	30	46	62	78	94			14	30	46	62	78	94
13	29	45	61	77	93	13	29	45	61	77	93			13	29	45	61	77	93
12	28	44	60	76	92	12	28	44	60	76	92			12	28	44	60	76	92
11	27	43	59	75	91	11	27	43	59	75	91			11	27	43	59	75	91
10	26	42	58	74	90	10	26	42	58	74	90			10	26	42	58	74	90
9	25	41	57	73	89	9	25	41	57	73	89			9	25	41	57	73	89
8	24	40	56	72	88	8	24	40	56	72	88			8	24	40	56	72	88
7	23	39	55	71	87	7	23	39	55	71	87			7	23	39	55	71	87
6	22	38	54	70	86	6	22	38	54	70	86			6	22	38	54	70	86
5	21	37	53	69	85	5	21	37	53	69	85			5	21	37	53	69	85
4	20	36	52	68	84	4	20	36	52	68	84			4	20	36	52	68	84
3	19	35	51	67	83	3	19	35	51	67	83			3	19	35	51	67	83
2	18	34	50	66	82	2	18	34	50	66	82			2	18	34	50	66	82
1	17	33	49	65	81	1	17	33	49	65	81			1	17	33	49	65	81
0	16	32	48	64	80	0	16	32	48	64	80			0	16	32	48	64	80

Figure 7 – Example Scheduler View with Leakage Detection Test Region

It should be noted that there are some differences in upstream scheduler processing rules for CLI/API-specified versus CMTS-determined lists of CMs. In CLI/API-specified list cases the cycle is not disrupted when CMs are not available to transmit (e.g., are offline). The unavailable CM is simply skipped in the cycle and its position is not granted at all. This preserves the timing of the cycle, which can be used algorithmically to determine when leakage is detected what CM was transmitting.

For CMTS-determined lists, the order of CMs is preserved but empty spots do not need to be maintained when a CM becomes unavailable. Conversely, if a CM becomes available during a test it can be inserted into the list by the CMTS. It is not expected that automatic determination of which CM transmission led to leakage is possible in this case. It is more likely that a manual detector strength-of-signal and proximity approach is used, or that the GPS location of the field detector at the time leakage is detected will be noted, and that further isolation will be required.

3.5. Leakage Detection Testing Metrics

In a perfect world there will be no leakage detected. Without test metrics, there is no way for the operator to be certain that grants were made to CMs and that granted CMs responded with test burst transmissions. For this purpose, the CableLabs specifications include metrics to count the number of OUDP test bursts granted per CM for leakage detection as well as the number of bytes received at the CMTS from the CM for these grants. There is also a “no burst received metric” available as an alternative to the bytes received metric. These metrics are shown in Table 1.

In combination with existing leakage detection verification mechanisms used for downstream leakage detection (e.g., service vehicle mileage and plant coverage records), these metrics can and will be used by operators in their reports to authorities, such as the FCC and CRTC, to prove the effectiveness and accuracy of their leakage testing operations.

Table 1 – Leakage Detection Test Metrics

Attribute Name	Units	Description
NumBurstsGranted	Grants	Count of grants made to a CM's OUDP Test SID during a leakage detection test session
NumBurstsNotReceived	Bursts	Count of bursts not received for bursts that were granted during a leakage detection test session
NumTestBytesReceived	Bytes	Count of bytes received for grants made to CM's OUDP Test SID during a leakage detection test session

3.6. Feature Interactions with Leakage Detection Testing

There are several feature interactions to consider with leakage detection testing. As previously mentioned, Probes need to be scheduled periodically for ranging-related functions such as determining CM transmit pre-equalizer coefficients, and for taking RxMER measurements. Probes consume minislots over the entire OFDMA channel and therefore cannot be granted when OUDP test bursts are taking place in the leakage detection test region of the channel at the same time. For this the CMTS implementation must either interrupt the leakage detection test session or make use of intra-CM and/or inter-cycle gaps. The decision is implementation specific. However, interrupting sessions where a CLI/API specifies the list of CMs to test can disrupt automatic determination of CMs causing leakage and is therefore not the best choice in such a scenario.

Similarly, the Active and Quiet Probe Proactive Network Maintenance (PNM) test relies on Probe grants to take measurements of underlying noise in an OFDMA channel at the CMTS. The same considerations as mentioned above apply to this PNM test.

OUDP test bursts were originally created to gather information on FEC performance or count CRC errors for a particular modulation profile. OUDP test bursts used for this purpose are generally to assign profiles to a CM for data transmission, which is typically done early in the lifetime of a CM's registration with the CMTS. In the same fashion as the Probe discussion, these bursts can either interrupt existing leakage detection tests or can be made to fit within intra-CM and/or inter-cycle gaps.

The DOCSIS specifications support a battery back-up mode for appropriately equipped CMs to transition into when they lose AC power. In battery back-up mode, CM functionality is taken down to a minimal viable subset of its full functionality to preserve battery life but at the same time keep the CM operational. It is implementation specific as to what the CMTS does in this case. On one hand, exempting the CM from leakage detection tests can save battery. On the other hand, if the CM transmissions are causing leakage it would be good to learn that.

Finally, DOCSIS Light Sleep and Energy Management Modes are supported in the specifications, even though not largely deployed. Having CMs that are otherwise idle transmitting OUDP test bursts as part of leakage detection testing clearly interacts with the CM's ability to sleep and conserve power. Once again, it is implementation specific how these features interact.

NOTE: Standards work on High-Split and Ultra-High Split leakage detection will continue as operators determine their strategies for deployment and field operations in the presence of these new capabilities.

4. Testing the Leakage Detection Solution

Significant progress has been made in demonstrating the viability of the OUDP test burst approach for leakage detection. Recently, successful radiated leakage detection was accomplished from an OUDP burst signal transmitted by a DOCSIS 3.1 CM. The test was significant and a milestone because the CM which generated the OUDP test burst was directed to do so by an R-MACPHY node using standard DOCSIS MAC Management signaling. This test was a natural progression from previous testing of the OUDP approach which was accomplished using test equipment which generated the OUDP test burst, or by using a diagnostic CMs instructed to burst directly via serial control.

4.1. Test Setup

The test setup is as shown in Figure 8. Using the R-MACPHY node's MAC Manager CLI, the OFDMA channel and test session configurations were input into the R-MACPHY node. The node then instructed the CM via DOCSIS MAC Management messages to generate OUDP test bursts according to the specified test session configuration. The coax output of the CM was connected to a reversed two-way splitter with one leg connected to node via the access network and the other leg connected to a transmit antenna to create a leaked RF signal over the air. The RF signal leakage was received by the field detector antenna – which measured the OUDP leakage signal level. The detector was configured and tested first in a lab and then vehicle mounted for drive-out testing.

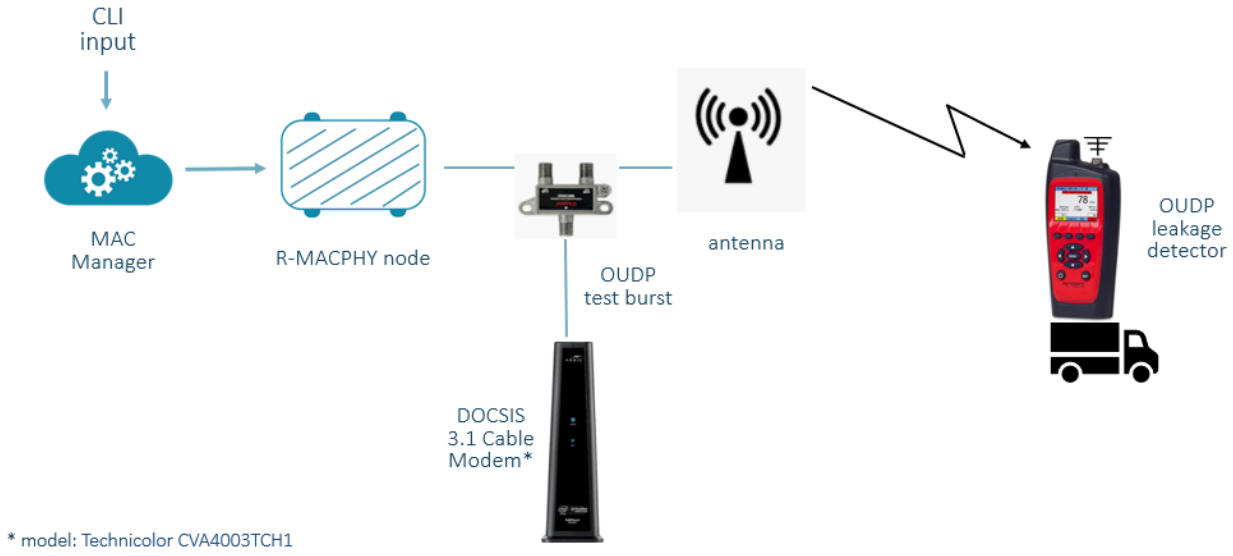


Figure 8 – Test Setup

4.2. Test Environment Configurations

In practice OFDMA channel and leakage detection test region configurations are an operational decision made by individual cable operators. We expect that common practice will be to define a narrow band test region of approximately 1.6 MHz located just above the aeronautical band in the OFDMA channel. Burst descriptors must be defined such that the desired burst characteristics will be used by the CMs when transmitting OUDP test bursts in the test region. These burst characteristics impact field detector sensitivity so must be chosen carefully. Field detectors must also be configured to match the test region frequencies and the expected OUDP burst characteristics used by CMs in the test region.

The OFDMA channel parameters of interest when configuring the OFDMA channel where the leakage detection test region resides are illustrated in Figure 9. Of these, the key variables are the size of the test region and Pilot Pattern configured in the burst descriptors defining the region, and number of symbols per frame and Cyclic Prefix of the OFDMA channel. It is desirable to use the least amount of spectrum needed for the test region. Our recommendation is to keep the test region to four minislots, or 1.6 MHz. The recommended Pilot Patterns for optimal field detector sensitivity are Pilot Pattern 4 for the 2K FFT and Pilot Pattern 11 for the 4K FFT. The number of symbols per frame needs to be balanced between optimal data throughput on the channel and field meter sensitivity. Fewer symbols per frame is better for field meter sensitivity. More symbols per frame may be better for data throughput. Cyclic Prefix length impacts the symbol duration and CMTS burst receiver accuracy. We show the values which are optimal for field detector sensitivity.

Parameter	2K FFT	4K FFT	Comments
Channel Start Frequency	108.50	108.50	Frequency of first active subcarrier in OFDMA channel (subcarrier# 74 | 148)
Channel Width	95	95	Range of active subcarriers within OFDMA channel (max 95 MHz)
Subcarrier Spacing (kHz)	50	25	
Symbols per Frame (K)	6	6	Range: 6-18 (2K FFT), 6-9 (4K FFT)
Cyclic Prefix	512	512	Value: 96, 128, 160, 192, 224, 256, 288, 320, 384, 512, 640
Pilot Pattern in Test Region	4	11	Configure in the Burst Descriptor for Test Region minislots (use IUC 13)
Test Region Start Frequency	138.10	138.10	Preferred frequency of first subcarrier of Test Region
Test Region Stop Frequency	139.70	139.70	Stop - Start should = multiple of 400 kHz (recommend 1.6 MHz, or 4 minislots)

Figure 9 – OFDMA Channel Configuration Key Parameters

Given an OFDMA channel with a configured leakage detection test region, the next step is to define the desired leakage detection test session configuration. When performing leakage detection testing on groups of CMs in a scheduling domain it is necessary to provide adequate burst durations per CM to achieve the desired field detector accuracy. It is also desired to use the minimum burst duration to enable a shorter overall time to cycle through the group of CMs. This is done to ensure that leaks can be detected by field detectors at vehicle speeds. The configurations shown in Figure 10 were selected such that sensitivity of the detection is sufficient to meet FCC signal leakage requirements when the CM is granted a minimal burst duration such that each CM within an Upstream Service Group can burst at least once every half second, as required for a robust confirmation of detection.

The CableLabs specification also specifies configuration for a gap between CM transmissions and between cycles through groups of CMs. These gaps are left to allow for data transmission when the number of CMs is low, and to allow for maintenance activities such as Probe transmissions.

Parameter	2K FFT	4K FFT	Comments
Burst Duration per CM (Frames)	8	4	
Gap Between Cable Modems (Frames)	0	0	
Gap Between Cycles Through List (Frames)	8	8	Leave for Probe processing; each CM gets a full symbol in Probe frames

Figure 10 – Leakage Detection Session Parameters

The OFDMA channel configuration and leakage detection test session configuration needs to be aligned with field detector configuration. Field detector configuration is vendor specific. Details of that are not provided here.

For the leakage detection test described in this paper we tested with a single CM which was granted a continuous stream of OUDP test burst opportunities by an R-MACPHY node. The OFDMA channel was from 108 MHz to 204 MHz with active subcarriers between 108.50 MHz and 203.50 MHz and 50 kHz

subcarrier spacing. The channel was configured with 18 symbols per frame, which was designed to test the worst sensitivity at the field meter. Cyclic Prefix was 512.

The test region was defined with burst descriptors to cover 135.025 MHz to 136.825 MHz. Pilot Pattern 4 was configured.

The spectrum of the OUDP test burst generated by the CM under test is shown in Figure 11.

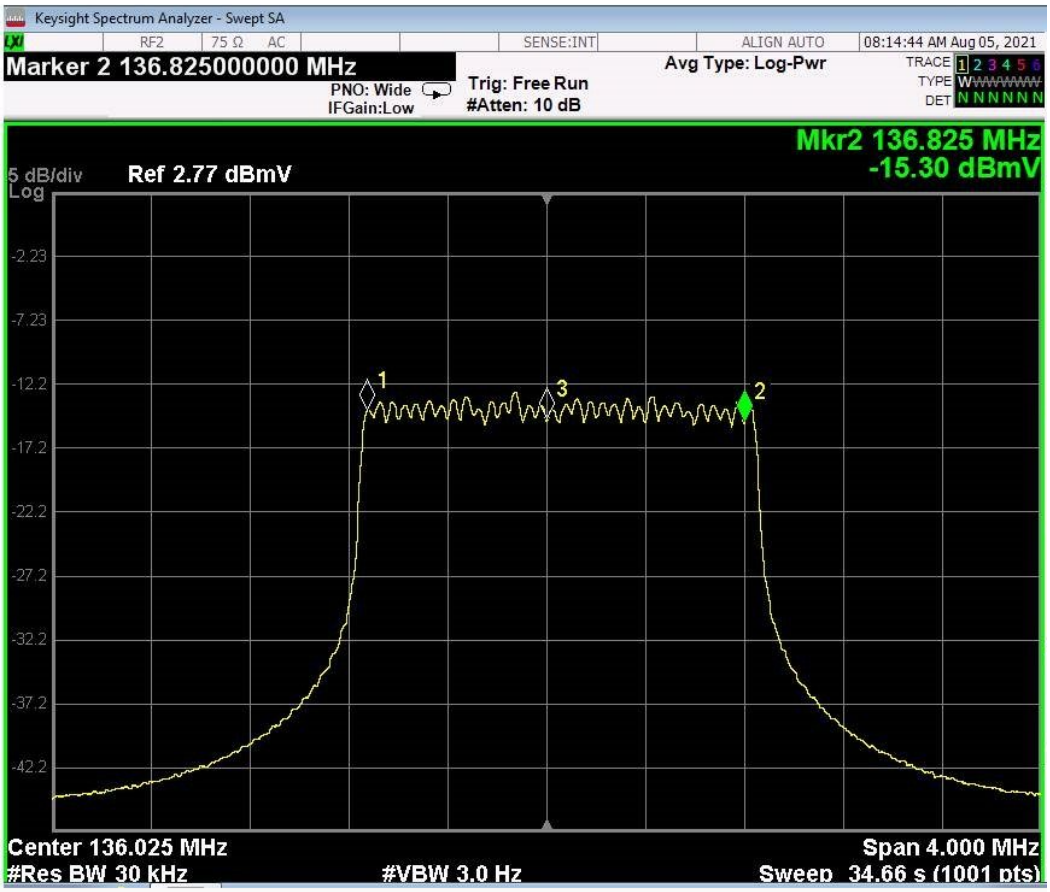


Figure 11 – Leakage Detection OUDP Test Burst Spectrum

After initial hard wired conducted tests were performed in the lab to confirm configuration alignment between the OFDMA channel, test region and field detector, a simple drive-out test was performed. A leak was generated outside of the building using an antenna tuned to the 136 MHz detection frequency. The antenna was connected to the CM output via a reversed two-way splitter as shown in the test setup. A vehicle was outfitted with a corresponding 136 MHz receive antenna, a GPS antenna, and an OUDP leakage detector. No leakage signal was detected until the OUDP burst command was entered into the R-MACPHY node's MAC Manager CLI. Once the CM began to burst, signal leakage was continuously detected, as illustrated in Figure 12.

Figure 12 – Leakage Detection in a Stationary Vehicle

The vehicle was then used to perform a short drive-out in the parking lot around the building, which is indicated in the bread-crumb trail on Figure 13. Each red dot shows a detection point in the one second leakage detection measurement interval. The test session was quite successful, and there was no trouble in detecting the OUDP test bursts. Since the transmitted signal level was a relatively large, leakage was recorded along the entire drive route. Detected signal leakage levels on the far side of the building were as expected, less than as compared to detection points in closer proximity to the leakage source.

Leak ID	Leak level, [µV/m]	Address	PSID	HUB	Node / device	Details	Manage	Ordered date
1	1000/-/-	185 Ainsley Drive, Syracuse, NY 13210	00000	OUDP Test Hub No.1	Not available	Info	Manage	
2	794/-/22	unnamed road, Duluth, GA 30097	00000	OUDP Test Hub No.2	Not available	Info	Manage	

Figure 13 – Leakage Detection in a Moving Vehicle

Future testing will be expanded to include multiple CMs in a group, along with the capabilities and use cases as described in this document. This type of testing will be enabled as vendors implement the functionality described in [3], and as operator plans for operationalizing leakage detection in their High-Split deployments become clear.

5. Conclusion

The OUDP test burst approach for detecting CM initiated signal leakage in DOCSIS 3.1 High-Split and DOCSIS 4.0 Ultra-High Split continues to look quite promising and viable. Excellent progress has been made with support of the OUDP test burst approach being specified in the [3] standard, which provides great flexibility to the operator community by enabling a variety of use cases such that the signal leakage process can be aligned with individual operator goals.

Significant progress has also been made with recent successful tests which prove that DOCSIS 3.1 CMs under the control of an R-MACPHY node can be instructed to and can in fact generate OUDP test bursts, and that the corresponding OUDP test burst (leakage signal) is able to be detected by an OUDP leakage detector installed in a vehicle while performing a drive-out.

As these advances morph into viable operations strategies the primary blocking factor for High-Split deployments in North America will be removed and operators will have at their fingertips the ability to offer 1 Gbps upstream services.

Abbreviations

AC	Alternating current
API	Application Programming Interface
CCAP	Converged Cable Access Platform
CLI	command line interface
CM	Cable Modem
CMTS	Cable Modem Termination System
CRC	cyclic redundancy check
CRTC	Canadian Radio-television and Telecommunications Commission
DOCSIS	data over cable service interface specification
FCC	Federal Communications Commission
FEC	forward error correction
FFT	Fast Fourier Transform
FMA	Flexible MAC Architecture
Gbps	Gigabits per second
gNMI	gRPC Network Management Protocol
gRPC	gRPC Remote Procedure Call
GPS	Global Positioning System
IUC	Interval Usage Code
kHz	kilohertz
MHz	Megahertz
MAC	Media Access Control
NETCONF	Network Configuration Protocol
OFDMA	orthogonal frequency-division multiple access
OUDP	OFDMA Upstream Data Profile
PHY	Physical layer
PNM	Proactive Network Maintenance
P-MAP	Probe MAP
RF	Radio Frequency
RxMER	Receive Modulation Error Ratio
SCTE	Society of Cable Telecommunications Engineers
SID	service identifier
YANG	Yet Another Next Generation

Bibliography & References

[1] DOCSIS® 3.1 Physical Layer Specification, CM-SP-PHYv3.1-I18-210125, January 25, 2021, Cable Television Laboratories, Inc.

[2] DOCSIS® 3.1 MAC and Upper Layer Protocols Interface Specification, CM-SP-PHYv3.1-I21-201020, October 20, 2020, Cable Television Laboratories, Inc.

[3] DOCSIS® 3.1 CCAP Operations Support System Interface Specification, CM-SP-CCAP-OSSIv3.1-I21-210716, July 16, 2021, Cable Television Laboratories, Inc.

[4] *Leakage in a High Split World – Detecting and Measuring Upstream Leakage Levels in a One Gbps Symmetrical High Split Hybrid Fiber Coax Network;* John Chrostowski, Greg Tresness, Dan Rice, Benny Lewandowski, SCTE EXPO '20

Lessons from Operating Tens of Thousands of Remote PHY Devices

A Technical Paper prepared for SCTE by

Jorge Salinger
VP, Access Architecture
Comcast Cable Communications
1701 JFK Blvd – Philadelphia, PA 19103
+1 (215) 439-1721
jorge_salinger@cable.comcast.com

Steve Sigman
Director, Access Architecture
Comcast Cable Communications
1701 JFK Blvd – Philadelphia, PA 19103
+1 (609) 685-3480
steve_sigman@cable.comcast.com

1. Introduction

Cable operators have been actively converging video and data services into a common converged cable access platform (CCAP) for many years now. This trend, which requires an evolution towards newer, more modern, and denser equipment, was intended to free up space in the headend.

However, as the success of high-speed data and on-demand services continues, the evolution of the access network progresses towards expanded capacity and/or ever-smaller service groups. For the former, the spectrum allocated to narrowcast services increases, driving operators to deploy Data-Over-Cable Service Interface Specifications (DOCSIS©) services including more SC-QAM (single carrier quadrature amplitude modulation) channels, more capacity for more or wider OFDM (orthogonal frequency division multiplexing) channels, as well as more narrowcast video services. For the latter, more CCAP ports are needed, which drives the deployment of more line cards and eventually more chassis. These expansion trends result in a continuous growth of headend equipment, which is already starting to exceed the capacity that headend facilities can support.

The above trends are now intractably linked to additional evolutions: distribution of components of the access network, implementing a distributed access architecture (DAA), and virtualization of the core network functions. By now, deployment of DAA devices, such as remote PHY (R-PHY) nodes, is accelerating to tens of thousands of remote PHY devices (RPDs), hosting millions of cable modems.

As a result of the experience, an agile approach to network and device management is now required to ensure that we can effectively monitor our networks, these distributed devices, and customer service status. Traditional proactive network maintenance (PNM) and operations support system (OSS) toolsets provide views into current system health, but unexpected behavior on the RPDs and the systems that work with them provide clear opportunities for new approaches to network and system monitoring and management.

This paper explores lessons learned the hard way, gaps discovered and filled, as well as processes developed to improve issue detection and lower mean time to repair (MTTR). The paper focuses upon factors to keep both maintenance technicians and customers happy, including:

- Whole system monitoring: Real-time telemetry, visual dashboarding tools, and error condition detection and alarming for not just RPDs, but the entire architecture, including R-PHY cores, engines, switches, and timing servers;
- Outside plant considerations, including plant powering;
- Provisioning management, including addressing challenges with having all required systems configured at time of node cutover;
- Hardware and software management, including how lab testing can lead to a better customer experience.

This paper begins by outlining the evolution of service provider networks, and then describes why and how the migration to a distributed access architecture is necessary and beneficial. The paper then expands into features that can be implemented with DAAs and discusses the topic of virtualization. Finally, the paper explores how the implementation of DOCSIS 4.0 could be implemented in DAA networks.

2. Cable Network Evolution

2.1. Typical HFC Networks Today

Most hybrid fiber/coax (HFC) networks have been designed with an upper frequency boundary of 750 MHz or 860 MHz, while some are designed to support 1 GHz and other newer networks are designed to support 1.2 GHz. For the more abundant 750 MHz or 860 MHz networks, if not already fully utilized, it is expected that use of their capacity will be increased to the point of exhaustion. This will happen as a result of 1) increased DOCSIS usage for even faster high-speed data (HSD) service tiers; 2) additional high-definition (HD) programs (for broadcast [BC] and especially narrowcast [NC] services, such as video on demand [VOD] and switched digital video [SDV]); and 3) new service additions such as Internet Protocol (IP) video and cloud-based digital video recorder (cDVR.)

In recent years the growth in, and demand for, HD programming has resulted in the need for allocation of large numbers of CTA (Consumer Technology Association) channels[1] for HD services, both for BC and NC, which has filled every available portion of the spectrum. This is especially true for BC, where large numbers of programs are offered in HD format, while simultaneously the need for distributing the standard definition (SD) version has persisted. This has resulted in the need for use of 3x to 5x the number of CTA channels than previously required. For example, a typical digital multiplex including 10 to 15 programs would require an additional three to five CTA channels for the HD-equivalent streams, even assuming the newer, more sophisticated multiplexing schemes available in the market. Of course not every program is available, or still sought by subscribers, in HD format. But very large numbers of them are, including 100 to 150 BC programs.

The above is also applicable to a great extent in systems utilizing SDV technology for content distribution. The difference is that the HD and SD versions of the program are not distributed unless a subscriber requests them, which reduces the marginal increase in capacity. Assuming that all programs are distributed in only one format, which is certainly a valid expectation for programs of low viewership, then the increase in capacity for a conversion from SD to HD would just be the increase in capacity required for the transmission of the HD program without requiring the simultaneous use of bandwidth for both formats.

Additionally, considerable spectrum is needed to deploy high-capacity narrowcast legacy video services, especially cDVR, and thousands of HD video-on-demand titles. For the former, initial observations suggest that network requirements for cDVR may be as high as 4x to 5x that of VOD, and that peak utilization overlaps, at least partially, with that of peak use for other narrowcast services, such as HSD.

Finally, the growth in HSD services continues. All network operators have offered increased service tiers and observed an increased use of broadband capacity for well over a decade now, as shown in Figure 1, which amounts to a constant year-over-year compounded growth. The applications have changed throughout this time, and the demand has continued to increase at the same relentless rate.

[1] 6 MHz-wide channel allocations used in North American and some other cable networks are defined in *CTA Standard "Cable Television Channel Identification Plan" CTA-542-D R-2018 (Formerly CTA-542-D), updated February 2019*.

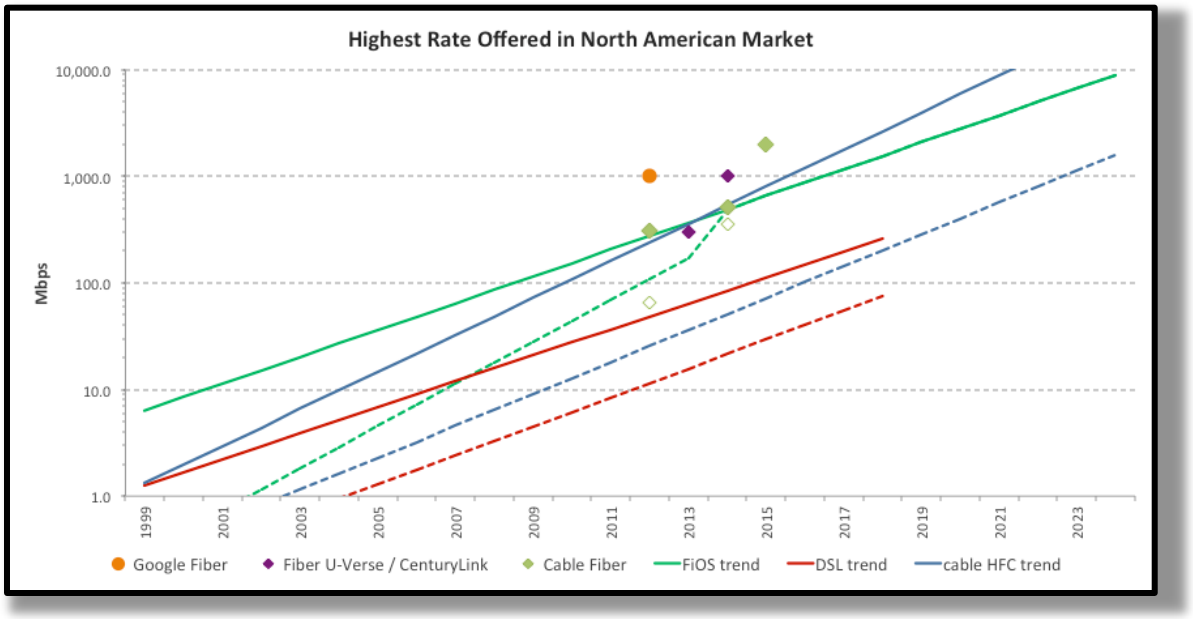


Figure 1 - Examples of HSD service tier capacity increase over time

How does this compare to other operators' data services and a longer period? Projecting an operator's HSD service growth back in time to when Internet services started 25+ years ago, services should have been about 100 bps. This coincides with the history of telephone modems from 110 and 300 baud modems from the mid-1980s, to 56 kbps/V.42, into ISDN (integrated services digital network) services.

This demonstrates that the growth seen in cable industry HSD services is typically over a much longer period of time, rather than an exception observed by operators in recent years.

2.2. Growth Projections

From all of the above, it follows that, should the usage growth pattern continue at the same rate as in the past, networks will be required to provide >1 Gbps HSD services within the next few years. This growth, coupled with the surge in HD video formats (8K TV sets are increasingly available, even as compatible content catches up), and more personalized narrowcast services, will result in a significant growth in narrowcast capacity, as shown in Figure 2.

To support this growth, cable operators have deployed bandwidth reclamation tools such as SDV for digital broadcast, digital transport adapters (DTAs) for analog service reclamation and the shift to all-IP, increased spectrum (1 GHz and above), or a combination of all. These tools have been extremely valuable to operators, and their operational complexity and cost remains well justified.

In the case of SDV, early predictions from industry analysts projected that the efficiency of SDV would reach 40% (e.g., programs requiring 10 CTA channels could be carried in 6). This has proven to be understated, since it was based on the use of SDV to reduce bandwidth required for existing services. As SDV's role in the network grew, the efficiencies have been even greater, especially as SDV expanded in scope to support niche service introductions with low initial viewership that would otherwise be difficult to deploy.

The benefit of DTAs has been just as, or perhaps even more striking. Cable operators deploying DTA devices are able to eliminate the need to distribute the analog channels in the network. Even if DTAs are distributed to top tier analog customers, such as only the traditional "expanded basic" subscribers, the move would reduce a channel line up from perhaps 50 6 MHz channels, or 300 MHz of capacity, dedicated to 50 analog programs to perhaps as little as four such 6 MHz channels (24 MHz) dedicated to transport the 50 programs in their digital-equivalent transport. Using the same comparison method as the above SDV case, this is a >90% efficiency. If extended to the entire analog tier, the efficiency gains are very significant.

Despite the availability of these tools, they are not universally applicable. With respect to SDV, in general it is not likely that all broadcast programs will be switched since experience shows that many broadcast programs are constantly viewed by someone in the service group during peak hours, which will leave a large portion of the spectrum still used for broadcast. Similarly, not all analog channels can be removed in the short term due to operational and/or cost constraints.

Additionally, while many cable operators will use some or all of the capacity tools, in general they won't be used by every operator for all applications because of a range of variables that are out of scope for this discussion.

Finally, there are also significant potential gains to be achieved from the use of advanced video compression standards (VCS), and variable bit rate (VBR) multiplexing. In the case of VCS, coding efficiencies of approximately 50%, depending on implementation and content type, can be obtained with H.264/MPEG-4 Part 10. Furthermore, with the release of the H.265 standard in April of 2013, it is possible to achieve a 50% improvement over H.264. The use of VBR promises to result in a capacity efficiency gain of as much as 70% versus constant bit rate (CBR) . The combined gains from using the above approaches for multiple services are even more significant.

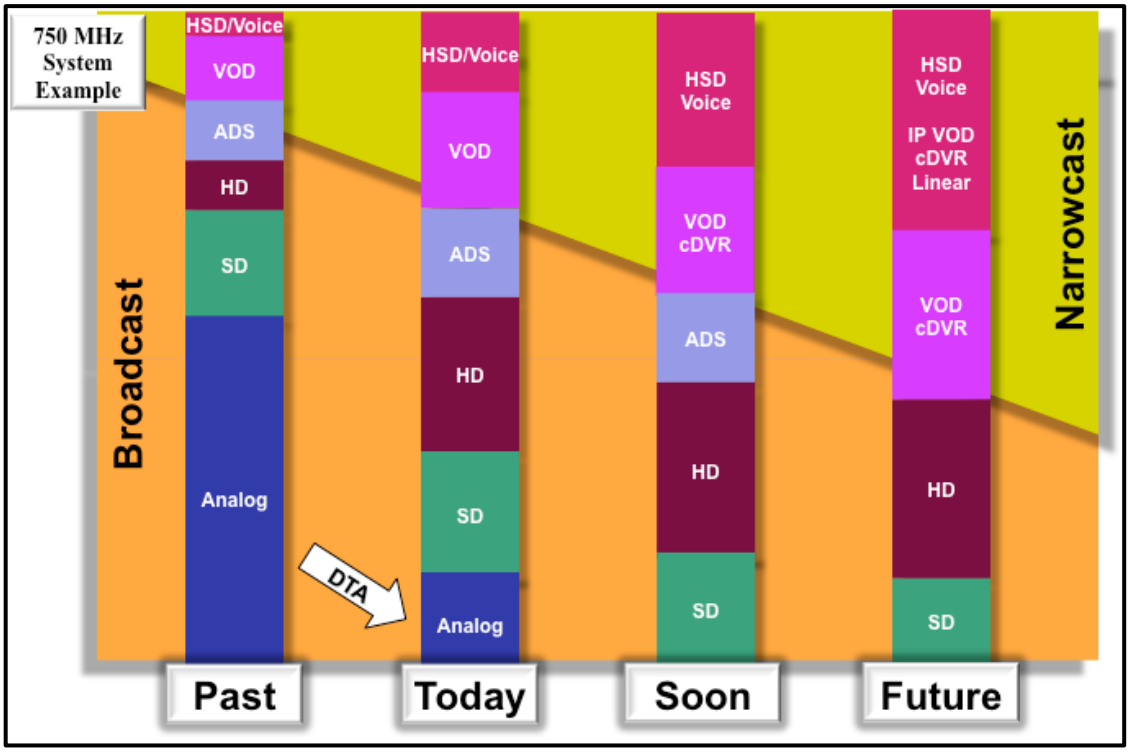


Figure 2 - Example of narrowcast service growth over time

However, these are difficult tools to take advantage of from a network perspective, because on a proportional basis, relatively few legacy set-tops will support all these technical advances, especially H.265. These tools are more likely to find significant support in equipment designed to handle newer, IP-video based services.

And, this approach will nonetheless require additional capacity from the network. This is especially true when considering that the deployment of these advanced video services will result in yet another simulcast of video programs, at least initially. This is because, realistically, advanced services will not, at least initially, replace the currently deployed service formats.

Furthermore, ubiquitous support for such devices would require considerable spectrum if the legacy services are maintained for an extended period – which is expected, given that legacy devices will continue to be deployed, for some amount of time. Moreover, this increase in simultaneous use of the more advanced IP video services while maintaining legacy services will be especially impactful over time, as the number of IP video services increases.

All of the above, coupled with the success experienced by operators in recent years with business services, homes security, etc., will likely require the deployment of IP capacity beyond what can be supported today. As well, it will require the development of tools for increased spectral efficiency and/or unleashing additional spectrum in the HFC network. The following sections of this paper enumerate some ways in which this can be achieved.

2.3. The Advent of DOCSIS 3.1

As it has been pretty well covered in the trade media, DOCSIS 3.1 deployment has been quite extensive. Most cable operators have deployed DOCSIS 3.1 across their markets, and several have even deployed DOCSIS 3.1 throughout their entire footprint.

The key motivation for the 3.1 version of DOCSIS technology is, in a nutshell, to scale DOCSIS more efficiently, both from cost and operations perspectives.

For the first 10 years of DOCSIS deployments, it was possible to offer Internet services and support its growth with just one downstream 6 MHz DOCSIS channel. Over the last five to 10 years, service speed demands increased to speeds of hundreds of Mbps, which require bonding of many 6 MHz channels. Therefore, the industry deployed multiple DOCSIS channels using DOCSIS 3.0, sometimes using 32 or more channels, and even requiring capacity beyond that supported by DOCSIS 3.0.

To that end, the three key goals and features of DOCSIS 3.1 were:

1. A much more efficient use of spectrum, with up to a 50% improvement in bandwidth efficiency (or bps/Hz). This is a result of more efficient forward error correction (i.e., replacing the older and less efficient Reed-Solomon approach with the far more efficient low density parity check [LDPC]), and the addition of the higher-order modulations1024- and 4096-QAM downstream, and 256- and 1024-QAM in the upstream.)

   These new modulation schemes provide two and four bits per second per hertz of improvement in both the upstream and downstream signal directions, while the use of the new forward error correction approach provides approximately 5 dB better RF performance. The end result is that networks are able to transport 1 Gbps of DOCSIS capacity in about 120 MHz of spectrum. For context, doing the same with DOCSIS 3.0, using SC-QAM requires about 180 MHz of spectrum.

2. Cost reduction, mainly by leveraging technologies commonly used in other transmission media. Specifically, the inclusion of OFDM, which is used extensively in wireless and wireline transmission media. The addition of OFDM for the downstream and OFDMA (orthogonal frequency division multiple access) for the upstream should enable operators to reduce costs by "packing" more bits in the HFC network more efficiently. As a result, these technologies will likely attract a larger supplier ecosystem, increasing innovation and fueling competition.

3. To enable a simple and orderly transition strategy. This applies doubly, in terms of compatibility with the previous generation of cable modem termination system (CMTS) and cable modem (CM) equipment, and simultaneously supporting an expanded spectrum capacity in the HFC network.

Specifically, DOCSIS 3.1 cable modems operate with DOCSIS 2.0 and 3.0 CMTS/CCAP equipment, enabling deployment of DOCSIS 3.1 CPE (customer premises equipment) as soon as it became available. Similarly, DOCSIS 3.1 CCAPs support DOCSIS 2.0 and 3.0 CPE, allowing operators to upgrade headend equipment without having to change any of the existing CPE. And, DOCSIS 3.1-based CM and CMTS equipment both support the currently required upstream and downstream spectrum, with expandability of the upstream to 85 MHz and beyond, and of the downstream up to 1.2 GHz.

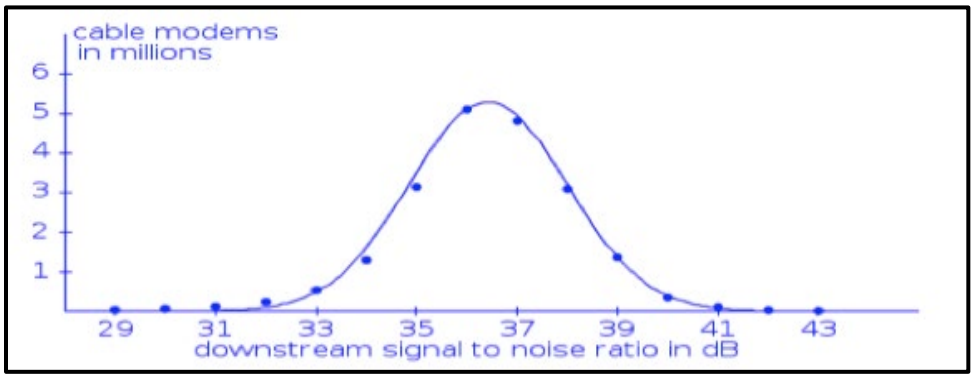


Figure 3 - Example of downstream SNR for a large population of cable modems

Figure 3 depicts the downstream signal-to-noise ratio (SNR) as reported by a very large population of cable modems. This data shows that many cable modems will be able to support the high-order modulation profiles included in DOCSIS 3.1.

Table 1 - SNR required for DOCSIS 3.1

Modulation	Signal-to-Noise Ratio
512-QAM	27 dB
1024-QAM	30 dB
2048-QAM	33 dB
4096-QAM	36 dB
8196-QAM[2]	39 dB
16384-QAM	42 dB

[2] 8196-QAM and 16384-QAM are included for future consideration in the DOCSIS 3.1 specifications.

Assuming an 8/9 LDPC coding ratio, Table 1 shows the required SNR for the modulation rates included in DOCSIS 3.1.

Applying the SNR requirements from Table 1 to the population of modems shown in Figure 3, we can see that a large population of cable modems would not achieve sufficient SNR to operate at 4096-QAM. Furthermore, if sufficient headroom is allowed to account for environmental fluctuations, the population of cable modems that would not receive signals with sufficient SNR to operate at 4096-QAM would be significant.

2.4. The Analog Intensity Modulated Forward Link in HFC Networks

As the name indicates, HFC networks use a fiber transport between the headend and the coaxial plant. This fiber link, intended to reduce the size of amplifier cascades, which improves performance, was originally developed with analog intensity modulated[3] lasers and compatible receivers in both signal directions, upstream and downstream.

Over time, the performance of the upstream link was improved by replacing the analog modulation with a digital transport. This improved performance significantly, and allowed for longer distances between the headend and the node. Different vendors implemented their own methods and technical capabilities to implement a digital transport, which resulted in incompatible systems and required the use of the same vendor's components for both the node and the headend.

However, the downstream link remained almost unchanged over time, with the only enhancements focused on improving distance and RF spectrum capacity. Performance has not really been an issue like it was in the upstream.

But more importantly, while the digital capacity of the upstream was limited to a few megabits per second, well under a gigabit of digital capacity which could easily be digitized and carried with Ethernet optics, the downstream digital capacity needed to transport the downstream spectrum has been considerably higher, reaching and even exceeding 10 gigabits per second.

Because of the above, analog forward links continue to be used. Even though headend equipment is currently capable of launching signals with >47 dB modulation error ratio (MER) performance, which is sufficient to generate and transport 16,384-QAM signals, analog lasers are limited to about 35 to 38 dB of MER performance, which would limit end-of-line performance to barely enough for 2,048-QAM or 4,096-QAM in short cascades in the best of the cases. In short, QAM density from headend equipment development outpaced the capabilities of the analog lasers we use.

2.5. Description of Options for Digital Forward Link

As time has gone by, technology evolution and certain developments as described below have enabled options for implementing a digital forward link. These include:

1. Evolution of edge-QAM modulators, which have gone from single and/or a few modulators to supporting 32, 64 or even more modulators,

2. Development of the CCAP, combining the functions of the video QAM modulator and DOCSIS into a single platform, and

[3] The majority of optical links deployed in HFC networks use linear fiber optic signal transmission, based upon analog intensity modulation. Many in the field use the term "AM fiber link" or similar.

3. Migration to digital video, either partially or completely.

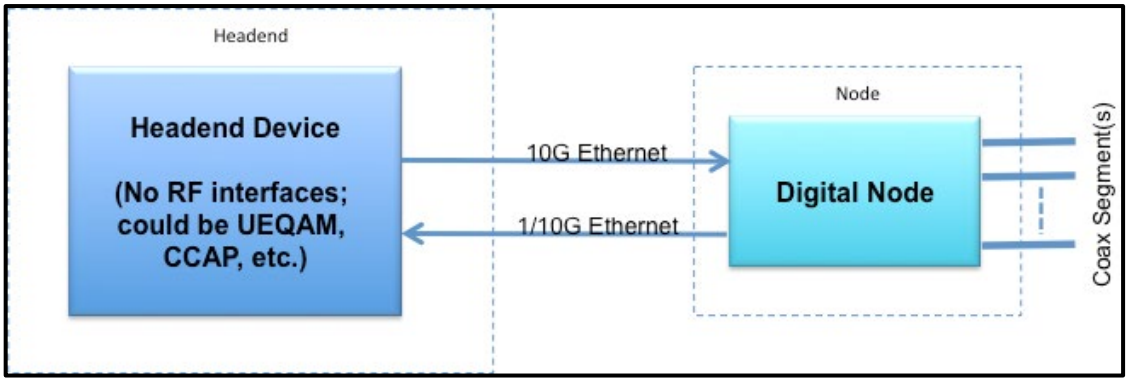


Figure 4 - Digital forward high-level architecture

With this technological evolution, it is conceivable to remove the RF combining network, and instead implement it digitally in the edge device, such as the CCAP. This evolution of the edge headend devices makes it possible to envision several options for digitizing the forward link.

Fundamentally, the migration to a digital forward includes the components included in Figure 4, and described as follows:

- The headend device, such as a CCAP, a high-density edge-QAM modulator supporting QAM for the entire spectrum,
- The node, which contains components normally implemented in the edge-QAM modulator or CCAP, to generate the RF signals,
- The link between the headend device and the node, comprising a digital interface, such as an Ethernet link.

There are the various approaches for how a digital forward link can be implemented to replace the currently used analog link. These approaches can be categorized into four groups, plus one option that would still leave RF generation at the headend device, as outlined in Table 2:

Table 2 - Categories of options for implementing a digital forward link

Option	Description and Approach
1. Maintain RF output in the headend	1.a Headend equipment remains unchanged 1.b Headend RF output is digitized, transported digitally, and RF is regenerated in the node
2. Remote the DAC from the PHY	2.a The DAC is removed from the headend 2.b Digital samples are transported digitally to the node where the DAC generates the RF signals
3. Partition the PHY and remote the lower portion of the PHY	3.a The PHY is split between the headend and the node 3.b The digital bit stream between upper and lower PHY is transported from headend to node
4. Remote the entire PHY	4.a The entire PHY modulation is moved to the node 4.b The MAC remains in the headend, and MAC frames are transmitted from the headend to modulator that resides in the node
5. Remote the entire PHY and MAC	5.a The entire PHY and MAC are removed from the headend device and placed in the node 5.b IP frames are transported from the headend to the node.

2.6. Comparison of Options for Digital Forward Link

There are pros and cons for each of the options. The following sections outline these tradeoffs.

2.6.1. Option 1: RF remains in the headend

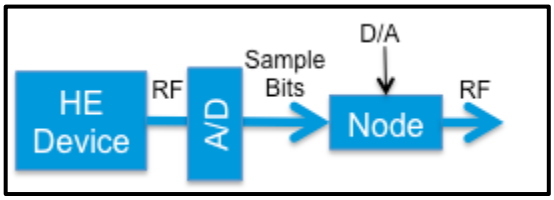


Figure 5 - Block diagram for Option 1

Equivalent to digital return, the RF output from the headend device is digitized, transported digitally, and converted back to RF in the node. This maintains HFC transparency. This option results in the highest bit rate over fiber; the capacity for multiple nodes would not fit into the available capacity of one 10 Gbps fiber. There is a loss of MER in the double conversion, so this option provides the least performance improvement. This option results in the least intelligence placed in the node, but an additional conversion stage is added in the headend.

2.6.2. Option 2: Digital-to-analog conversion is moved to the node

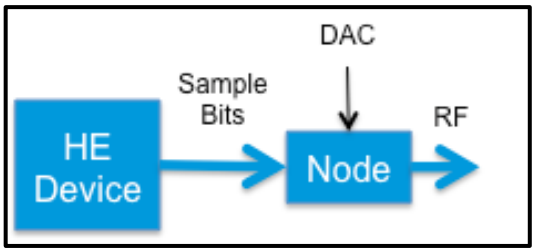


Figure 6 - Block diagram for Option 2

This option requires separation of the digital-to-analog conversion from the modulator. Together with Option 1, it results in the least intelligence in a node. It has a similarly high bit rate over fiber as Option 1; capacity for multiple nodes would not fit into the available capacity of one 10 Gbps fiber.

2.6.3. Option 3: Lower PHY is moved to the node

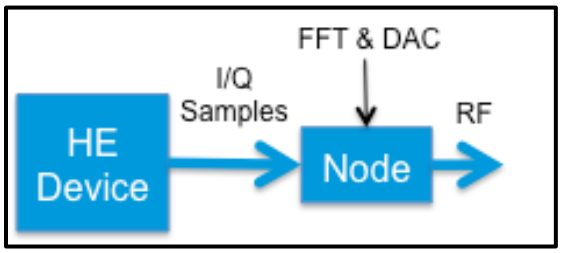


Figure 7 - Block diagram for Option 3

In this option, the PHY layer needs to be split into two components: upper and lower. This implements more node intelligence than in either of the previous options. Although it offers a lower bit rate over fiber than the previous options, it is still reasonably high. However, this option would require an industry proprietary point-to-point link between the headend port and the node, to transport the I and Q samples. Also, implementing this option would require the definition of interfaces which have never been defined in previous versions of the DOCSIS specifications, which in turn would result in modification of the silicon used and/or planned to date, and therefore results in the highest implementation complexity.

2.6.4. Option 4: Entire PHY is moved to the node

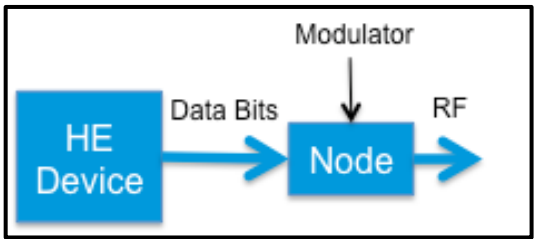


Figure 8 - Block diagram for Option 4

Moving the entire PHY layer into the node places more intelligence in it than with all previous options. This option results in the lowest bit rate over fiber; multiple nodes fit into the capacity of a 10 Gbps fiber. It enables a packet-based link between the headend and node, which results in significant benefits (which are outlined later in this paper.) It could use existing/planned silicon devices, and thus may be the easiest and quickest to implement. Lastly, this option offers the best MER performance improvement over analog.

2.6.5. Option 5: Move PHY and MAC to the node

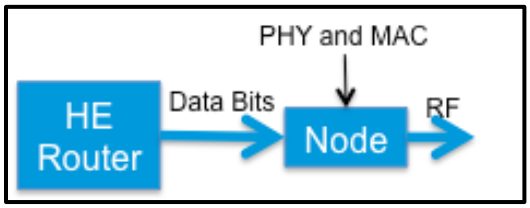


Figure 9 - Block diagram for Option 5

This option puts the most intelligence in the node. The data rate between the headend and the node is equivalent to the actual data transmitted, except for the addition of ancillary network data. It offers the same packet-based network benefits, and the same (highest) MER performance, as Option 4.

2.6.6. Comparison of Options and Implications

Any of the five options described above accomplish a migration away from the analog intensity modulated forward link and into a new era, where the link between the headend and the node becomes a digital link. And, while any of the above approaches accomplish a migration to a digital link, over the years, options 4 and 5 have received the most attention because of their relative implementation simplicity versus options 1, 2 and 3. We now call these options remote PHY and remote MAC-PHY, and we call the devices that implement them RPDs and RMDs (remote MAC-PHY devices), respectively.

As we migrated towards the implementation of a digital link, and separated either the physical layer in a remote PHY implementation, or also migrated the MAC in a remote MAC-PHY implementation, we stepped into the era of DAA. In these distributed access architectures, the remainder of the CCAP in the headend no longer needs to be implemented in an application-specific hardware design. Instead, the remainder of the CCAP in the headend can be implemented entirely as software running in general purpose compute platforms, which we now call a virtualized headend platform.

3. Benefits of Distributed Access Architectures

We will now focus on the benefits of a distributed access architecture, discuss some of the features of DAA, and outline network evolution strategies.

There are many benefits from the implementation of DAA. The following sections of this paper describe them.

3.1. Improved performance

Improvements on performance are achieved in multiple ways, including:

- Improved SNR (and MER) characteristics
- Longer link distances
- Higher reliability
- Better use of capacity

As described in the above sections of this paper, one key benefit of digital forward link is the improved performance resulting from the migration from an analog to a digital link. This gain varies depending on the characteristics of the analog link being replaced, but can be generalized as an improvement of 5 dB in SNR at the end of the line. This gain will result in higher capacity/Hz as it will be possible to run the higher order modulations as shown in Figure 3 and Table 1 for more of the cable modems in the network. This will enable significantly higher transport capacity for customers in the HFC network.

In addition, the digital forward link will enable longer distances between the headend and the node. This is because digital interfaces, such as an Ethernet link, are designed to operate over much longer distances while carrying the designated capacity. Extending the distance between the CCAP and the digital node would enable operators to move their CCAP devices back in the network to more centralized facilities, leaving the hub or OTN free of CCAP equipment. The benefit of such change could be very big for some operators, especially as segmentation of the network continues towards smaller service groups, for which additional CCAP equipment needs to be deployed.

A third benefit from the digital forward link is improved reliability of the optical link. It is well known that analog links require periodic maintenance and are subject to the effects of environmental changes. By contrast, Ethernet optical links are far more stable across a wider range of environmental conditions, and require little to no maintenance. The impact of this benefit could be very significant to operators.

Finally, the data transmitted through the link can be used more efficiently. One key example of such efficiency is the case where one link is used for multiple remote devices. As shown in Figure 10, one link from the headend CCAP device can be used to transport broadcast services once for multiple remote devices. This is achieved by using multicast addressing, whereby each of the remote devices uses the same lineup for each of the respective service groups. In doing so, a single link from the headend CCAP can be used for all the remote devices without exhausting the transport link capacity.

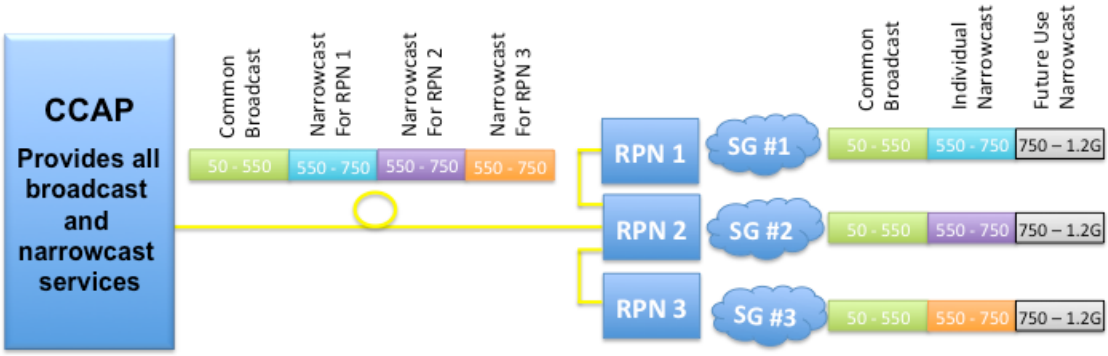


Figure 10 - Reuse of broadcast capacity across multiple RPNs

3.2. Increased Headend Equipment Density

The implementation of DAA makes it possible to improve the density of CCAP devices in several ways.

First, while CCAP devices are normally implemented via separate upstream and downstream line cards, a DAA line card could implement both upstream and downstream. This, in effect, doubles the capacity of a CCAP chassis.

In addition, a typical CCAP downstream line card will house eight or perhaps 12 RF ports, as defined by the printed circuit board space consumed by the components required for RF modulation, plus the sheer connector spacing required. However, Ethernet connectors can be placed considerably closer to one another, allowing a similar line card to easily house 16 to 24 ports. This additional density gain once again doubles the capacity of a CCAP chassis.

Finally, it is possible to consider "daisy chaining" remote PHY nodes (RPNs) off of a single CCAP Ethernet port. This is because, on the one hand, the capacity of a 10 Gbps Ethernet link would support the capacity needed for a single RPN. Plus, it is possible to generate an RPN "channel lineup" by transmitting the broadcast content once to multiple RPNs. As depicted in Figure 10, the data stream transmitted from the CCAP could contain a single "copy" of the broadcast line-up content, plus individual versions of the narrowcast content for each of the RPNs. The RPNs would then reuse the broadcast lineup content to recreate the individual RPN channel lineup. In this way, each service group served by the CCAP port would contain the same broadcast lineup while allowing for its own unique narrowcast line-up.

Then, as the narrowcast lineup capacity grows over time, CCAP ports would be segmented to support fewer RPNs, akin to the way service groups are split today to provide more narrowcast capacity as it is required.

As summarized in Table 3, the combined effect of the three factors described above is very significant, ranging from 8x to 18x of headend capacity gain. From a space and power perspective, this can facilitate huge savings.

Table 3 - DAA headend density gain

Density Factor	Density Gain
Combined US/DS line card	2x
Greater number of ports per line card	2x to 3x
Multiple RPNs per CCAP port	2x to 3x
Combined capacity gain	8x to 18x

But, just how meaningful is this headend density gain?

Consider: A migration from an HFC architecture with an average of N+5 (meaning an optical-to-RF node followed by five cascaded amplifiers) to N+0 would require about 10x the number of nodes, and the headend density benefits resulting from the DAA would neutralize the potential increase in CCAP equipment.

It is then quite clear that from a space and power savings, distributed access architectures take the benefit of CCAP to a whole new level.

3.3. Integration of HFC and Fiber Services

One of the largest areas of growth for operators is business services. Cable operators have deployed business services via both cable modems and fiber-based infrastructure. The fiber-based services are either point-to-point, using Ethernet and wavelength division multiplexing (WDM), or point-to-multipoint, using PON (passive optical network) technologies (either EPON [Ethernet passive optical network] or GPON [gigabit passive optical network]).

This duality results in the existence of two parallel networks. One of them, the HFC infrastructure, uses fiber from the headend to the node via an analog intensity modulated link for the forward direction and either analog or proprietary digital return, followed by coax infrastructure from the node to the home. The other consists of digital fiber from the headend to the customer, which is often used for commercial services.

Given the use of a digital fiber in both the forward and the return for the RPN, and especially because this digital fiber is based on Ethernet technology, it is possible to collapse both of these networks into a single infrastructure. Even without fully collapsing the Ethernet network better utilization of physical fiber can be enabled by the move to common DWDM wavelengths and multiplexers.

Therefore, the implementation of RPNs with an Ethernet interface between the CCAP and the RPN would make it possible to implement a PON interface at the RPN.

The benefits from this integration include:

- Reduction of the optical link for PON to the distance between the node to the customer premise
    - The typical distance from a node to a customer premise in an N+0 architecture is 1 to 2 kilometers. This would virtually eliminate any distance limitations for PON, making it possible to implement the largest possible densities.
    - In addition, this shortened distance would enable the use of lower power optics, which can translate into significant savings – especially for 10 Gbps optics, and for the upstream, which results in significant savings in the ONU.

- Leverage a single network for multiple services, which will reduce maintenance and increase operational efficiencies.

3.4. Migration Strategy

One of the more concerning issues to operators is the migration strategy when going to DAA.

Any migration that requires synchronized cutovers, or which requires changes in multiple locations to execute, is problematic, and usually results in a barrier to adoption. Therefore, it is very important that the migration to DAA allows for unsynchronized changes.

Ideally, the migration to DAA allows for opportunistic changes in the network. For example, one such change would be to migrate a single node, such as would be the case in a multiple dwelling unit (MDU) to increase capacity.

As it turns out, DAA enables such gradual, unsynchronized and opportunistic changes in the network. What follows is an overview of the steps and components involved in the migration to DAA.

Starting with the components of the network on both sides of the DAA, neither the back-office nor the various components in the customer premise need to be modified in any way. All back-office components are unaffected by the migration to DAA, and any additional MIBs (management information bases) for management and/or commands for configuration can be added well before the first CCAP line card or node is deployed. Customer premise devices are not affected in any way when deploying DAA, and any enhancements that are made possible through the introduction of DAA would be implemented in CPE that can be introduced before or after the migration to DAA.

The critical portion of the network where changes need to be made are in the headend and the plant.

To begin with, the changes required in the headend are primarily in the CCAP platform. The CCAP architecture was specifically designed to support multiple technologies simultaneously, which makes it possible to install regular RF upstream and downstream line cards and DAA line cards in the same chassis. While some operators may choose to deploy a separate CCAP platform for DAA, it is certainly possible to support both types of line cards in the same chassis. Of course, these DAA line cards can be installed at any time prior to beginning the migration in the plant, and any removal of RF upstream or downstream line cards can follow the deployment of any number of DAA line cards or nodes.

Turning our attention to the plant, it is similarly possible to migrate regular nodes to DAA nodes in any sequence. As an example, what follows is a sequence of steps where a single node is gradually converted from standard HFC to DAA.

Figure 11 depicts a single HFC node connected to a CCAP device.

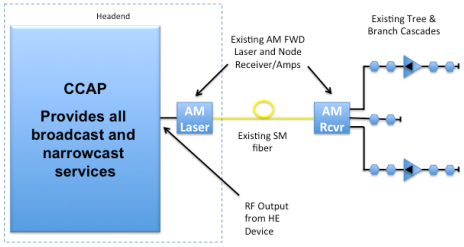


Figure 11 - Single traditional HFC node

Figure 12 shows how the HFC node would be converted to RPN while the rest of the HFC network remains unchanged. The DAA line card in the CCAP would have been deployed in the headend in a prior activity, and even the RPN could have been deployed before the day of the cutover. Then on the day of the change the fiber cable could be swung in the headend from one AM laser to the CCAP DAA card, and in the field from the HFC node to the RPN. Of course it is not necessary to perform the migration in such a fashion, but it would be possible if desired.

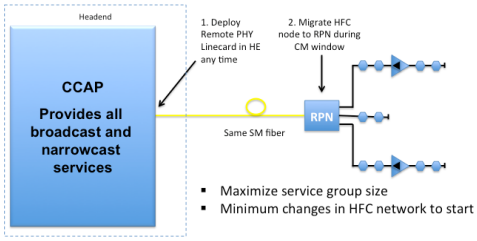


Figure 12 - RPN deployment step 1

Figure 13 depicts a possible step 2 in the process, whereby additional RPNs are installed to segment the original service group further. These additional RPNs could be daisy chained from the original RPN by taking advantage of the broadcast reuse feature, minimizing complexity in the deployment process.

NOTE: The example depicted is one in which fiber is run to every amplifier station. However, a more efficient segmentation scheme would include optimal placement of RPNs in an N+0 HFC architecture with some turnaround of passive components.

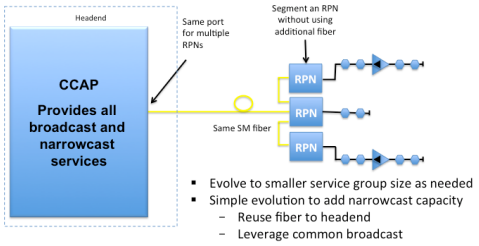


Figure 13 - RPN deployment step 2

Figure 14 shows how further segmentation could take place by replacing the remaining amplifiers in the network with RPNs.

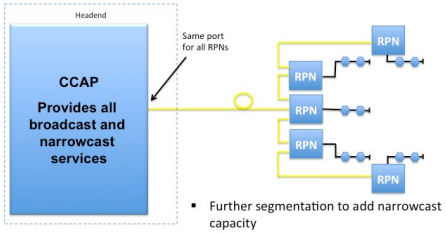


Figure 14 - RPN deployment step 3

Figure 15 shows that the RPN service group depicted above is segmented as additional narrowcast capacity is required. In this example, two of the RPNs from the DAA service group shown in Figure 14 are split into separate service groups using separate CCAP ports.

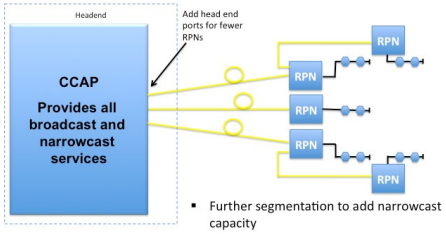


Figure 15 - RPN deployment step 4

Eventually each of the RPNs could be connected to an individual CCAP RPN port as shown in Figure 16. This would provide up to 10 Gbps of capacity to each RPN. This could, for example, be desirable to provide both RF and PON services from the RPN.

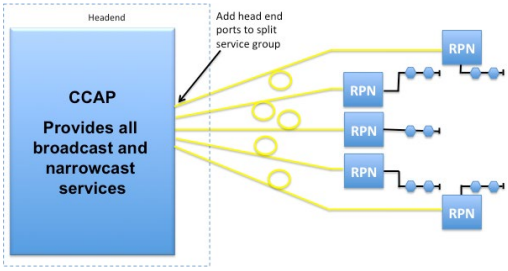


Figure 16 - RPN deployment step 5

Similarly, the DAA line card in the CCAP could be upgraded to support even more capacity as such capacity is needed and becomes cost effective. For example, the Ethernet link from the CCAP to the RPN could eventually be upgraded to 40 or 100 Gbps, both of which are already commercially available.

3.5. From Today to Virtual CCAP

As the network has to continue in operation through the transition, virtualizing the CCAP requires careful planning and a sensibly staged process. As with roads, where cars must be kept moving during any lengthy highway reconstruction, in the network customer traffic must continue flowing day after day. In a sense, while road work is visible to car drivers, in a network the modifications remain invisible to the end user.

One way to do so is to migrate individual functions, one at a time. So, one must develop a list of the functions that would be virtualized, and this list would be prioritized, such as on the basis of complexity of implementation and benefit. Those features with the lowest implementation complexity and the highest benefit would be prioritized higher in the list, and consequently implemented first.

In DOCSIS 3.1, one of the functions that would rise to the top of any such list is modulation profile management (MPM). This is because MPM will take time to be implemented by vendors in a CCAP chassis, but implementing externally via virtualization could be quite simple. In the process, its benefit to operators is quite significant since it would enable better efficiencies from DOCSIS 3.1.

Over time, implementing virtualization of the various functions of the CCAP would lead to a completely virtualized CCAP platform. Such a platform would be more easily scalable than CCAP

platforms are today, where segmentation of service groups requires the addition of more chassis in a linear relationship fashion.

In addition, and perhaps more importantly, virtualizing the CCAP will enable the development of additional functionality, and improvements to such functionality, to occur much more rapidly than it is possible to do today.

4. DAA Components, Use Cases and Generations

As operators move forward with the implementation of DAA, the evolution of DAA components, implications of the various use cases, and the generational aspects of DAA have become important to understand and track.

4.1. DAA Components

As the link from the headend to the node is converted from analog intensity modulated forward to digital, using Ethernet as the transport, several approaches can be taken for the implementation of the remaining headend components. In this section we examine one approach in some detail, for which a key goal is to convert all required components into functionally individual software pieces implemented independently.

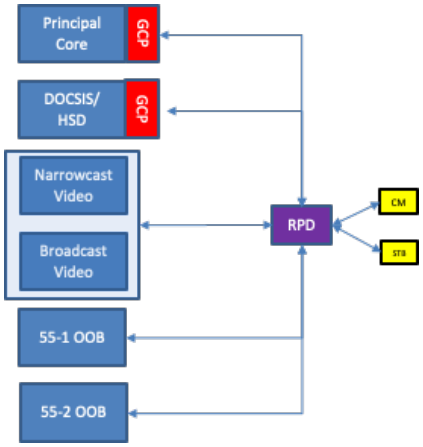


Figure 17 - DAA implementation components

4.1.1. Advantages and disadvantages of discrete components

As shown in Figure 17, an implementation approach for DAA is to develop discrete software (SW) components for each of the various DAA components. Some of the advantages for doing so include:

- The implementation can consist of a multi-vendor platform, where each component can be developed by a different party.
- By having smaller functional components, their implementation tends to be simpler.
- Time to market also tends to be reduced.

However, implementation of smaller discrete components has its downsides, such as:

- There is an implied requirement to more tightly specify the behavior of each component to ensure that the overall system will operate as intended.
- Interface specifications between the various components is required.
- Management of the various components, including their configuration and upgrade is generally more complicated, and requires more elaborate orchestration.

4.1.2. Key DAA discrete components

The key components depicted in Figure 17 include:

- The principal core, which is the first component that the RPD will contact after receiving an IP address.
    - The principal core is implemented such that it will configure all the RPD functions except DOCSIS channels and behavior, which will be implemented by the DOCSIS CMTS.
    - As depicted in Figure 17, the principal core communicates with the RPD using the GCP protocol, for which it is known as the Generic Control Protocol Principal, or GCPP for short.
    - Included in the GCP principal core are all the non-DOCSIS command and control functions for the RPD, including configuration, management and reporting.
- The DOCSIS core, which is the second component that the RPD will contact in the network.
    - The DOCSIS core also communicates with the RPD using GCP.
    - The DOCSIS core provides all configuration, command and control for DOCSIS channels, both downstream and upstream.
- Narrowcast and broadcast video engines
    - Implemented as separate components, the narrowcast and broadcast video engines provide all the content services for the various RPDs in the network.
    - Neither the narrowcast nor the broadcast video engines communicate with, nor have knowledge of, the RPDs.
    - Services are configured statically in the narrowcast and broadcast video engines upon their bring-up, and are multicasted to all RPDs, which listen for these services as configured by the GCPP.
    - The narrowcast and broadcast video engines could be implemented separately, but they could be operated together as a single functional system.
- Out-of-Band engines or cores
    - The OOB (out of band) functions are implemented separately from the video engines.

- Given that video systems are implemented using a single encryption and command/control technology, only one (i.e., either SCTE 55-1 or SCTE 55-2) of them is deployed in any one system.

- The OOB function may or may not implement GCP for communicating with the RPD. When GCP is implemented the OOB server is a core, and it will configure the OOB downstream and upstream OOB channels in the RPD. However, when GCP is not implemented the OOB server is an engine, and the GCPP will configure the downstream and upstream OOB channels.

- Finally, not depicted in Figure 17 is a very important component: the timing server.

  - The timing server, also known as the grandmaster, provides the critical timing synchronization for all the DAA components.

  - Each of the DAA components will include a timing client, which will communicate with the timing server to maintain timing synchronization.

  - While timing synchronization is not absolutely critical for video services, it is imperative for DOCSIS service to operate. Therefore, video services may be initiated before timing synchronization is achieved, but DOCSIS services will not.

4.2. Key aspects of DAA interoperability

The base implementation of a DAA system is generally simple. However, significant complexity is introduced when interoperability with different vendors' components is introduced.

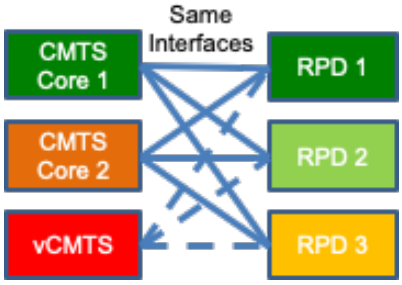


Figure 18 - Functional CMTS-RPD interoperability matrix

As depicted in Figure 18, the number of combinations of interoperable components increases geometrically as additional components are added on either side of the interoperability matrix. Having a single CMTS interoperate with multiple vendors' RPDs is complex and requires a lot of careful planning and implementation. If the number of CMTS implementations is increased to two or three, the interoperability complexity doubles and triples respectively.

When considering the overall CCAP system, the complexity to achieve multi-vendor interoperability is even larger. For example, if multiple GCPPs and/or multiple video engines and/or multiple OOB engines/cores are introduced into the mix, the amount of complexity and work required for testing and interoperability results in increases by orders of magnitude.

Therefore, a multi-vendor RPD deployment coupled with a single headend implementation is a sensible approach to an interoperable DAA ecosystem.

4.3. Use Cases

In the same way as there are different kinds of nodes for different HFC network applications, there are RPDs with different characteristics that are best suited for each of the specific HFC network use cases. Similarly, while there are use cases for RPDs in the outside plant, there are also applications for RPDs in headends, or the "inside plant" as it is frequently called, which will have different implementation characteristics. The following sections cover the key scenarios.

4.3.1. Outside plant

The environmental characteristics of RPDs developed for outside plant make the design of such devices very different than for RPDs developed for inside plant. The key characteristics for outside plant RPDs are as follows:

- Designs must conform to very tight space availability requirements inside of a node enclosure
- RPDs must support an environment where heat dissipation without the use of fans is critical
- Powered from quasi-square wave power supplies used in HFC networks
- Minimize power consumption to the extent possible given the limited amount of power during normal operation and especially during stand-by power mode

In addition, and perhaps more importantly, there are different kinds of nodes for different HFC network applications, which will drive varying designs for RPDs for outside plant, as follows:

- Traditional HFC networks include cascades of multiple amplifiers and cover a plant footprint of a few hundred homes. In such cases the network capacity offered by the RPD should be maximized, such as including multiple downstream and multiple upstream ports.
- Newer HFC networks are built with fewer, or even no amplifiers, and are targeted to cover smaller network footprints. In such cases it is not necessary for the RPD to support much more than a single downstream port, with either a single or dual upstream ports.
- Finally, given that scaling is needed as in any other network application, it should be possible to support greater capacity over time to the extent possible. For example, while initial deployment may only require a single downstream and/or upstream, over time service group segmentation may require additional downstream and/or upstream ports in a single node. For that purpose it is usually a design requirement that multiple individual RPDs fit into a single node enclosure.

Given the above, RPDs that are built with a single downstream and a single upstream (frequently called 1x1) or a single downstream with dual upstreams (1x2), such that operators can place a single one in the node or place additional units when capacity demands require it. Newer silicon designs include more capacity at lower power levels, making it possible to develop RPDs that

contain multiple downstream ports and multiple upstream ports, such as 2x2 and 2x4 designs in a single RPD.

4.3.2. Inside plant

In contrast to the outside plant environmental characteristics, RPDs developed for inside plant have other constraints that make the design of such devices very different than for RPDs developed for outside plant. The key differences in the design of RPDs for inside plant are as follows:

- Rather than the physical volume of the allocated space, the layout is a primary concern so that dense set-ups in a rack are possible, including cabling distribution in the front and/or back of the rack.
- Designs must support an environment where forced air is used for heat dissipation, requiring airflow from front-to-back or back-to-front, sometimes allowing airflow from side-to-side and/or in a vertical direction.
- Powering frequently requires DC power supplies, but AC power supplies are used in other cases.

In the case of inside plant RPDs, these are frequently implemented in one of two different form factors:

- Modular, where RPDs are individually removable in a chassis-based design, or
- Fixed, where the entire set of RPDs are part of a monolithic device

The modular design is generally used in larger headends where the ability to replace a defective individual unit is a paramount concern. The fixed design is targeted for a smaller facility, or even a cabinet, where space is the primary concern.

4.4. Generational considerations

One additional consideration is the evolution of the DAA to support new generations of equipment.

The initial implementation of DAA components included support for DOCSIS 3.1. The main component that is specifically developed and implemented for DOCSIS 3.1 is the RPD, which incorporates application specific integrated circuit (ASIC) devices which support up to DOCSIS 3.1, but will not support DOCSIS 4.0 functionality.

As newer parts of the DOCSIS 4.0 specifications are implemented, such as full duplex (FDX) DOCSIS, the RPDs will have to be swapped out in order to expand their support for DOCSIS 4.0. This process is akin to what had to be done with CCAP line cards in the past, where either upstream and/or downstream line cards are swapped over time as new versions become available. And, as part of this upgrade, the older equipment is reused in other locations where the newer equipment is not yet needed.

However, for the remaining DAA components, if these are implemented in software on general purpose compute platforms, these should be upgradeable to support newer DOCSIS specifications

such as DOCSIS 4.0 and/or other enhancements by simply expanding the functionality implemented in software and downloading it to the platforms in which they run.

In fact, the process for upgrading the DAA components to support changes in the functionality including enhancements to DOCSIS, becomes easier than ever before given the nature of the DAA platform, especially when the DAA implementation includes a minimalistic RPD at the edge of the HFC network and virtualized components for the remainder of the DAA components.

5. Deployment experiences and lessons learned

As operators move forward with broad deployment of DAA, including tens of thousands of RPDs in hundreds of locations, there are many useful lessons learned. The following sections of this paper describe findings in four key areas: monitoring system approaches, plant powering, provisioning processes and approaches, and hardware and software management techniques and experience.

5.1. Monitoring System Approaches

As the number of DAA sites and components increases rapidly, the use of tools such as Simple Network Management Protocol (SNMP) become exceedingly difficult to operate and scale. With just a few thousand RPDs, operators would have more end points from which to gather data than they have CMTS or CCAP systems in total. Having deployed multiple tens of thousands of RPDs, and projecting that deployments will pass the hundreds of thousands in just a few years, streaming telemetry approaches become much better suited for monitoring such vastly distributed systems, and perhaps the only viable approach.

Open source tools like Grafana can be very easily used to plot streaming data collected from network components. And, gathering data from various sources, such as those outlined in Section 4.1, allows for quicker correlation of issues in the network.

As an example, Figure 19 depicts the status of RPDs in a particular area of the network. The view in the figure shows the status of RPDs as seen by a vCMTS, including the number of active and inactive RPDs and the corresponding number of cable modems in those RPDs. The vCMTS is one of the network cores in the DAA, and similar views are available as seen by other network cores as outlined in Section 4.1.

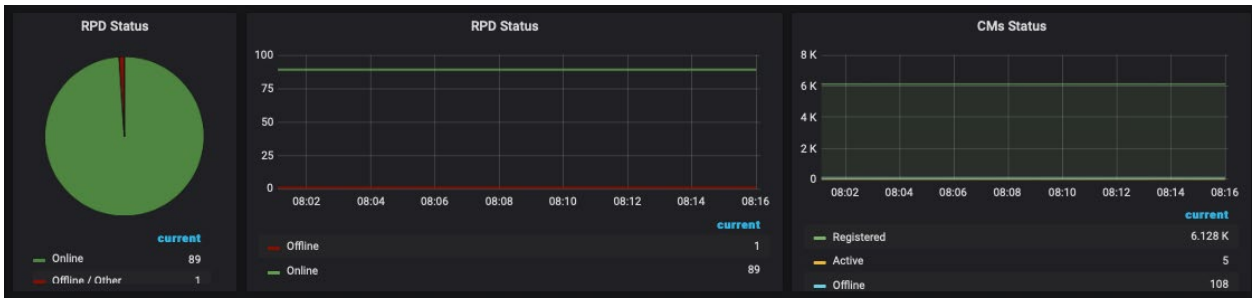


Figure 19 - RPDs and cable modems in a particular section of the network

Streaming telemetry can be easily obtained from modern compute resources, such as those used to virtualize the CMTS and other DAA components. However, streaming telemetry is not currently

available from RPDs. Instead, an RPD supports polling telemetry through protocols such as GCPP, which is used in the RPD for other purposes, including configuration.

To further expand on the use of streaming telemetry to RPDs, work is currently underway in an effort led by CableLabs, and supported by cable operators and suppliers alike, to incorporate streaming telemetry in the remote PHY specifications. This specification development effort includes a prototype being developed between CableLabs, Comcast and the support from vendors.

5.2. Plant powering

Plant powering of RPDs has become surprisingly challenging, and as a consequence represents a rich set of "lessons learned." This section provides some relevant background, describes the challenges encountered, and the solutions applied to date.

5.2.1. HFC network topology and component powering

Network power supplies have always provided power to multiple HFC plant devices. Historically this included nodes, amplifiers and line extenders. In addition to HFC plant devices, network power supplies provided power for any other device connected to the plant, such as Wi-Fi access points, cellular microcells, etc. In the past, the plant powered some individual CPE devices used for specific applications, such as telephony (which are largely removed from cable networks now).

The footprint covered by an individual network power supply, which we might refer to as the power serving group, depended on the specific characteristics of the HFC network, especially related to the density of the area in question. In areas where traditional N+x HFC architecture is deployed, an individual power supply may provide power to a single node and a collection of amplifiers and line extenders, plus whichever other devices may be present in that power serving group. In the case of passive HFC plants, such as fiber deep and/or N+0 plants, a single network power supply might power a number of nodes and any other devices that may be present in the power serving groups.

5.2.2. Traditional network maintenance activities

Before the introduction of DAA, the HFC plant devices became operational almost instantly upon receiving power from a network supply. By contrast, other devices that may be present in the power serving group, such as a Wi-Fi access point or microcells, would take several minutes to become operational once network power is supplied.

Over the years, the normal process of a maintenance technician included some interruptions of network power for troubleshooting purposes. This might include procedures intended to identify sources of signal interference or impairments, or replacement of HFC components that might be causing signal interference or impairments, such as taps, splitters, power inserters, and any other passive component, all of which could be done very quickly. In doing so, power would be interrupted for a matter of seconds, or even fractions of seconds, which would cause a very brief loss of power for an amplifier or line extender, which CPE devices would not perceive, and therefore not cause an interruption of service beyond the brief moment of the power interruption itself.

5.2.3. Impact on DAA network devices

As DAA components have been deployed, those nearly imperceptible power interruptions described above became much more impactful.

For example, while a traditional node, amplifier or line extended would recover from a power interruption nearly instantly when power was restored, the same does not apply to an RPD. Instead, the RPD, which has a processor (or multiple, actually), requires considerable time to boot (in the minutes range), and then go through a power-on self-test (POST) process, which takes additional time to complete. This was followed by processes such as time-synchronization, which take even more time to complete.

Therefore, while a second or sub-second power interruption would have been imperceptible to most users in a traditional HFC network, it now causes multiple minutes of interruption until an RPD becomes operational, and even longer until the CPE in the home becomes operational once again.

To make matters worse, in cases where a serving group power supply provides power to multiple RPDs, such as is the case in passive networks, the service interruption would affect an even larger footprint.

5.2.4. Solutions

Several actions can be taken to overcome these problems, including training, shortening the power cycle process, and using power holding devices.

The most immediately available solution to minimize power interruptions is training, to prevent unnecessary power interruptions. As all cable operators' maintenance team leaders well know, this is much easier said than done. Training requires time, practice and repetition, all of which takes months or years to implement. Furthermore, there are many cases when interrupting power becomes impossible to avoid, such as when component replacement is necessary for impairment resolution.

Shortening the power cycle process requires, in most cases, substantial changes in the SW running in an RPD and the related systems, and in their integration. For example, it is relatively easy to simplify the POST process and/or to shorten the bring-up time. It requires more effort to reduce the configuration time, such as by keeping and reusing configurations whenever possible. Other aspects of the power-up process are more difficult to shorten, such as the time synchronization, which can only be shortened, and even eliminated, with a reboot from active operation that did not follow a power cycle.

An additional approach for eliminating short power loss events is to implement power holding devices within the RPD or in the node housing the RPD. Such devices, which are based on hardware, take even more effort to implement and have a linearly increasing cost (i.e., each unit has a cost, and the cost does not decrease substantially with the deployment volume). In addition, power holding devices can't hold power for considerably long periods of time, so they only prevent short power interruptions (i.e., power holding devices only prevent power interruptions lasting a few seconds).

The combination of all of the above approaches is probably the most effective combination.

5.3. Provisioning process and approaches

The installation, configuration, bring-up and provisioning process for an RPD requires information related to all the services provided by cable operators, plant characteristics and operational requirements. The process is complex given the various sources of information, and requires automation and verification. The following paragraphs provide an overview of those topics.

5.3.1. RPD Installation

There are two fundamental types of RPDs: those that are installed in nodes and those installed in headends (shelf-based). While the two types are functionally the same, the use case is different, and therefore some aspects of the installation and configuration are different.

For example, for node-based RPDs, there is an individual RPD to identify, configure and verify (unless two RPDs are placed inside of the same node). In the case of shelf-based RPDs, the identifier label used is typically for the entire shelf instead of being specific for an individual RPD, and identifying the specific RPD that is being provisioned requires an understanding of the shelf architecture.

While RPD identification may seem trivial, actually identifying an RPD that is part of a shelf presents several complications. Finding the identifier label may be difficult, especially if the shelves are installed immediately above or below another shelf (i.e., without any space between them in the rack). Additionally, the media access control (MAC) address of the RPD, which is a fundamental component of the configuration process, may be derived from a base MAC address for the shelf, instead of included on the label as it would be for an individual RPD in the node. While these may seem inconsequential aspects of the installation process, they become significant complexities operationally, especially when speed of deployment is a key goal.

An additional twist to the installation process is the replacement process. When an RPD needs to be replaced, the process has to involve some form of removal plus an addition. Given that the RPD configuration is specific to the RPD, the MAC address of the RPD being replaced needs to be removed, and the MAC address of the new RPD needs to be used in its place.

Finally, the installation process needs to take into account the isolation of known defective RPDs, a return material authorization (RMA) and repair, and a manufacturing, purchase and installation identification for tracking devices with problems and defects, all of which has to be taken into account as part of the process.

5.3.2. Information and Sources

The information that is used in the configuration and provisioning process includes plant characteristics such as power levels, tilt, leakage detection tones, pilot tones, etc. It also includes service information, such as broadcast and narrowcast video services' frequencies and channels, DOCSIS channel frequencies and modulation characteristics, and OOB signals, to name the most common.

The above information comes from different sources within the operator's organization, and therefore various individuals determine what's the correct information and use various types of source data. This can naturally lead to conflicting information and configuration.

In addition, there is added complexity in having multiple configuration sources for an RPD. For example, in the DAA described in this paper (see Section 4.1), the RPD receives configuration information from two or three different sources within the network, which include the GCPP, followed by the vCMTS auxiliary core, and the OOB engine auxiliary core for 55-2 encryption networks. Having multiple sources of configuration information can easily lead to conflicts that are hard to identify.

Finally, parameters change over time for a multitude of reasons on a regional or per-node basis. Some of the changes will apply to a number of nodes simultaneously, such as changes in channel line-up and/or addition of DOCSIS channels. Other changes will be unique to a node, such as changes in channel configuration to optimize capacity usage through profile management application (PMA) tools.

Given the multitude of information sources and configuration components, it is almost inevitable that errors will occur. One example involves the specific frequency assignments for either of the configuration components which might be incorrectly captured or communicated, including frequency overlaps. Given that the RPD does not have the inherent intelligence to understand which overlaps are intended and which are errors, it is possible to create a configuration that will not work.

5.3.3. Configuration and Verification Tools

Resolution of configuration errors, such as overlaps and conflicts, could be considered to be relatively straightforward. Configuration tools can certainly be developed to capture the information, and automation can be implemented to minimize errors, and even prevent them in some cases.

However, there are factors that make the detection of errors more challenging, such as when the information arrives at a central configuration tool from different data sources. In that case, an additional tool would have to consolidate all the configuration information before the final configuration is created, which is especially the case if the configuration is created “on the fly.” In such cases, the change in the configuration could have been made a priori, perhaps hours or even days before the configuration is applied, and therefore the person or system that commanded the change would no longer be notified of the conflict.

Other information and/or changes are much harder to verify, such as configuration of pilot tones or leakage tones, either frequencies and/or signal magnitude. Only experts in the specific field might understand the information that is being entered, and therefore determine that a different value would have been appropriate.

Perhaps the ultimate approach to verify the configuration is to conduct a field configuration test. Tools have been developed for quite some time to measure RF characteristics of a channel line-up, including signal levels, channel content, etc. More recently, tools have been expanded to read an entire channel lineup, and compare it to a known, “good” lineup, to determine if there are any errors. Such tools are very useful because they verify the actual RF signals being generated by the RPD in their entirety. However, such testing does take resources, time and additional cost.

5.4. Software and hardware management

As with any device containing a processor, RPDs require software upgrades. While that is very much expected, how frequently changes are required may be surprising. In addition, hardware changes and

use cases have been even more frequent than anticipated. The following paragraphs provide an overview of some of these situations found through field experience.

5.4.1. Software upgrades

Everyone involved in access networks has experienced their fast-paced evolution. Changes not only include the expansion of capacity, which is, of course, a very big challenge, but also of functionality, uses, and most recently of configuration, including dynamic changes.

The most typical of such changes involves incorporating new or modified functionality. For example, almost all operators deployed downstream channel bonding before upstream channel bonding, or downstream OFDM channels before upstream OFDMA channels. Similarly, lots of functionality included in the DOCSIS specifications is deployed over time. As a result, equipment manufacturers implement such functionality progressively, as it becomes necessary and useful to the operator. Therefore, software upgrades are required to make such additional functionality available.

Less predictable are new or different use cases. For example, while initially an operator may deploy RPDs in passive networks, they may over time expand to using RPDs in a more traditional N+x HFC network. Therefore, while the operator may not have needed to support pilot tones initially, such functionality may be required later on in the deployment lifecycle. As with any other functionality that is not needed initially, it may not have been incorporated into the initial software load and then deployed over time, requiring software upgrades.

An additional type of software change involves making configurable information dynamically possible. For example, initial deployments of DOCSIS channels may not require that modulation profile configuration changes be made without rebooting the RPD, such changes may need to be made dynamically later on, without requiring the reboot of the RPD, and even without requiring cable modem reset or even re-ranging. As with the above cases, such changes in the RPD operation require software upgrades.

Finally, as with any other software, bugs are almost inevitable. Therefore, when bugs are found they have to be identified and fixed, which of course requires software upgrades.

Upgrading RPDs that have been deployed requires the use of software deployment tools. This is important when the number of RPDs that have been deployed is in the thousands, and it becomes especially important when the number of RPDs increases by one, two and possibly three orders of magnitude.

In addition, it is inevitable that RPDs are manufactured over time, bought over even longer periods of time, and then placed in warehouses until these are deployed. Therefore, RPDs are almost always upgraded upon initial deployment, which generates the need to manage multiple software versions and upgrade processes.

Moreover, RPDs become even more challenging when their variety increases, for example when multiple vendors are deployed in a network, and especially when multiple models from each vendor are deployed. Performing field tests becomes necessary, requiring handling of exceptions, and gradual or partial deployments are inevitably necessary to support operational requirements. All this makes the need for a sophisticated software management and deployment tool.

5.4.2. Hardware upgrades

Just like having to support multiple software versions, over time multiple hardware versions are also inevitable. This can happen because of product evolution by manufacturers, new use cases by the operator, and industry technology evolution. Let's examine a few examples of each.

Product evolution is probably the least appreciated, but the most real form of hardware evolution. Manufacturers will almost always evolve their products to reduce costs, resolve component obsolescence, or achieve better performance. Therefore, multiple hardware versions are almost always necessary and inevitable. In some cases, it might be necessary to keep track of different hardware versions for purchase tracking requirements, especially when costs change. But keeping track of hardware versions is also important for other reasons, such as for certain troubleshooting activities, and even to understand and support certain software differences, such as when certain newer versions support features not available to older hardware versions.

Less predictable is the evolution of use cases. For example, an operator may initially deploy RPDs with fewer downstream and/or upstream ports than may be required over time. While it may be possible to double the use of some RPDs on certain nodes (i.e., start with one RPD in the node and later move to two RPDs in the node), in other cases it may be necessary to deploy a different type of RPD (i.e., start with one model of RPD in the node and over time move to use a different model of RPD in the node). In most cases the evolution would require one to keep track of the number or modem of RPDs in each node. Yet another example of change in use case would be the initial need for node-based RPDs, which eventually changes to incorporate the use of shelf-based RPDs, or vice versa.

Finally, the almost certain hardware evolution scenario is that the technology used initially evolves over time. For example, initially RPDs were developed to support DOCSIS 3.1, but more recently their design has been modified to support FDX. Cable operators have dealt with the evolution of access network technology for decades, migrating from pre-standards to DOCSIS 1.0, then to DOCSIS 2.0, followed by DOCSIS 3.0 and eventually DOCSIS 3.1. Equipment management, including shifting equipment between locations, is a process that operators have mastered and will continue to apply to support technology evolution.

6. Conclusions

Demand for more narrowcast service capacity has driven many changes, allowing operators to reclaim spectrum. Splitting nodes to reduce the number of homes passed in an HFC node's footprint is one key method used to provide more narrowcast capacity. Enabling narrowcast services on this reclaimed and newly added spectrum requires more equipment in the headend. Given the trajectory of growth, using traditional equipment technology creates an increased demand for more headend space. Deploying DAA enables the required growth without expanding facility footprints.

The deployment of the highest modulation orders available with DOCSIS 3.1 required higher MER out of the node. Transitioning to a digital forward link, which became possible with the newer DAA, is a key enabler to achieve this demand. Moving the entire PHY layer to the node has become the standard method to implement a digital forward link.

Deploying DAA in the outside plant allows for more efficient infrastructure utilization between services. DAA and commercial services can share physical fiber due to commonly used DWDM wavelengths and spacing. There is also the possibility of a converged Ethernet switching network to serve both DAA

nodes and commercial customers. Having Ethernet inside of a DAA node allows for efficient PON deployments sourced from the same node.

Moving from specialized CCAP hardware to virtualized components leads to a more scalable system, where additional capacity can be added without the step function of adding new CCAP chassis. Furthermore, feature velocity can be improved on a virtualized system. DAA is implemented with discrete components that allow for a multi-vendor platform and includes smaller functional components.

On the flip side, implementing discrete components requires well behaved interfaces that adhere to a well-defined set of specifications. In the DAA implementation described in this paper, key functions include the GCPP core, the DOCSIS core, the narrowcast and broadcast video engine, the SCTE 55-1 OOB engine or SCTE 55-2 OOB core, PTP timing distribution, etc.

RPDs are available for both inside and outside plant applications, implemented as individual units or in shelves containing multiple RPDs, and including single or multiple services groups. Power level and segmentation capability/capacity are key attributes of outside plant RPDs. Density and serviceability are key attributes of inside plant RPDs.

Even though current generation RPDs lack support for FDX operation, other DAA components may have the ability to be upgraded for FDX operation using the current hardware. This enables an easier transition to FDX on an as-needed basis, by swapping hardware on desired nodes while utilizing the existing platform.

The experience acquired over the course of the last five years, especially the last two years, leads to several key lessons learned that are outlined in this paper. One such lesson is the need for flexible and scalable monitoring systems, using streaming telemetry and open source reporting tools. Another is the impact of plant powering considerations, which resulted in the development of software systems and hardware components to alleviate the effects of power interruptions during normal maintenance activities.

Having provisioned tens of thousands of RPDs already, a well thought out provisioning system process and approaches for configuration verification are very important to ensure smooth, rapid and correct deployment. It is especially important to consider the variety of information and the number of sources, and to create a process for gathering and applying the information that minimizes errors, some of which are easier to foresee than others.

Finally, given the need to deploy RPDs of multiple types and from multiple suppliers, software and hardware upgrade management becomes very important. Both foreseen and unforeseen changes require the use of flexible version management tools.

Abbreviations

A/D	analog-to-digital
AC	alternating current
AM	amplitude modulation
ASIC	application specific integrated circuit
BC	broadcast
CBR	constant bit rate
CCAP	converged cable access platform
cDVR	cloud digital video recorder

CM	cable modem
CMTS	cable modem termination system
CPE	customer premises equipment
CTA	Consumer Technology Association
DAA	distributed access architecture
DAC	digital-to-analog converter
dB	decibel
DC	direct current
DOCSIS	Data-Over-Cable Service Interface Specifications
DS	downstream
DTA	digital transport adapter
DWDM	dense wavelength division multiplexing
EPON	Ethernet passive optical network
FDX	full duplex [DOCSIS]
Gbps	gigabits per second
GCP	generic control plane
GCPP	Generic Control Protocol Principal
GHz	gigahertz
GPON	gigabit passive optical network
HD	high definition
HFC	hybrid fiber/coax
HSD	high speed data
Hz	hertz
I	in-phase
IP	Internet Protocol
ISDN	integrated services digital network
LDPC	low density parity check
kbps	kilobits per second
MAC	media access control
MDU	multiple dwelling unit
MER	modulation error ratio
MHz	megahertz
MIB	management information base
MPEG	Moving Picture Experts Group
MPM	modulation profile management
MTTR	mean time to repair (sometimes mean time to restore)
NC	narrowcast
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
ONU	optical network unit
OOB	out-of-band
OSS	operations support system
OTN	optical transport network
PHY	physical layer
PMA	profile management application
PNM	proactive network maintenance
PON	passive optical network
POST	power-on self-test

Q	quadrature
QAM	quadrature amplitude modulation
RF	radio frequency
RMA	return material authorization
RMD	remote MAC-PHY device
RPD	remote PHY device
R-PHY	remote PHY
RPN	remote PHY node
SC-QAM	single carrier quadrature amplitude modulation
SCTE	Society of Cable Telecommunications Engineers
SD	standard definition
SDV	switched digital video
SNMP	Simple Network Management Protocol
SNR	signal-to-noise ratio
SW	software
US	upstream
VBR	variable bit rate
vCMTS	virtualized cable modem termination system
VCS	video compression standards
VOD	video on demand
WDM	wavelength division multiplexing

Lessons Learned: Embedding AI in Cable Customer Experience to Better Serve Agents and Customers

An Operational Practice prepared for SCTE by

Rachel Knaster
Chief Product Officer
ASAPP
One World Trade Center, 80th Floor, New York, NY 10001
(212) 658-0990
Click here to rachel@asapp.com.

1. Introduction

As we all know the last two years drove massive change and presented new challenges for people and business. From a technology standpoint there was a renewed focus on artificial intelligence (AI) and its application— particularly as call wait times skyrocketed and customer frustration grew. The cable and telecom industry adapted, as it has done during its long history, and there has been an incredible opportunity for business. So, I wanted to share some of the lessons learned and best practices around AI for the cable and telecom customer experience (CX) that can bring immense value, efficiency, productivity, scale and brand loyalty.

Digital transformation is underpinned by technologies such as artificial intelligence and 10G— fundamental technologies to scale productivity, operations and generate better financial returns. Large enterprise businesses have petabytes of data, but major systemic CX inefficiencies that technology has not been able to solve to date. How cable companies interact with their customers is changing. Customer expectations are changing, and contact center agents can't scale without a move to digital. This touches on a fundamental challenge of rules-based technology: it is systemically inefficient because it relies on people programming rules constantly. Managing these rules is not scalable. Machine intelligence can be far more efficient in significantly improving productivity gains, financial performance, along with the customer and employee experiences. If implemented well, artificial intelligence can take months, not years, to implement and deliver the outcomes desired by the business.

With that said, people play a key role in defining successful customer experience and driving organizational success. Despite what many technologists would have you believe; AI is simply not sophisticated enough to replace all people on the frontlines of customer service. There are myriad applications for AI to help drive successful customer experience. To merely equate AI to chatbots does the field of AI research and development a disservice.

Shifting how an organization views and adopts artificial intelligence technologies like machine learning, natural language processing and speech recognition should not be seen solely through a prism of automating people out of the equation. Don't get me wrong automation is a key component but putting people at the center of thoughtfully designed AI tools will radically increase directed and independent automation, efficiency, productivity and make agents better at their jobs. By automating a significant amount of repetitive day-to-day processes, AI-driven solutions help increase the cost savings a company is seeking while creating a stellar customer experience.

2. Determine the Business Goals

A focus on business outcomes is important from the outset. Artificial intelligence has the capability to transform an entire cable company and its operations. Thinking big and thinking different is key. Nonetheless, the size, scope and steps within a major AI project should not be so wide that it creates a generic and impersonal customer experience, or so narrow that it misses part of the audience. Once goals are established, it's critical that the data used to build bespoke AI models isn't skewed such that models end up with poor predictions.

Healthcare IT supplier, McKesson recognized that if you want to earn customers, you put their interests first. That meant reducing hospital expenses in non-labor areas, such as medical supplies, one of the highest costs outside of labor. Once the goal was determined, McKesson looked beyond the traditional challenges and visible symptoms inside their operations. They looked across the entire supply chain to uncover the true problems and bottlenecks —challenging existing people, processes, and technology

solutions[1]. Using data insights learned from its product, location and transport data from scanners and sensors, the company was able to cut inventory by $1 billion when needed and identified opportunities for cost reductions and time-savings. [2] Artificial intelligence, when applied smartly, is capable of (as one executive at the time put it) "making the invisible, visible."

For the cable and telecom industry, if we consider customer satisfaction (CSAT) as a business goal, the competition isn't other large enterprises, rather it's the most popular consumer apps such as Facebook, TikTok, Instagram, even Uber or Lyft that both agents and customers are experiencing frequently. These apps provide best in class software that offers versatility and an intuitively designed user experience.

3. Measure What Matters

Over 20 years ago, Bill and Melinda Gates put $20 billion into their foundation to tackle global health challenges, spending $1 billion each year. Former CEO Patti Stonesifer said: "*Sometimes… we were probably measuring the wrong thing. You have to pay close attention to whether your data is getting you to the ultimate goal. We were learning so fast that sometimes we had to change data sets midstream. Say you had a seed that would double production of yams, and you were focused on that number. But then it turned out that nobody would use the seed because the yams took four times longer to cook at night."* [3]

And this is one of the big challenges with measuring what matters—not only can it be difficult to determine the most actionable metrics to focus on, but it can also feel more instantly gratifying to track vanity metrics. "A vanity metric is like a distorting mirror– you might like what you see in it but it's not a true picture, it's just your vanity." [4] These vanity metrics can be thought of as those that don't cleanly map to business objectives.

Take the example of an e-commerce business: it's certainly tempting and fun to track page views and to keep seeing that graph of page views going up and to the right. However, if the goal is sales, then "page views" is not the most clear or efficient metric to drive the desired behavior. In this example, a team might be encouraged to create superfluous or "click bait" links to the site to keep driving up page views, rather than focusing on meaningful engagement that drives sales The same is true of all of us trying to solve different problems in the CX space and in the cable industry.

When you're evaluating using artificial intelligence in CX operations, this industry needs to consider evolving the way it thinks about measurement. Take containment or conversion rate as examples. These are industry standard measurements established with rules-based technology.

In the case of conversion, an overly simplistic view is also problematic. Take your retail sales team, telephone-based sales team, and your digital-based sales team. All these groups will have varying conversion rates. Operating by established approaches, a retail-based team will always have higher conversion rates than telephone or digital agents. Does that mean companies should go out and build more brick-and-mortar stores? No.

[1] https://www.dcvelocity.com/articles/25482-who-are-this-year-s-rainmakers

[2] 'Datisim' Steve Lohr 2015

[3] 'Measure What Matters,' John Doerr (2018)

[4] Lean Startup Series: Vanity Metrics vs. Actionable Metrics (2018)

Shifts in buying habits, technological advancements and socio-economic changes are upending the way consumers engage with brands that have profound ramifications on how sales organizations capture, retain, and grow their customer base. This requires sales leaders to transform their organizations' approach for engaging consumers, particularly digital-first millennials, and Gen Zers who want to engage via digital.

So, it's important to review what you're measuring.

For example, contact center management has spent decades measuring average-handle time, concurrency, containment, and abandon rates among other measurements. But what about considering labor-per-hour, or revenue-per-labor hour as a new metric? Measuring conversion rate might not be the best metric, especially when a company wants to scale customer services. Take sales conversion rate in digital. If it's lower than retail, does that mean you abandon digital? No, it means we need to look at things differently to better understand the throughput of an organization.

Consider agent efficiency as throughput of customer issues: the number of issues handled per agent per unit of time. The industry can improve throughput both by decreasing agent handle time (AHT) for single issues and by increasing concurrency as well as by automating more conversations. As agents' use of AI expands, handle time decreases – and agents can manage more concurrent conversations, leading to higher throughput.

To encourage a shift to digital, cable companies need to encourage customers with awesome and intuitive experiences, reimagined with the help of AI-driven tools. Further, a rethinking of service level agreements for all channels is likely needed. Long term, digital provides better analytics and automation rates with a more convenient customer experience that is staffed more cost effectively with people who are operating more asynchronously.

Another example I see a lot is focusing on containment, meaning keeping conversations "contained" by automated systems that don't require reaching an agent. In the case of containment, it can only be treated as a proxy for success in limited situations because complex issues can rarely be handled effectively without human involvement. So, tracking containment as a "success" metric doesn't account for the significant number of customers who still require an agent to solve their problem. But what if the conversation wasn't successful? What if the automation did half of the work for the interaction so the agent no longer needed to do that half of the work? Under a metric like containment, any agent involvement is considered failure, but if we take a more holistic view of total throughput, we'd discover that we're significantly increasing the number of conversations that an agent can handle in the day.

The cable industry's CX operations have an opportunity to look more holistically at measurements and understand their organizational throughput, which requires a degree of change management to follow.

It's critical for organizations to measure what matters and update thinking by looking at what new metrics will deliver. Looking at the right measures can deliver results that will truly represent the customer experience overall.

4. Evolve Workforce Management Practices

Several years ago, the Associated Press announced that the majority of U.S. corporate earnings stories for its business news report would be produced using automation technology. In a relatively short amount of time, AP increased its earning coverage 14-fold delivering over 4,000 more public company earnings

quarterly, because of AI. [5] It altered the way their 248 bureaus globally reported news. The technology didn't replace journalists, rather it freed them up to tackle more work and interpretative reporting of earnings.

When cable companies want to use AI to maximize automation, productivity, and scale in its business, it's critical to understand where changes in the workforce management processes and operations will be required. With a change in how a company measures success of contact center operations, there needs to be an accompanying change in the way the contact center operates. To start it's important that you have the right people on the team. For example, it's important to have people who know and think about your customers and the accompanying journey mapping, operations, integrations, or the professional services that support a company.

There's a very interesting paradigm shift that needs to happen within the groups responsible for implementing or overseeing AI systems—the best results with AI are born out of allowing the systems to do their own work in processing vast amounts of data, not from configuring all the rules. Configuring rules gives a sense of control without allowing teams to achieve scale. I've observed across many use cases and many industries a learning curve in adapting to that flexibility of AI systems coming from the configuration mindset.

A specific example comes up with call routing. There's an urge to optimize by defining specific and narrow groups of queues, classifying calls into those queues, and setting up granular rules for when a queue is full or when it has availability. A downside of this is that defining rigid rules with such specificity can result in an imbalance of distribution—yes, if there's no constraints for equal availability and equal demand everywhere, then it can work, but the reality is often much messier. Depending on what is happening, which agents are staffed, how many customers have one question vs another, these types of policies can backfire. If instead we preserve larger groups of agents—but allow the models to learn about which agents do well with which types of customers, which types of intents, at which points in the day—AI can bring higher levels of global optimization vs sub-optimization of more distinct areas. [6]

Training for agents and how applications are permissioned for agents will need to evolve. Some agents will be able to handle two, three or more customers at a time asynchronously depending on their skill and experience level.

With the introduction of AI, an organization must be willing to invest in evolving the way it operates and willing to react flexibly as it encounters new challenges on the frontier of chartering a modernized customer experience and redefining its KPIs.

To drive digital adoption new SLAs will need to be designed, which could impact non-digital efforts like voice to succeed. When customers use a digital channel, they are retrained on how to approach the company in the future. This will in turn require an organization to look at intentional staffing, that might require labor sourcing or retraining existing agents that have good typing skills and response speed.

Attrition rates of digital agents are lower than attrition rates of voice agents so hiring practices could also change because of AI.

[5] 'A leap forward in quarterly earnings stories,' AP blog June 30, 2014

[6] 'Optimizing Each Part of a Firm Doesn't Optimize the Whole Firm' January 2016 Harvard Business Review

5. Cooperate, not Compete with Other Departments

Large companies often have competing departments, which can undermine each other's efforts and create contradictory visions. If these challenges aren't addressed, it can lead to a demotivated and siloed workforce that loses sight of the common goal and how they can contribute.

While it's difficult to accurately measure the impact of one department vs another, a company can measure output such as revenues, ROI and speed-to-market. Measuring behavior doesn't work but measuring results and considering the degree of cooperation will help departments to continually shape workflows to nurture better cooperation. Therefore, not only defining the best business goals and measuring the right things is important, but also ensuring those goals permeate through different teams, aligning incentives for different departments.

Take for example: A digital self-service sales team that competes with a chat-supported sales team. One group is trying to lock customers into the self-service model, while the other team is proactively getting consumers to talk with agents to sell cable packages. Both groups should be working together, rather than operating in silos or competing for customers – having people be integrators between both teams is essential, as well as ensuring both teams KPIs and incentives are directed at the same overarching goal.

Additionally, consider how often different groups have ownership over different channels, some own the website, another the mobile app, another social channels, another asynchronous channels, and a completely different department owns the voice stack. While there are different skills necessary for some channels, and the goals may be slightly different, customers don't ultimately care—they are contacting a company and expect consistency regardless of the channel. Ensuring that all the different teams are tying back the customer experience to a holistic goal vs. to individual channel goals is crucial.

6. Know Your Tech Stack

Many large customer service organizations are dependent on a frankenstack of technologies and tools that have been layered on top of each other—sometimes as a result of mergers, acquisitions, or years of accumulation. These digital towers of Babel were built as sources of data for rules-based programming and are often a patchwork of incompatible systems and data formats. Achieving personalization and productivity gains within the constraints of rules-based technology is a challenge, so it's critical to ensure that the technology stack is the right one for a digital transformation.

For AI to work well with a technology stack, the APIs need to be ready. Far too often bad APIs or poor API documentation is a major tripping hazard for the success of AI. The more systems you glue in the more difficult it becomes to untangle and experiment. However, when more integrations are built and established, AI-Native technology can simplify workflows and the tech stack, while caching customer details to feed routing, autosuggest and multiple parts of the conversational experience for a more impactful, personalized experience.

Not allowing silos due to organizational structure or vendor strategy to seep into the user experience is critical. As companies look towards using AI-powered solutions to users, the need to simplify the stack in order to simplify the experience is increasingly important.

7. Design AI For People

As Martin LeBlanc famously said, “A user interface is like a joke, if you have to explain it, it’s not that good.” And this becomes ever more critical with AI. Artificial intelligence alone won’t solve the challenges of cable without human agents who are key to delivering a successful customer experience. For the industry to reap the benefits of AI, companies need to think about the various users from agents through to customers—each has different goals, expectations, and metrics for success.

When AI is designed to be a part of the business process, and a thoughtfully designed product that specifically caters to its users, it makes for more productive and efficient agents in their day-to-day operations. It can make agents better at what they do more rapidly than any other technology.

AI designed for the end customer experience can provide a smooth, simple engagement that meets customers where they are—whether it’s a digital or voice interaction. The design of AI can empower cable companies to have real-time insights and reporting on every single customer interaction.

When AI is designed for agents, you have significantly more visibility into how agents are solving issues in their journey. For example, with auto pilot greetings and data inserts in autosuggest, AI can speed up and improve quality of the greetings while allowing the agent to focus on getting up to speed with the customer’s account and queries.

Customer conversations can be a mix of automation and conversation. For example, when a customer is having a service issue that can only be resolved with an onsite cable technician, AI can take over for the agent and automate the appointment scheduling, freeing up the agent to troubleshoot other more complex issues. Automation can summarize the conversation and confirm everything with the customer. If the customer ends up needing more help, the agent can seamlessly jump back into the conversation where a human touch will make the difference. AI can provide the flexibility to leverage automation without leaving the customer stuck if the conversation veers off track.

Customers and agents, like all of us, are used to using the absolute best consumer software every day such as Uber, or Facebook. That’s the competition, not other enterprise software providers. If you, as a consumer, downloaded an app that required you to switch through multiple applications and web pages to complete a payment, would you do that? Or look for the product you were attempting to purchase from another source with a more streamlined checkout flow. The interface of AI is crucial to make users feel supported in their jobs, to making your customers enjoy interacting with the company. Treat all your users like consumers, the agents, the supervisors, the data analysts-- build tools that make them want to do more work and better work, this can only be done if they are at the center of the research and development process.

8. Conclusion

With customer agents working from home, a lot of the infrastructure around management, service and quality that previously relied on in-person interaction need to adapt quite significantly. With access to peers and management limited in some instances, artificial intelligence can be a major enabling tool for employees to improve their efficiency and productivity wherever they are. When AI is designed for people, the AI models learn what works well for agents, to improve future actions and automations—and most importantly business outcomes, even in this new environment. The right AI models can help supervisors and managers adapt, allowing them to see what’s happening on an agent’s journey in real time, provide real-time coaching and support where it’s needed to deliver the best customer experience.

Designed well, AI should make an agent's job better. It can identify where agents are getting stuck, or when a customer's mood is shifting in a conversation in a good, or bad way, so that in the future AI can help move those conversations along in a positive way where the agents feel supported.

Think big. Think different and the cable industry's CX operations can lead the way artificial intelligence transforms business operations in customer experience.

Machine Learning and Proactive Network Maintenance: Transforming Today's Plant Operations

A Technical Paper prepared for SCTE by

Brady Volpe
Founder and CEO
The VolpeFirm and NimbleThis
3000 Old Alabama Rd. Suite 119-434, Alpharetta, GA 30022
+1-404-954-1233
brady.volpe@volpefirm.com :: brady.volpe@nimblethis.com

Berk Ottlik, Intern, NimbleThis LLC

0. Introduction

Proactive Network Maintenance (PNM) aims to proactively determine issues in a network so higher quality service can be provided and service impairments can be fixed before subscriber's experience issues. PNM can leverage updates in Data Over Cable Service Interface Specification (DOCSIS), an international telecommunications standard that enables high-bandwidth data transfer through existing cable television systems. The introduction of many PNM related test metrics has made it possible to pinpoint the root causes of issues in an HFC network. Full band capture (FBC) data allows operators to have visibility into all downstream RF signals anywhere DOCSIS 3.0 or 3.1 modems are deployed. This eliminates the need to bring spectrum analyzers to customer homes and perform inspection. Through PNM, downstream RF signals can be monitored 24x7x365 just using the subscriber's cable modem, which leads to better performance and impairment resolution. Issues can be identified and located faster, leading to greater cost savings and improved subscriber experience.

A challenge for operators is manually analyzing the FBC data from thousands or millions of modems. Further, the cable operator must be able to determine if RF impairments in FBC data are associated with a single home or multiple homes. When an impairment impacts a single home one can usually assume sending a technician to the individual home is the correct action. However, when multiple homes see the same impairment, sending a technician to a single home is almost always the wrong answer as the impairment is in the outside plant. In this scenario, rolling a truck to a single home for an outside plant impairment wastes time, money, extends MTR and annoys the subscriber.

This is where the power of machine learning and PNM shine. Machine learning can quickly analyze the data of thousands or millions of modems in just minutes. Then it will lead the end user to determine if there are impairments and if so, where the impairments are located.

This paper will discuss the type of RF impairments observable by PNM. Next it will discuss how machine learning is used to analyze impairments using an unsupervised model. Then it will look at how machine learning is combined with CableLab's spectral impairment detector (SID) to substantially improve on SID's impairment classifiers. Finally, the paper will look at how the author is using gamification to use feedback from end users to migrate to a supervised learning model.

Enjoy.

1. Types of Impairments

It is important to provide a brief understanding of the typical types of impairments that are generally found using FBC. Rather than using the standard impairment chart kindly produced by Larry Wolcott of Comcast, this document will demonstrate the same impairments, but with new charts. These charts are taken from live cable operator plants using a PNM application.

In the next sections, we will be focusing on the following impairments: adjacencies, suckouts, resonant peaks, rolloff, standing waves, and tilt.

1.2.1 Adjacencies

Adjacencies are essentially misalignments of radio frequency (RF) channels where adjacent channels have a large delta in channel power. This can result in packet loss, video tilting, freezing, and black screens. These impairments can often affect multiple cable modems (CMs) downstream, meaning that

clustering and especially localized clustering could find common impairments to better be able to identify the root cause of an adjacency [7].

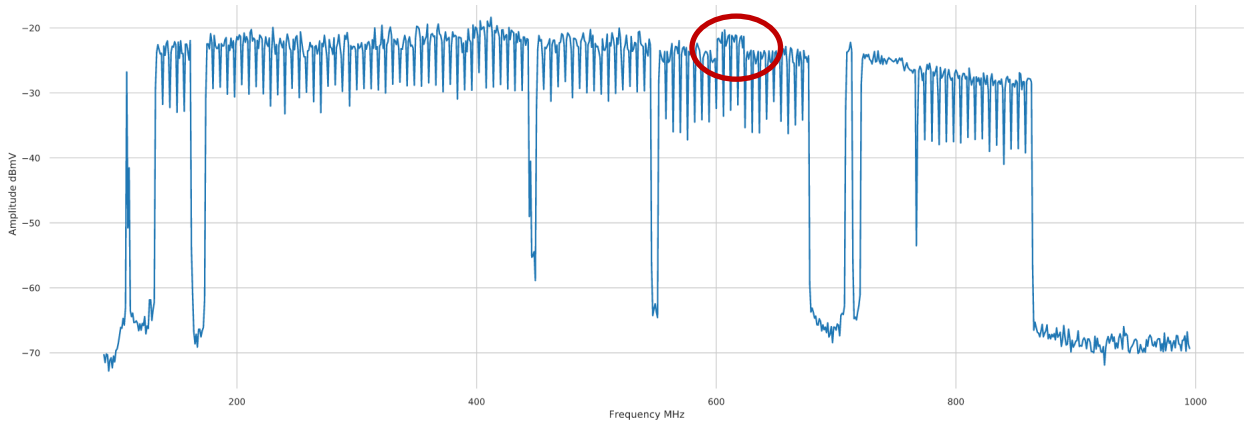


Figure 1.1 Example of Adjacency at around 600 MHz

1.2.2 Suckouts

Suckouts are another type of RF impairment that span multiple channels. They dip down to a certain depth to make a V-shape in the signal. The depth and width of these impairments determine the severity and effect and may often not have any significant impact on customer performance [7].

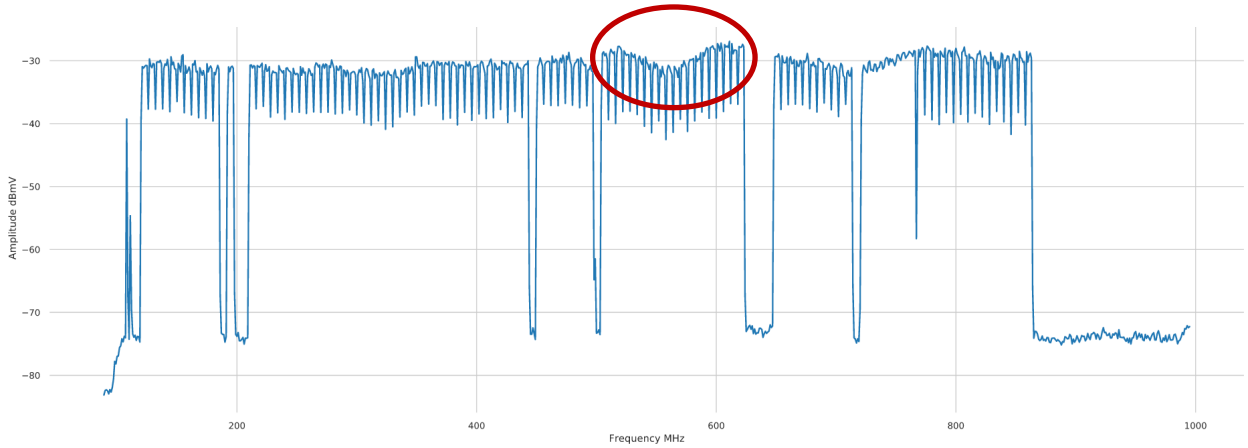


Figure 1.2 Example of a Suckout at around 550 MHz

1.2.3 Resonant Peaks

Resonant peaks are another impairment that usually spans multiple RF channels. They look like inverse suckouts, forming mountain-like peaks in the signal. They can be quite sporadic, forming and disappearing quickly, due to factors such as temperature and can have a wide range of performance impacts including packet loss, tiling, and freezing [7].

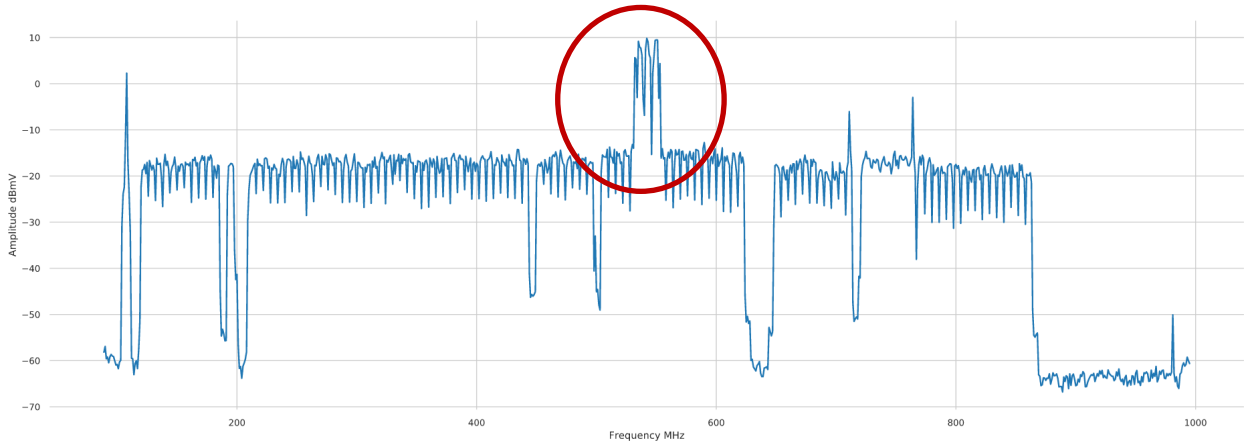


Figure 1.3 Example of a Resonant Peak at around 520 MHz

1.2.4 Roll Off

Roll off is an impairment characterized by a gradual, non-linear, exponential looking decrease in amplitude and power. Roll off can have many causes including old cables being used or individual elements in the network being configured incorrectly. It can cause freezing or tiling of video channels and is, unfortunately, one of the more common RF impairments [7].

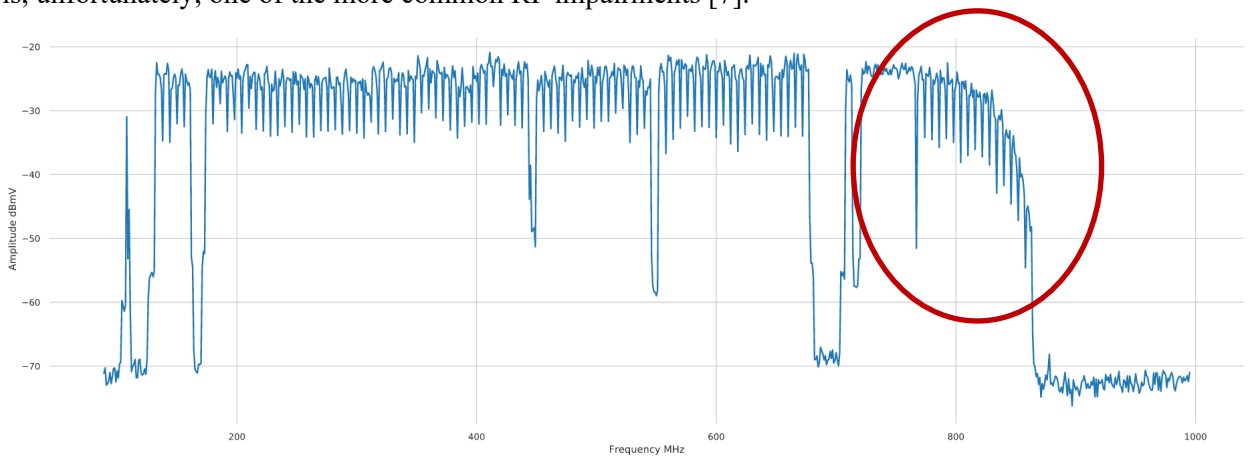


Figure 1.4 Example of Roll Off at around 820 MHz

1.2.5 Standing Waves

Standing waves are RF impairments which affect the entire spectrum. They are usually caused by an impedance mismatch in the signal and appear as waves seen at the peaks of the signals [7].

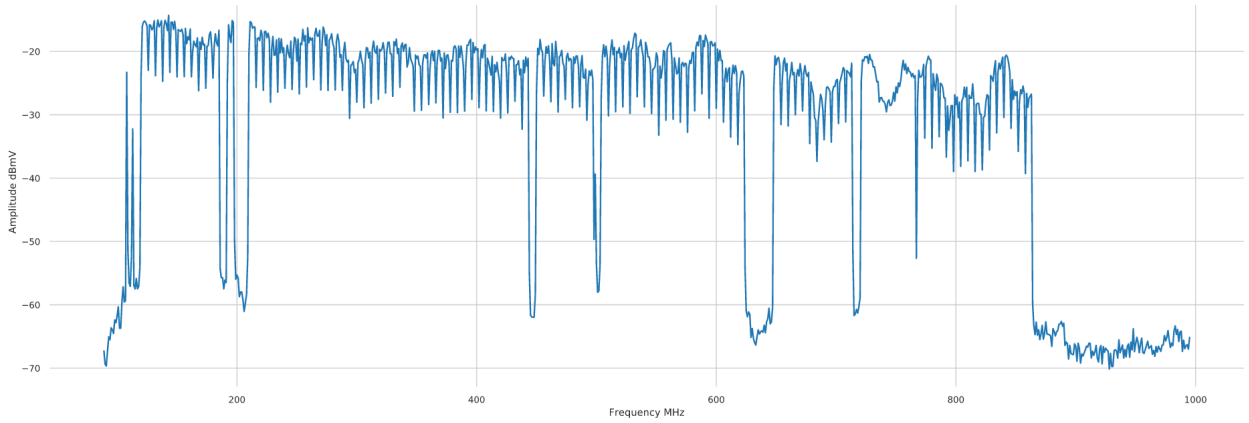


Figure 1.5 Example of a Standing Wave

1.2.6 Tilt

Finally, tilt is an impairment that is characterized by amplitude differences between higher and lower frequencies. There can be a positive or negative slope to a CM. This impairment does not always cause issues for customers [7].

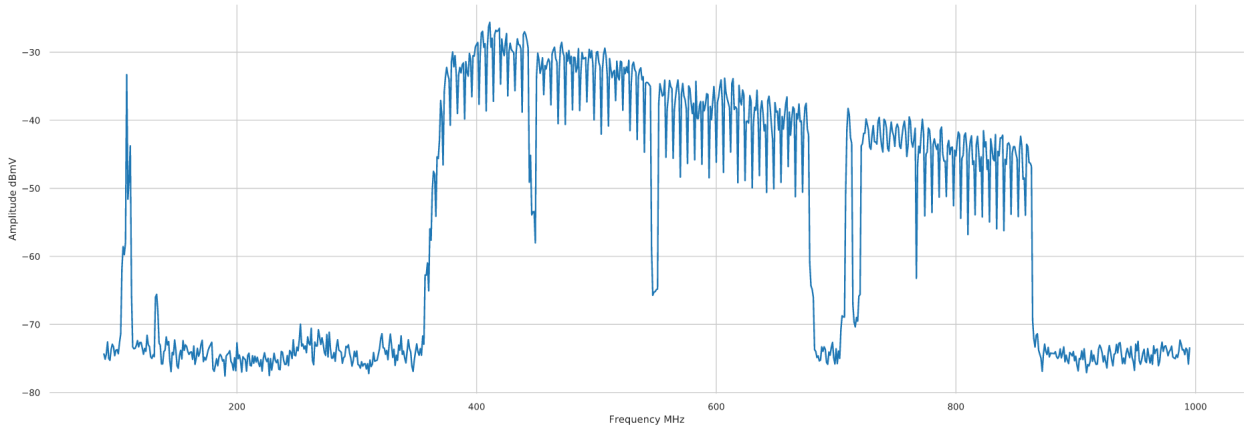


Figure 1.6 Example of Tilt

1.3 Purpose of Clustering

The purpose of clustering is to be able to find shared impairments between CMs in an automated method. This allows us to determine if an impairment affects one cable modem (CM) or if it affects multiple CMs so that cable operators can better pinpoint issues in their systems. As mentioned previously, the focus on clustering is to determine; do we roll a truck to the subscriber's home or to the outside plant. Getting this right results in immediate time and cost savings.

The objective is to find both global and local clusters of impaired modems. Global clusters are clusters where the entire signature of cable modems matches tightly while local clusters have similar signatures or

impairments in localized regions. Examples of localized clusters may be things such as a shared suckout between multiple cable modems.

1.4 Purpose of SID Overlays

The final step is to overlay CableLabs spectral impairment detection (SID) impairment labels to validate SID outputs and generalize SID predictions to more CMs. SID is software that can identify and localize common RF impairments in FBC data. It was created by the CableLabs PNM working group and has many thresholds that must be tuned to optimize performance. This does mean however that it often makes mistakes in correctly identifying impairments such as suckouts, standing waves, etc.

By overlaying SID detections on existing clusters, it is possible to validate the SID impairment labels and generalize them to the rest of the CMs in the cluster. Further, if a certain percentage of CMs in a cluster share a common SID label, then these SID overlays can be applied to both global and local clusters. Once the ML model has identified a high correlation of modems having impairments detected by the ML engine and by the SID engine, it is possible to label the impairment to the end user with a high degree of confidence.

2. Technical Approach

It is assumed that the reader has some knowledge of machine learning from previous SCTE or other papers on this topic related to ML and FBC. For this reason, topics such as unsupervised learning, supervised learning, features, and general machine learning terminology will not be covered. It will be up to the reader to review currently available documents as referenced at the end of this document.[1][5][11]

The next sections discuss how FBC data must be manipulated prior to any machine learning analysis. From a development standpoint, this is where the most work occurs in a machine learning exercise. It is often said that machine learning is easy, but it's the preparation of the data that is hard. Meaning bad data in means bad data out. To get to a good machine learning model means lots of pre-work.

While the following sections may initially appear a bit intimidating, it is important to note that the work being shown is fully automated in a PNM application. The user need not know any of the mechanics behind the machine learning. From the end user's perspective, the result is displayed data which is easily actionable because now meaningful data is being presented. Machine learning just did the hard work of sifting through piles of data for the user.

Now, let's look at the technical approach for giving the end user a meaningful experience.

2.1 Pre Processing

Preprocessing steps are necessary to manipulate FBC data to have an optimal performance with clustering.

2.1.1 Downscale Data

The raw FBC data has a varied sample rate and a varied number of data points per cable modem. The clustering algorithms that were used however require all the data to have the same dimensions. Also, it makes the code simpler and faster since it allows for NumPy arrays to be used for many operations [8]. The data is downscaled to one datapoint per integer frequency from 89 to 996 MHz. Simply put, the

corresponding amplitude for the sampled frequency closest to each integer frequency is kept and stored to end up with 907 data points per cable modem.

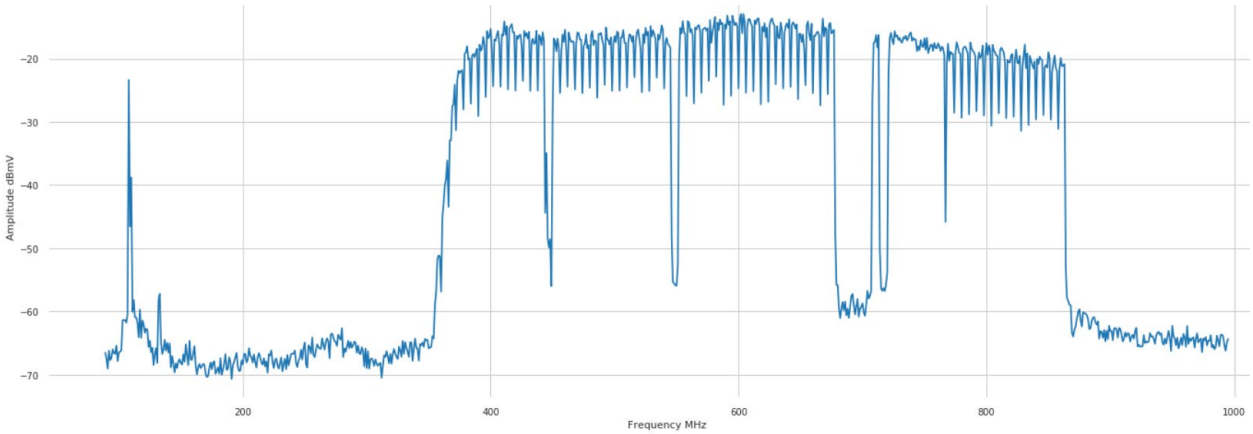


Figure 2.1 FBC Data Scaled Down

Additional downscaling was tested by using the data points closest to every other frequency from 89 to 996 MHz since most impairments did not lose any resolution and instead only some noise in the signal was eliminated.

2.1.2 Rolling Median

A rolling median is simply a median calculated for a certain window size passed over an entire signal [6]. This can remove very small and unimportant variations in the signal which would otherwise introduce unnecessary noise and produce incorrect clusters. Additionally, the advantage of a rolling median over something like a rolling mean is that it can filter out the gaps between channels. A rolling mean is not resistant to outliers meaning those dips would have a large impact on the surrounding regions. The median was calculated from the center meaning that n number of data points on the left and n number of data points on the right were used to find the median for the center point.

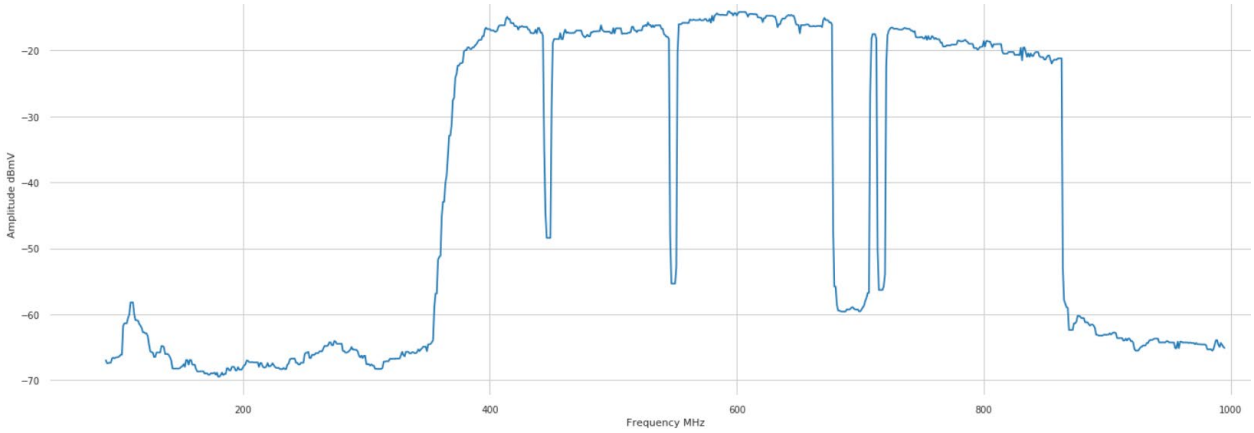


Figure 2.2 FBC Data Scaled Down with Rolling Median

2.1.3 Transforming to a New Center

When creating clusters, the power of a CM is irrelevant because impairments come from the deviations in a signal. Therefore, the signal was transformed up to 0 dBmV for the average of a certain region of the signal. The area of the signal was chosen to be 820-850 MHz because this results in transforming modems with rolloff much higher than other modems and therefore easily filtering them out (as seen with the blue CM in figure 2.3). Note that this arbitrary center frequency must be adaptive. For example, some plants may not have any signals between 820-850 MHz or there may be large impairments in this band. The transformation to a new center is used to find a clean and flat center where machine learning can be achieved. Finally, it was determined that rejecting the FM band (88-108 MHz) improved results because in nearly all cases FM ingress occurred in or near the subscriber home.

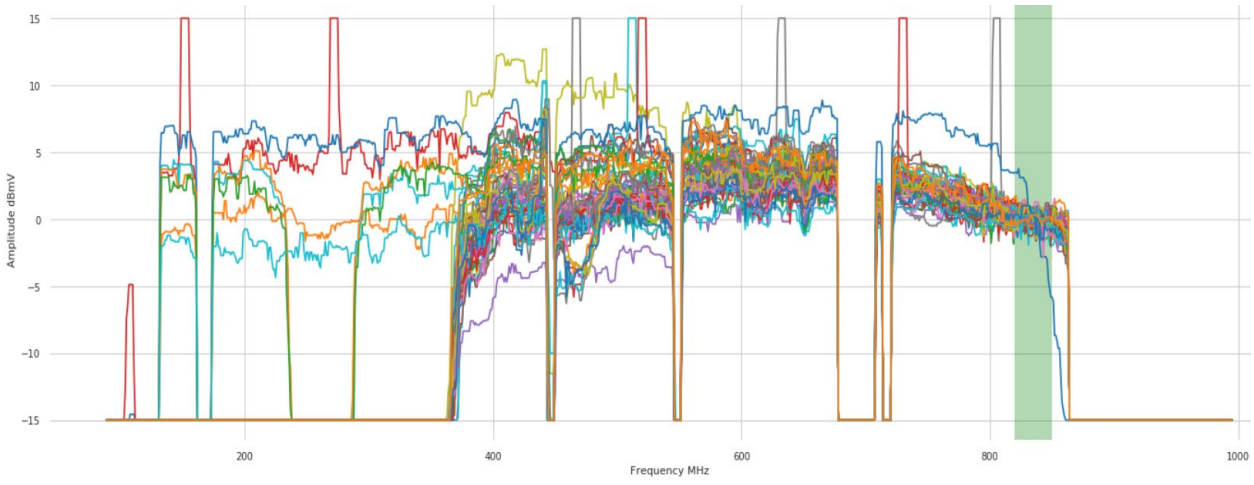


Figure 2.3 FBC Data Scaled Down with Rolling Median and Transformed and Clipped

2.1.4 Normalize Data

Data is normalized from -1 to 1 before any machine learning is performed. Since all data that is clustered on is clipped before clustering, all the data is simply divided by the clipping amplitude maintaining the

same shape of the data but reducing the amplitude range to -1 to 1. The purpose of normalization is to have a standard range of the data so that any predetermined parameters work optimally [5]. Data normalization is a very common machine learning practice.

2.2 Modem Health Classifier

The first step in clustering aims to classify each modem in a node as impaired or not impaired. These classifications are done in an unsupervised manner as no accurate ground truths currently exist on which to create a supervised learning model. For this system to work properly, an assumption is made that the majority of CMs in a node are healthy. This is a bold assumption, but one which must be taken until 1) proven otherwise by inspection or 2) a supervised machine learning model is available. Note that as discussed later in this document, by using SID data, it is possible to overcome 1) above with the use of SID data. This is because SID data will provide the necessary information that all modems in a node have some type of impairment. Once a cluster is created, if all modems in the "healthy" cluster are shown to have SID impairments, then by inspection of the SID data it can be determined that the cluster is not healthy.

2.2.1 Region Identification

The first step of classifying modems as healthy or not was to identify the sections of RF spectrum that occupied by video or QAM channels. For the purpose of this document, occupied spectrum are named regions. Identifying regions is a necessary step because otherwise, classification could be made on modems with varying sections of used and unused spectrum which would be more quickly classified as an outlier and impaired than any modem which has an actual impairment. In Figure 2.5, an example of unused spectrum is the spectrum below about 375 MHz, where the blue line drops to -1.0. For someone familiar with the industry, it is apparent that a data only filter is in use in Figure 2.5. The data only trap causes all signals below 375 MHz to be attenuated. The signals that are still present below 200 MHz are FM ingress signals (88-108 MHz). As previously mentioned, the FM band is omitted from machine learning in the current model, so these low frequency signals will be ignored.

The process starts by roughly identifying all the used frequencies of every modem in a node (see figure 2.5 and 2.6).

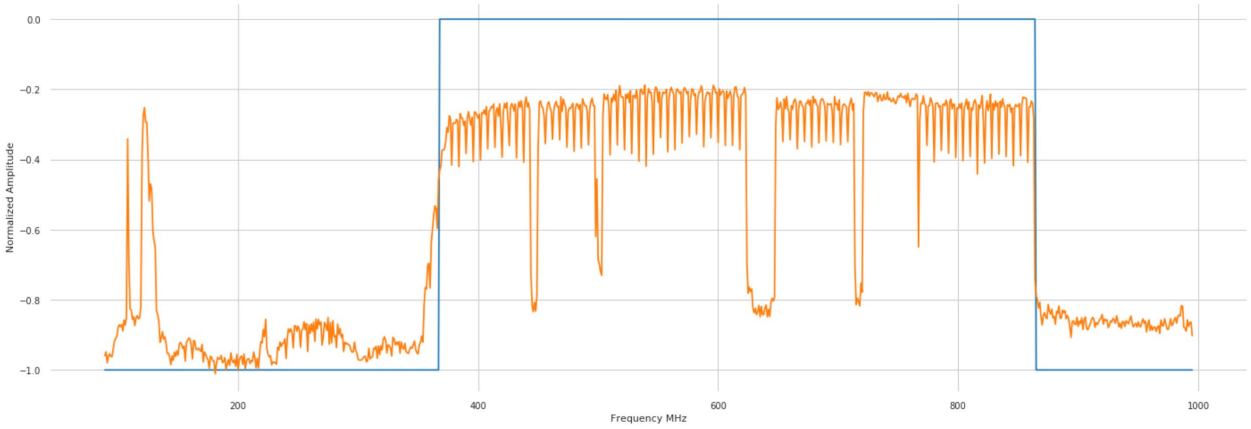


Figure 2.5 Used Frequencies Identified on Modem

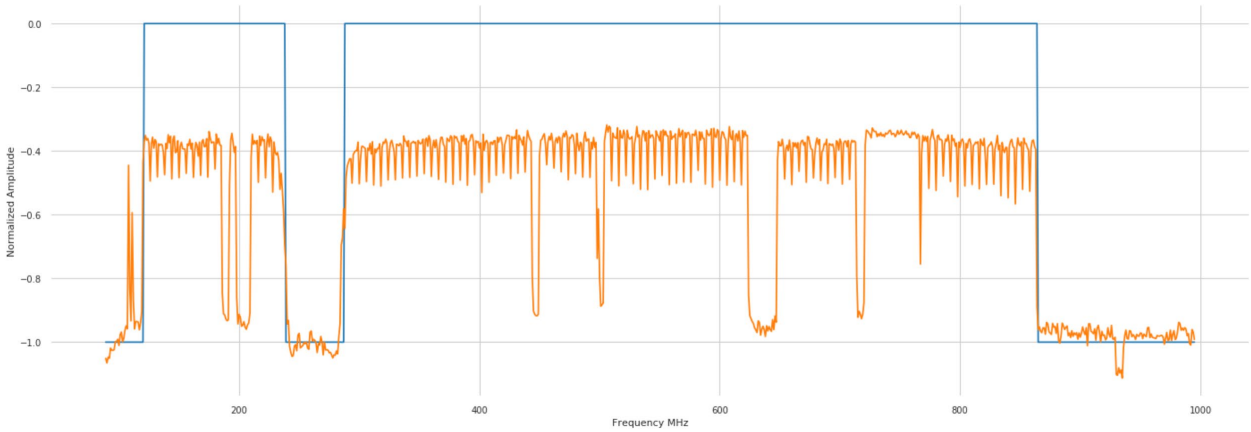


Figure 2.6 Used Frequencies Identified on Modem with Band Stop Filter

In Figure 2.6, it is evident that the pre-processing can identify unused frequencies, but this time it is not due to the presence of a data only filter. In this case, there are channels missing between 275-300 MHz. The blue line indicates pre-processing is eliminating these frequencies from the model. While there are unused frequencies higher in the band around 625 MHz, these do not the minimum width for the classification engine to exclude the band.

The used spectrum is found by first passing a rolling median with a large window size over the normalized, transformed, and clipped signals from section 2.1.3 (see figure 2.7). Then, at any frequencies where the normalized amplitude is greater than -1, those regions are labeled as used (blue line at 0) and any frequencies where the normalized amplitude is -1, those regions are labeled as unused (blue line at -1).

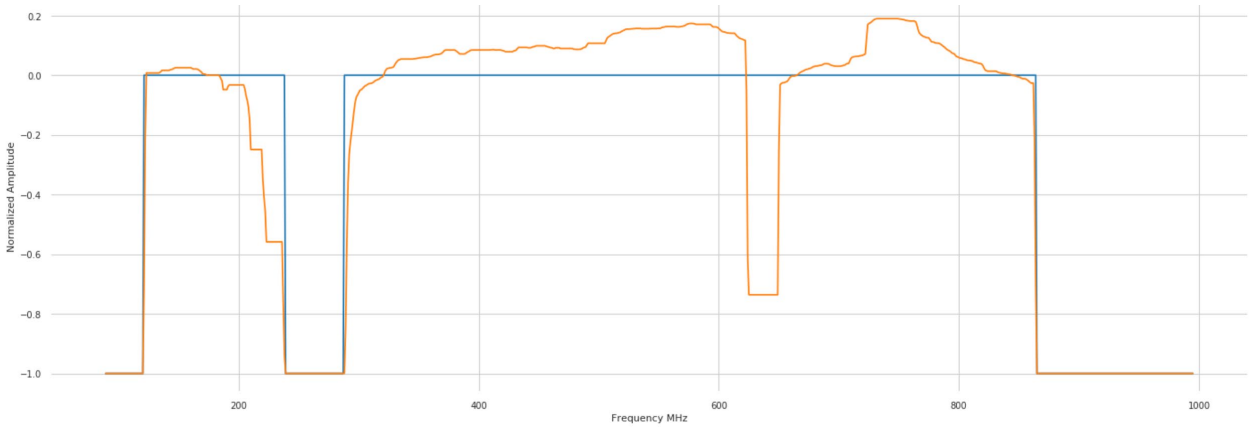


Figure 2.7 Used Frequencies Identified on Smoothed Modem with Band Stop Filter

Once all the modems have their used regions extracted, regions can be extracted representing the various regions of commonly used frequencies between modems. This works by first identifying the regions with the most used spectrum of all modems (see figure 2.8 for used spectrum of all modems). In the case of this node, that region is from around 380-850 MHz. Then, health classification is done on this region as described in region 2.2.2. Following this, all modems will be identified as impaired or not impaired for that region. This region is then removed as being a used spectrum from all modems which have used spectrum in that region. Following this, the process repeats for the next region with the highest number of commonly used frequencies by classifying and then removing that region from the used regions as well. This process is repeated until there are no more used regions in the spectrum. At this point all modems have been classified as impaired or not impaired based on regions where modems have the same used and/or unused spectrum.

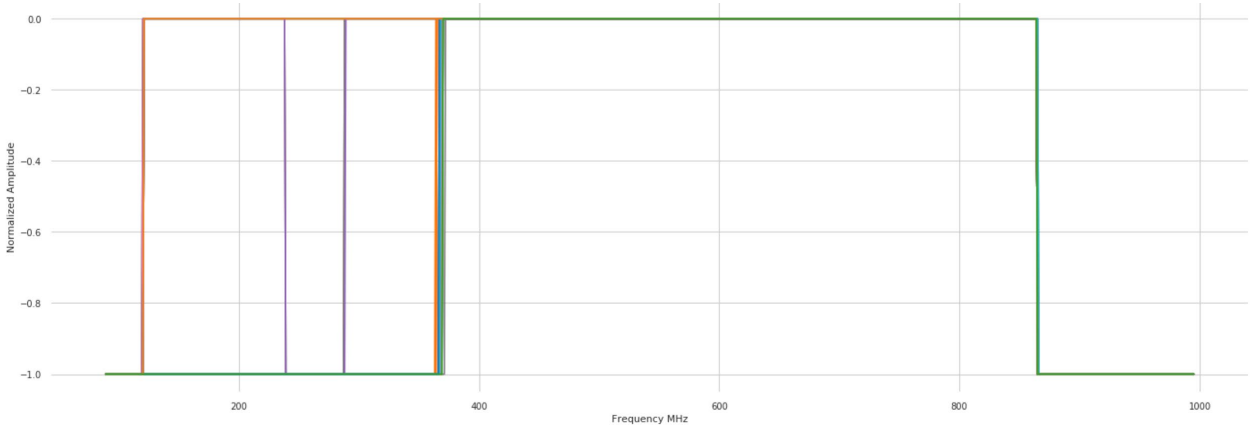


Figure 2.8 Overlapped Used Frequency Regions

2.2.2 Health Classification on Extracted Regions using Local Outlier Factor

Once common regions of used frequencies are extracted, outlier detection is done on modems which have the regions. This outlier detection utilized the local outlier factor (LOF). LOF is an unsupervised (well, semi-supervised) machine learning algorithm that uses the density of data points in the distribution as a key factor to detect outliers, LOF roughly works by calculating a standardized distance to n number of

neighbors and labeling data points with a large distance as outliers [1]. In turn, this can filter out modems which deviate from the rest of the modems due to impairments.

2.3 Global Clustering

The next part of the clustering is to cluster together the impaired modems (outliers from clustering) to determine if there are any similar global impairments between cable modems. Global clustering requires that signatures of CMs match throughout the entire frequency range, meaning that it cannot cluster together smaller local clusters, but it can find modems with common signatures.

Global clustering is useful in finding modems that have large impairments such as standing waves or tilt, which impact the entire spectrum.

2.3.1 DBSCAN Clustering

In machine learning there are many algorithms for grouping or clustering common sets of data together. One of the most used is K-Means. K-Means clustering works very well, however it is non-optimal for noisy data, such as FBC data. The current implementation in this paper is using a model called DBSCAN.

DBSCAN stands for density-based spatial clustering of applications with noise. The algorithm works by first selecting a random point in the data. Then, if there are minimum points (a specified parameter) number of data points within the radius of Epsilon (EPS, a specified parameter) distance or the Euclidean distance, straight-line distance to the original point, it is labeled as a cluster. Then this process repeats for every point that was in the original cluster, if the points have at least minimum points number of points within their EPS distance, then the points are labeled as core points. However, if a data point does not have the minimum points number of data points, it is labeled an outlier, unless it is within the EPS of a core point. If there are no more data points nearby, then a new random point is chosen until all the data has been clustered [3].

2.3.1 Global Impairment Clustering

Global impairment clustering was implemented purely using DBSCAN. The minimum points parameter is lowered to two to allow for very small clusters and the EPS is slightly lowered also to ensure clusters are tight. Figure 2.8 demonstrates the use of global clustering but also highlights some of the weaknesses.

Global clustering can cluster these modems and conclude that they are all impaired and have both a common impairment and rest of signature. What the impairment is or where the impairment is located cannot be determined purely using global impairment clustering (this is where SID overlays are needed).

The issues that arise from this approach is that if that same suckout were to be seen around 470 MHz on a different modem with used spectrum below 350 MHz, then they would not be clustered together. This is when local clustering is effective.

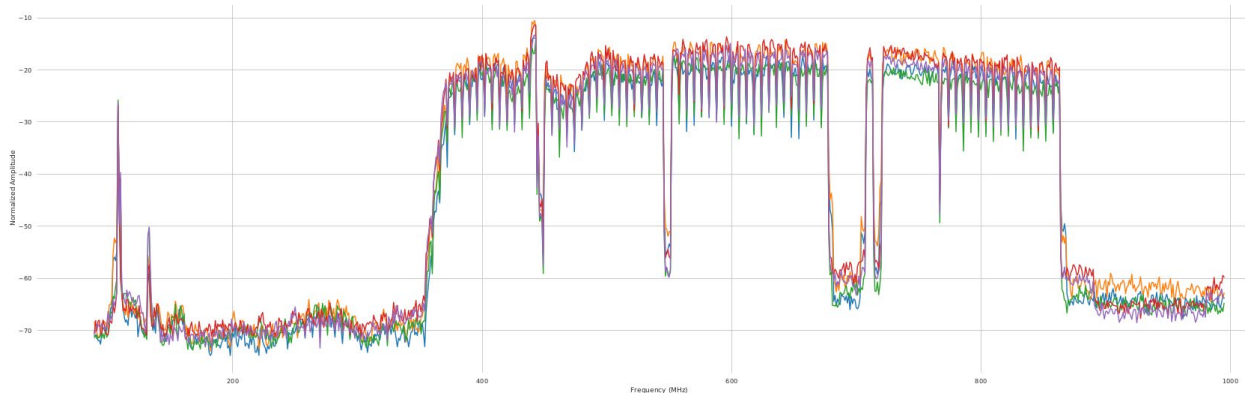


Figure 2.9 Global Cluster on Impaired Modems

2.3.2 Global Impairment SID Overlay

SID labels are overlain on global clusters to determine the location and type of impairments seen in a cluster. If a certain percentage of modems have an impairment, then that impairment can be generalized to the rest of the modems. Additionally, if that impairment is found in a local region for multiple CMs, then it can be generalized that those modems all have that impairment at that specific location. This is beneficial for two reasons:

1. The data being used is unsupervised data, which means it is not known if the FBC data is impaired or not impaired, but overlaying SID data, it is now possible to determine not only if the FBC data is impaired, but also the type of impairment (i.e., suckout, standing wave, etc.).
2. SID impairments are often inaccurate. For example, suckouts and adjacencies are often mis-labeled by the SID engine. By classifying many modems with the same SID overlay, it is possible to improve the accuracy of SID classifications through scale. If many modems show the same SID impairment at the same frequency, then the probability that SID is accurate is high.

SID overlays also allow the algorithm to discard SID impairments that are rarely found in the cluster. In figure 2.10 it is apparent that only 20% of the CMs in the cluster were labeled as having a resonant peak by SID. From this one can conclude that the modems in the cluster most likely do not have a resonant peak at around 620 MHz.

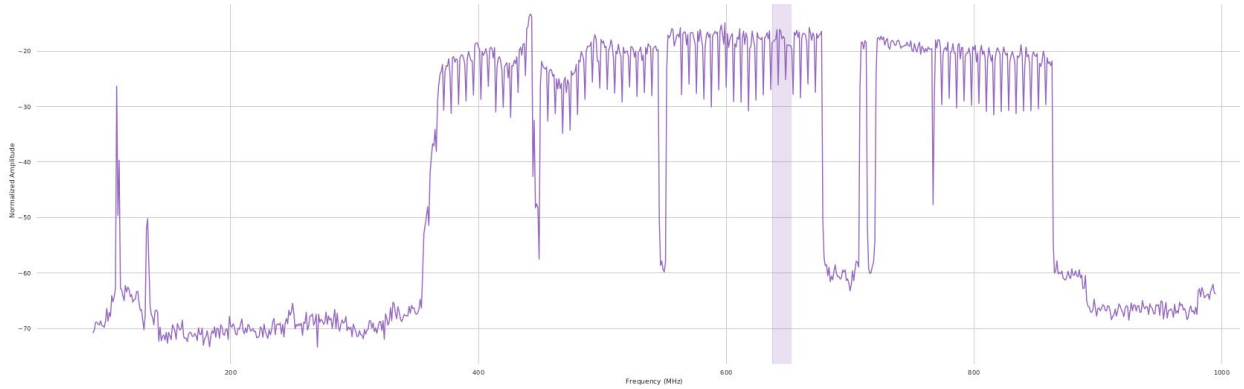


Figure 2.10 Resonant Peak Identified at 20% in Global Cluster shown by vertical purple bar

On the other hand, SID labels can be verified based on the percentages that they occur at and if the location that they occur at overlap. In Figure 2.11, it is visible that 100% of modems are labeled as having a standing wave from around 100-900 MHz. Therefore, one can accept this label as being most likely true. This can also apply to other labels which are not seen in 100% of modems however such as the adjacency labeled in 40% of modems for this cluster around 650 MHz. This can be generalized to all modems in the cluster if the threshold of the percentage of modems in a cluster that need to have a SID label for a certain region is met.

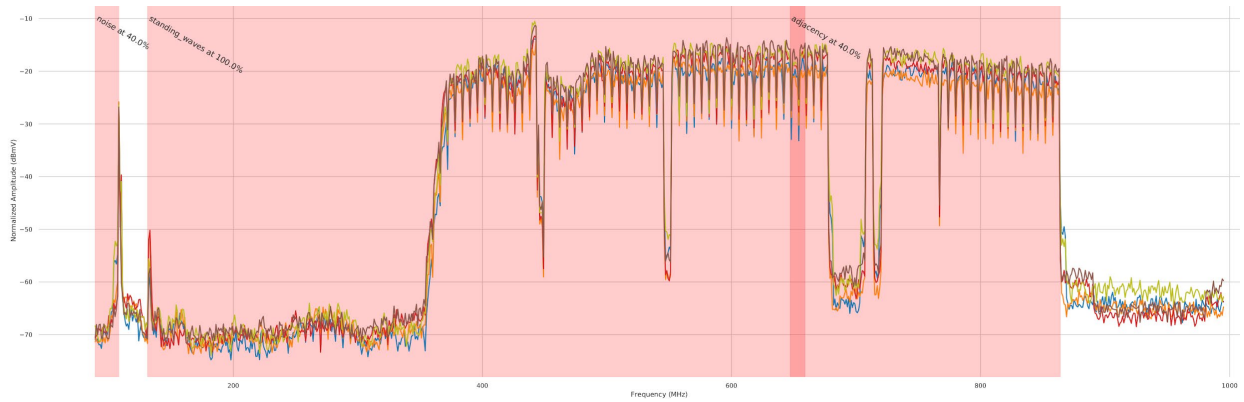


Figure 2.11 SID overlay with 40% Threshold

2.4 Local Clustering

To cluster together similar local impairments, different clustering techniques need to be used which only look at local regions. This is because CMs may share one common impairment, while not sharing a whole separate range of impairments and signatures which result in them being placed in different global clusters. Again, this is very important for narrow impairments such as suckouts and adjacencies.

2.4.1 Local Impairment Clustering

To find local impairments, clustering was done using DBSCAN on a certain window throughout the spectrum. This window slid over the entire FBC spectrum with a certain step size and clustered together

all the modems in the node using DBSCAN [10]. The FBC spectrum was preprocessed the same way as with variability-based outlier removal (section 2.3.2) so that the FBC data was straightened and centered at 0. This was important because DBSCAN would not work properly unless the data was in the same shape and located at the same amplitude because that is the only way for the Euclidean distance between sections of the FBC data to be small and therefore clustered together.

Then if one of the clusters formed only contains impaired modems, it is a localized impaired cluster in that region. If a cluster contains modems that are not impaired, it means that the cluster found similarities that are not an impairment, and the cluster is therefore not considered a localized impairment cluster. Additionally, if a cluster contains more than a certain percentage of the modems in a node, it was removed from the local impairment clusters because it most likely is clustering on something that is not an impairment.

In figure 2.11 we can see that the localized impairment clustering was able to find local clusters due to the preprocessing steps taken. Preprocessing manipulated the data to be in the shape seen in figure 2.12. Here DBSCAN can easily find clusters in certain windows even though the orange modem has many used channels under 400 MHz that the blue modem doesn't and that the blue CM has tilt but the orange one doesn't. This resulted in a important breakthrough, which was the ability to detect similar impairments on modems with radically different spectrum usage due to bandpass and band stop filters.

Figure 2.12 Localized Impairment Clusters on Impairment Modems

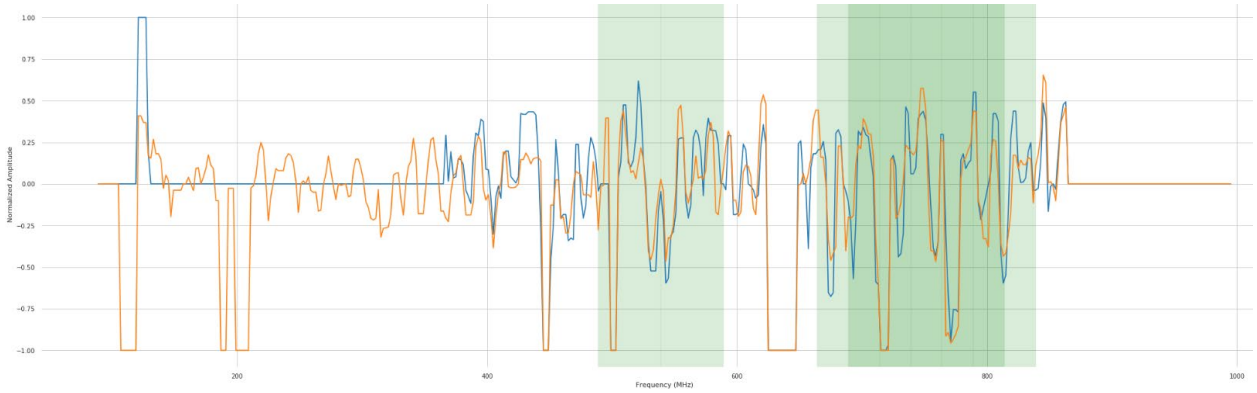


Figure 2.13 Localized Impairment Clusters on Impairment Modems Pre-Processed

2.4.2 Local Impairment SID Overlay

The last step with the localized clustering was to overlay the SID impairment labels for localized impairments (i.e., suckouts and adjacencies). This allowed us to draw conclusions about the accuracy of the SID labels and see if the labels intersected with the local cluster regions. In all the figures below, the green highlights indicate local cluster regions while the red indicates the global cluster regions.

In figure 2.13, we can see that SID labeled that both modems in the same local cluster had standing waves and that the labels intersected with regions of local clusters, meaning that we can conclude that both these modems have standing waves in the regions of overlap between the SID label and local cluster regions. We cannot generalize and say that both modems have standing waves on all regions from 100-850 MHz however because since this is local clustering, the regions could have nothing to do with each other.

As seen in figure 2.13, local impairment clustering can also be used to generalize impairments that are not detected in all modems such as the resonant peak around 530 MHz. Even though only 50% of the modems in the cluster had this label, it can be generalized to apply to both modems because it falls under a local clustering region. Figure 2.14 contains more visuals of a different local impairment cluster.

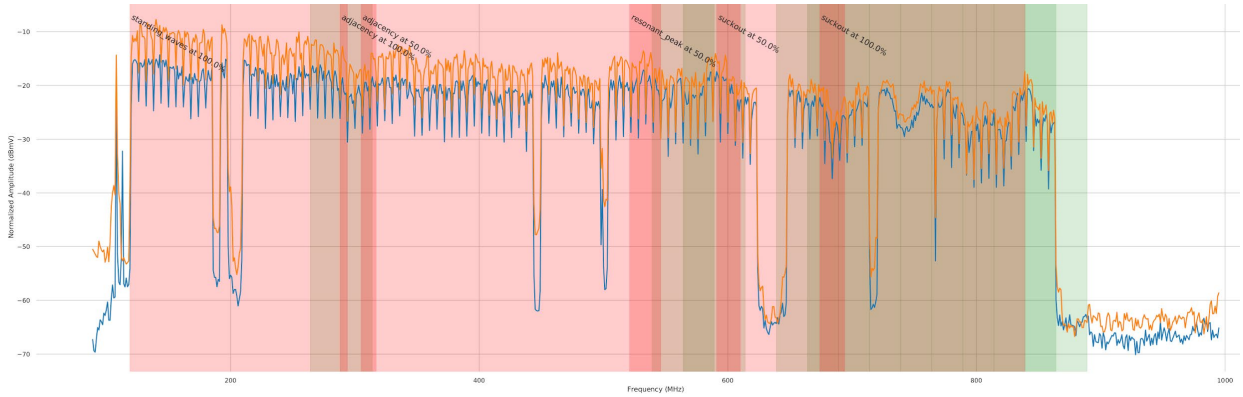


Figure 2.14 SID Overlay on a Local Cluster with Threshold of 40%

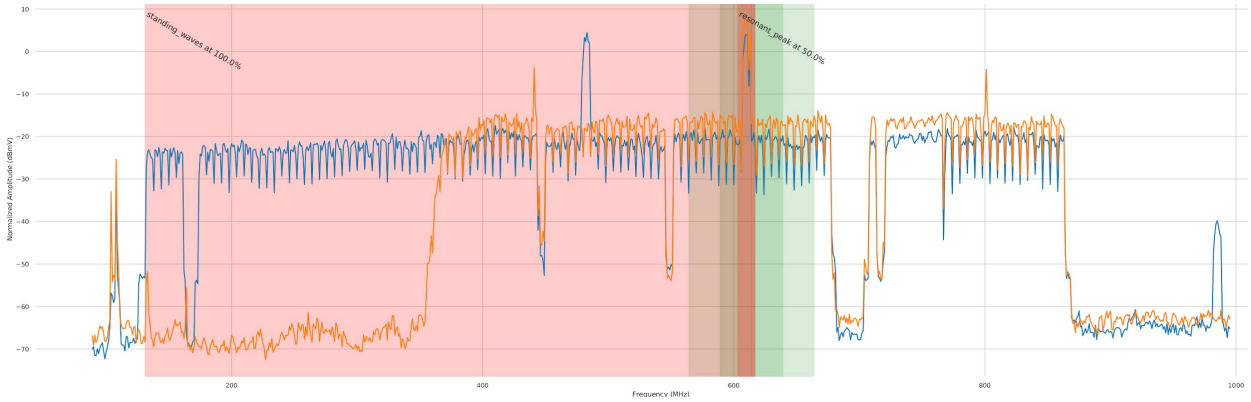


Figure 2.15 Resonant Peak SID Overlay

3. Results and Discussion

This section will analyze some of the performance, nuances, optimizes and observations identified during the use of machine with full band capture data. As this technology is continuously being improved upon but not only the author of this paper, but others in the industry, it is the hope that some of the findings in this paper will be used by others to build upon this and a collaboratively shared for the betterment of the industry.

3.1 Modem Health Classifier

3.1.1 Pre Processing Performance

The preprocessing steps were vital to the clustering. By clipping and centering the signals, small local impairments such as resonant peaks and suckouts become much more influential in the data and are therefore easily labeled as outliers only using DBSCAN.

Preprocessing does occasionally run into issues, however. Occasionally, minor impairments such as the red spike around 600 MHz in figure 3.1 are run over in preprocessing as seen in figure 3.2. This is most likely a result of the spike being very thin and the rolling median, therefore, discarding it. The spike is most likely the result of RF interference that could impact customers and was missed.

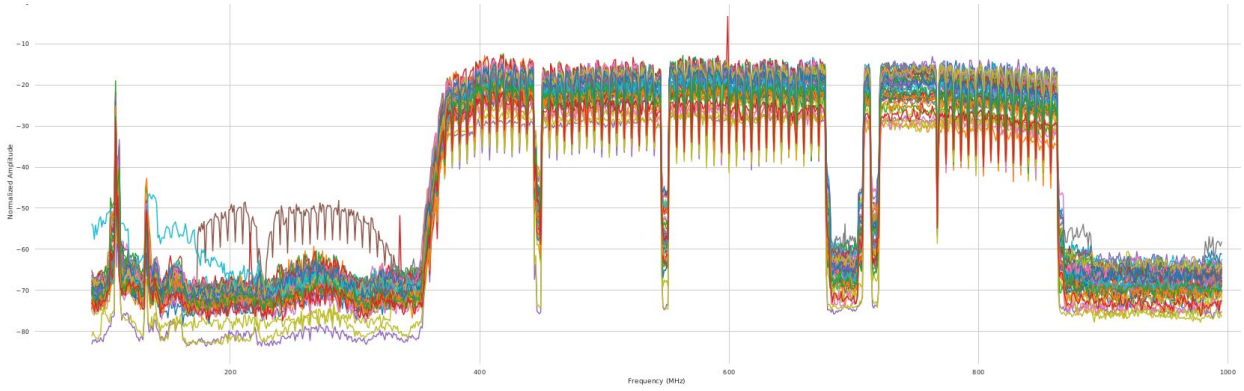


Figure 3.1 Raw FBC Cluster

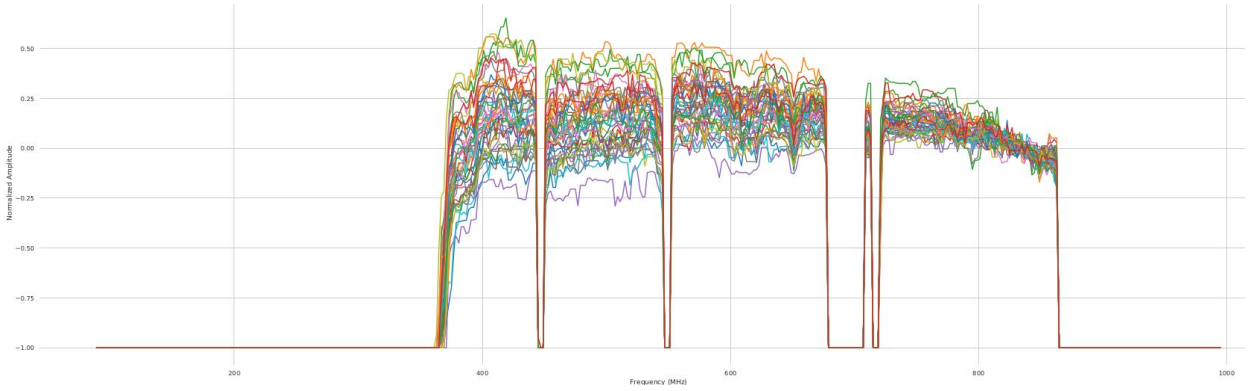


Figure 3.2 Pre-Processed FBC Cluster

3.1.2 LOF Performance

LOF proved to be optimal for this use case because it accurately and efficiently was able to identify outliers with impairments. Finding the optimal number of neighbors was a difficult task and will continue to be a difficult task when applying this software to various CMTSs and cable operators. Figure 2.14 shows LOF finding outliers on the highlighted regions while figure 2.15 shows the modems not labeled as having outliers.

Figure 3.3 shows a cluster of modems which has been continuously showing up across cable operator systems since this algorithm has deployed. Notice the spikes appearing roughly 20 dB above nominal RF spectrum. When examined more closely, these spikes are roughly 8 MHz in bandwidth. The author of this paper has visited several subscriber locations where the signal is present. The signal was not observable using traditional swept-spectrum analyzers. Further, once the subscriber modem was replaced the modem was replaced, the signals were no longer present. Suspect modems were collected and are under current evaluation.

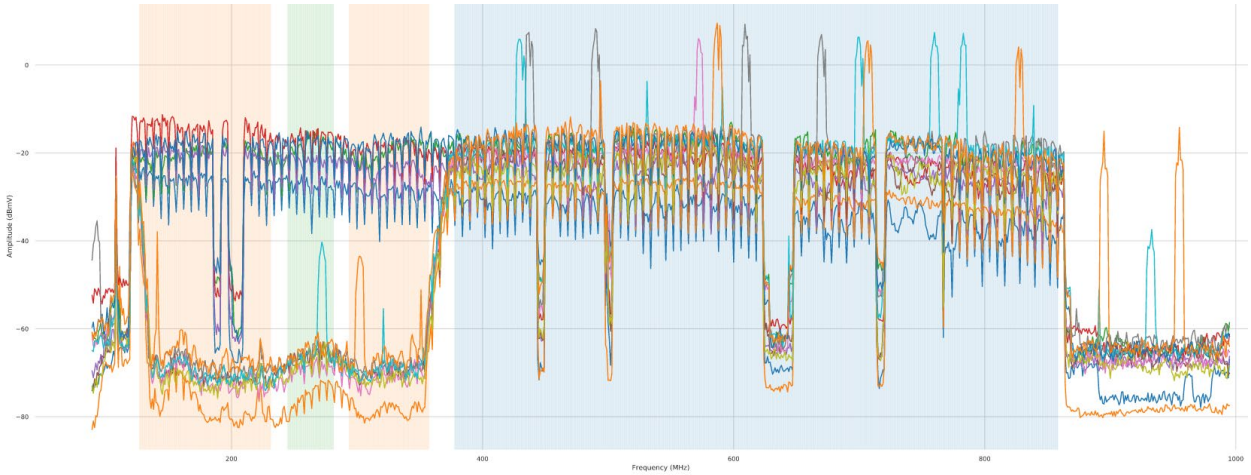


Figure 3.3 Impaired Modems with Used Regions Indicated

Figure 3.4 shows what is expected for a healthy cluster of modems on the same node. Here we see many cable modems not showing the spike in figure 3.3. The point of this example is that although root cause is not yet determined, machine learning combined with PNM was successful in identifying anonymous activity in modems which were customer impacting. Each customer with spurious activity had open tickets for downstream video or DOCSIS issues which where un-resolved.

Figure 3.4 Healthy Modems with Used Regions Indicated

3.1.3 Fixes for Band Stop Filters

Band stop filters filter out ranges of frequencies in a signal [2]. They are used by cable operators to restrict certain channels from customers. In the data we used for this paper, band stop filters were occasionally found in CMs as seen in figure 3.3. By using the different regions with commonly used frequencies, modems with Band Stop filters were not automatically classified as impaired, but instead accurately classified based on the rest of the used frequencies and how they related to the rest.

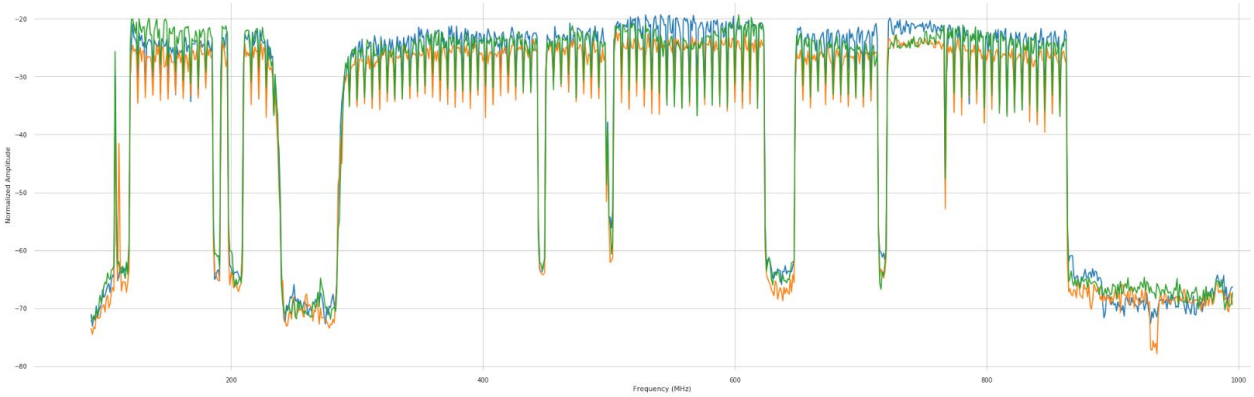


Figure 3.5 Modems with Band Stop Filters

3.2 Global Clustering

Global clustering was effective for its purpose of finding impaired modems with similar global signatures.

3.2.1 Global Clustering Shortcomings

One issue seen with both global and local clustering is when validating SID impairments when there are few modems in a cluster. If there are only 2 CMs in a cluster and one has an incorrect SID label, if the threshold for the correct SID label to be generalized is 50% then the incorrect impairment is accepted. However, if the threshold is at over 50%, correct SID labels may also be overlooked.

3.3 Local Clustering

Local clustering was effective in finding local regions with similar signatures. Occasionally regions that just happen to only be found in impaired CMs are labeled as local clusters when they are completely healthy regions of the spectrum that just have signatures not found in the rest of the healthy CMs.

3.3.1 Local Clustering Shortcomings

Shortcomings include times local clusters were identified but no SID labels were to be found in those regions and when common SID labels were found in regions and no local clusters were identified. Using both the clusters and SID labels, however, allows one to build greater confidence in the SID labels even if the system is not 100% accurate.

Local clustering also has the same issue when there are few cable modems in a cluster as seen with global clustering in section 3.2.1.

3.4 Optimization of Parameters

One factor seen across the board is that there are many parameters with this approach. Optimizing parameters also takes lots of processing and lots of time. Every cable operator with different hardware and different severities may need different parameters and perhaps even different CMTSs. There is no numerical way to find the optimal parameters as this is an unsupervised setting with no ground truths, meaning that someone must look through as much data as possible and look through many variations of parameters to find which fits the given data the best while not overfitting.

There are also many parameters that impact each other such as the window size parameter for local clustering. If the window size is changed, the EPS parameter in DBSCAN must also be changed to adjust. This makes optimization even more complex.

For this reason, real-time optimization parameters are added to the application which enable the end user to modify DBSCSAN parameters. These optimizations on performed on the cable operators' network to ensure DBSCAN is optimized based on the system's channel lineup.

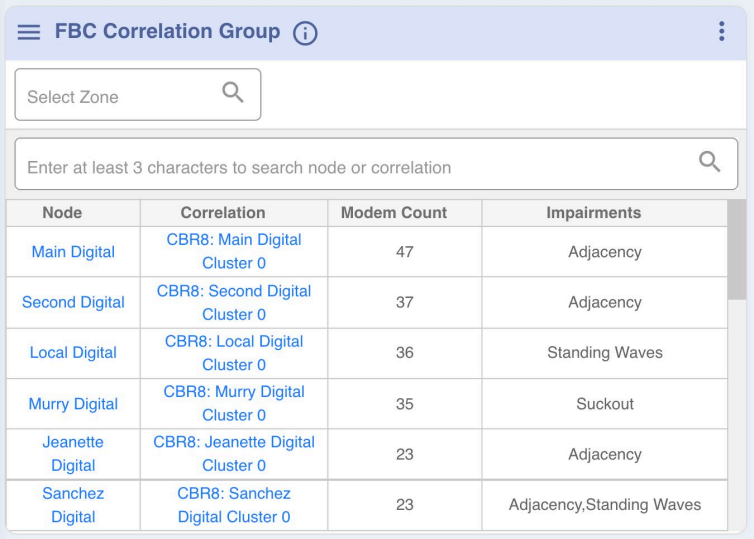
4. Conclusion

This project is effectively able to identify common impairments between CMs and is also a step towards replacing SID with a more intelligent system. Extensive preprocessing and clustering on FBC signatures were able to reveal shared impairments between CMs in a node. Additionally, SID labels were then overlaid to clusters to verify the accuracy of SID labels. This model was further modified and applied to RxMER data of OFDM channels in DOCSIS 3.1 downstreams. This had value in identifying outside plant impairments impacting multiple subscribers with DOCSIS 3.1 modems.

4.0 Operationalizing the Plant Maintenance

PNM and machine learning begin to show their complimentary value when it comes to operationalizing plant maintenance. Before applying machine learning, it would be up to the end user to manually view many fullband capture images and attempt to make mental correlations. This was a tedious process and relied on human to first do the work and second be effective at doing the job. The job being to determine if an impairment was impacting one home or many. Machine learning will automate this task by automatically clustering FBC data. It is up to the application programmer to make the data available to the end user.

One example can be seen in Figure 4.1 where a widget is made available to the end user with a list of FBC correlation groups (i.e. cluster groups), the node each correlation group, the number of subscribers impacted by the impairment and the impairment type. Clicking on the blue hyperlink takes the user to a visual representation of the impairments (shown in Figure 4.2) on a map, where action can be taken.



Node	Correlation	Modem Count	Impairments
Main Digital	CBR8: Main Digital Cluster 0	47	Adjacency
Second Digital	CBR8: Second Digital Cluster 0	37	Adjacency
Local Digital	CBR8: Local Digital Cluster 0	36	Standing Waves
Murry Digital	CBR8: Murry Digital Cluster 0	35	Suckout
Jeanette Digital	CBR8: Jeanette Digital Cluster 0	23	Adjacency
Sanchez Digital	CBR8: Sanchez Digital Cluster 0	23	Adjacency,Standing Waves

Figure 4.1: FBC Correlation Widget Groups by Node, Modem Count, and Impairment Type
(Image Courtesy: NimbleThis)

Figure 4.2 the machine learning-based results of FBC correlated modems as plotted on a map. The blue modems on the map are associated with the FBC data on the right-hand side. A user may select a modem on the left, a MAC address on the right, or a trace bottom right to interact with the data. The actionable data for the end user is that this section of coax plant has a system-wide standing wave. Fixing the standing wave by visiting a subscriber's home is not a good choice. This is an outside plant problem which must be addressed as such.

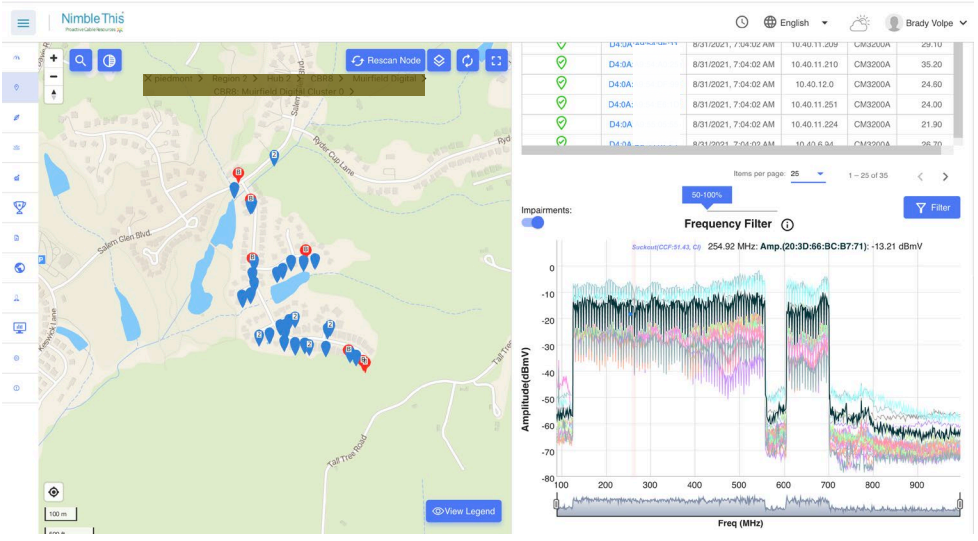


Figure 4.2: Representation of FBC correlation group on map (left) with respective FBC impairments (right)
Image courtesy: NimbleThis

As indicated, this same algorithm is easily adapted for RxMER data in DOCSIS 3.1 OFDM channels. As with FBC data, it is useful to present data at a high-level first as a widget, shown in Figure 4.3.

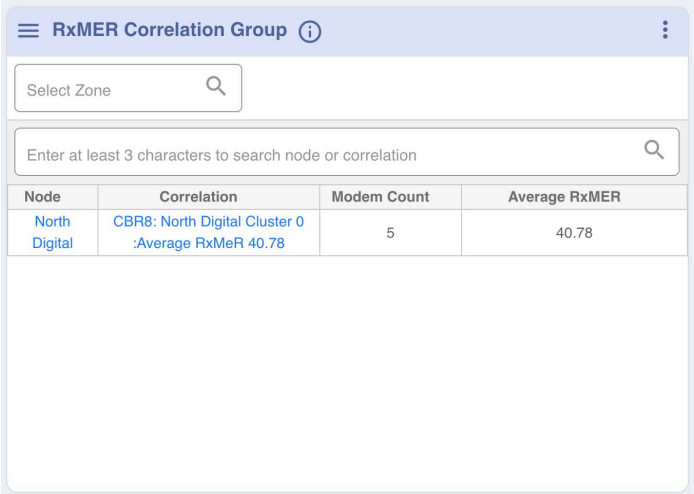


Figure 4.3: RxMER Correlation Widget Groups by Node, Modem Count, and Average RxMER

(Image Courtesy: NimbleThis)

The operational value of the RxMER correlation widget is that an end user can quickly identify clusters with low RxMER. Clusters with low RxMER will operate at a low OFDM modulation resulting in low or no data speed to subscribers. The low RxMER in a cluster is a result of outside plant impairments, so these can be addressed by outside plant techs.

Figure 4.4 shows the mapping of the clustered data. This view results when clicking on the correlation group in the widget of Figure 4.3.

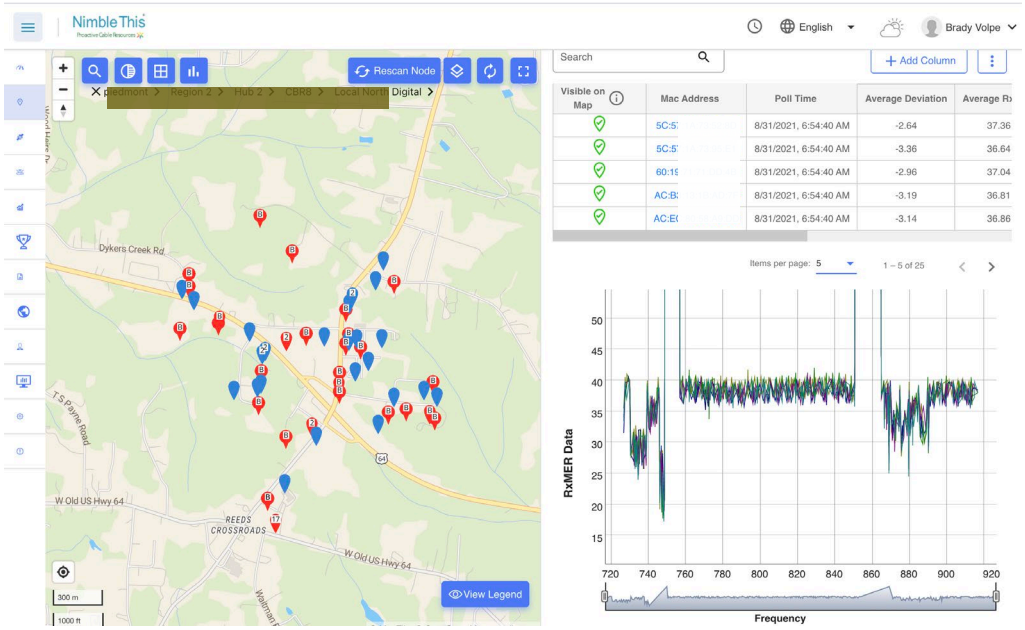


Figure 4.4: RxMER Clustered data on map (left) with clustered RxMER data (right)
(Image Courtesy: NimbleThis)

Figure 4.4 shows the actual RxMER clustered data on the right-hand side. As can be observed, there are many locations in the RxMER data where the MER drops below 35 dB. An ideal OFDM channel would have its RxMER data above 39 dB across every data point to support 4096-QAM.

On the left-hand side of Figure 4.4, the blue modems indicate which DOCSIS 3.1 modems are part of the cluster group and are impacted by low RxMER. In many networks today there is not 100% penetration of DOCSIS 3.1 modems, so it is quite common to have many DOCSIS 3.0 and DOCSIS 2.0 modems that are around the cluster but are not impacted because they do not use the OFDM channel.

Again, the value of Figure 4.4 is that a technician can quickly observe that the downstream impairments in the OFDM channel are common to every subscriber. Visiting an individual subscriber home will not fix the impairment. This is an outside plant problem which must be addressed in the outside plant.

4.1 Future Research

4.1.1 Supervised Learning

The current implementation, while very powerful, uses unsupervised machine learning. This means the ML engine has no knowledge if the FBC data is impaired or not impaired nor does it know the impairment type once it is classified as impaired. The classification comes from SID data, which is only somewhat accurate. The next level is to achieve supervised learning, which is a machine learning engine whereby the engine already has knowledge about FBC impairment types. This requires a lot of work from end users to label these existing impairments and build a database through which the ML engine can be trained.

There are two approaches in process to do this. First, a built-in gaming feature that encourages its users to label impairments when fixing problems has been applied. Fix a problem and label an impairment and get points. Users with the most points get on the leader board. This is used by cable operators as an incentive program for their technicians. An example of this is shown in Figure 4.5.

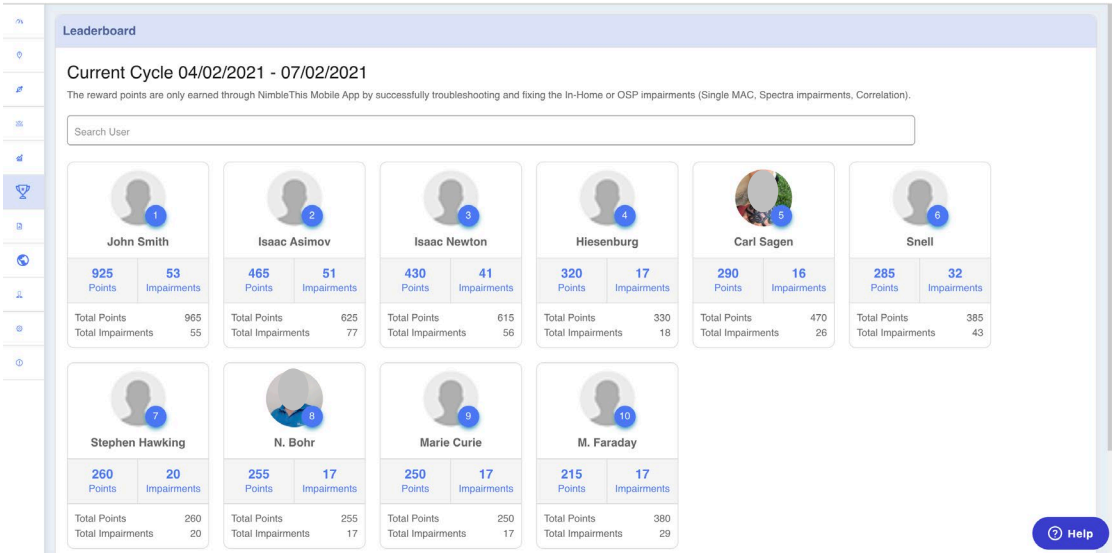


Figure 4.5: Leaderboard used for incentivizing users and collecting labeled data
(Image Courtesy: NimbleThis)

Second, CableLabs is working to have several expert users, including the author of this paper, to label a set of FBC data. It is hopeful that either or both two methods will generate a large enough dataset to be used for a true supervised learning model which can further improve upon the existing model.

4.1.2 New Impairment Detector

In the future, one could analyze ways to create a reliable way of identifying and localizing impairments to bypass the need for SID and clustering in the first place. Analysis/clustering on accurate labels of impairments would be able to find the same impairment in multiple modems but do so with greater confidence and accuracy. To be able to identify and localize all these impairments there needs to be large datasets of labeled data which are currently not accessible. Models for prediction could be made for each impairment or one large model could be made to make predictions about all impairments. Possible avenues to investigate include Convolutional Neural Networks (CNNs), perhaps some like ones seen in computer vision such as YOLO (You Only Look Once) to both classify and localize impairments [4, 9].

This starts to move into the arena of artificial intelligence (AI). Which the author of this paper chooses to use with great care. Today machine learning is being used and often times AI is used as a marketing gimmick. However, given enough data, a true AI model can be developed with the help of technicians. Lots of technicians feeding accurate data into PNM applications. Once this level is achieved, ML and/or AI models can be developed which will look at a single or multiple FBC images and not only inform the user of what the impairment is (i.e., suckout or standing wave), but further it can make very accurate suggestions of what the most probably repair for the impairment may, such as “85% probability of a bad drop cable”. This technology is not years away, but something we expect to realize within the next 1-2 years and will change the technicians interact with PNM technology.

Abbreviations

CM	Cable Modem
CMTS	Cable Modem Termination System
CNN	Convolutional Neural Network
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DOCSIS	Data Over Cable Service Interface Specification
EPS	Epsilon parameter in DBSCAN
FBC	Full-Band Capture
FEC	forward error correction
HD	high definition
Hz	hertz
LOF	Local Outlier Factor
ISBE	International Society of Broadband Experts
PNM	Proactive Network Maintenance
RF	Radio Frequency
SCTE	Society of Cable Telecommunications Engineers
SID	Spectral Impairment Detector Released by CableLabs
YOLO	You Only Look Once

Bibliography & References

[1] Breunig, M., Kriegel, H.P., Ng, R., & Sander, J. (2000). LOF: identifying density-based local outliers. In ACM sigmod record (pp. 93–104).

[2] elktros. "Band Stop Filter Circuit Design and Applications." Electronics Hub, 26 Jan. 2019, https://www.electronicshub.org/band-stop-filter/.

[3] Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Proc. KDD (pp. 226–231).

[4] Kiranyaz, Serkan & Avci, Onur & Abdeljaber, Osama & Ince, Turker & Gabbouj, Moncef & Inman, Daniel. (2019). 1D Convolutional Neural Networks and Applications: A Survey.

[5] Lakshmanan, Swetha. "How, When, and Why Should You Normalize / Standardize / Rescale Your Data?" Towards AI — Multidisciplinary Science Journal, 16 May 2019, https://towardsai.net/p/data-science/how-when-and-why-should-you-normalize-standardize-rescale-your-data-3f083def38ff.

[6] Pandas.Core.Window.Rolling.Rolling.Median — Pandas 1.0.5 Documentation. https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.core.window.rolling.Rolling.median.html. Accessed 7 July 2020.

[7] PNM Best Practices Primer: HFC Networks (DOCSIS 3.1). CableLabs, 6 May 2020.

[8] Python Lists vs. Numpy Arrays - What Is the Difference?: IST Advanced Topics Primer. https://webcourses.ucf.edu/courses/1249560/pages/python-lists-vs-numpy-arrays-what-is-the-difference. Accessed 7 July 2020.

[9] Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection (cite arxiv:1506.02640)

[10] "Window Sliding Technique - GeeksforGeeks." GeeksforGeeks, 16 Apr. 2017, https://www.geeksforgeeks.org/window-sliding-technique/.

[11] “Proactive Network Maintenance using Fast, Accurate Anomaly Localization and Classification on 1-D Data Series”,(July 2020), Jingjie Zhu, Karthik Sundaresan, Jason Rupe CableLabs, Louisville, CO, U.S.A

Maximizing Returns on the Path to DOCSIS 4.0

A Technical Paper prepared for SCTE by

Mike Darling
Principal Engineer
Shaw Communications
2728 Hopewell Place NE, Calgary AB T1Y 7J7
mike.darling@sjrb.ca

1. Introduction

As wireline subscriber demand for increased broadband speed grows, cable operators will have to upgrade their networks to accommodate. Cisco estimates that the average fixed broadband speed in North America will increase 2.5-fold from 2018-2023 [1]. In Canada, strong competition from telco fibre-to-the-premises (FTTP) networks adds to the need for upgrades. Shaw Communications is in the fourth year of a five-year upgrade cycle of its hybrid-fibre coax (HFC) network. This upgrade has allowed Shaw to offer a 1.5Gbps downstream by 100Mbps upstream service tier to most of its footprint and has driven congestion to near zero, all during a global pandemic that has seen network utilization spike by 30% in the downstream and 60% in the upstream. Despite its strong position, Shaw is already planning the next upgrade cycle, which will follow quickly on the heels of the current upgrade. This paper will discuss a strategy to manage network congestion and tier offerings, leveraging available HFC network architectures.

2. HFC Network Status

HFC networks traditionally operate in a frequency division duplex (FDD) mode with upstream spectrum carried below downstream spectrum, separated by a cross-over region. The Data over Cable Service Interface Specification (DOCSIS) supports several different upstream spectrum splits, including 5-42MHz, 5-85MHz, 5-204MHz, and higher, otherwise known as sub-split, mid-split, high-split, and ultra-high-split respectively. Each subsequent version of DOCSIS expands its support of upstream frequencies, always with backwards compatibility as depicted below.

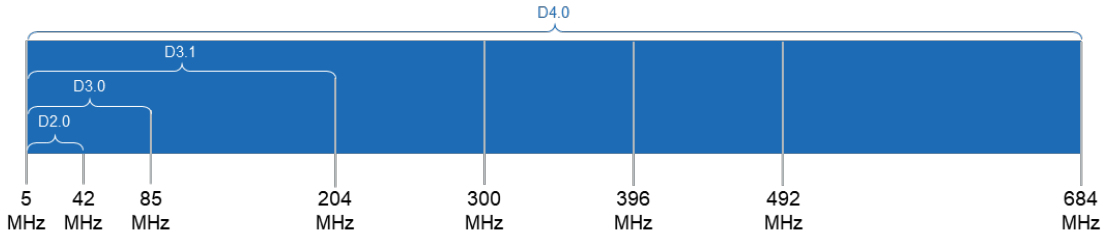


Figure 1 – Upstream Spectrum Band Plans

In the downstream, DOCSIS supports up to 857MHz for DOCSIS 2.0, 1002MHz for DOCSIS 3.0, 1218MHz for DOCSIS 3.1, and 1794MHz for DOCSIS 4.0 FDD [2] – [5]. For simplicity, plant types will be referred to as mid-split, high-split, and DOCSIS 4.0 FDD, describing band plans of 1002/85MHz, 1218/204MHz, and 1794 with an ultra-high-split upstream option.

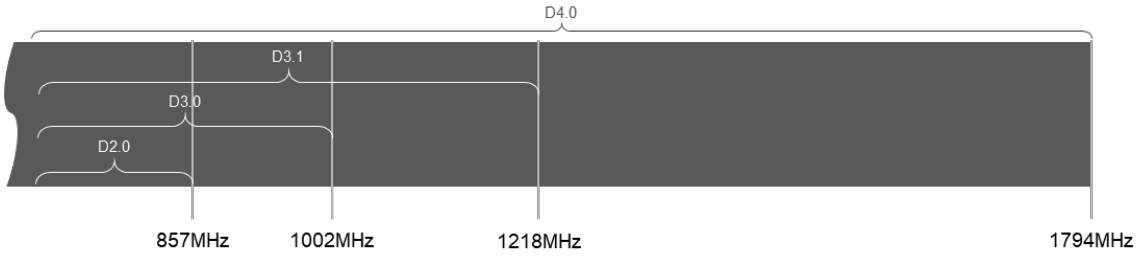


Figure 2 – Downstream Spectrum High Frequencies

Shaw operates an HFC network, that at the time of writing was over 90% mid-split with upstream signals carried from 5-85MHz and downstream signals carried from 108-1002MHz. The approximate spectrum breakdown is shown in Figure 3.

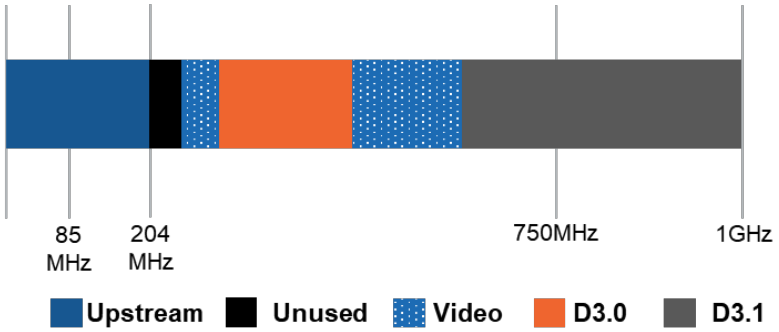


Figure 3 – Shaw Mid-split Spectrum

The HFC network uses analog optical transport from the hub site to the optical node where the transition to coax happens. Plant architectures are measured both by the allowable number of amplifiers in cascade as well as the link budget or spacing between amplifiers. The cascade length is referred to as N+X where N refers to the optical node and the X to the maximum cascade depth. Amplifier spacing refers to the maximum designed frequency, as in 750MHz or 1GHz plant. Shaw's outside plant contains a mixture of architectures, but the average cascade depth is N+4 and, with some exceptions, is designed to 1GHz. A high-level depiction of an HFC network is shown in Figure 4.

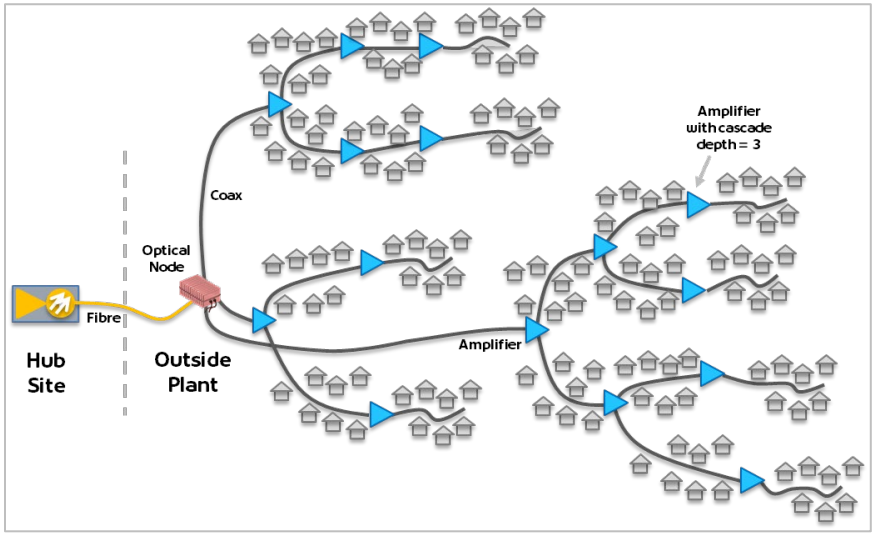


Figure 4 – Hybrid-Fibre Coax Network

2.1. COVID-19 Impact

The COVID-19 pandemic caused a shift in subscriber behaviour unlike anything previously experienced by broadband providers (Figure 5). Restrictions on movement increased subscriber usage of their in-home broadband services in a short time span. In the downstream, this growth was quickly mitigated through a reduction in video bitrates by streaming video providers such as Netflix [6].

Figure 5 – Relative Downstream Traffic During COVID-19

The growth of specific categories of traffic can be analyzed using deep-packet inspection (DPI) systems. In Figure 6, the height of the bars represents the relative amount of added consumption with traffic categories on the left adding the most consumption and categories on the right adding the least consumption. The percentage on top of each bar represents the increase in consumption of that traffic category compared to pre-COVID-19 levels. As can be observed, streaming media, which includes Netflix and YouTube, grew by 30%, which was lower than some of the other traffic categories but accounted for the bulk of increased consumption.

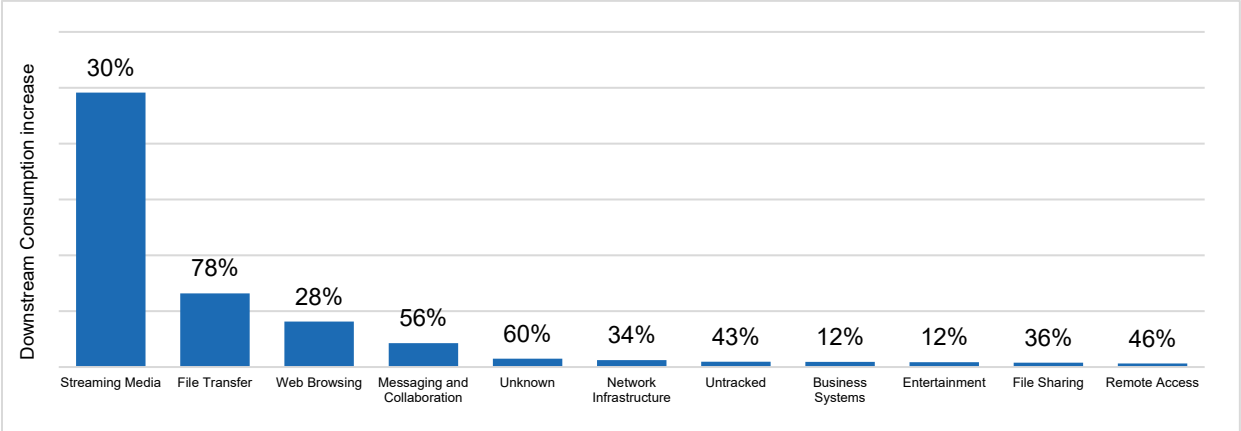


Figure 6 – Downstream Consumption Increase by Application

In the upstream, the shift to remote learning, working from home and the increased use of video calling applications kept upstream traffic higher for a longer period compared to downstream traffic (Figure 7).

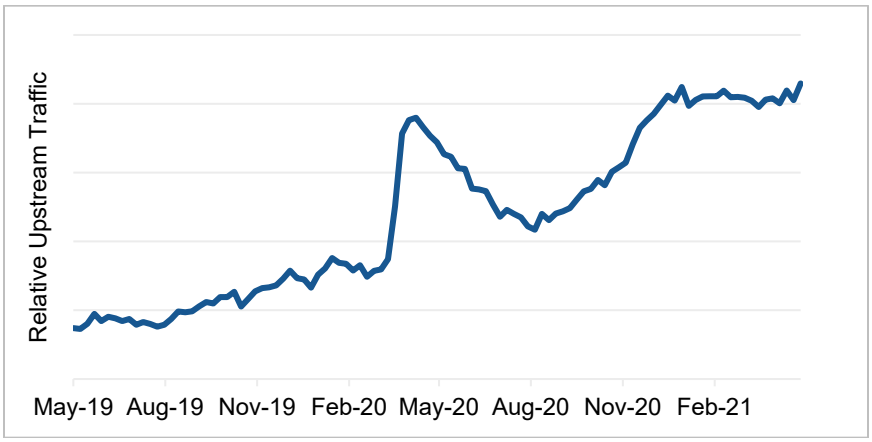


Figure 7 – Relative Upstream Traffic During COVID-19

The largest amount of added consumption, as well as the largest percent increase, came from the messaging and collaboration category, which includes FaceTime and Skype (Figure 8).

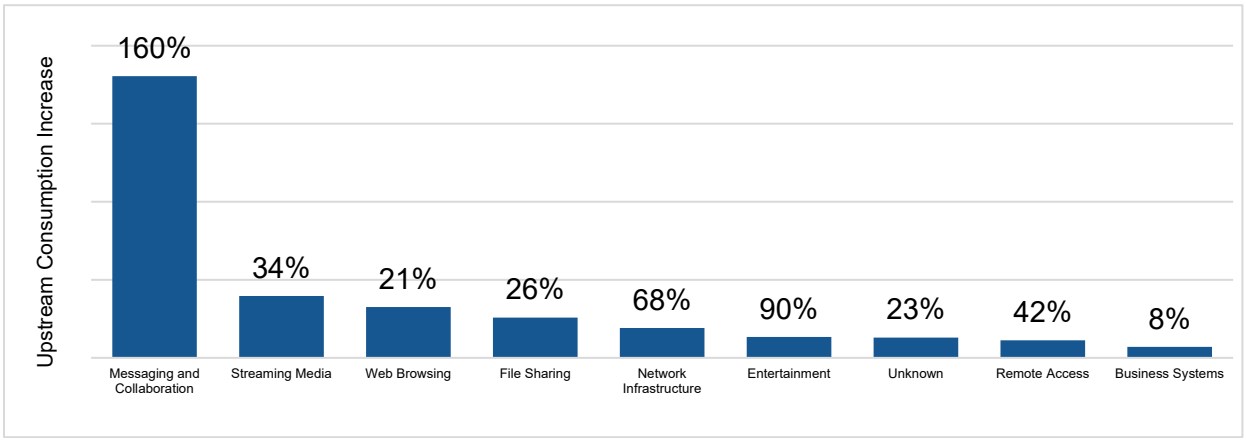


Figure 8 – Upstream Consumption Increase by Application

2.2. Mid-split Journey

Prior to the mid-split upgrade, Shaw's HFC network operated to a high frequency of 750MHz with a sub-split band plan. Increasing demand for broadband services had driven the transition from analog video to the more efficient digital video, freeing up spectrum for DOCSIS carriers. While this provided more spectrum to allocate to the downstream, it did not increase the spectrum allotted to the upstream.

At the time, the downstream-to-upstream ratio was much lower than it is today and there was a concern that upstream congestion would become unmanageable. The decision was made to reserve the downstream spectrum from 54-108MHz for a mid-split upgrade—a decision that was not easy at the time given that downstream spectrum was highly desired for additional broadcast video services.

Mid-split testing began in earnest to ascertain the level of effort required to upgrade the network to 85MHz in the upstream and 1GHz in the downstream. Video set top boxes that used out-of-band communication in the 72.75MHz range had to be re-tuned to a new frequency where possible or retired from the network if necessary. Interference testing was undertaken to see if upstream transmissions in the 54-85MHz range would harm the function of customer premises equipment (CPE) in the same or

neighbouring house, and performance of optical transmitters was measured to ensure proper functionality. Lastly, research was undertaken to understand whether the upgrade could be done on a drop-in basis, or whether amplifier respacing was required.

In 2013 the first upgrade of a production node was completed and used to conduct further testing. In the meantime, the network was experiencing unprecedented growth in both the upstream and downstream, driven by increased demand for broadband services. In 2012 upstream year-over-year growth reached a high of 55%, while in 2013 downstream year-over-year growth reached a high of 85%. This demand created increased levels of network congestion, for which a solution was required.

Analysis was undertaken to explore network levers which could be pulled to increase network capacity. Potential solutions included a digital-to-IP video transition to make room for more downstream DOCSIS carriers, node splits in congested areas, an HFC upgrade to N+0 or N+few, and a mid-split upgrade. In comparison to the first three options, a mid-split upgrade was found to be cost effective and could be executed in a relatively short timeframe.

3. Competitive Environment

Canadian telcos are aggressively upgrading their Digital Subscriber Line (DSL) networks to FTTP. As of the end of 2020, Bell's network was 57% FTTP and Telus' network was 81% FTTP [7], [8].

Telus, which operates in much of the same footprint as Shaw, offers a max tier of 1.5Gbps downstream by 940Mbps upstream in its FTTP footprint using Gigabit Passive Optical Network (GPON) technology, as shown in Figure 9. This technology operates at a physical layer (PHY) rate of 2.488Gbps downstream and 1.244Gbps upstream [9].

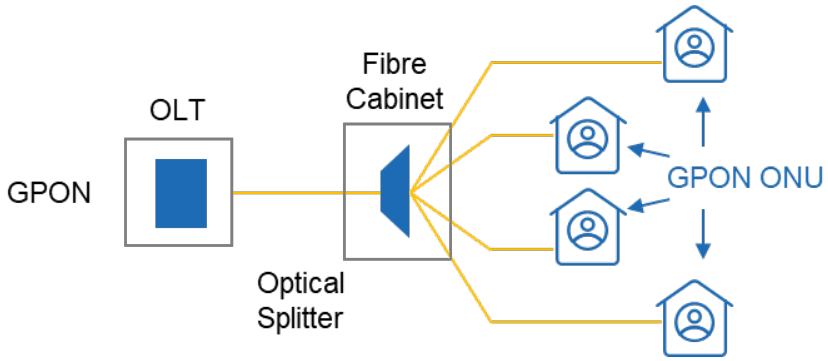


Figure 9 – Gigabit Passive Optical Network

As shown in Figure 10, this Passive Optical Network (PON) architecture allows for the coexistence of GPON and Ten-Gigabit Symmetric PON (XGS-PON) with the addition of Optical Line Termination (OLT) line-cards and additional splitters.

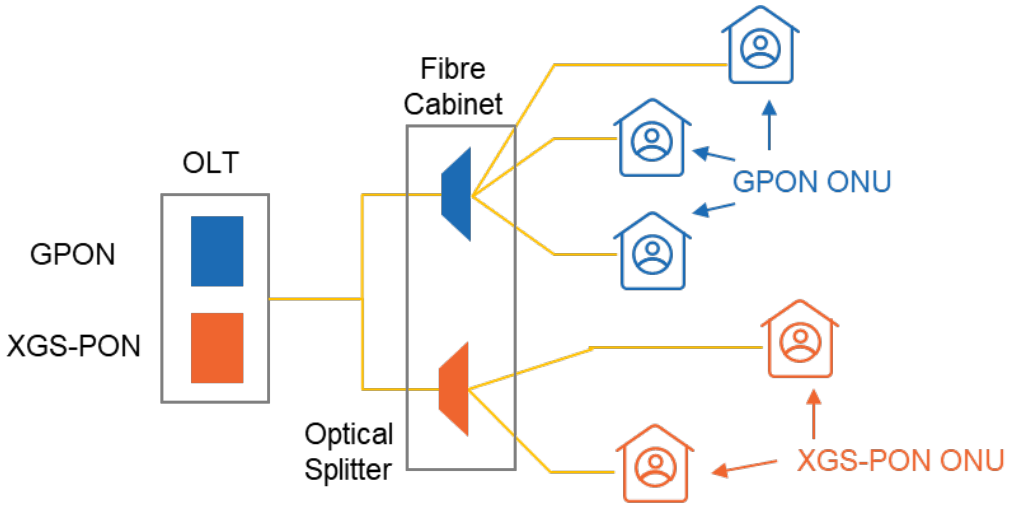


Figure 10 – Passive Optical Network Coexistence

XGS-PON is 10Gbps capable [10] and the max tier supported will depend on the split ratio and how aggressive the operator chooses to be. At the time of writing, Telus launched new 2.5Gbps symmetric services, giving an indication of initial XGS-PON tier offerings. It is important that any upgrade strategy have a roadmap for supporting similar tiers.

4. Upgrade Needs

A mid-split HFC network has similar or greater capacity when compared to a GPON network in the downstream, which allows Shaw to offer a 1.5Gbps downstream tier with options for higher tiers in the future. In the upstream however, a mid-split network has a lower capacity compared to GPON. An HFC network upgrade would be required in order to match the upstream tier.

When network traffic is observed, it is highly asymmetric, with downstream-to-upstream ratios above 15:1 for residential consumers and in the range of 5:1 for business subscribers (Figure 11), even during the COVID-19 pandemic that saw a disproportionate increase in upstream usage, as discussed earlier.

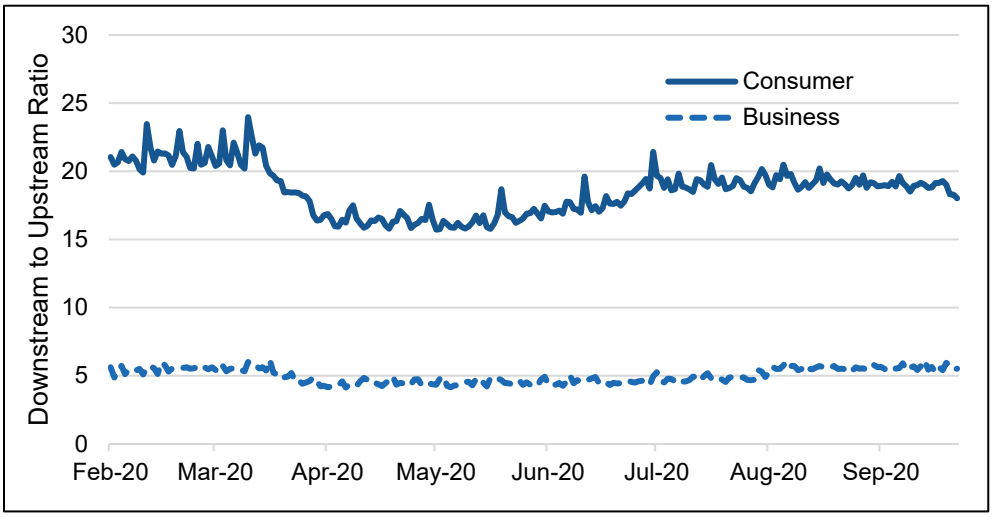


Figure 11 – Downstream-to-Upstream Ratio

Given this relationship, there is a question as to whether subscribers can fully utilize gigabit upstream services, and if so, whether their experience is improved. In looking at the 1Gbps downstream by 100Mbps upstream tier, two interesting phenomena can be observed. Using DPI systems, the max speeds a subscriber is able to reach over the span of a month can be measured. As seen in Figure 12, a significant number of subscribers, likely connected over Ethernet, use their full 1Gbps tier. Many subscribers, however, have a max speed in the 400-500Mbps range due to in-home Wi-Fi limitations.

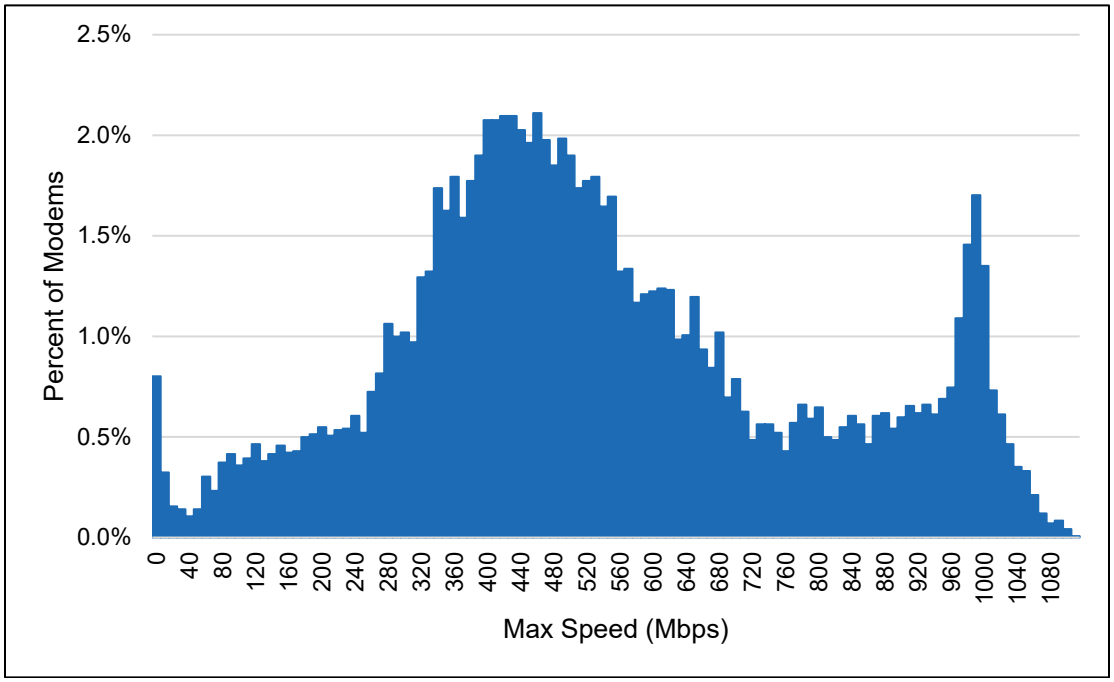


Figure 12 – Downstream Max Speed Distribution (1Gbps Tier)

Wi-Fi 6E will remove current limitations as it introduces new spectrum that will enable gigabit speeds [11]. However, the process to upgrade CPE and end-user devices to Wi-Fi 6E compatible equipment will take time.

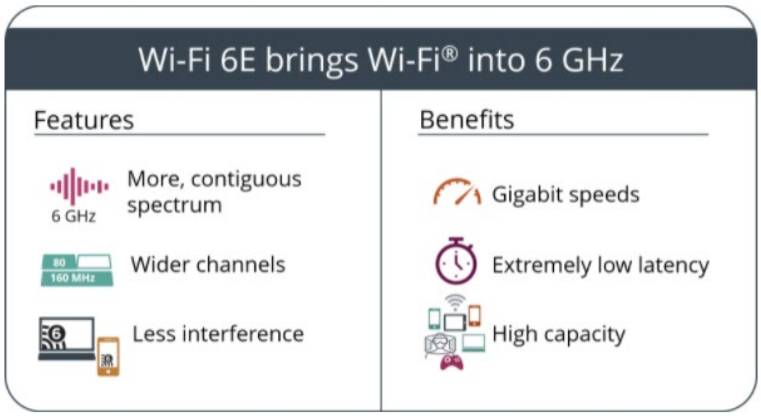


Figure 13 – Wi-Fi 6E Benefits (Source: Wi-Fi Alliance)

The types of applications used during high-speed upstream bursts can be observed using DPI systems. The data shows that the higher the speed, the more the traffic is dominated by backup applications operating in the background, not necessarily noticed by the user (Figure 14).

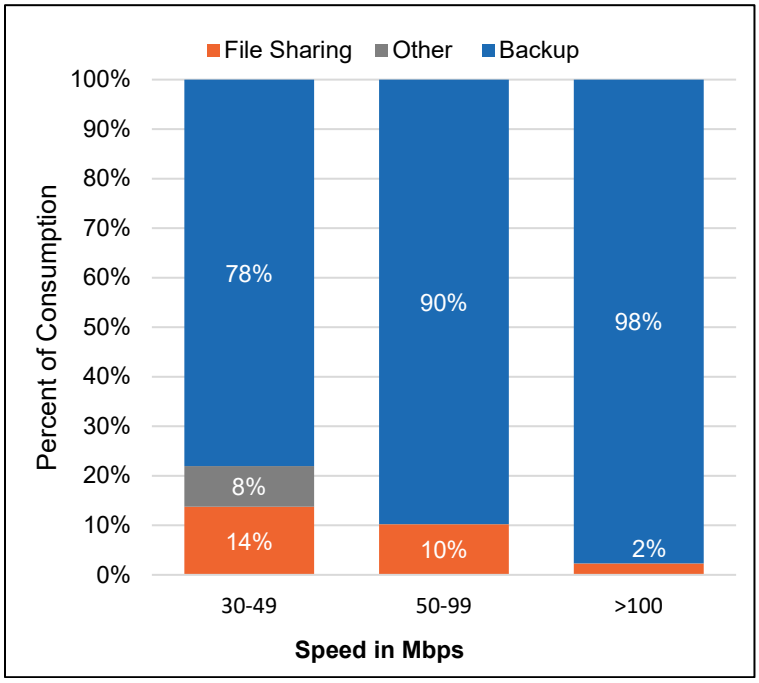


Figure 14 – Median Upstream Consumption by Speed

As mentioned earlier, Shaw is experiencing near zero congestion as the mid-split upgrade cycle approaches completion. Congestion forecasts were produced to estimate when a subsequent upgrade will be required for capacity reasons. Forecasts assume 15% downstream compound annual growth rate (CAGR) in addition to increased IPTV traffic due to customers migrating from older quadrature amplitude modulation (QAM) video hardware to new IPTV video solutions. In the upstream 25% CAGR is assumed. These numbers are in line with recent growth.

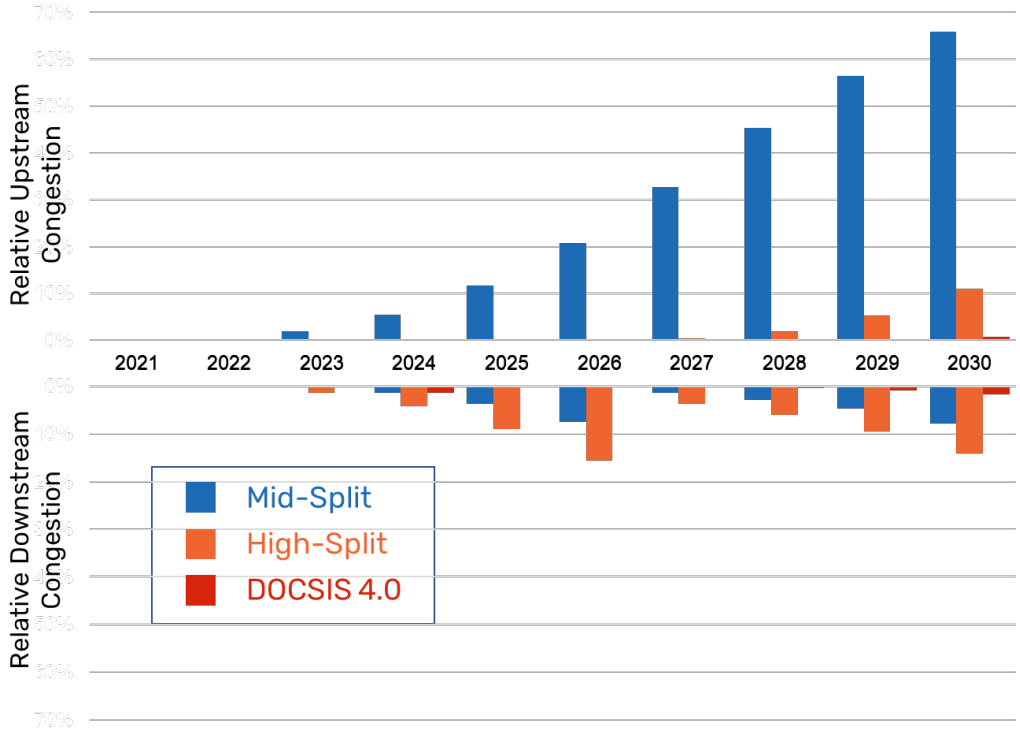


Figure 15 – Congestion Forecast

In Figure 15, congestion is plotted for mid-split, high-split, and DOCSIS 4.0 FDD networks, with upstream in the upper half and downstream the lower half. Low levels of congestion can be resolved with targeted node splits, while higher levels require network-wide solutions. As can be observed, concerns do not arise with mid-split capacity until the latter half of the decade.

In the downstream it is assumed that a high-split plant is operated only to 1GHz, resulting in a net decrease in downstream capacity when compared to mid-split. Downstream congestion grows in 2025 and 2026 but is reduced in 2027 when the transition from QAM to IP video is expected to be complete and more spectrum allocated to DOCSIS traffic. These two topics are discussed further in the following sections. In contrast to the congestion forecasts for mid-split and high-split, a DOCSIS 4.0 FDD upgrade provides enough capacity for the network to run nearly congestion-free through 2030 and beyond.

5. Upgrade Options

With our mid-split upgrade nearly complete, the next network upgrade is being planned. The options currently available are high-split, DOCSIS 4.0 FDD, DOCSIS 4.0 Full-Duplex DOCSIS (FDX), and FTTP.

5.1. High-split

An upgrade of the HFC plant to high-split, with upstream spectrum from 5-204MHz and downstream spectrum from 258-1218MHz (Figure 16), represents a modest increase in spectrum but has the advantage of the required equipment being available now.

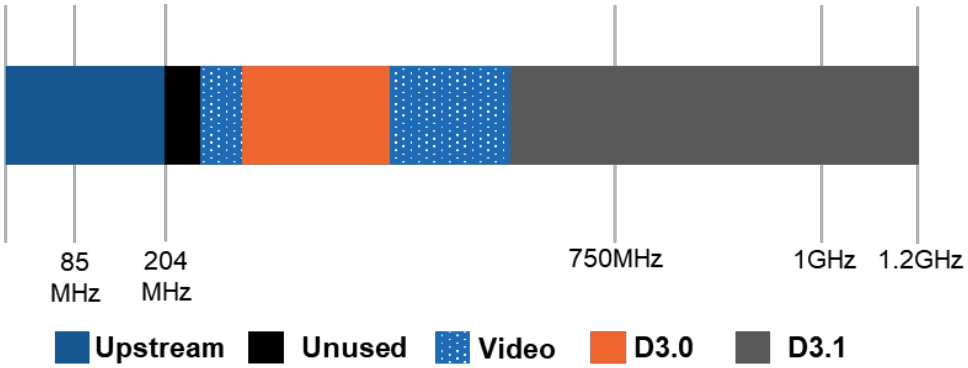


Figure 16 – High-split Spectrum

As displayed in the table below, the increase in spectrum is heavily weighted toward the upstream with 149% additional spectrum, while the spectrum allocated to downstream increases by 7%.

Table 1 – High-split Spectrum Allocation

	Mid-split	High-split	Percent Increase
Upstream Spectrum (MHz)	80	199	149%
Downstream Spectrum (MHz)	894	960	7%

As discussed earlier, the strategic reason to upgrade to high-split is to match the tier capability of GPON, which requires additional upstream spectrum. Another advantage of a high-split upgrade is that the technical considerations are similar to those recently explored with a mid-split upgrade.

Shaw set top boxes that do not have DOCSIS modems use the ANSI/SCTE 55-1 standard for communication. The downstream out-of-band (OOB) transmission frequency range is 70-130MHz [12] in spectrum that transitions to upstream when upgrading to a high-split network. While there are options available to continue OOB operation [13], the decision was made to remove the OOB and all video equipment reliant on ANSI/SCTE 55-1 in the case of a high-split upgrade. Operationally, this means that this equipment must be removed from a service area before upgrade but does not require all equipment to be removed at the system level.

Industry Canada mandates that cable distribution networks do not interfere with international emergency frequencies and that they are monitored for compliance [14]. Monitoring is done by measuring the signal level of an analog television channel chosen by Industry Canada. In a high-split plant these frequencies switch from downstream to upstream and a new method for monitoring is required. There are options available for leakage monitoring in a high-split plant [13], but until commercial solutions are available that range of spectrum will be left unused.

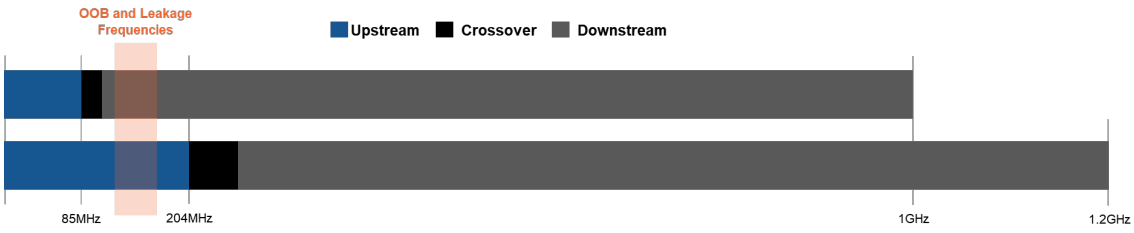


Figure 17 – Out-of-Band and Leakage Frequencies

Another issue to consider in the implementation of high-split is the in-home coexistence of DOCSIS and Multimedia over Coax Alliance (MoCA) signals above 1GHz. MoCA devices use in-home coax wiring for transport of multimedia content, as seen in Figure 18. At Shaw, MoCA is used with whole-home video solutions, allowing a single MoCA gateway to transport video content to MoCA portals over the in-home coax network.

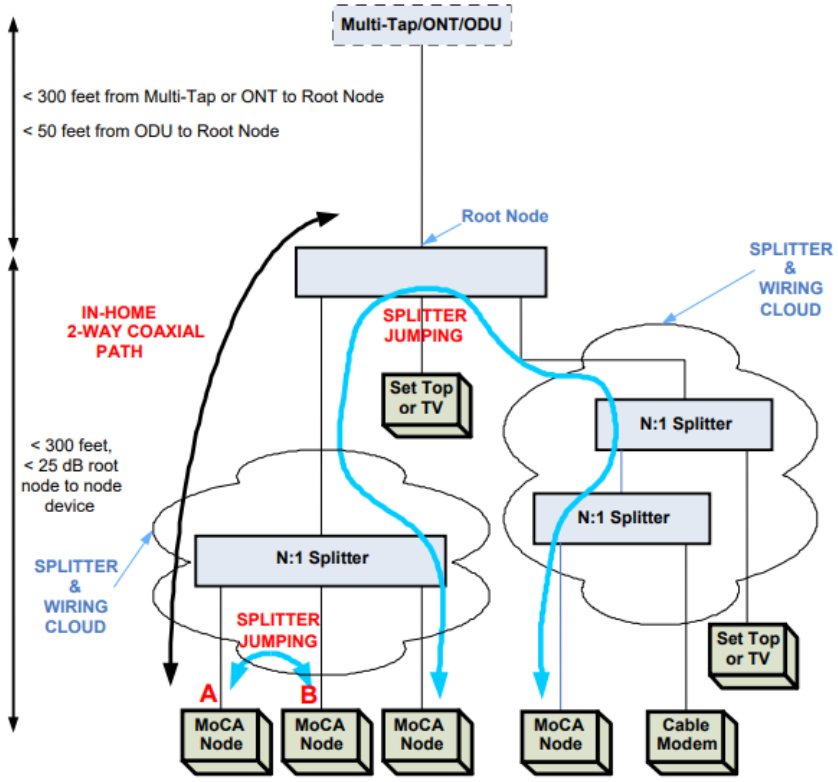


Figure 18 – MoCA Network (Source: MoCA Alliance)

MoCA 1.1 devices operating in an HFC network use 50MHz wide channels in the D-band which is located from 1125MHz to 1525MHz [15], as illustrated in Figure 19. Devices will move to a different channel if received signal strength is not adequate at the current channel.

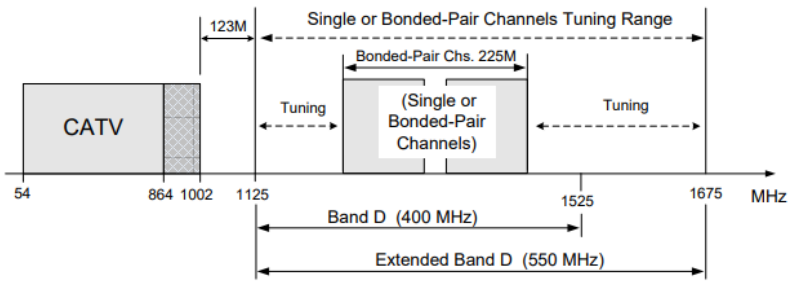


Figure 19 – MoCA Spectrum Use (Source: MoCA Alliance)

A sample of MoCA devices were polled to see what channel they were using. As shown in Figure 20, 46% of devices were using the lowest channel with a centre frequency of 1150MHz, but a significant percentage of MoCA devices had moved to higher spectrum.

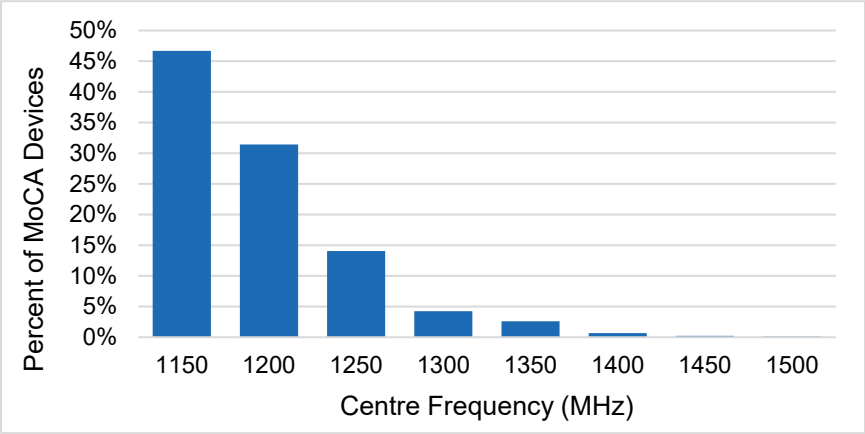


Figure 20 – MoCA Centre Frequencies

Research is required to determine whether MoCA devices can be forced to higher frequencies either through back-office commands or by occupying spectrum to 1218MHz with DOCSIS carriers. If this is not possible, then MoCA spectrum would have to be left unutilized until MoCA devices are removed from the network. In a worst-case scenario, if high-split frequencies are limited to 258-1002MHz downstream, the high-split upgrade would result in 17% less downstream spectrum when compared to mid-split (Table 2). While this is not a preferred outcome, downstream spectrum can also be freed up by the more efficient transmission of video content, or by removing a portion of that content altogether.

Table 2 – High-split Spectrum Allocation when Limited to 1GHz

	Mid-split	High-split (MoCA)	Percent Change
Upstream Spectrum (MHz)	80	199	149%
Downstream Spectrum (MHz)	894	744	-17%

As previously mentioned, a decision was made to remove set top boxes reliant on the OOB for communication. These boxes also happen to be the only ones in the network that do not support MPEG4 video and thus their removal allows for MPEG2 video signals to be carried more efficiently via MPEG4. The amount of spectrum that can be freed up is sufficient to prevent downstream DOCSIS capacity loss during a high-split upgrade.

A high-split upgrade provides the opportunity to move to a Distributed Access Architecture (DAA), although it is not strictly required. DAA allows for scalability and preserves space in hub sites as the number of optical nodes increases. This eliminates the analog optical link as well as any testing that would otherwise have been required to ensure upstream signals could be carried at high modulation rates over the 5-204MHz range. The difference in performance between DAA and analog nodes can be observed by comparing signal-to-noise ratio (SNR) data. A group of approximately ten thousand modems in each type of plant were polled for SNR. As shown in Figure 21, the average downstream SNR of modems in analog nodes was 38.8 decibels (dB) versus 41.4dB in DAA nodes, a difference of 2.6dB.

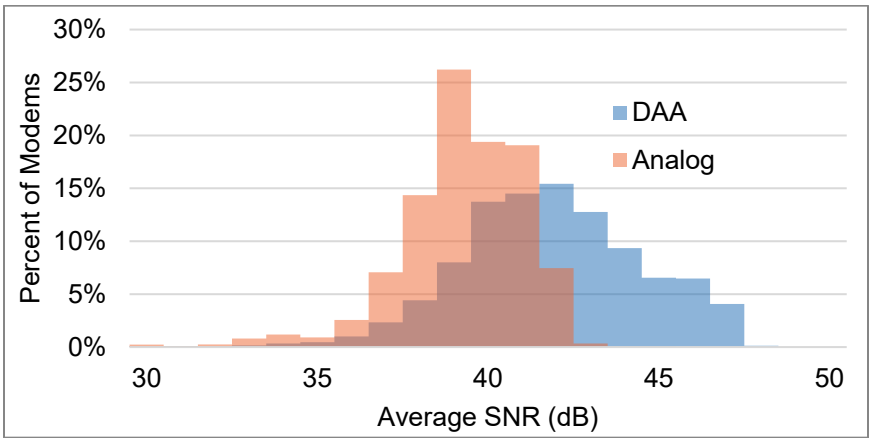


Figure 21 – Downstream SNR DAA vs Analog

As depicted in Figure 22, the average upstream SNR of modems in analog nodes was 33.1dB versus 37.3dB in DAA nodes, a difference of 4.2dB. By using DAA nodes, more modems will be able to make use of the highest modulation profiles, thus increasing capacity.

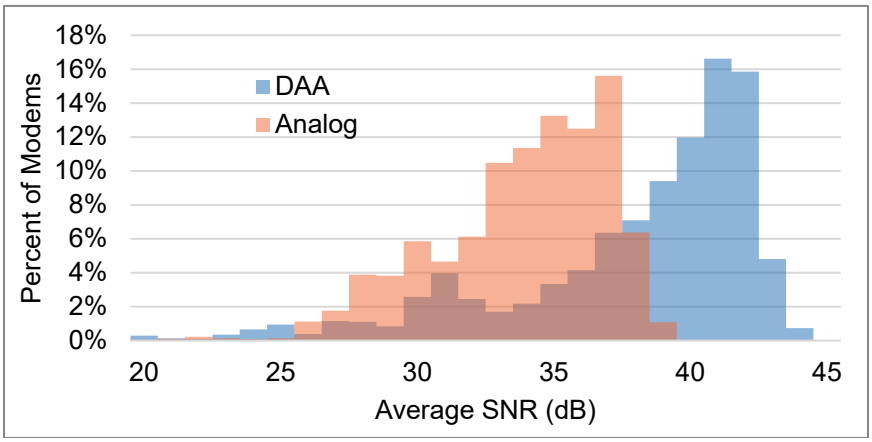


Figure 22 – Upstream SNR DAA vs Analog

As with mid-split, the preferred method for a high-split upgrade is to replace amplifiers in a drop-in fashion. This eliminates the need for plant redesign and amplifier respacing, which would complicate the upgrade. The risk with this method is that signal quality degrades with each additional amplifier in cascade. While adding fibre to the network to reduce cascades to a chosen maximum depth would solve this problem, the time and capital required for such an upgrade is substantial. DOCSIS technology can help in this regard as DOCSIS 3.1 changed the way we think about node capacity. DOCSIS 3.1 introduced the concept of profiles used by orthogonal frequency division multiplexing (OFDM) and orthogonal frequency division multiple access (OFDMA) signals. These profiles allow cable modems with higher signal quality to transmit and receive at higher data rates compared to cable modems with lower signal quality. While the data rate of a cable modem located at the end of line may be lower than other cable modems on the node, there is no issue so long as the data rate is sufficient to enable the service tier. In the case where cable modems at the end of line cannot support service tiers, either tier offerings in that location can be reevaluated, or fibre optics can be deployed to reduce cascade lengths.

During Shaw's mid-split upgrade all passives that were sub-1GHz were upgraded. 1GHz passives could be retained in a high-split upgrade as the top frequency for DOCSIS 3.1 was chosen with 1GHz passives

in mind because they were found to continue operation beyond 1GHz. Optionally 1GHz passives could be replaced with 1.8GHz or 3GHz models, which would provide decreased loss from 1-1.2GHz for high-split plant and pave the way for DOCSIS 4.0 FDD.

5.2. DOCSIS 4.0 FDD

Upgrading the HFC plant to DOCSIS 4.0 FDD with a high frequency of 1794MHz represents a large increase in spectrum that comes with several options for upstream frequency range (Table 3). As upstream spectrum increases, downstream spectrum decreases. In deciding which band plan is optimal it is important to look at the downstream-to-upstream ratio and the maximum service tier that can be offered in both directions. DOCSIS 4.0 FDD can materially outperform GPON and compete effectively with XGS-PON, which is capable of 10Gbps symmetric speeds. Depending on which band split is selected, DOCSIS 4.0 FDD can have greater capacity than XGS-PON in the downstream but remains lower in the upstream.

Table 3 – DOCSIS 4.0 FDD Spectrum Allocation

	Mid-split	High-split	UHS-300	UHS-396	UHS-492	UHS-684
Upstream Spectrum (MHz)	80	199	295	391	487	679
Downstream Spectrum (MHz)	894	960	1422	1302	1188	960
DS:US Spectrum Ratio	11.2	4.8	4.8	3.3	2.4	1.4

For capacity planning purposes it is beneficial to have a downstream-to-upstream capacity ratio similar to the downstream-to-upstream demand ratio, which as discussed earlier, is greater than 15:1 for residential subscribers. For competitive purposes however, it may be beneficial to offer symmetric or near symmetric tiers.

Another consideration is the high frequency of current DOCSIS cable modems and video set top boxes. In Shaw's network 100% of CPE currently has a maximum high frequency of 1002MHz. As shown in Figure 23, ultra-high-split options would reduce the available spectrum to current CPE.

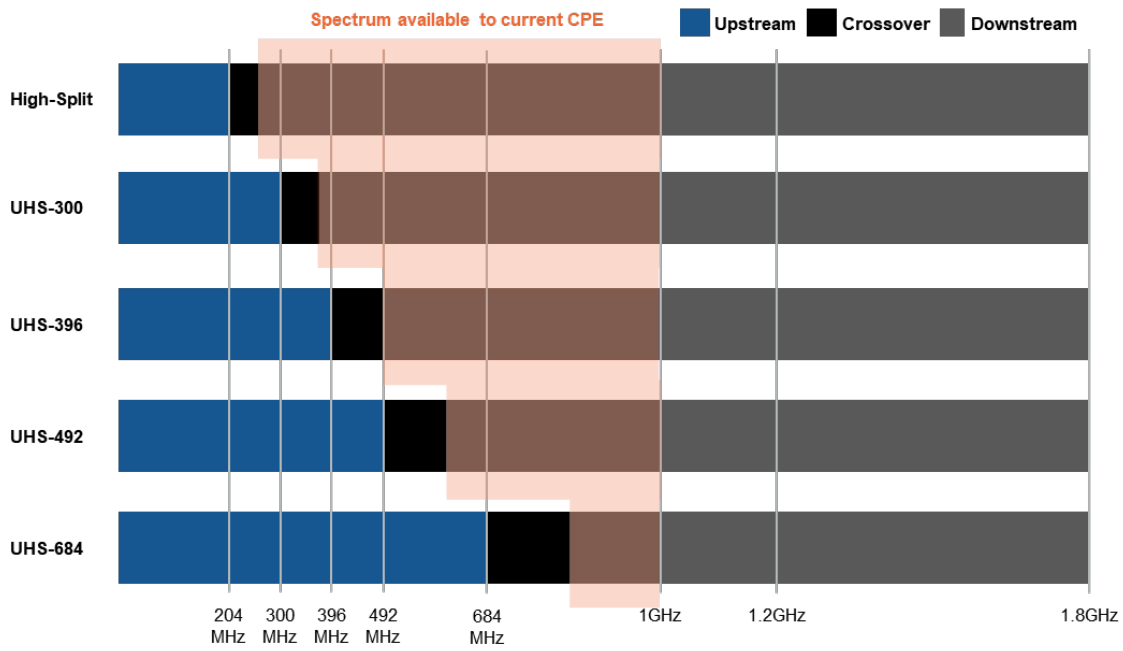


Figure 23 – Spectrum Available to Current CPE

Reducing the capacity available to the installed base of CPE could create congestion until such time as either enough DOCSIS 4.0 FDD cable modems are in the network or QAM video can be retired. One potential strategy is to initially run DOCSIS 4.0 FDD with a high-split band plan and switch to an ultra-high-split band plan in the future, perhaps when the transition to IPTV is complete. UHS-396 is a potential band plan that balances traffic engineering with the need for competitive service tiers.

The mid-split upgrade process at Shaw involved replacing amplifier modules when possible, and the entire amplifier with housing only when necessary. Amplifiers that were 1GHz/42MHz were bench upgraded to 1GHz/85MHz in a properly controlled lab environment as it was not found to be practical to swap diplex filters in the field. DOCSIS 4.0 FDD was designed with multiple band plans in mind, and it is important that new amplifiers have a practical method of changing between them. There are many options for accomplishing this change, ranging from field-swappable diplex filter boards to centrally controlled and dynamic systems that remotely switch band plans depending on instantaneous demand.

As with a high-split upgrade, issues with ANSI/SCTE55-1, leakage detection, and MoCA coexistence arise with a DOCSIS 4.0 FDD upgrade. MoCA coexistence issues are more acute as the overlapping spectrum is greater. The MoCA D-band is 400MHz and the extended D-band used by MoCA 2.0 is 550MHz, which is too much spectrum to forego for long.

Initial tests have been performed on plant segments to confirm operation to 1.8GHz, but as with all HFC networks, signal performance will degrade with increased cascade length. Similar to high-split, the desired upgrade path for DOCSIS 4.0 FDD is a drop-in upgrade. It is expected that a drop-in upgrade will operate in current plant cascades and spacings with performance gains realized as fibre is deployed deeper into the network. Unlike with high-split, all passive devices will require upgrading to 1.8GHz or greater. If available, swapping passives or their housings to 3GHz could future-proof those devices and make any future upgrade easier.

5.3. DOCSIS 4.0 FDX

FDX allows spectrum to be used in both upstream and downstream directions simultaneously, effectively doubling the spectral efficiency for those frequencies (Figure 24).

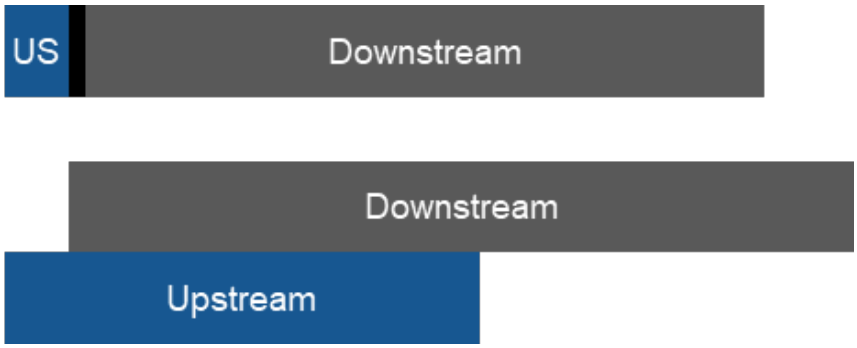


Figure 24 – FDX Spectrum Overlap

The FDX specification supports a high frequency of 1218MHz, and the frequency range from 108-684MHz may be used bidirectionally, leading to a large increase in spectrum (Figure 25).

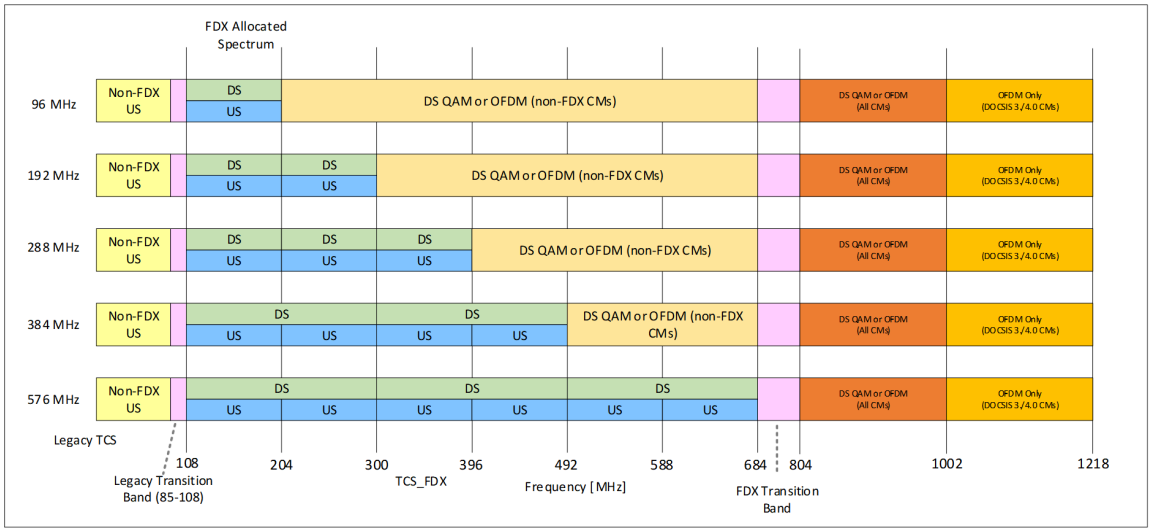


Figure 25 – FDX Band Plans (Source: CableLabs)

The amount of spectrum allocated to upstream and downstream can be delineated by the amount of bidirectional spectrum, as shown in Table 4.

Table 4 – FDX Spectrum Allocation

	Mid-split	96MHz	192MHz	288MHz	384MHz	576MHz
Upstream Spectrum (MHz)	80	176	272	368	464	656
Downstream Spectrum (MHz)	894	1098	1098	1098	1098	1098
DS:US Spectrum Ratio	11.2	6.2	4.0	3.8	2.4	1.7

The quantity of downstream spectrum remains the same for different amounts of bidirectional spectrum. The frequency range between 108-684MHz not used bidirectionally, however, is only available to non-FDX Cable Modems (CMs), as shown in Figure 25.

FDX is designed to be operated in passive, or N+0, plant. This requires the removal of all amplifiers and fibre deployed deep into the network (Figure 26).

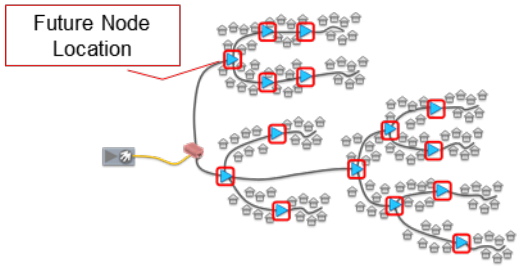


Figure 26 – N+0 Build

In Shaw's network, N+0 represents a small fraction of homes passed, generally in newer areas or especially in Multiple Dwelling Units (MDUs). In these areas, a deployment of FDX is a matter of swapping out the node. In networks with amplifier cascades, however, fibre deployment is required.

The cost and time required to upgrade to N+0 is heavily influenced by the type and ownership of the infrastructure. Owned aerial infrastructure is the easiest to upgrade, with un-owned underground infrastructure the most difficult, with estimates showing upwards of ten times the difference in cost and build time between the two. The mix present in an operator's network will dictate the average cost per home passed as well as the time required to complete the upgrade.

5.4. FTTP

As with FDX, an upgrade to FTTP is heavily impacted by the infrastructure and has the additional requirement to deploy fibre drops, either at the time of upgrade or in a success-based fashion at a later time. One important consideration in moving to FTTP is the lack of backwards compatibility, requiring CPE to be replaced. Depending on infrastructure access agreements there may not be the option to overbuild and transition customers slowly. In a worst-case scenario, the HFC network is removed, a PON network deployed, and all subscribers transitioned within a short period of time. The logistics of such an upgrade would have to be managed diligently.

6. No Regrets Investments

Deploying fibre into the HFC network to decrease coax cable distances and amplifier cascade depths is a no regrets investment beneficial for each upgrade option. Additional capacity is achieved both by decreasing the number of subscribers per serving area and by increasing the spectrum and signal quality available to that group of subscribers.

Outside plant networks can be differentiated by what percentage of the network is fibre optics, with FTTP on one end and N+X on the other. Upgrade strategies that deploy fibre all the way to the subscriber or deep into the network have the advantage of creating a large surplus of capacity. A pragmatic upgrade strategy would be to deploy just as much fibre as is necessary to relieve current network deficits. There are also strategies to reduce the maximum cascade depth in the network, or deploy fibre to an intermediate depth, in areas where congestion is expected. There are reasons to pursue all the above

strategies depending on network context. The benefit of a deep fibre strategy is that when work is completed in an area, no network augmentations will be required for a significant amount of time. Any permitting or approval process is done only once, and logistics are concentrated. The drawback of this strategy is that in an environment where capital and time are constrained, the time delay between the first and last subscriber upgrade is significant. A broad fibre strategy, which reduces amplifier cascades to a set depth, would allow capital and time to be spread more evenly throughout the network, increasing the capacity to all subscribers. This approach can be useful in upgrading the network to compete with FTTP offerings in a relatively short time frame.

As shown in Figure 27, upgrade strategies can be placed on a plot of capacity versus fibre depth. Capacity is considered as the sum of upstream and downstream, while fibre depth is subjectively defined, with current HFC networks considered low fibre depth and FTTP considered high. This is a relative judgement as a very high percentage of current HFC networks are fibre.

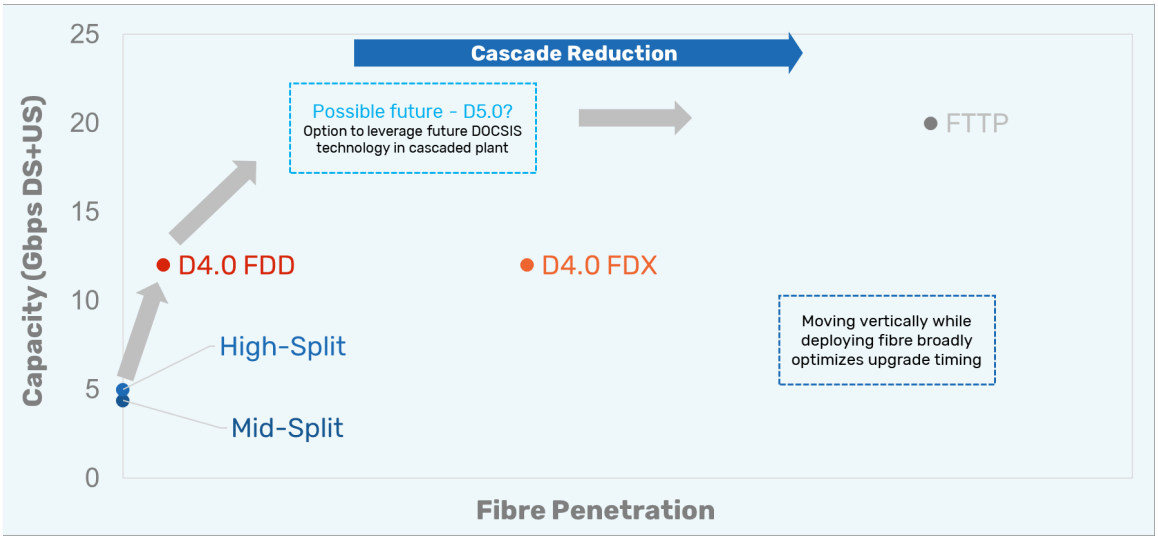


Figure 27 – Capacity vs Fibre Penetration

A deep fibre strategy moves diagonally to the top right on this plot, enabling FDX or FTTP by increasing capacity and fibre depth. A broad fibre approach moves in a near-vertical fashion, reducing cascades over time while using drop-in upgrades to create more capacity. The strategic direction will be influenced by future predictions of the direction of technology, fibre deployment costs, and subscriber demand.

A broad fibre strategy benefits if DOCSIS versions beyond DOCSIS 4.0 FDD are created that continue to operate in cascaded plant. If the direction moves toward passive operation, the deep fibre strategy is optimal. It is best to deploy fibre as quickly as possible if future fibre deployment costs, which are heavily weighted toward labor, increase significantly. An advantage of the broad fibre strategy is that capacity growth can more closely match demand growth, moderating or accelerating as needed.

Another no-regrets investment is the transition of video from QAM to IP delivery. As serving group sizes get smaller, a QAM broadcast video model becomes less and less efficient and the capacity increase achieved by moving to IPTV more compelling.

Lastly, pre-seeding the network with future-proof CPE will ensure that new capacity is available to subscribers as soon as an upgrade is complete. Switchable diplex filters can be used to ensure there is no interference from new CPE.

7. Potential Path

One potential path is to begin upgrading the plant to high-split as soon as the mid-split upgrade is complete. This upgrade would not be undertaken to reduce congestion as mid-split capacity is forecasted to satisfy customer demands for the next few years. Instead, a high-split upgrade would demonstrate HFC's ability to achieve gigabit upstream tiers and could be used to match or exceed an FTTP competitor's high service tier of 1.5Gbps downstream and 940Mbps upstream. A high-split upgrade can also pave the way for future DOCSIS 4.0 upgrades through the development of leakage detection, DAA, QAM reclaim, and other foundational components required for both solutions.

As soon as DOCSIS 4.0 FDD equipment is commercially available, the switch can be made from a high-split upgrade to a DOCSIS 4.0 FDD upgrade but operated with a high-split band plan. In parallel, programs to reduce cascade lengths and retire QAM video hardware will increase the capacity available to subscribers. When QAM video is retired a significant amount of spectrum will be freed up, which can be used for additional downstream capacity or to increase upstream capacity via a change to one of the ultra-high-split band options.

In new developments, FTTP will be the technology of choice, slowly becoming a larger fraction of the overall network. As fibre is deployed into the network to reduce cascades, the amount of work required to upgrade to N+0 or FTTP will be reduced, making such a move more practical in the future.

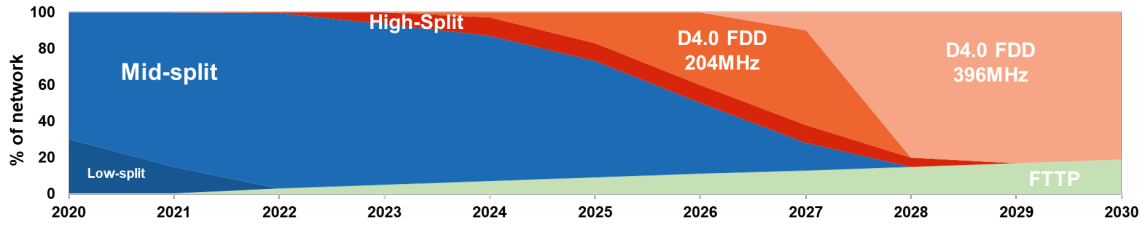


Figure 28 – Potential Network Path

8. Conclusion

Broadband subscriber demand continues to grow, as does competitive pressure, obliging network providers to upgrade their networks. The upgrade strategy network providers choose will impact capital intensity and competitive position. One strategy is to upgrade the HFC network in a drop-in fashion, allowing for a broad increase in capacity while minimizing time and capital spent. This allows time for a measured fibre deployment, slowly bringing the network toward N+0 or FTTP.

While DOCSIS 4.0 FDD is not currently available, starting an upgrade with high-split allows for gigabit symmetric tiers and requires operators to tackle some key issues such as leakage, DAA, QAM reclaim and MoCA coexistence. Once DOCSIS 4.0 FDD is available, focus can shift, initially to a DOCSIS 4.0 FDD upgrade with a high-split band plan and eventually to an ultra-high-split band plan. This strategy maximizes return on investment while efficiently competing with FTTP challengers.

Abbreviations

CAGR	Compound Annual Growth Rate
CM	Cable Modem
CPE	Customer Premises Equipment

DAA	Distributed Access Architecture
dB	decibel
DOCSIS	Data over Cable Service Interface Specification
DPI	Deep Packet Inspection
DS	downstream
DSL	digital subscriber line
FDX	full duplex
FTTP	fiber to the premises
Gbps	gigabit per second
GHz	gigahertz
GPON	Gigabit Passive Optical Network
HFC	Hybrid Fibre Coax
IP	Internet Protocol
IPTV	Internet Protocol Television
Mbps	megabit per second
MoCA	Multimedia over Coax Alliance
MPEG	Motion Picture Experts Group
ms	millisecond
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
OLT	Optical Line Termination
OOB	out-of-band
PHY	physical
PON	Passive Optical Network
QAM	Quadrature Amplitude Modulation
SNR	signal-to-noise ratio
Tbps	terabit per second
UHS	ultra-high-split
US	upstream
XGS-PON	Ten Gigabit Symmetric PON

Bibliography & References

[1] Cisco – Annual Internet Report (https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html), updated March 9, 2020.

[2] CableLabs - Data-over-Cable-Service Interface Specifications DOCSIS 2.0 Radio Frequency Interface Specification

[3] CableLabs – Data-over-Cable-Service Interface Specifications DOCSIS 3.0 Physical Layer Specification

[4] CableLabs – Data-over-Cable-Service Interface Specifications DOCSIS 3.1 Physical Layer Specification

[5] CableLabs – Data-over-Cable-Service Interface Specifications DOCSIS 4.0 Physical Layer Specification

[6] CBC News – Netflix, Bell Media reduce video quality to lower internet bandwidth use (https://www.cbc.ca/news/entertainment/streaming-services-reduce-quality-1.5512596) Mar 27, 2020

[7] Bell – 2020 Annual Report (https://www.bce.ca/investors/AR-2020/2020-bce-annual-report.pdf)

[8] Telus – 2020 Annual Report (https://assets.ctfassets.net/rz9m1rynx8pv/RhDeVJUMvjqrFwVycTU5L/9465279e0154bccef0d00c926b3794bf/TELUS_2020_annual_report-acc.pdf)

[9] ITU-T G.984.2 Gigabit-capable passive optical networks (GPON): Physical media dependent layer specification

[10] ITU-T G.9807.1 10-Gigabit-capable passive optical network (XGS-PON)

[11] Wi-Fi Alliance - https://www.wi-fi.org/discover-wi-fi/wi-fi-certified-6

[12] Digital Broadband Delivery System: Out of Band Transport Part 1: Mode A – SCTE-ISBE Standards 2019

[13] To High Split & Beyond: The New Frontier in Leakage Detection – Kyle Hohman, Shaw Communications, SCTE Fall Technical Forum, October 2020

[14] Industry Canada - Spectrum Management and Telecommunications Interference-Causing Equipment Standard – Cable Distribution Networks, June 2015

[15] Multimedia over Coax Alliance – MoCA 2.0 Specification for Device RF Characteristics, 2015

Measuring DOCSIS 3.1 & 4.0 Capacity: It “HERTZ”!

An Operational Practice prepared for SCTE by

Claude Bou-Abboud
Sr. Director
Comcast
1800 Bishops Gate Blvd.,
Mt. Laurel, NJ 08054
609-685-3782
claude_bouabboud@comcast.com

Priyan Sarathy, Sr. Manager, Comcast

Ganesh Chandrasekaran, Principal Engineer, Comcast

Alexandru Tufescu, Sr. Engineer, Comcast

Santosh Dadisetti, Sr. Engineer, Comcast

1. Introduction

In the late 1990s, when the only number after the DOCSIS (Data Over Cable Service Specification) acronym was "1.0," measuring capacity and throughput, i.e., speeds, was a straightforward matter, limited to one channel in each direction: 40 Mbps Downstream and 10 Mbps Upstream. Now, with DOCSIS 3.1 and 4.0, we can bond multiple channels to support amazing multi-gigabit upstream and downstream speeds – enabling network capabilities that will benefit broadband consumption categories spanning the Internet of Things (IoT) and virtual reality, interactive video conferencing, health care and remote learning, among others.

However, with advanced performance comes advanced complexity, and that's especially true with respect to how we measure broadband capacity and throughput. This is especially noteworthy as underlying technologies become more deliberately dynamic, able to shift from one carrier or subcarrier to another, when connectivity becomes impaired or is more optimal in an adjacent or nearby spectral location. Measuring capacity and throughput in today's advanced stages of DOCSIS is trickier and more complicated because each bonded channel's frequency and modulation type can vary, based on spectrum availability and HFC network reliability.

This paper will address the new challenges in capacity and throughput measurements. It will examine how one operator measures capacity now, versus how those measurements are made or will be made, in more dynamic spectral configurations. It will also explore the capacity measurement methods and requirements for Full Duplex (FDX) and Extended Spectrum DOCSIS (ESD) techniques.

2. The Increasing Complexity of Capacity Management

2.1. Spectrum Range and Capabilities

Coaxial Cable lives on spectrum frequencies (measured in Hertz) located between 5 MHz and a range of upper boundaries (750 MHz, 1 GHz, 1.2 GHz, and future 3 GHz configurations are envisioned.) Figure 1 shows a capacity and services timeline for the cable industry, dating from the pioneering systems of the late 1940s, to today's growing service mix of voice, video, data and beyond. Capacity-wise, in the 1960s, the maximum frequency reached on coax was 240 MHz; now, cable systems operate at 1218 MHz (1.2 GHz) and beyond. The spectrum is traditionally divided into 6 MHz channels, a now-vestigial segmentation associated with analog television signals. In the beginning, there were three networks carried on three channels: ABC, CBS and NBC. After that, the name of the game was capacity expansion, to transmit dozens of analog TV channels one way, downstream, to homes.

In the 1990s, and coincident with return path / reverse spectrum activation, cable modems were introduced, in part to replace dial–up telephone connections to what was then a very new world wide web. It was those cable modems, and the broadband connections they enabled, that gave rise to so many "broadband native" video streaming, online shopping, and all things done over Internet Protocol (IP) connections. In the 2018 timeframe, for instance, over-the-top (OTT) video, social media services and online gaming drove substantial growth in broadband consumption and catalyzed our industry of connectivity providers to advance our capacity, throughput, and latency. Now, DOCSIS 3.1 enables multi-gigabit broadband speeds over cable spectrum upto 5 Gbps and can combine multiple 6 MHz channels into one large bond, which is what enables multi-Gig services. DOCSIS 4.0 further increases the downstream RF frequency range by using two methods: Full Duplex DOCSIS (FDX) and Expanded Spectrum DOCSIS (ESD). Newer Orthogonal Frequency Division Multiple Access (OFDMA) techniques

designed to accompany the rollout of FDX, ESD and DOCSIS 4.0-based components will provide additional bandwidth up to 10 Gbps. OFDMA will also advance how we measure overall throughput and capacity, which will be discussed later.

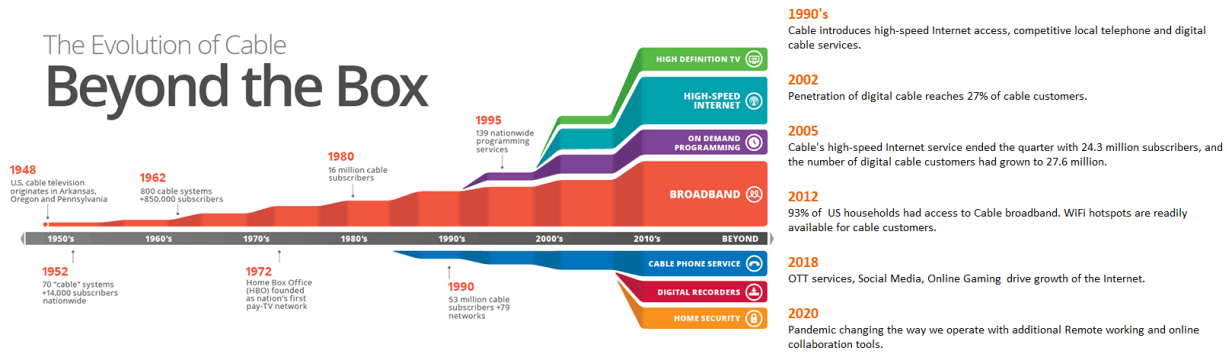


Figure 1 – Evolution of Cable (CCTA)

2.1.1. Coaxial Cable and the Frequency Landscape

Even though fiber is faster than cable capacity wise, coaxial cable is the faraway winner, in terms of capacity to cost ratio – the fire hose to the drinking straw, relative to telco twisted-pair. That said, coaxial does carry its own nuances. The upstream signal path in particular is perpetually nuanced by signal/noise that can overlap with downstream frequencies -- especially as the upstream path is eventually widened to accommodate mid-split, high-split and other upper spectral boundaries beyond the traditional top of 42 MHz.

Expanding the upstream band also allows a buffer for noise reduction and allocates more spectrum for higher speeds -- with frequencies ranging between 5 MHz - 396 MHz (from between 5 MHz – 42 MHz), while the downstream could reach up to 1.8 GHz or 3 GHz, depending on the use of Full Duplex DOCSIS (FDX) or Extended Spectrum DOCSIS (ESD.) Current industry plans to increase the upper boundary of the downstream spectrum to 1.2 GHz allows FDX to play a major role in simultaneous bidirectional transmissions on a signal carrier.

One view of a spectrum progression is shown in Figure 2, below. The top line illustrates a traditional 750 MHz configuration, with a 5-42 MHz return path. Below it, in yellow, the downstream spectrum above 750 MHz is illustrated as potentially usable for downstream transmissions, and the upstream is widened to a mid-split from 5 MHz – 85 MHz. The third line illustrates a high-split return path, from 5 MHz – 204 MHz, guard band between 204-258 MHz, and a combination of video, VOD, DOCSIS 3.0, DOCSIS 3.1, and reserved/unused bandwidth above 1 GHz. The fourth line, with the cutout detail below it, shows an ultimate spectral allocation with simultaneous bidirectional traffic moving between 300 MHz and 644 MHz. In that zone, traditional digital video, VOD and some DOCSIS traffic moves over the same carriers at the same time as FDX and DOCSIS 4.0 traffic. Clearly, as spectral configurations grow and shift to accommodate the ceaseless growth of broadband consumption, our mechanisms for measuring capacity must adapt and grow as well.

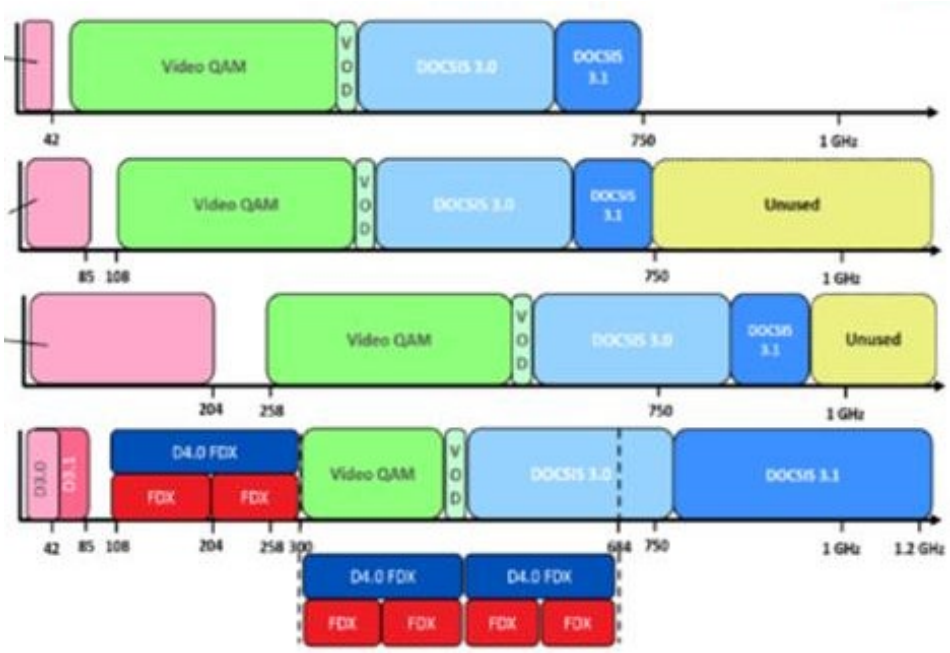


Figure 2 – HFC Spectrum

2.1.2. Hybrid Fiber Coax (HFC) Plant Issues

Attenuation, or the loss of signal strength, is a factor when measuring network capacity. The outside hardline coaxial cables typically attenuate at a rate of 2.4 dB/100 ft at 1.8 GHz, and 3 dB/100 ft at 3 GHz. By contrast, drop cables connecting to homes, and in-home RJ-6 cables, can attenuate at a rate of 12 dB/100 ft at 3 GHz. In general, the older the cable, the less it can carry, in terms of higher frequencies and long distances. In some cases, taps, amplifiers and nodes will need to be replaced to support an upper spectral boundary of 1.8 GHz. In other cases, the amplifiers could be replaced with fiber deeper nodes (such as to achieve N+1 or N+0 amplifier configurations).

2.2. Upstream and Downstream Frequencies

2.2.1. Upstream Boundaries

The upstream frequency band from 5 MHz to 42 MHz typically carries 4 Single-Carrier Quadrature Amplitude Modulators (SC-QAMs) that can deliver an aggregate 100 Mbps. Implementing an upstream mid-split to 85 MHz increases the aggregate capacity to 8 SC-QAM channels, which can deliver up to 200 Mbps that are usable by D3.1 modems.

2.2.2. MAC Layer vs. Physical Layer Capacity

One measurable difference between the Media Access Control (MAC) Layer and the Physical (PHY) Layer is the DOCSIS overhead. The Profile Management Application (PMA), which was developed by Comcast in 2019, and consequently shared with the industry, is vital for generating the D3.1 profiles for OFDM channels. In a capacity measurement context, PMA enhances part of the overhead efficiency to increase the usable bandwidth by ~15%, on average, based on a mix of modem types. The dynamic changes in the profile range between 10% to 25% in improvement. Table 1 is a PMA example using 6.4 MHz channels, where profile 271, when dynamically experiencing changes such as improved plant conditions/reduced noise, or a reduction of the number of connected modems, will proactively shift to profile 251, thus gaining an additional 5 Mbps of capacity.

Table 1 – Profile Management Application (PMA) example

Profile Template for 6.4 MHz Channels

251	256	261	266	271
25.6 Mbps 64-QAM short: 97/2, burst=5 long: 247/4 snr at 1% uccw = 22.8 dB	24.5 Mbps 64-QAM short: 91/5, burst=5 long: 239/8 snr at 1% uccw = 22.1 dB	23.3 Mbps 64-QAM short: 105/10, burst=6 long: 229/13 snr at 1% uccw = 21.2 dB	22.5 Mbps 64-QAM short: 99/13, burst=6 long: 223/16 snr at 1% uccw = 20.5 dB	20.7 Mbps 64-QAM short: 99/13, burst=6 long: 121/16 snr at 1% uccw = 20.3 dB

Based on the table 1 we observe ~9% improvement for Profile 266, ~12.5% for Profile 261, ~18% for Profile 256 and up to ~24% improvement for profile 251. In actual deployments we have observed ~15% improvement on average when PMA has been deployed in the field.

2.2.3. Frequency Bands, Modulation and Spectrum Allocation

One current approach in capacity is to increase the HFC spectrum to 1.2 GHz, allowing FDX combined with D3.0 and D3.1channels to achieve a D4.0 bandwidth supply of up to 4.2 Gbps for the US and 7.6 Gbps for the DS. The modulation and spectrum enhancements intrinsic to FDX and ESD will allow even higher Multi-Gig services. This is depicted in Figure 3.

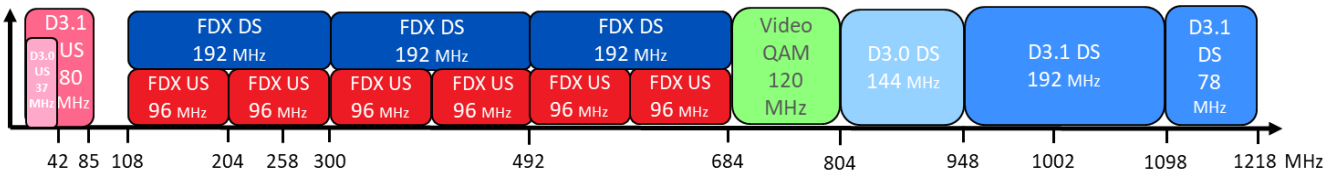


Figure 3 – D4.0 FDX HFC Spectrum

2.3. Multi-Gig Services That Will Likely Impact Capacity Measurements

Demand for higher speeds is increasing, as it has since the very onset of cable modems and broadband connections, in the mid-1990s. The current amount of bandwidth consumed per device is continuously

on the rise. Many applications and services are likely to require multi-Gig speeds, such as those within the Internet of Things (IoT), Virtual Reality (VR), Augmented Realty (AR), cloud gaming and streaming. Further discussion of these anticipated growth engines is out of scope for this paper, other than to note that anything that impacts the demand for network capacity must be considered and modeled, to the extent possible, when measuring throughput and capacity.

Downstream speed tiers reaching 1 Gbps using D3.1 are expected to expand to 10 Gbps with D4.0, which will yield ample bandwidth that can be applied to both foreseen and unforeseen applications.

2.4. Capacity Measurement Methods

As mentioned previously, in the "good old days" of early DOCSIS networks, with static channel widths and fixed modulation mechanisms (variations on QAM, from 64-QAM to 256 QAM in the downstream, and QPSK in the upstream), it was reasonably straightforward to predict capacity readings: The modulation rate on a particular channel width yielded a specific number of bits/Megabits per second.

As our broadband networks continued to advance and were deliberately engineered to stay ahead of consumer and business demand, the "simple math" of capacity measurement became part of a far more complex assessment. Now, multiple channel widths and a wider range of advanced modulation makes capacity measurement a trickier proposition.

2.4.1. Capacity Measurement Methods for OFDM Carriers

Currently, and because OFDM channels are primarily available on the downstream signal path, we have been collecting the percent utilization values reported by the CMTS based on demand. For SC-QAM, D3.1 modems can access both SC-QAM channels and OFDM, so they benefit from measurably higher throughput.

When high speeds are required for certain applications, or to run a speed test, the DOCSIS 3.1 modem can utilize both SC-QAM and OFDM spectrum. Since there is no means to determine the combined throughput between the two metrics, the actual available capacity can be determined based on the available capacity -- after deducting the 98th percentile SC-QAM used bandwidth from the total capacity, to remove anomalies, as well as the 98th percentile OFDM utilization. This is achieved by converting it to bandwidth based on the most prevalent profile used during a designated polling hour, then deducting that amount from the capacity of that profile. The sum of these two provides the available bandwidth during peak hours – in a sense, an erlang for broadband capacity. This matters to ensure that we plan for augments and have enough capacity to support our advertised speeds. This will be discussed later.

2.4.2. Orthogonal Frequency-Division Multiplexing (OFDM)

Combining multiple 6 MHz channels into one larger band (192 MHz) avoids Inter-Symbol Interference (ISI) and improves Signal-to-Noise ratio (SNR), which impacts capacity measurement in a beneficial (to capacity) way. OFDM traffic combined within SC-QAM channels provides an ample amount of BW to allow multi-Gig transmissions.

2.4.3. Orthogonal Frequency-Division Multiple Access (OFDMA)

OFDMA uses time slot and frequency allocation for D3.1 modems and higher. Each modem will be given a time slot or group of subcarriers to send the data. Excellent coverage of the topic of OFDMA for

upstream capacity gains can be found in the 2020 SCTE Expo paper titled "Field Experiences with US OFDMA and Using US Profile Management." As well, substantial detail about improvements to the upstream signal path are expected as part of this year's Fall Technical Program. For those reasons, we will say only that from a capacity management perspective, the inclusion of OFDMA adds another layer of complexity to the capacity measurement task.

It represents an active area of exploration.

2.5. Tools and Methods to Measure Utilization

Currently, we measure capacity by regularly polling the CMTS for network activity from channels modulated with SC-QAM, as well as channels or bonded channels modulated with OFDM. In the case of the upstream signal path, the four SC-QAM channels are polled every five minutes, and the results aggregated. The total value for those four channels is then adapted from a daily or weekly view, depending on report type, and the last two percent removed, to get to a 98th percentile reading. The top 2% is removed as a countermeasure against network anomalies.

2.5.1. Cable Modem Consumption

Measuring hourly consumption in bytes per service flow at the cable modem level helps in identifying "top talkers" and the contribution during peak hours. Peak hours are based on the highest consumption of the averaged top 20 hours per month. Figure 4 shows the US and DS speed tier consumption at peak hours by service type. Modem consumption by speed tier is crucial to evaluate the growth and projection for multi-Gig speeds.

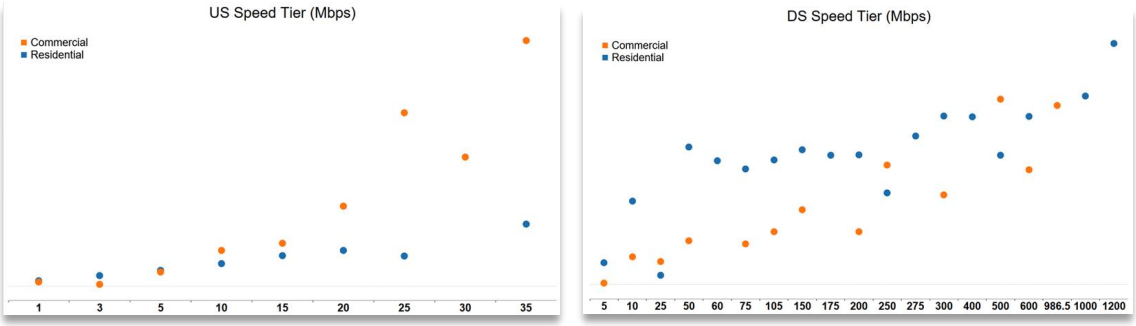


Figure 4 – BW Consumption During Peak Hours by Speed Tier

2.5.2. Traffic Aggregation and Concurrency

Aggregating hourly consumption during the top-twenty hours per month, per modem, at the CMTS Service Group (SG) and CMTS Bonding Group (BG), provides the behavior by demographics and shows any heavy usage by SG and BG. This method helps in better augmentation approaches and predictive analysis. It also aids in capacity measurements.

2.5.3. Reporting Frequencies and Metrics Needed

We use the iCMTS/vCMTS pollers to collect the average 5-minute polls, daily, over a month period, to calculate the 98th percentile values. The reported values are based on daily, weekly, and Monthly periods at a SG/BG level.

The metrics collected are as follows:

- In-bps, out-bps converted into percent utilization
- Up-channel utilization
- OFDM percent utilization
- OFDMA percent utilization
- PMA profiles and capacity
- Interfaces' speeds

Cable modem pollers provide the following metrics:

- Cable modem type
- Bootfile specifications
- CMTS mapping
- SNR, MER

Additional metrics collected are:

- HHP per node, per zip code
- Node-to-CMTS mapping

The combination of these metrics is utilized to address network augmentations, forecasting, and DOCSIS 4.0 capacity.

2.5.4. Future of Measurements

More trials and higher D4.0 penetration will be needed, in addition to methods to report on FDX. Profile types by interface are necessary to collect 5 minutes utilization for OFDM/A.

2.6. DOCSIS 4.0, Full Duplex DOCSIS (FDX) and Extended Spectrum DOCSIS (ESD)

2.6.1. Boundaries of D3.1

DOCSIS 3.1 modems max out at 1.2 GHz, where one (relatively large) OFDM channel could be utilized. This sets the downstream speed at 2 Gbps and the upstream speed at 300 Mbps (and potentially 600 Mbps where OFDM/A is applicable.)

2.6.2. Bandwidth Boundaries and CM Populations (D2, D3,…)

Currently, our network, like that of most operators, operates a range of cable modems, some based on DOCSIS 3.0 and others on DOCSIS 3.1. Currently, most of our footprint uses D3.0-based CMs, with the rollout of D3.1-based modems poised to shift the balance in a reasonably swift fashion. To fully and efficiently utilize an advanced, future-state network. outfitted with OFDM/A and FDX/ESD, and to benefit from the additional BW, that same trajectory will be required, to shift the balance again, from D3.1-based CMs to D4.0-based CPE.

Plus, because the maximum capacity of D2.0 is limited to 40 Mbps DS and 30 Mbps US, modems based on that DOCSIS version will ultimately need to be replaced.

For D3.0-based modems, the number of DS SC-QAM channels varies from 8 to 32, resulting in a capacity range from 320Mbps to 1.2 Gbps. Similarly, US capacity ranges from 30 Mbps to 120 Mbps. This is summarized in Table 4, below.

Table 2 – Modem Types and Bandwidth

Modem Type	Max. Speed	SC-QAM Bonding	OFDM/A Capable?
D2.0	40 Mbps DS 30 Mbps US	1 DS 1 US	NO
D3.0	1 Gbps DS 120 Mbps US	8 DS 4 US	NO
D3.1	10 Gbps DS 1 Gbps US	32 DS 8 US	YES
D4.0	10 Gbps DS 6 Gbps US	44 DS 8 US	YES

2.6.3. Differences between D3.1 and D4.0

A very important reason why measuring broadband capacity is sometimes an "It Hertz" proposition can be seen in Figure 5, below, which shows spectrum ranges and modem capabilities. The expansion of spectrum frequency is the only means to obtain additional throughput:

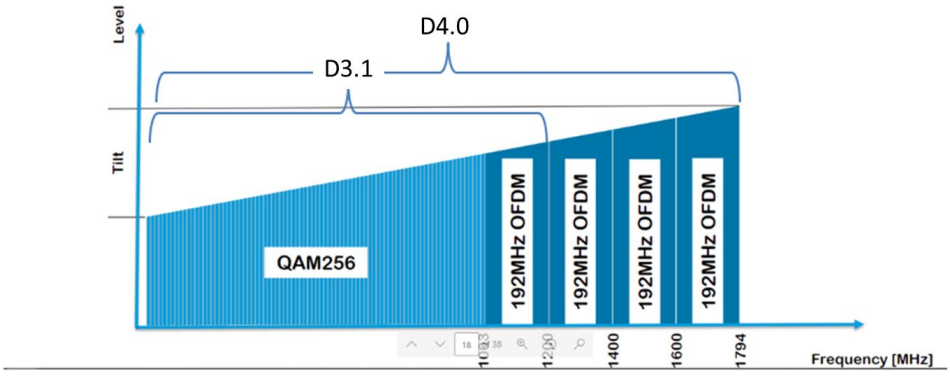


Figure 5 – Frequency Expansion

DOCSIS 3.1 can function at a maximum frequency of 1.2 GHz, with OFDM channels, while DOCSIS 4.0 can utilize higher spectrum up to 1.8 GHz, thus combining multiple OFDM channels to reach multi-Gig speeds.

2.6.4. Why is D4.0 needed?

Higher speeds, like 10 Gbps downstream and 6 Gbps upstream, are only possible with DOCSIS 4.0, also known as Full Duplex DOCSIS (FDX). This will coincide with plans by operators to drive optical fiber deeper into their networks, go higher spectrally, and/or to provide sufficient capacity to the Remote PHY Devices (RPDs).

Although this paper didn't go into substantial detail, it is worth repeating that more bandwidth, wider spectrum, and higher efficiency (lower latency and faster speeds) will be required to meet the demand for known and unknown broadband services and service categories (e.g. IoT, the "metaverse," and so on.)

3. Conclusion

Measuring network capacity is an activity that is vital to strategic planning, especially as broadband consumption continues to steadily grow. In the earliest days of broadband, such calculations were a relatively straightforward matter related to modulation type and channel width. As network capabilities advanced, however, the very technologies that enabled the advancements simultaneously increased the complexity for capacity planners.

DOCSIS 4.0 traffic, for instance, is especially convoluted – so, from a complexity standpoint, "it Hertz". In addition, the dynamic nature of the Profile Management Application (PMA) introduces additional challenges, in that it is vastly more challenging to count when channel widths and modulation types are hopping around, on the fly, in pursuit of optimal transmission paths. While many of the DOCSIS 3.1 techniques of measuring traffic apply, an aggregation of the OFDM interface metrics is required for DOCSIS 4.0.

Upgrading the network to 1.2 GHz and 1.8 GHz by replacing taps, amplifiers and nodes is a major effort, as is stairstepping D2.0 and D3.0 modems with D3.1 and D4.0 gear. With FDX and ESD on the near-term horizon to further increase the capacity of the access network, it is paramount that we make sure our capacity measurement techniques keep pace.

The good news is, over time, these network improvements will facilitate the periodic measurement of a balanced demand at SC-QAM, and OFDM/A levels combined. Tracking Latency/Jitter and SNR/MER proactively will further eliminate network issues. Along the way, we will need to also consider how to blend WAN, CPU, Memory, and other metrics beyond DOCSIS into the capacity monitoring mix, as all are or will be essential to the overall capacity planner's work.

Abbreviations

AP	Access Point
AR	Augmented Reality
bps	bits per second
BW	Bandwidth
CMTS	Cable Modem Termination System
CPU	Central Processing Unit
D2.0	DOCSIS 2.0
D3.0	DOCSIS 3.0
D3.1	DOCSIS 3.1
D4.0	DOCSIS 4.0
dB	Decibels
DOCSIS	Data Over Cable Service Interface Specification
DS	Downstream
EoHFC	Ethernet over HFC
ESD	Extended Spectrum DOCSIS
FDX	Full Duplex DOCSIS
FEC	Forward Error Correction
GHz	Giga Hertz
Gpbs	Gigabits per second
HD	High Definition
HFC	Hybrid Fiber Coax
Hz	Hertz
IoT	Internet of Things
ISBE	International Society of Broadband Experts
ISI	Inter-Symbol Interference
MAC	Media Access Control
Mbps	Mega Bits per Second
MER	Modulation Error Rate
MHz	Mega Hertz
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	OFDM Access
PMA	Profile Management Application
SC-QAM	Single-Carrier Quadrature Amplitude
SCTE	Society of Cable Telecommunications Engineers
SG	Service Group
SMB	Small Medium Business
SNR	Signal-to-Noise Ratio
US	Upstream
VR	Virtual Reality
WAN	Wide-Area Network

Bibliography & References

CommScope: Extended Frequency Performance of Coaxial Cable

Corning/Broadband SP: Get Ready – 'Cause here it Comes DOCSIS 4.0 (D. Kozischek, J. Burton)

CCTA: History of Cable

CableLabs: DOCSIS® 4.0 Technology Realizing Multigigabit Symmetric Services (D. Jones)

RF Wireless World: Difference between DOCSIS 3.0 DOCSIS 3.1 and DOCSIS 4.0

QualComm: VR and AR pushing connectivity limits

The Workshop- Mangiante: VR is on the Edge: How to Deliver 360° Videos in Mobile Networks

Google Stadia: Internet usage (https://support.google.com/stadia/answer/9607891?hl=en#zippy=)

StreamLabs: Q3 2020 Live Streaming Industry Report (https://streamlabs.com/)

ANSI C63.5-2006: *American National Standard Electromagnetic Compatibility–Radiated Emission Measurements in Electromagnetic Interference (EMI) Control–Calibration of Antennas (9 kHz to 40 GHz)*; Institute of Electrical and Electronics Engineers

The ARRL Antenna Book, 20th Ed.; American Radio Relay League

Code of Federal Regulations, Title 47, Part 76

Reflections: Transmission Lines and Antennas, M. Walter Maxwell; American Radio Relay League

Message Queuing Telemetry Transport (MQTT) For IoT Devices: Less is More

A Technical Paper prepared for SCTE by

Sweety Bertilla
Senior Android Engineer
Comcast Cable
518 Highland Ave, Cherry Hill, NJ, 08002, USA
+1-267-785-3798
sweetybertilla_francisxavier@cable.comcast.com

Kristopher Linquist
Principal II Engineer
Comcast Cable
1050 Enterprise Way #100, Sunnyvale CA, 94089
+1-408-940-5747
kris_linquist@comcast.com

Robert Farnum
Principal II Engineer
Comcast Cable
13029 Titus Court, Austin, TX, 78732, USA
+1-512-860-6166
robert_farnum@comcast.com

1. Introduction

Message Queuing Telemetry Transport (MQTT) is a well-designed, lightweight messaging protocol that can be used for communication between mobile clients, microservices, and IoT devices. Unlike HTTP (Hypertext Transfer Protocol) and other messaging protocols, MQTT is a low bandwidth, low latency alternative for IoT device transmissions, which is far more suitable because these devices may operate within tiny bandwidth, power, and transmission footprints. MQTT uses publish/subscribe operations to exchange data between client and server -- meaning an IoT device (or any other client) "subscribes" to a topic and asynchronously receives messages when data is published on that topic.

Also, unlike HTTP, this method saves a substantial amount of time previously spent on polling, which makes updates occur more quickly and smoothly. The lightweight nature of MQTT helps the end user – or customer - receive messages even when they are in low bandwidth situations, such as when traveling in areas with limited connectivity. Quality of Service (QoS) features supported by MQTT help clients opt into the level of service based on network reliability.

In this paper, attendees will learn how and why Comcast opted to adopt the MQTT protocol in customer-facing applications such as the Xfinity Application when connecting and bridging communications with IoT devices such as Philips Hue, LIFX, Ecobee, August door looks, and other Zigbee devices (via a Residential Gateway) using the AWS IoT Core as the MQTT message broker.

2. History of MQTT

MQTT was created in 1999 by Andy Standford-Clark (IBM) and Alen Nipper (Arcom, now Eurotech) as part of the IBM MQ series of products. The goal was to invent a new protocol for connecting oil pipelines over unreliable satellite networks. In 2011, IBM and Eurotech donated MQTT to the proposed Eclipse project called Paho. In 2013, MQTT version 3.1 was submitted to OASIS (Organization for the Advancement of Structured Information Standards). On October 29, 2014, MQTT 3.1 was approved. Since then, MQTT 3.1.1, currently the most supported version, was ratified in 2016, and MQTT 5.0, which supersedes MQTT 3.1.1, was ratified in 2019. MQTT is also published as an ISO/IEC 20922 standard reference.

A frequent question concerning MQTT that arises is, "What does the acronym stand for?" The answer is debatable. Given its source, the "MQ" in "MQTT" is historically based on the name of the IBM "MQ" Series product line. Others say "MQ" means message queueing - but this is a misnomer, as MQTT is not a message queue, but *can* queue messages for clients. Unfortunately, the OASIS technical committee named itself "OASIS Message Queuing Telemetry Transport Technical Committee". This is likely the source of the acronym's more common definition. In any case, the messaging protocol is no longer really an acronym, but is simply recognized by the letters "MQTT."

Given the maturity and design of MQTT as a lightweight messaging protocol, it has been widely adopted for MTM (machine to machine) communications for industrial, messenger, and IoT applications. MQTT's ability to keep bandwidth requirements to a minimum -- while dealing with high latency, unreliable networks, small footprint devices, and low power consumption -- has made it an excellent choice for communications over a variety of networks between clients and/or devices with high cardinality to cloud-based platforms.

Other protocols, such as, HyperText Transfer Protocol (HTTP), Advance Message Queuing Protocol (AMQP), eXtensible Messaging and Presence Protocol (XMPP), or Websockets, while often considered as potential solutions, have fallen short of meeting some of the basic selection criteria. MQTT continues to edge out other options because of its nature as a lightweight, asynchronous, bi-directional, secure, efficient, reliable, publish/subscribe, and data-agnostic messaging protocol.

3. Advantages of MQTT

3.1. Asynchronous bi-directional communication

Asynchronous APIs matter in that they can provide significant performance and responsiveness in client-server communications. This mechanism is inherently *non-blocking,* which means that an application does not need to wait for a response in order to proceed. In the context of an application showing the status of IoT devices, a list of devices can be shown – and updates to their status will be received as devices report them, rather than pausing to wait for all devices to respond.

Messages can also be pushed directly to the client without an explicit request. Downstream microservices can choose to directly send updates to a client application – an ability they may not have had if they communicate only to an upstream synchronous API. This is useful for status reporting, user messaging, and more.

For asynchronous APIs to work, the client must maintain a persistent connection with the server.

3.2. MQTT Features that make it an ideal protocol for async messaging

3.2.1. Publish/Subscribe

MQTT uses a publisher / subscriber messaging pattern to provide a framework for exchanging messages between publishers (producers) and subscribers (consumer clients). A publisher can send a message and it will automatically be routed – via the MQTT message broker – to any client subscribed to that topic.

3.2.2. Topic structure

MQTT uses a topic hierarchy to send and receive messages. This can be thought of as the equivalent of a path or route in a HTTP API, except that they do not have to be pre-defined.

MQTT topics can consist of wildcards and topic separators. An example topic may be:

platform / customerId4 / house1 / lights / lamp4

MQTT topic levels are separated by a forward slash (/) and subscribers can use two different types of wildcards when subscribing:

+ A plus represents a single topic level wildcard.

A hash is a multi-level wildcard that can only be at the end of a topic subscription.

If a client wants to subscribe to all of customerId4's devices located in house1, they will subscribe to:

platform / customerId4 / house1 / #

If a client wants to subscribe to all of customerId4's lights in ANY of their houses, they will subscribe to the topic:

platform / customerId4 / + / lights / #

These patterns allow for significant flexibility across many types of services.

The MQTT protocol has no limits on the number of subscriptions for a given connection.

3.2.3. QoS

Quality of Service (QoS) is a key feature of MQTT. An MQTT Client gets an option to select the level of service that works for their application, based on network reliability. MQTT manages the re-transmission of messages and guarantees delivery, even when the underlying transport is not reliable. QoS makes communication in unreliable networks a lot easier. There are 3 different levels of QoS:

QoS 0 is also known as "Fire and Forget service". An MQTT client sends the message to an MQTT broker and doesn't wait for an acknowledgement, so there is no confirmation if the message is received by the end client. However, this is considered the fastest message delivery service.

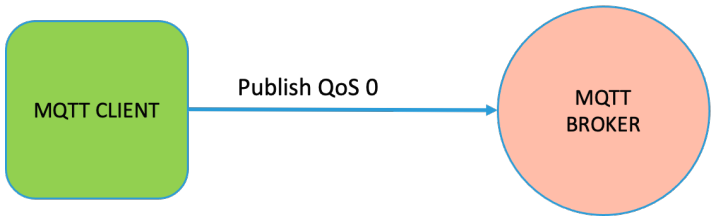


Figure 1 – Quality of Service 0(QoS 0)

QoS 1 makes sure that the MQTT message is delivered at least once to the end client. The publisher sends the message to MQTT broker and waits for an acknowledgement. If the publisher doesn't receive the acknowledgement within the given time, it publishes the message again with a duplicate flag (DUP). The MQTT broker sends the message to end clients and responds back with PUBACK packet to the publisher. The only drawback of this service is the possibility of end clients receiving the same message more than once.

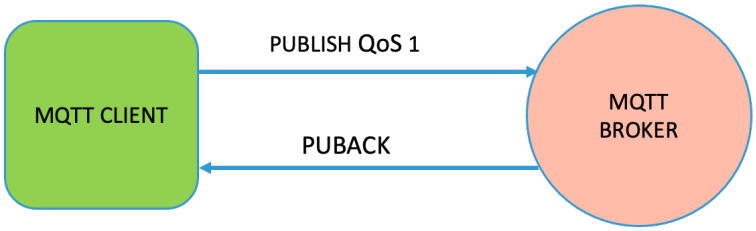


Figure 2 – Quality of Service 1 (QoS 1)

QoS 2 is also known as the safest and slowest message service. The publisher (MQTT Client) sends a QoS2 message to an MQTT broker and waits for acknowledgment. Once the publisher receives an acknowledgement (PUBREC) from the MQTT broker, it sends a PUBREL message

to the MQTT broker. The MQTT broker only sends the message to the end clients after it receives the acknowledgment (PUBREL) from publisher and sends back PUBCOMP to the publisher. Simply put, there is a four-way handshake that happens between a publisher and an MQTT broker to make sure the end client receives the message only once.

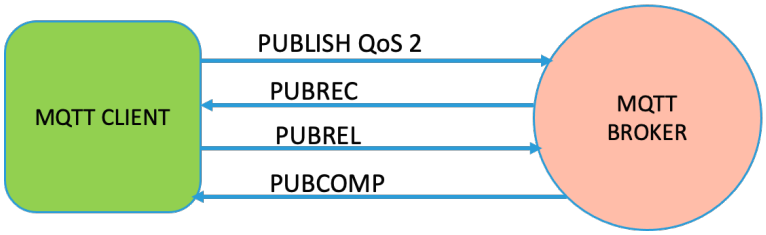


Figure 3 – Quality of Service 2(QoS 2)

The AWS IoT Core (MQTT broker) only supports QoS0 and QoS1 but not QoS2.

3.2.4. Last Will and Testament

Last Will and Testament (LWT) is a feature of MQTT that can notify other clients about a client that is disconnected without notice, often due to a lost network connection.

During the connect phase, a client registers the LWT with the broker. The broker then stores the LWT and publishes it if the client disconnects ungracefully.

LWT contains several parameters:

1. Last Will Topic (topic that the LWT message will be published to)
2. Last Will QoS (QoS of the LWT message)
3. Last Will Message (LWT message itself)
4. Last Will Retained (boolean) (whether the last lost connection information will be retained)

One example of usage in IoT devices is a *connected* flag. Immediately after connecting to a broker, an IoT device could send a *connected: true* message **and** set a LWT message that contains *connected: false*. If the IoT device's network connection is lost, the MQTT broker would send *connected: fals*e on the device's behalf, giving other clients real-time information as to whether it can expect a device to receive a command or update its status.

4. How to use MQTT for mobile or web application client

The MQTT message broker acts as medium for communication between the clients. Clients can be mobile or web application or an IoT device or microservice. Let's take a look at how a mobile application client can use MQTT for communication with IoT devices. Key actions include connect, disconnect, subscribe, unsubscribe, and publish. Key components include the MQTT broker, client(s) and topic(s).

Clients connect to the IoT services and devices through the MQTT broker using the connect method. Next, the client subscribes to topic(s) to receive messages. When a message is published to a topic, all clients that are subscribed to the topic will receive the message asynchronously, as depicted in Figure 4, below:

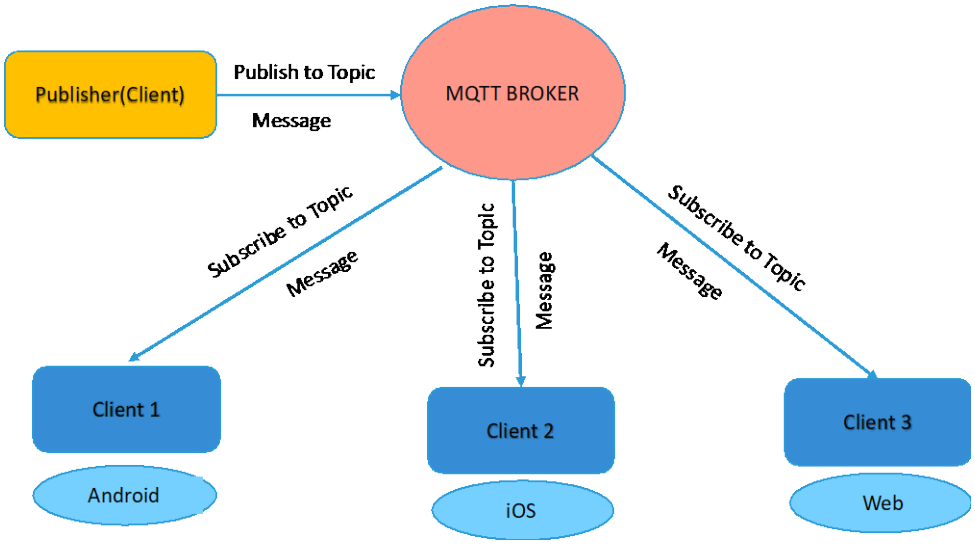


Figure 4 – High level diagram of MQTT

4.1. MQTT Connect

Application clients can establish a secure, persistent connection with an MQTT broker. Let's take a look at one of the MQTT brokers (used in Xfinity app) within AWS's IoT Core. IoT Core provides an open source SDK which can be used by the mobile client to integrate and consume the MQTT broker. AWSIotMqttManager, which is part of the SDK, provides an interface to make an MQTT connection. The AWSIotMqttManager requires the region, ClientID, and account endpoint, to which it uniquely establishes an MQTT connection.

The client ID provided by the application client is a unique identifier that represents the connection (one example could be a concatenation of a client account ID and app installation ID). "Region" is the AWS region identifier to which to connect (such as us-east-1 or us-west-2). The topic namespace is limited to an AWS account and namespace. The string accountEndpointPrefix indicates the specific customer endpoint. Example:
[prefix].iot.[region].amazonaws.com

AWS IoT Core allows a client to connect to the MQTT broker in four different ways:

1. Connect with keystore (secure location for storing cryptographic keys) and port number
2. Connect with AWS credentials provider
3. Connect with proxy host and proxy port.
4. Connect using custom authorizer. The IoT client of Xfinity Application uses custom authorizer to connect with MQTT broker. This custom authorizer is used to validate the customer against an internal identity provider.

The AWSIotMqttClientStatusCallback provides the status of the connection. Here is a sample of status call back details from a mobile client:

Connecting -> Trying to establish the connection Connected -> Successfully connection is established Reconnecting -> Trying to reconnect after a connection error ConnectionLost -> No longer client is connected to MQTT broker ConnectionError -> Error in the connection due to network or attempting to connect to a invalid topic

The attributes set as defaults can be altered by the application client. By default, the autoReconnect attribute is set to TRUE, which indicates that a reconnect attempt will be established when there is a connection error reported by the MQTT broker. A client can set the minimum and maximum reconnect time in seconds and maximum reconnect attempts. Another key attribute used is *clean session*. By default, clean session is set to TRUE, which means a non-persistent connection is established, which will not store any subscription information or undelivered messages for the client. By setting clean session to FALSE, a client can establish a persistent connection which will store subscription information and undelivered messages are delivered with QoS 1.

Table 1 – Clean Session with QoS

CleanSession Flag	QoS	Message received after Reconnect
TRUE	0	NO
TRUE	1	NO
FALSE	0	NO
FALSE	1	YES

4.2. Publish and Subscribe to Topic

Clients who subscribe to a topic will receive messages published to the topic, after a secure connection is established through MQTT broker. QoS level can be set to subscribe and publish topics. A client can also unsubscribe to topics before they disconnect. Using the AWS IoT Core, a client can publish messages as a JSON string or simple string.

4.3. MQTT and AWS Shadow State

The AWS Shadow State service, provided by the AWS IoT Core, is a persistent cache that can be used to store IoT device state. Reserved MQTT topics are provided to update this cache and to receive notification of create, update, and delete events. The two types of state in shadow state messages are *desired state* and *reported state*. Desired state lets the client send a request through an MQTT broker to the Shadow State service, where the Shadow State service calculates a *delta state*

(difference between desired and reported) and forwards the request to a device via an AWS IoT Core Rule. The device then attempts to honor the request, resets the desired state, and updates the reported state accordingly. The reported state is in turn returned via the device's MQTT topic to which the client is subscribed.

Here is a sample json payload that demonstrates desired and reported shadow state messages:

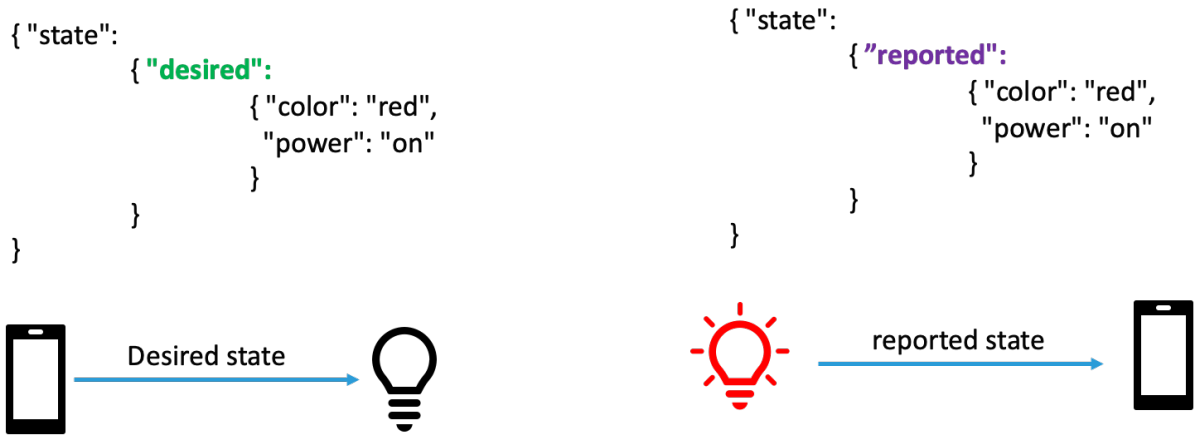


Figure 5 – desired and reported state messages

4.4. Message payload

Since MQTT is a lightweight protocol, the message payloads cannot extend in size. AWS's IoT Core supports a maximum payload of 128kb. In the Xfinity Application, the client receives the payload of one device at a time. MQTT messages also have no guaranteed ordering. Due to these limitations, we included s*equence number* and *total messages expected* to our MQTT message payloads to ensure all messages are received and ordered properly.

Shadow Reported state message payloads contain a property named *version*, which gets incremented with each change in shadow state. If the client receives a payload with version 5 and then a payload with version 4, the client can ignore the payload with version 4 as it is considered stale. Also, the desired state payload request sent by the client either needs to be same or higher version than the reported state, otherwise the AWS IoT Core MQTT broker rejects the desired state message with a *409 version conflict*. This feature, called optimistic locking, ensures that the client has the latest device status prior to trying to update the state.

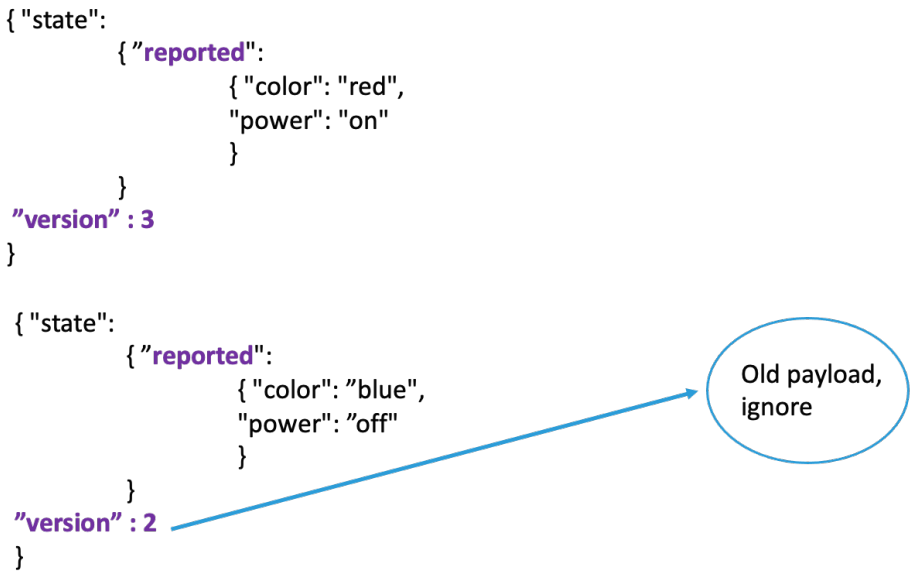


Figure 6 – reported state with old payload

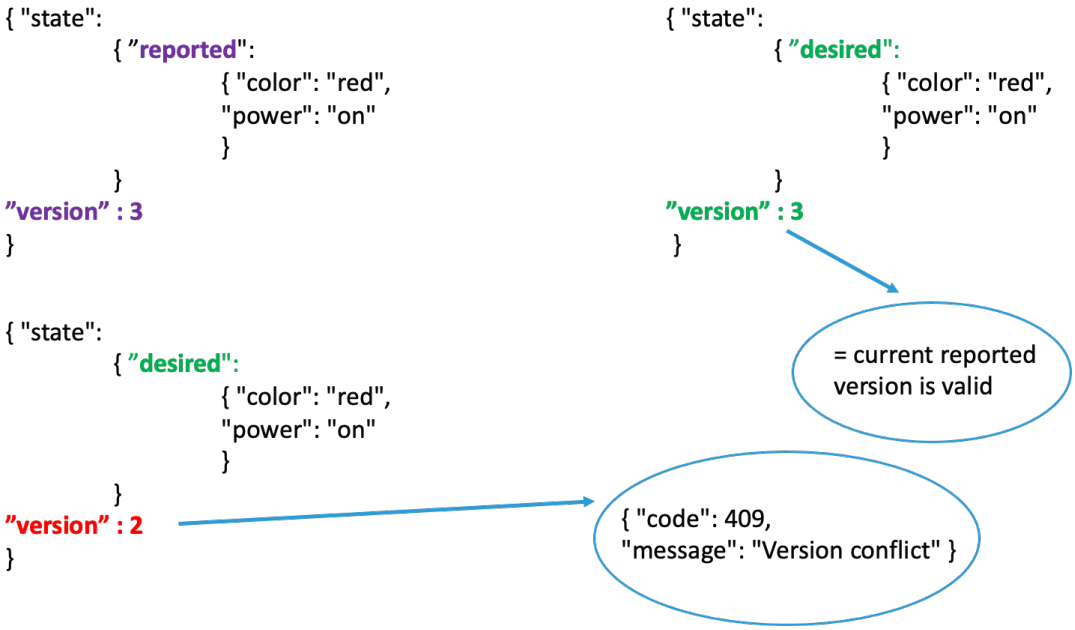


Figure 7 – desired state with previous version number

5. Security

The security of MQTT connections for a given subscriber (using mobile/web clients and residential gateways) is of the highest importance. To achieve fine-grained security controls and minimize blast radius (total impact of a security event), Xfinity leverages the capabilities provided by the AWS IoT Core MQTT Message Broker and its support for custom authorization using Java Web Token (JWT) and Just in Time Registration (JITR) with x509 Certificates. The authorizers generate AWS IoT Core Policies that allow MQTT connect, publish, subscribe and receive actions at the finest level of granularity - the MQTT topic.

5.1. Mobile Client

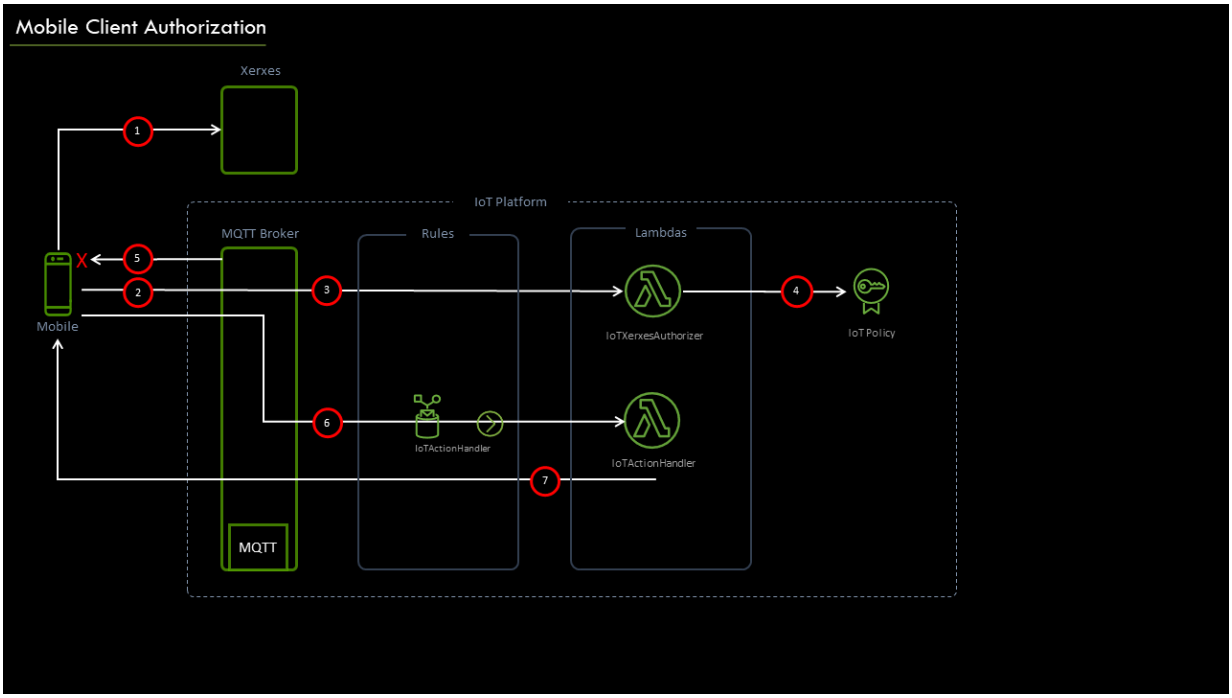


Figure 8 – Mobile Client Authorization Flow Diagram

The Xfinity Application uses an Enterprise Federated Identity Management system, known as Xerxes, to integrate with syndicated partner IdPs (Identity Providers). This system provides authentication and authorization via a Java Web Token (JWT) based Identity Token. Within the JWT there are principal attributes critical in supporting a fine-grained security model. These attributes are:

- Partner Identifier (PID) – Syndicated partner identifier for the Multi System Operator (MSO).
- Account Identifier (AID) – Identifies the subscriber's account id
- JTI – Java Token Identifier (JTI) attribute

The AWS IoT Core MQTT message broker provides a custom authorization capability which is used to verify the JWT and authorize the Mobile Client MQTT connect operation. The JWT is verified via Signature, audience attribute (aud), and issuer attribute (iss) match. Once

verified, the JWT principal identity attributes are used to generate an AWS IoT Policy that allows the MQTT Mobile Client to:

- Connect – allow the MQTT connection with a unique MQTT Client ID (CID) generated from the principle information within the JWT. This information includes the PID, AID, and the Java Token Identifier (JTI). An MQTT Client ID must be unique amongst all connected MQTT clients or the prior MQTT Client is disconnected. An example MQTT Mobile Client ID is as follows:

  ```
  xfi:[PID]_[AID]:clt:[JTI]
  ```

  This authorizes only the MQTT client matching the validated JWT to establish a connection to the MQTT Message Broker. An MQTT client with a JWT that is expired or is invalid is denied connection.

- Publish – allows the MQTT client to publish messages to MQTT topics which are name spaced based on the PID, AID and JTI. An example service request message topic is as follows:

  ```
  c/xfi/[PID]_[Account ID]/[SID]/[CID] /[RID]
  ```

  where,

  c is the root of the MQTT topic

  xfi represents the Xfinity Application part of the MQTT topic

  [PID] represents the Syndicated Partner

  [SID] represents a unique Service Identifier as the destination of the request

  [CID] represents the MQTT Client Identifier as the source of the request

  [RID] represent the type of request. This limits the ability of the MQTT Client to publish only to topics containing its principle identity and to services to which it has been granted permission to send action request messages.

- Subscribe/receive – allows the MQTT client to subscribe/receive messages to/from MQTT topic(s) which are name spaced based on the PID, AID and JTI. An example service request response topic is as follows:

  ```
  c/xfi/[PID]_[AID]/[CID]/[SID]/[RID]/[success|failure]
  ```

  where,

  c is the root of the MQTT topic

  xfi represents the Xfinity Application part of the MQTT topic

  [PID] represents the Syndicated Partner

[SID] represents a unique Service Identifier as the destination of the request

[CID] represents the MQTT Client Identifier as the source of the request

[RID] represent the Response Identifier or type of response. This limits the ability of the MQTT Client to subscribe and receive action response messages only on topics containing its principle identity and from services to which it has been granted permission to receive action response messages.

5.2. Residential Gateway

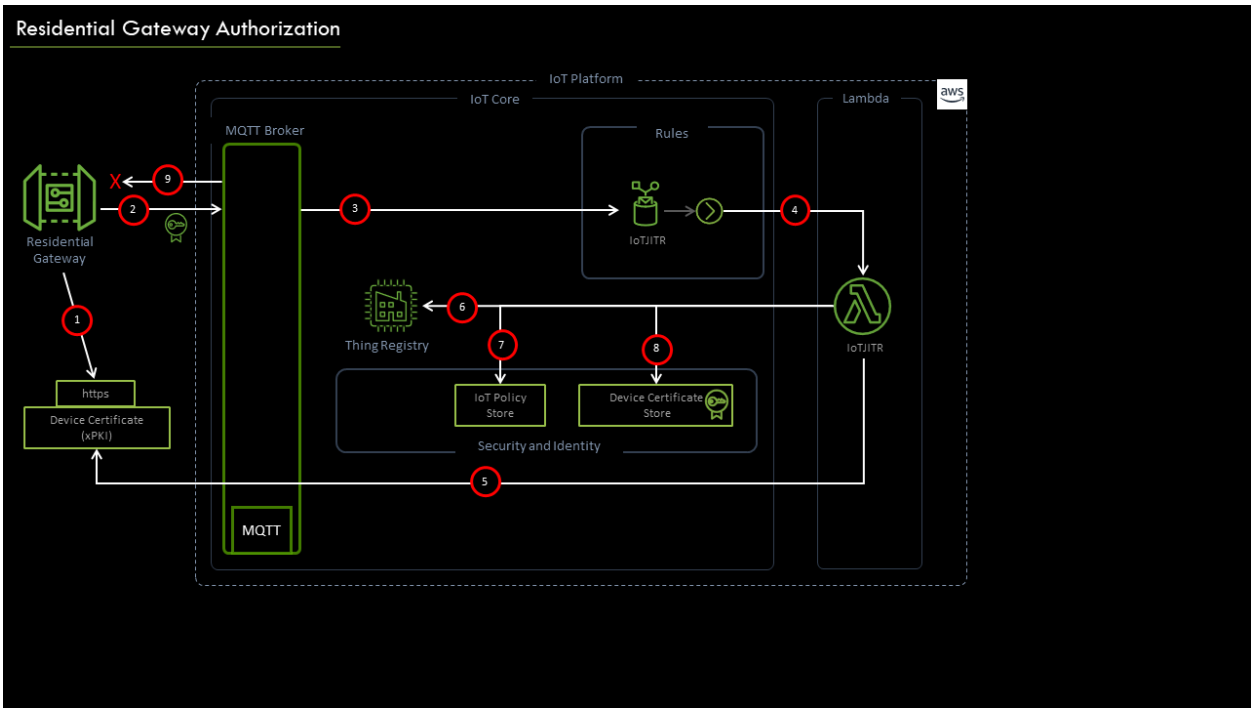


Figure 9 – Residential Gateway Authorization Flow Diagram

The Xfinity Residential Gateway enables advanced IoT Bridge capabilities for Zigbee, Thread, and WiFi-based devices. To provide a secure MQTT connection to the IoT Platform, the Residential Gateway performs a CSR (Certificate Signing Request) and is issued a device certificate by Xfinity Public Key Infrastructure (xPKI). Within the Certificate there are critical principle attributes that provide for a fine-grained security model. These attributes are:

- Partner IDentifier - syndicated partner identifier for the Multi System Operator (MSO). Stored in the SAN (Subject Alternative Name) -> Dir Name (DirName) -> Org Name (O) attribute.
- Account Identifier (AID) - identifies the subscriber's opaque account id; Stored in the SAN (Subject Alternative Name) -> Dir Name (DirName) -> Common Name (CN) attribute.
- Installation Identifier – the device identifier as a UUID (Universal Unique Identifier); Stored in the Subject -> UID attribute as the fourth part of the form [Model]:[Serial Number]:[AID]:[Installation ID]

- Mac Address – Device CMAC address. Stored in the SAN -> URI of form *urn:dev:mac:[CMAC]*. CMAC is of valid EUI-64 format.

The MQTT message broker provides certificate-based authorization support. The AWS IoT Core exposes this capability as part of the Just In Time Registration (JITR) feature, which is used to verify a pending device certificate and authorize the MQTT connect operation. The certificate issuer chain is verified and an OCSP check is performed. Once verified, the certificate's principal identity attributes are used to generate an AWS IoT Policy that allows the MQTT Residential Gateway to:

- Connect – allow the MQTT client to connect with a unique MQTT Client ID (CID) generated from the principle information within the Certificate. This information includes the PID, AID, and IID. An MQTT Client ID must be unique amongst all connected MQTT clients. An example MQTT Residential Gateway Client ID is as follows:

iot:[PID]_[AID]:adp:xhf:[IID]

  This authorizes only the Residential Gateway MQTT client matching the validated Certificate to establish a connection to the MQTT Message Broker. An MQTT client with a certificate that has expired or has been revoked is denied the connection.

- Publish– allow the MQTT client to publish messages to MQTT topics which are namespaced based on the PID, AID and IID. An example service request message topic is a follows:

c/iot/[PID]_[Account ID]/[SID]/[CID] /[RID]

  where,

  c is the root of the MQTT topic

  iot represents the IoT Device part of the MQTT topic

  [PID] represents the Syndicated Partner

  [SID] represents a unique Service Identifier as the destination of the request

  [CID] represents the MQTT Client Identifier as the source of the request

  [RID] represents the type of request. This limits the ability of the MQTT Client to publish action request messages only on topics containing its principle identity and to services to which it has been granted permission to send action response messages.

  This limits the ability of the MQTT Client to publish only to topics containing its principle identity and to services to which it has been granted permission to send action request messages.

- Subscribe/receive – allow the MQTT client to subscribe/receive actions to/from MQTT topics which are name spaced based on the PID, AID and CID. An example service request response topic is as follows:

c/iot/[PID]_[AID]/[CID]/[SID]/[RID]/[success|failure]

where,

c is the root of the MQTT topic

iot represents the IoT Device part of the MQTT topic

[PID] represents the Syndicated Partner

[SID] represents a unique Service Identifier as the destination of the request

[CID] represents the MQTT Client Identifier as the source of the request

[RID] represents the type of response. This limits the ability of the MQTT Client to subscribe and receive action response messages only on topics conatining its principle identity and from services to which it has been granted permission to receive action response messages.

This limits the ability of the MQTT Client to subscribe and receive action response messages only on topics that contain its principle identity and from services to which it has been granted permission to subscribe/receive action response messages.

6. MQTT versus other messaging protocols for client-platform communication

How does MQTT stack up against other messaging protocols? This section seeks to answer that question by first defining the key criterion upon which the messaging protocols can be fairly evaluated. This criteria can then be weighed upon each messaging protocol. These messaging protocols include HTTP, Websocket, AMQP, XMPP, and MQTT. The pros and cons of each messaging protocol is then considered, and finally this evidence can be used to inform a decision.

6.1. Evaluation Criteria

It is important to first define the evaluation criteria used to determine how the messaging protocols fit the needs of an IoT-based messaging system.

Asynchronous bi-directional messaging is necessary to provide support for enhanced Message Exchange Patterns (MEP). This includes asynchronous request/response messages and events using a publish/subscribe messaging pattern. Support for either UDP/IP or TCP/IP bi-directional communication is required.

The ability to support different type of data models for message exchange, whether binary, XML, JSON, etc, provides the flexibility to choose the data type most appropriate to publish a message specification, and enable message validation using client libraries and tooling that is widely available and adopted by the industry. JSON Schema support is required given its wide adoption and support within our organization.

Given the heterogenous nature of program languages and platforms across mobile clients, web clients, devices and services, wide support for client libraries must be available. At a minimum, client libraries must be supported from the likes ofAndroid (Kotlin), IOS (Swift), web clients (Javascript), devices (C/C++), and services (Java/Golang) ..

The security of an IoT system is of upmost importance. It is absolutely necessary to provide a secure communication channel between the mobile/web application clients, devices, and services of the IoT Platform. To this end, TLS 1.2 is required for inflight encryption, support for port 443 to avoid router/firewall traversal issues, and support for token-based auth for mobile/web clients and certificatebased auth for devices.

The performance and efficiency of the messaging protocol and overhead of the client/device libraries must be considered. The solution must offer minimum resource overhead, given factors including the negotiation of the initial connection and connection keep-alive, and the need for efficient messaging (small size on the wire), reduced power consumption for all devices and lightweight implementation (i.e code footprint, memory, etc). The solution must minimize latency related to network protocol negotiation and message broker and router traversal.

The reliability of the communication protocol is important as it offers a way to mitigate errors that may arise from unreliable mobile and/or home networks. The ability for the messaging protocol to deliver in-order messages, to acknowledge receipt of messages, and retransmit messages, if necessary, is required to build a reliable IoT system.

The solution should be cost-sensitive. Some of the key factors to consider that drive cost are the number of mobile clients and devices connected over a period of time (i.e. one day), the size and frequency of the messaging payloads, and the cost of vertical/horizontal scalability of the solution. Note that the extent to which the mobile and web clients are typically connected is a fraction of the time that the gateways are connected, which is to say 24x7.

6.2. HTTP (HyperText Transfer Protocol)

HTTP semantics has dealt quite well with brokering synchronous request/response messaging between the client and backend services, via URL paths. But HTTP has fallen short of satisfying the need to support asynchronous notifications across the convergence of application clients, gateways, and backend service(s). The use of multiple HTTPS long or short polling mechanisms, which mimic support for asynchronous responses, requires the mobile and web clients to connect and poll numerous backend services simultaneously. A better mechanism is required, and hence the need to extend the HTTP API semantics to include asynchronous notifications, using a publish/subscribe MEP. HTTP simply does not support publish/subscribe.

Also, it is not always easy to provide a single secure HTTP API connection endpoint for HTTPS client(s). Different services, within a vast organization, typically expose their own HTTPS endpoints. This necessitates a common strategy to expose a single endpoint across an HTTP API Gateway deployment. and in turn. fan-out the requests across the backend services. Horizontal scaling and high availability of this solution is required. Sections 6.2.1 and 6.2.2 detail the pros and cons of HTTP.

6.2.1. Pros

- Supports primary request/response but can lend itself to pub/sub with some work using PUSH_PROMISE vs. Long-polling, but please no short-polling!
- Reasonably small with HPACK for header compression required
- IETF standards are preferred by some vendors
- An HTTP API gateway supports millions of connected clients
- Rsed in mobile applications and most web applications that do not require pub/sub
- Support for many programming languages

6.2.2. Cons

- No guaranteed delivery (retry required
- No “last will & testament” feature for IoT devices that suddenly lose connectivity
- Slightly larger message size due to headers
- Primarily a pull technology:
    - short polling is not real-time (request timer driven) and if the time between requests is short, it can be server resource-intensive; most recommend never to use
    - long polling is immediate but is more complex and server resource-intensive; typically used for responses to requests for information in cases where a server is doing work in the background
- Server push requires work to implement pub/sub and asynchronous responses via PUSH_PROMISE

6.3. Websockets

Websockets provide an extension to application clients and web browser support of HTTPS. Where HTTPS offers only a request/response message pattern, websockets offer a very raw form of bi-

directional packet-based communication. It is the lack of built-in MEP and the number of potential connections that is the primary distractor of using websockets directly. Sections 6.3.1 and 6.3.2 detail the pros and cons of websockets.

6.3.1. Pros

- Works over port 443
- Supported by numerous mobile clients and web browsers (implementation on modern browsers is less of a concern)
- Supported by many web servers such as NGINX and Apache

6.3.2. Cons

- Client reconnection implementation on top of websockets is required
- A bit too raw even though it supports two-way asynchronous client/server communication. There is no built-in support for asynchronous request/response and pub/sub message exchange patterns. This would require additional work.
- Using raw websockets would still suffer from having to maintain a connection per client and potentially backend services. This approach would be very resource inefficient from both a client and backend perspective. It would lead to having to scale the web servers to meet the needs of all client/service connection needs, especially when websocket endpoints are not shared across services.

6.4. AMQP (Advanced Message Queueing Protocol)

AMQP is a great protocol for communicating between applications. It meets many of the needs of an asynchronous publish/subscribe and guaranteed message delivery system which supports numerous programming languages. But, as a message-oriented, middleware-based solution, it has not really found its home in the IoT space. Sections 6.4.1 and 6.4.2 detail the pros and cons of AMQP.

6.4.1. Pros

- Richest set of message scenarios (patterns)
- Asynchronous
- Supports queuing
- Mobile client support; web browser support via web sockets

6.4.2. Cons

- Uses port 5672 (not 443), which would lead to potential router and firewall issues
- Requires RabbitMQ on AWS EC2 instances; not a managed AWS service
- Larger protocol
- Used in IoT space by very few vendors

6.5. XMPP (eXtensible Messaging and Presence Protocol)

Extensible Messaging and Presence Protocol is an open communication protocol designed for instant messaging, presence information, and contact list maintenance. Based on XML, it enables the near-real-time exchange of structured data between two or more network entities.

6.5.1. Pros

- Support for message read, write, and consume
- Extended messaging and presence protocols
- Used for two-way chat and push notifications

6.5.2. Cons

- Uses port 5222 (not 443) which would lead to potential router and firewall issues
- Requires an XMPP/Jabber server on EC2 instances; not a managed AWS service
- It's XML-based, not JSON

6.6. MQTT

MQTT, as described in this paper, is a mature, stable, and feature rich messaging transport system. It delivers on all the key evaluation criteria. Sections 6.6.1 and 6.6.2 detail the pros and cons of MQTT.

6.6.1. Pros

- Provides a backend service the ability to publish notifications without the added complexity of implementing a long polling or websocket mechanism, as well as message routing within the context of each individual service. Aggregating and pushing the responsibility to the MQTT Message Broker decouples this problem from the client and underlying services. Note: MQTT also handles session persistence by delivering missed messages to the client.
- Provides a means to multiplex notifications across 1 or more topics grouped by backend service responsibility; the MQTT Message broker inherently supports message routing via topics.
- Only one secure client connection to the MQTT Message broker is required, allowing for simpler configurations for clients, load balancers, security policies, etc
- JWT and X.509 certificate-based authorization support
- Support for port 443
- Supports asynchronous message patterns (i.e. pub/sub)
- Assured delivery (3 QoS levels) and retained messages which provide flexible options for Client/Server.
- Supports the “last will & testament” feature that notifies other clients of an ungraceful disconnected of gateway-based devices
- Multiple subscriptions are ‘multiplexed’ over one connection
- Smaller size on the wire - minimum compressed header
- Great for a resource-constrained device, low bandwidth conditions, and high latency networks
- Brokers support millions of connected devices
- Easy to route messages to topic subscriber(s)
- Used in mobile applications, web clients, and devices for numerous IoT-based applications

- Lightweight implementation for embedded clients
- Power-efficient due to no-polling, shorter messages
- Less resources consumed compared to long polls on server.
- Support for many programming languages

6.6.2. Cons

- MQTT is based on the OASIS standard (OASIS-open.org), which is not preferred by some vendors
- Fixed headers and options apply if extensibility is required

6.7. Decision

In the end, after numerous discussions, careful consideration, and weighing the options based on the selection criteria, MQTT was the clear choice. Our Mobile/Web Client and Gateway teams decided that MQTT was the preferred communication protocol for messaging to/from the IoT Platform.

7. Conclusion

The Xfinity Mobile and Platform teams worked together to implement data models and topic structures for MQTT communications that are standardized and extensible across any device type and payload size. Its robust security, enabled by custom authorizers, ensures that mobile/web clients and devices can only publish and subscribe to topics relevant for the operator and subscriber's account.

By contrast, API semantics based on HTTP and other communication protocols have fallen short when it comes to satisfying the need to support secure, asynchronous notifications across the convergence of application clients, gateways, and backend service(s). The use of MQTT, in addition to the use of HTTP API semantics, fulfills the need to extend the API semantics to include asynchronous notifications using a Publish/Subscribe MEP.

We chose MQTT as a protocol for mobile-to-platform communications as well as Residential Gateway-to-platform communications. Its maturity as a lightweight, scalable, and robust protocol has proven itself in our production deployment of millions of clients.

8. Abbreviations

AID	Account IDentifier
AMQP	Advance Message Queuing Protocol
AWS	Amazon web services
CSR	Certificate Signing Request
HTTP	HyperText Transfer Protocol
IdP	Identity Provider
JTI	Java Web Token Identifier
JWT	Java Web Token
KB	Kilo Byte
MSO	Multiple System Operator
MQTT	Message Queuing Telemetry Transport
PID	Partner IDentifier
PUBACK	Publish acknowledgement
PUBREC	Publish received
PUBREL	Publish release
PUBCOMP	Publish complete
QoS	Quality of Service
RID	Request/Response IDentifier
SCTE	Society of Cable Telecommunications Engineers
SID	Service IDentifier
XMPP	eXtensible Messaging and Presence Protocol

9. References

MQTT client programming concepts

https://www.ibm.com/docs/en/ibm-mq/9.0?topic=telemetry-mqtt-client-programming-concepts

IoT Core Developer's Guide: MQTT

https://docs.aws.amazon.com/iot/latest/developerguide/mqtt.html

Metadata and Telemetry Support to Enable Telecom for Healthcare Opportunities

A Technical Paper prepared for SCTE by

Dr. Sudheer Dharanikota
Managing Director
Duke Tech Solutions Inc.
111 Fieldbrook Ct. Cary, NC 27519
+1-919 961 6175
sudheer@duketechsolutions.com

Jason Page
Principal Engineer
Charter Communications
6360 Fiddlers Green Circle, Denver, CO 80111
+1 720 699 6236
Jason.page@charter.com

1. Executive summary

The delivery of healthcare is undergoing seismic shifts. From remote consultations to remote patient monitoring, the healthcare of tomorrow will look and feel different than it does today. Numerous technological advancements have enabled these transformations, but each comes with a series of challenges to solve. These challenges include issues such as connecting devices to a local network, guaranteeing optimal bandwidth and latency, ensuring that services are being properly delivered by providers, and ensuring that patients are adhering to treatment plans. Cable operators are well-positioned to assist in developing standards to address these obstacles.

In this paper, we describe at a high level some of the services that tomorrow's healthcare will offer, the stakeholders that interact with these services, and the needs of each stakeholder according to the service. We then define four types of data categories and how they address the series of challenges that tomorrow's healthcare services face. We also provide recommendations on how cable operators can support these data needs and high-level architecture of a potential implementation for data collection, access, and analysis.

2. Introduction

The Healthcare industry is going through a major transformation to modernize the infrastructure, reduce the cost and increase the quality of care. In a series of articles, we have suggested how the Telecom industry can assist the Healthcare industry [1][2][3]. We call this inter-industry collaboration Telecom for Healthcare (T4H). Even though the T4H opportunity is not limited to these two major intersection points, we focus on Aging in Place (AIP) and Telehealth use cases to illustrate our thoughts on the end-to-end T4H architecture. (Refer to [4] for six different opportunities that a Telecom operator can address through the T4H architecture covered in this paper.) The SCTE Data Standards Subcommittee, in which the authors are members, is actively working on T4H solutions for the AIP and Telehealth areas in working groups three [5] and four [6]. The current paper on telemetry and metadata can be reviewed with the companion paper on the end-to-end T4H architecture published in the 2021 SCTE Expo [7] for a detailed understanding of the solution.

Figure 1 provides an intuition for the data needed from a T4H solution from per platform users (including the payor, who are not extensively considered in this paper) per type of services point of view. These data needs are categorized into four groups in the following section for further detailed purpose-driven analysis from telemetry and metadata points of view.

The players involved in the T4H solutions include the users (for AIP these are the elders, and for Telehealth[1], these are the patients who want to use the platform being developed in [5][6]), the service providers (such as the doctors, caregivers, etc.), and the other stakeholders (such as the family, legal guardians, trusted circle, etc.). These players use the platform for the following services:

- Use the platform for **communicating with different T4H players** (unified collaboration and communication (UCC)), and **communicate different information** relevant for the success of

[1] The US Department of Health and Human Services (HHS) defines telehealth as “the use of electronic information and telecommunications technology to support and promote long-distance clinical health care, patient and professional health-related education, and public health and health administration”. It encompasses everything from video calling to text messaging. In the realm of reimbursement, insurances may cover different levels of telehealth. The level of reimbursement has also changed after changes that came with COVID. Within telehealth here are four other broad categories: Synchronous, Asynchronous (store-and-forward), Remote Patient Monitoring, and Other Services such as mobile health (mHealth).

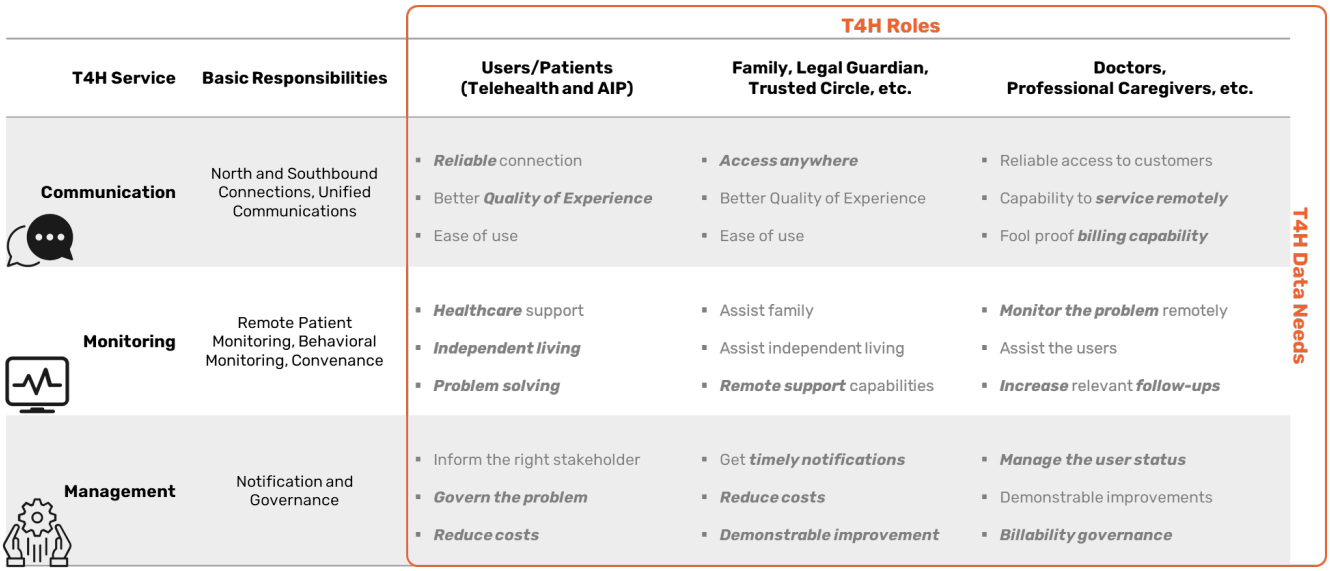


Figure 1 Understanding the T4H services, different roles, and their data needs

T4H services. For the users to adopt these services, they need to have a reliable connection with quality of experience and that is easy to use. For the stakeholders, in addition to the needs of the users, they should be able to access the services from anywhere. For the providers, the platform shall provide reliable remote service to the customer and an accountable billing capability.

- The players in the T4H space are very interested in **the monitoring capabilities** of the platform. The users are interested in healthcare and non-healthcare (independent living) needs. The incentive for them to use this platform is its problem-solving capabilities using the data that they assemble from different sources. The stakeholders, in addition to supporting their loved ones with their needs, have to provide these monitoring services remotely. The service providers shall be able to effectively monitor the problems remotely, improve the quality of care and reduce the overall cost of care.
- Use the platform to **manage timely notifications** and **govern the condition** of the user. To enable such functionality, the user depends on the timely assessment of the problem and reduces the cost through the management infrastructure. The stakeholders depend on timely notification, cost reduction, and most importantly the demonstrable improvement. The providers, on the other hand, in addition to the user status management, shall be able to claim the billability of the services.

In the next sections, we elaborate on T4H data categories based on the high-level incentives that we discussed in this section, expand on each of these categories in the following section, summarize the findings and propose the next steps.

3. Telecom for healthcare data categories

By analyzing the T4H users, stakeholders, and service providers, we categorize the metadata or telemetry data collected from the T4H platform, as shown in Figure 2, into four categories. (Note that the data collected in the metadata and telemetry cases are the main data streams provided by the devices, but are the supplemental information provided to assist the T4H players). They are:

- **Quality of Experience (QoE)**: The questions to be answered by the T4H platform include - are T4H services meeting stakeholder experience needs and are users engaging with services? The data collected here such as bandwidth availability, latency behavior, platform experience for back-office, and individual connections represents the customer adoption.
- **Monitoring**: The goal of this set of data is to evaluate if we are providing useful T4H services? The types of data collected here are healthcare and non-healthcare related streams, sensor and UCC flow information, network and device status, etc.
- **Connectivity**: As opposed to the quality of the service, reliable and highly available communication infrastructure is essential for providing these emotionally sensitive AIP and Telehealth services. The connectivity metrics offer these parameters. They are measured using platform availability, SLA (Service Level Agreement) guarantees, service up times, zero-touch provisioning, etc.
- **Accountability**: Are the service providers delivering expected results? Are the users adhering to the treatment plan? Is the provider's billing in line with the services they offered? These accountability metrics include – Quality of Care, timely resolutions, problem resolution rate, timely status notifications, etc.

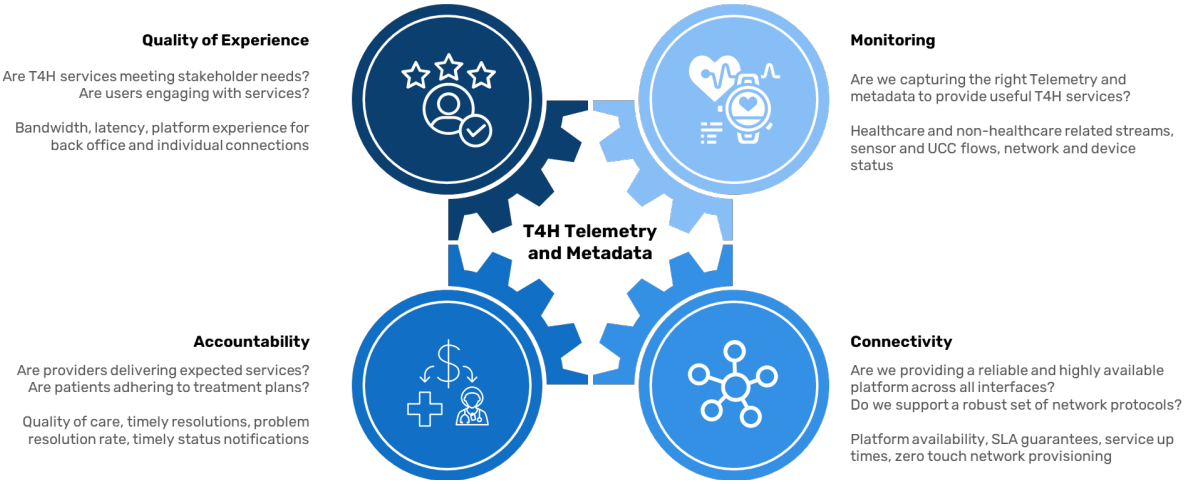


Figure 2 T4H telemetry and metadata categories

4. Metadata details

Figure 2 provides an end-to-end T4H architecture based on the framework proposed in [4][7]. There are multiple locations, as shown in the figure, where the T4H related metadata and telemetry information can be gathered. These data collection points include:

- In-home device interface: This interface is used to monitor the in-home T4H sensor and interactive devices. The QoE metadata can be monitored from this data collection point.
- In-home network interface: The north-bound interface of the sensor network gateway can be used for the aggregated in-home information such as broadband connectivity-related data and per session-related monitoring.
- T4H sensor network interface: The reliability and availability metrics can be monitored from the T4H sensor network infrastructure.
- T4H service back-office interface: This interface provides the overall service level QoE metrics, aggregate service level information, Quality of Care analysis metrics, governance metrics, etc.

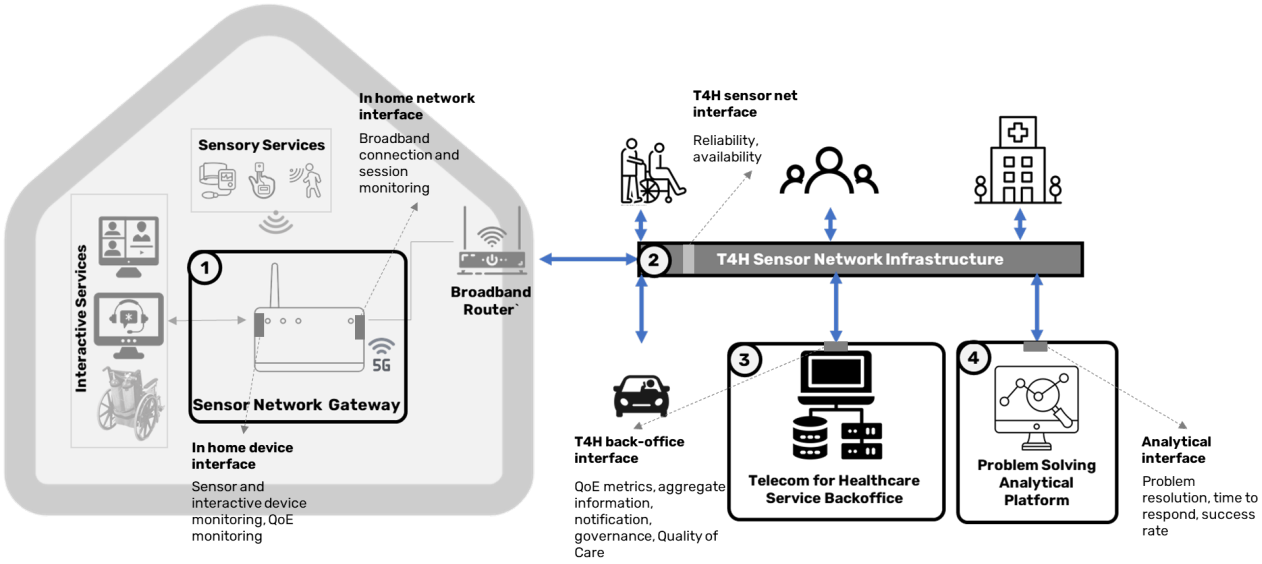


Figure 3 Different monitoring points in Telecom for Healthcare architecture

- T4H analytical interface: This interface provides the responsiveness, accuracy, and success rates of the problem-solving analytical infrastructure.

The information collected from the above interface must be securely collected and shall follow HIPAA privacy compliance [10]. We will not elaborate on these requirements in this document. In the following sections, we will highlight some of the details behind the proposed metadata and telemetry classes of information.

4.1. Quality of Experience related data

Understanding the QoE needs of different applications [8] and measuring them to see if the platform is meeting the needs, is essential for the adoption of T4H services. The applications used in the T4H environment are sensor and interactive applications. These application's QoE is measured at in-home for individual usage and at back-office service infrastructure that hosts the applications for aggregate usage.

Class of applications*	Throughput sensitive	Loss sensitive	Delay sensitive
Onetime measurements	X	X	
Video monitoring	X		X
Sensor monitoring	X	X	
Video communications	X		X

Figure 4 Quality of Experience needs of T4H applications

Generic metadata required to support QoE metrics:

- Sensor instrument-related metadata: Sensor id, type, group, priority (critical, high, medium, low), location in the house, vendor information, etc.
- Interactive applications-related metadata: UCC id, UCC type, application location (home, caregiver, provider, family), UCC vendor information, application experience (e.g., 5-star scale), etc.

QoE specific metadata: As discussed in [8], and as shown in Figure 4, the application QoE can be grouped into throughput, loss, and delay sensitivity.

- Throughput sensitivity data: This includes bits per second (Peak, Min, Average) metric. The monitoring locations include sensor hub and back-office hosting. The level granularity of this metric should be per session and per aggregated (per sensor or UCC id) stream.
- Latency sensitivity data: This includes one-way delay (Peak, Min, Average) metric. The monitoring location should be at the back-office hosting at a granularity of per session. To analyze latency the relevant communicated information should be timestamped.
- Loss sensitivity data: The data points lost per minute (Peak, Min, Average) will give the performance of the underlying communication infrastructure. This data needs to be monitored at the back-office hosting location at the granularity of per session. To measure the loss statistics the data needs to be time sequenced.
- Experience-related data: As an overall experience metric for the application, the user experience rating (from 1 to 5) per usage may be monitored. This can be tracked at different aggregation points based on the scope.

4.2. Monitoring related data

As shown in Figure 5, the applications that are offered and hence are monitored are the sensor and interactive applications. They can support both healthcare and non-healthcare applications. The idea of these data needs does not include the core data streams such as the temperature from a thermometer, but the additional metadata/telemetry that supports the players.

Generic metadata required to support monitoring data:

- Additional sensory applications metadata to support monitoring capabilities: Status (up or down), start time, healthcare or non-healthcare related, etc.
  Additional interactive applications metadata to support monitoring capabilities: Start and end times of the sessions, type of interactive application use (video, audio, video +audio), number of sessions, etc.

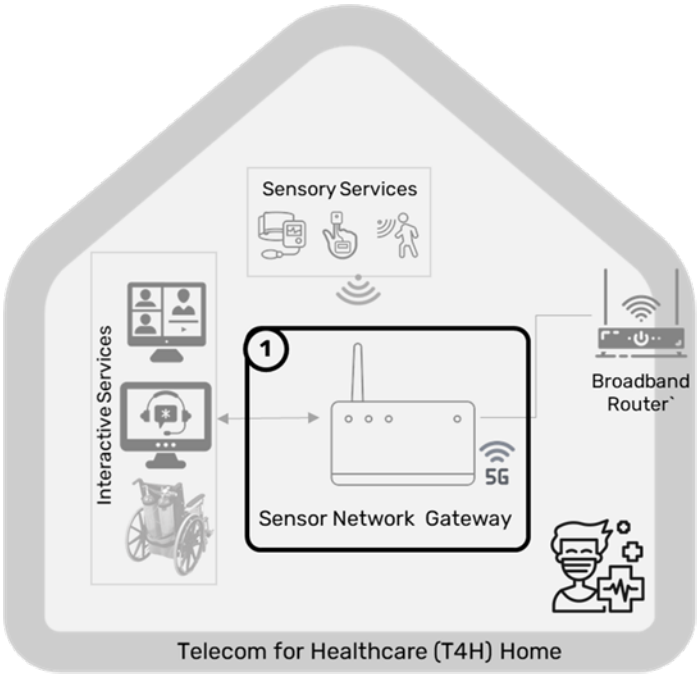


Figure 5 T4H monitoring services

Monitoring specific metadata: To provide a responsive platform to different healthcare and non-healthcare needs, we need to capture different metadata from the monitoring streams. The data can be monitored at the aggregation point in the home (sensor network gateway, as discussed in [7]) or at the hosted service back office. These metadata
include:

- Senor monitoring data: Priority of the
  sensor, privacy level of the data (generic, provider-specific, stakeholder, the user alone, etc.), urgency level of the notification (such as threshold crossing alarms)
- Interactive services monitoring data: Session related (number of legs, number of streams, etc.), stream related (QoE measures, transcriptions, metadata, etc.)

4.3. Connectivity related data

Connectivity focuses on providing a highly available service platform with five 9s reliability. These measures are very important to support highly emotional and sensitive subjects of healthcare and elderly care. In addition, providing ease of configuration (zero-touch configuration) is essential for T4H adoption. The reliability is measured reliability of the devices, connections, and the platform. The availability of the end-to-end services is the uptime and Service Level Agreement (SLA) guarantees for these time-sensitive T4H services. Also, due to the number of devices and solutions that will be integrated with the T4H services, we include the ease of configuration as part of connectivity.

Generic metadata required for connectivity data:

- Additional sensor metadata to support connectivity services: Sensor uptime, sensor loss of connectivity, sensor reliability

Connectivity-specific metadata: Providing a reliable platform to demonstrate the capabilities of T4H services as a differentiator is essential for cable operators to enter this market. These metrics depend on both the device and the platform's reliability and availability.

- Reliability data: Device reliability metrics, service reliability metrics
- Availability data: User device uptime (primary connection, secondary connection), server uptime, percentage availability, service availability
- Zero-touch configuration assessment: Number of service calls during installation, failed self-installs, in-home installation percentage, and average installation duration per service offering
- Other connectivity data: Availability SLA adherence

4.4. Accountability related data

The accountability of the T4H environment is used to evaluate the Quality of Care provided by the service providers and platform providers. These assess the improvement, the timely notifications that can be provided by the platform, and the billability of the services offered by the platform (and hence the provider). In this hyper-competitive and very expensive T4H environment, demonstrating the value of the service is a critical differentiator. Also, the healthcare industry is moving towards a pay-per-performance model, which constantly checks on the accountability of the providers. Hence cable operators need to showcase how they can provide a platform to assist providers with their accountability goals (such as improve notifications to different stakeholders). Providing corroborative information to bill appropriately is an essential service that cable operator T4H should offer to make their solutions attractive.

Generic metadata to support accountability data:

- Analytical platform metadata: Efficacy of the algorithms (problem resolutions rate), speed of analysis
- Notification infrastructure metadata: Notification statistics, notifications per type of problem

Accountability specific metadata:

- Quality of care data: Time to resolve the issue, condition improvement, reduce the number of missed appointments, cost reduction (for user, stakeholder, provider)
- Notifications related data: Response time (average, peak, minimum), per problem, per provider,
- Billability related data: Session context (duration, reason, parties involved, provider information, etc.), stream context (devices, device performance, potential transcription, additional notes)
- Other related data: Other stakeholder accountability, a payor accountability measure

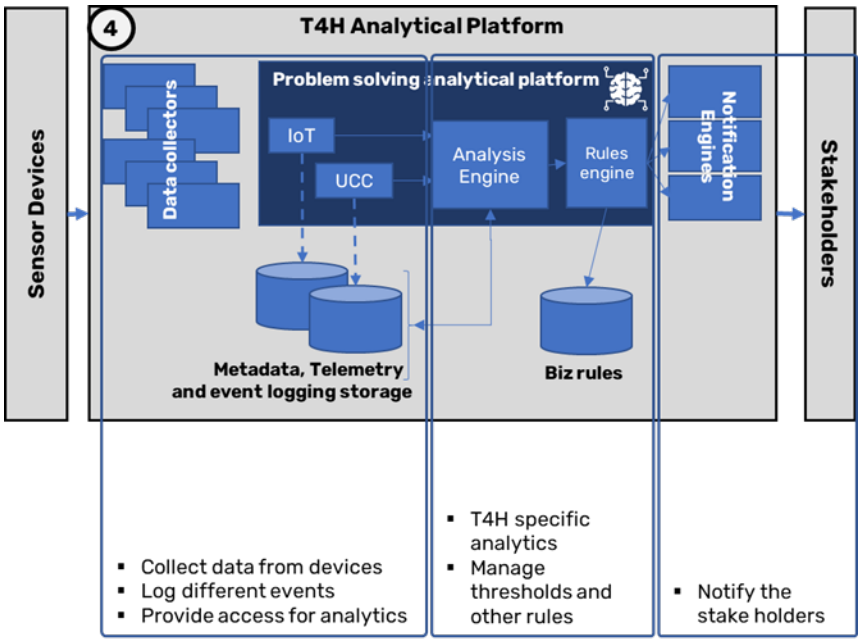


Figure 6 Using metadata to solve T4H analytical problems

5. Interaction of metadata with the analytics framework

The topic of metadata is not complete without understanding how the data is used. In [7], we explain the end-to-end T4H architecture including the analytical platform, as shown in Figure 6. The four categories of data, as discussed in the previous sections, are stored and used in the problem-solving analytical platform. The analytical platform provides the interface to different metrics as discussed in the previous sections, and assists in providing timely notifications to the stakeholders. These analytical functions that are presented in Figure 6 can be centralized or distributed in or closer to the home for availability purposes. These discussions are not in the scope of this paper.

Component	Status in MSO	Comments
Data collectors	Existing for IoT and other service info.	Need to repurpose for T4H data
Analysis engine	Existing for IoT engines	Need additional development for T4H
Rules engine	Potentially new function	Need solutioning
Notification engines	Existing with service assurance tools	Need to extend to T4H
Data privacy	Existing for PII	Need to extend to PHI
Performance	**Status in MSO**	**Comments**
Scalability	Device level alarms	Need to extend to per sub per stream
Security	SNMPv3 based	Need to validate if this is enough
Privacy	PII after collection	Need to validate if we need to anonymize at the collection points
Reliability	Reliable communication	No additional changes in our opinion
Responsiveness	Good for current use	Crucial for the success

Figure 7 Understanding the gaps in the end-to-end analytical platform

As presented in Figure 7, many of the analytical components relevant for the T4H solutions are already developed for the current Cable operator solutions. As provided in the comments section, these solutions need to be extended for the T4H requirements. Also note that the performance of the platform, which is currently tuned for the network device level needs to be scaled to per-stream level information gathering. The current Telecom solution data security and privacy constructs shall be validated against the needs of T4H needs. Although the responsiveness of the current solutions is good for the Telecom needs, using this platform for the time-critical and highly responsive T4H solutions calls for a fresh look at the data architectures. Further analysis on these architectural constructs will be conducted in the SCTE working groups [5][6].

6. Conclusion and next steps

In this paper, we have highlighted the transformations taking place in healthcare delivery and some of the challenges they pose. We have described the different needs of the various stakeholders and how they can be met with additional data and analytics. These needs include metrics to assess the quality of experience; monitoring applications; reliable, highly available, and easily configurable connectivity services; and notification and billing accountability systems. Many of these data sources and services already exist and simply need to be exposed to authorized third parties. In other instances, new infrastructure and standards are needed.

Cable operators have a unique opportunity to play a foundational role in the transformations taking place in the healthcare industry. We can take the lead in establishing standards for data transit, storage, access, security, and analysis. We can assemble a coalition of device manufacturers, inter-industry partners, and healthcare providers to ensure wide adoption. Regardless of what cable operators choose to do, these changes are coming to healthcare. By capitalizing on this opportunity cable operators can design a suite of new products to keep them relevant into the future. Failure to capitalize ensures that over-the-top solutions will emerge and eventually render them a dumb pipe.

7. References

[1] Sudheer Dharanikota, Ayarah Dharanikota, *Why are cable operators natural fit to support Telehealth – An inter-industry perspective*, 2020 SCTE Expo, available here

[2] Sudheer Dharanikota, Ayarah Dharanikota, Dennis Edens, Bruce McLeod, *Aging in Place business case for cable operators*, SCTE Journal, June 2021, available here

[3] Sudheer Dharanikota, Ayarah Dharanikota, Dennis Edens, Bruce McLeod, *Telehealth business case for cable operators*, SCTE Journal, September 2021, available here

[4] Duke Tech Solutions market research, *Telehealth market report – A Telecom based opportunity analysis*, available here

[5] Data Standards Subcommittee, Working Group 3, Aging in Place, available here

[6] Data Standards Subcommittee, Working Group 4, Telemedicine, available here

[7] Sudheer Dharanikota, Clarke Stevens, *End to end Telecom for Healthcare architecture – a cable operator perspective*, 2021 SCTE Expo, available here

[8] Sudheer Dharanikota, *Understanding the Basics of Quality of Experience – Gaming Focused*, DTS white paper, available here

[9] Sudheer Dharanikota, *What are the impacts of changing consumption patterns on bandwidth usage?* DTS white paper, available here

[10] HHS.gov, *Health Information Privacy*, available here

Mission Critical Microgrids

Securing a Better Energy Future through the Power of Choice

A Technical Paper prepared for SCTE by

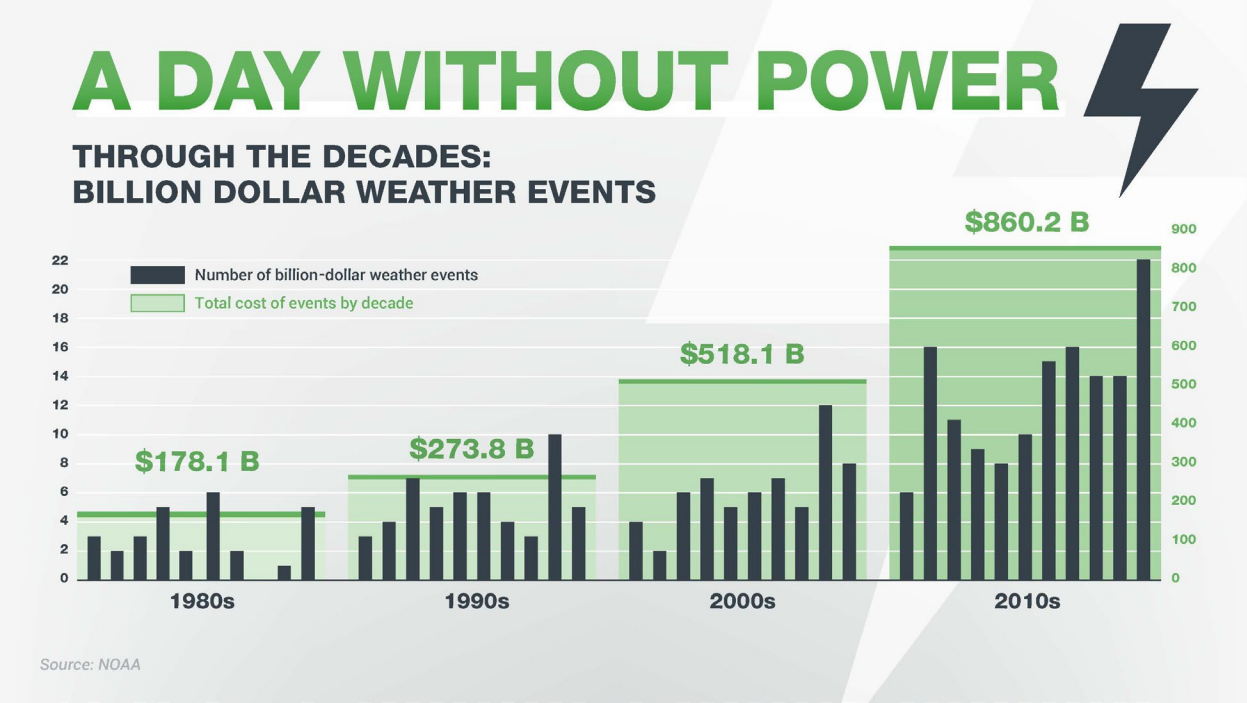
Chris Ball
Senior Manager, Product Marketing
Bloom Energy Corporation
4353 North First Street, San Jose, CA 95134
(408) 543-1709
Chris.Ball@bloomenergy.com

Kathryn McAuliffe
Product Marketing Associate
Bloom Energy Corporation
4353 North First Street, San Jose, CA 95134
Kathryn.McAuliffe@bloomenergy.com

1. Introduction

A day without power is hard to imagine. Electricity has become a basic necessity, powering everything from our homes, to our businesses, to our transportation systems. But in a post-climate change world, extreme weather events are happening more frequently and becoming increasingly severe, threatening our access to the 24 x 7 power so integral to our everyday lives. This has caused a shift in thinking from, "the cost of power" to "the cost of not having power."

Demands of our digital economy coupled with the escalating consequences of our changing climate have created an unprecedented risk landscape. Energy challenges are on the rise and pressure is mounting for our electric system to get cleaner, faster. The sheer financial cost of climate impact is nearly outpacing the billions of dollars utilities are spending to improve the electric delivery system.



During 2020, there were 22 separate billion-dollar weather and climate disaster events across the United States, breaking the previous annual record of 16 events that occurred in 2017 and 2011.

This is the risk landscape for all of us, but without proper safeguards in place, the consequences of those risks are higher for those whose power choices are directly tied and vital to the functioning of their organizations. Mitigation requires a highly strategic approach to energy management.

Fortunately, we are now in a time where there are multiple options for sourcing and delivering electricity. Distributed generation has completely shifted the energy paradigm, providing a clear path forward for those seeking to gain more control of their electricity supply.

This paper explores this crucial paradigm shift. Its objective is to educate and mobilize individuals, businesses, communities, and policymakers around the importance of resilient power – the challenges we

face, the risks we can mitigate, and how leaders can gain control of their energy future through distributed generation and microgrid solutions.

2. Centralized Present, Distributed Future

Over the course of the 20th century, the U.S. electric grid was built as a one-way value chain from fuel supply to end-user consumption. Such infrastructure design created cascading vulnerabilities whereby a failure of any one component can result in disruption of service to end users. Grid hardening programs are underway in every region, but upgrading such a complex system is expensive, and takes years to properly execute.

Some might say it's too little too late – that these complexities have left society too heavily reliant on an electric delivery system that has simply not kept pace with the evolution of its surrounding environment. This aging centralized power grid that is inherently prone to failure, now faces heightened demands of digitization, a rising frequency and intensity of natural disasters and a dangerous cyber-threat landscape.

As a result, companies with mission critical facility requirements must invest heavily in back-up and conditioning devices to ensure a constantly available, high-quality energy stream to power their equipment. However, the few traditional solutions available are prone to their own failures.

But consider the implications of developments in the energy space that have come from the rapid digital transformation over the last few decades. The conventional way of delivering electricity is experiencing a fundamental and promising shift – we are moving away from our centralized present towards a distributed future.

3. Energy as a Keystone Metric

Depending on the industry, energy can be a large component of a company's cost structure and a complicated operational issue. It touches every piece of a business' value chain.

Companies are increasingly opening the aperture on their energy strategies to meet their corporate objectives, with each company's approach reflecting critical business needs. These critical business needs may be substantially different depending on industry and facility requirements.

Retailers face different challenges than data centers or manufacturers. In some cases, business needs primarily reflect rising energy costs and their impact on other operating margins. In other cases, business needs may be more closely tied to reliability of critical infrastructure and ensuring that the business can continue to operate effectively during a prolonged grid outage.

The critical challenge for businesses in building a cohesive energy strategy is identifying their full scope of electricity-related risk factors. Recognizing risk is one thing. Having a strategy to mitigate it is another. Doing so requires a deeper understanding of the issues and how to address them from the top down.

4. The Reliability Spectrum

Every building has a distinct energy "fingerprint" based on the needs and requirements of its routine operations. When drilling down to the facility level, understanding the difference between demand and consumption is key to defining its energy characteristics. Fluctuation of demand and the usage that results from that demand make up one's load profile.

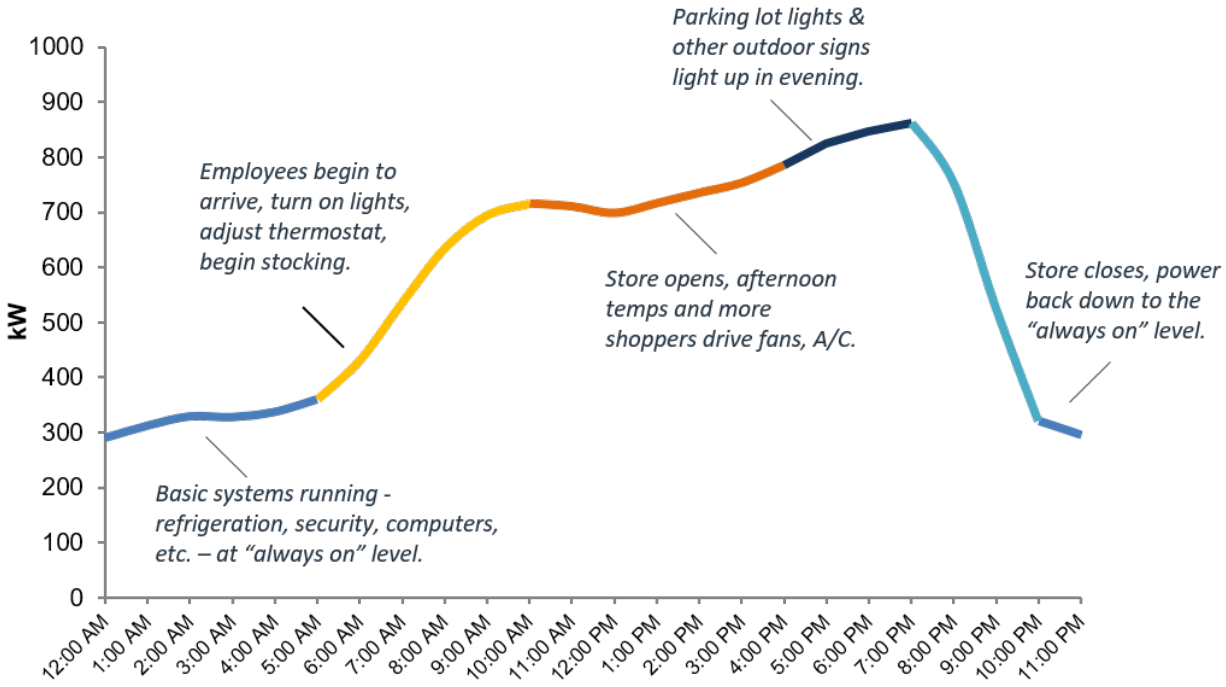


Figure 2: Snapshot representation of energy use for a typical day of business

A 'critical load' is a portion of electricity supply that powers infrastructure directly related to an organization's ability to operate. Infrastructure that warrants this status must either be kept running when main power supply fails or be powered down in an orderly manner to prevent system crashes, data loss or corruption, and life shortening hardware damage.

The term 'mission critical' is defined as something that is vital to the functioning of an organization. Operating under these conditions necessitates stronger emphasis on criteria that ties directly into the structural elements of a building's design.

Figure 3: Key considerations when operating a mission critical facility

Special focus must be placed on meeting minimum requirements across several key energy components – utility power supply, back-up generation, uninterruptible power systems (UPS), and cooling systems – which can come at great cost. Emerging technologies and business models represent a key opportunity to manage costs and mitigate risks. However, they also present an optimization challenge in terms of structuring those options to maximize the economic, reliability, and sustainability benefits.

5. The Resiliency Challenge: Eliminating Tradeoffs

As businesses look to address their critical resiliency needs, growing ambitions of a clean energy future and impacts of rising energy costs have elevated the importance of choice when it comes to power.

Most distributed energy resources (DER) are self-sufficient, but they are not one-size-fits all. Resiliency is just one benefit among many that microgrids provide, but resiliency decisions should not be made at the expense of environmental concerns and sustainability decisions should not ignore the importance of reliable energy supply. Diesel generators have been the status quo solution for power disruptions for decades. However, they are monolithic machines without inherent redundancy and they produce over 40 toxic air contaminants, including a variety of carcinogenic compounds during operation. What's more, since they are idle assets, they needlessly consume fuel while testing to ensure they can be available when needed. Further, the availability of diesel fuel during an extended outage and the reliability of diesel engines to operate continuously for long periods of time are both risks to the traditional design.

Technologies like solar and wind are great for their renewable profile but due to their inherent intermittency, cannot practically solve resiliency challenges. Their very nature requires some sort of energy storage and today's technology does not cost effectively support the massive load shifts from day to night, and certainly not from summer to winter. Further, the majority of solar (66%) and wind (~100%) are utility scale projects, meaning they still rely on the vulnerable, above ground, transmission and distribution system.

Fuel cell distributed generation projects are uniquely positioned to help eliminate tradeoffs and solve these compounding challenges. A fuel cell microgrid provides a distributed solution that is onsite and

always on. It is not a UPS or generator sitting idle waiting for an outage event – it's an active asset that produces clean, highly reliable power 24 x 7 without particulate emissions.

6. Fuel Cell Building Blocks

Fuel cells provide a critical foundation for microgrids of varying complexity and can provide significant benefits to the communities, businesses, and utilities they are part of. This type of technology targets a customer's 24 x 7 energy usage whereas technologies like solar or battery storage are intermittent. Fuel cells were invented over a century ago and have been used in practically every NASA mission since the 1960's, but until now, they have not gained widespread adoption because of their inherently high costs – a challenge that once plagued a developing solar industry decades earlier.

There is a range of fuel cell technologies that exist today, however solid oxide fuel cells (SOFCs) are the market's most efficient and practical form. SOFC's provide an electrochemical pathway to convert fuel directly to electricity without combustion. The cell itself consists of three parts: an electrolyte, an anode (-), and a cathode (+). The electrolyte itself is a solid ceramic material and the anode and cathode are made from special inks that coat the electrolyte. As oxygen ions interact with fuel in the cell, the resulting electrochemical reaction is able to produce electricity without combustion. Because of its extremely high conversion efficiencies, they are able to produce twice as much electricity as conventional combustion generators using the same amount of fuel.

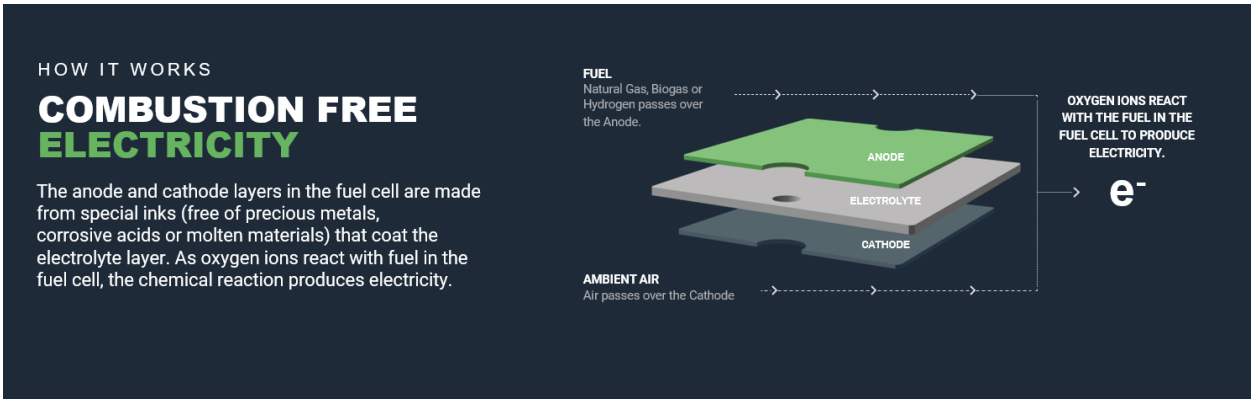


Figure 4: How a solid oxide fuel cell works

7. Microgrid System Architecture

Fuel cells are extremely versatile when comprised within a larger system. The process starts with a single cell which produces 25W, roughly enough to power a light bulb. The cells are then stacked within the system and assembled into 50 kW power modules – modules that can function independently from each other. This modular flexible architecture design allows for any number of modules to be clustered together, in various configurations, to form solutions from hundreds of kilowatts to many tens of megawatts.

The standard system architecture design interconnects with your facility in a grid parallel configuration. This means the system leverages a current mode inverter replicating the frequency and voltage of the grid and will not impact site power quality. Load changes by the site will similarly not impact output; the site will essentially have two utility feeds and anything needed above the fuel cell output will be supplied by the electric grid. A grid parallel interconnection is subject to IEEE-1547, similar to solar, such that when

the grid is unavailable or out of IEEE-1547 specifications, the system is required by code to stop exporting power until the grid returns.

A mission critical design topology allows the fuel cell system to continue operating during an electric grid outage. The design requires a segmented load to be wired directly into a second set of voltage mode inverters that will be added to the system's architecture.

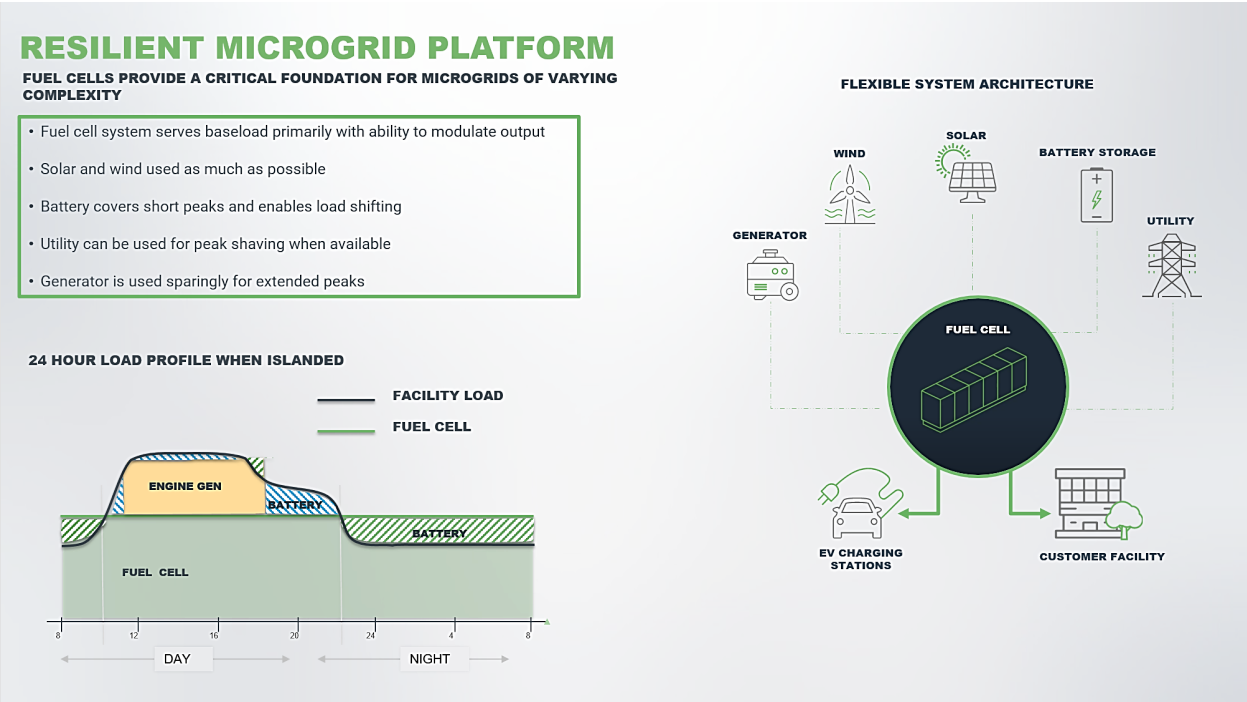


Figure 5: Illustrative view of fuel cell microgrid system architecture

8. Key Differentiators for Mission Critical Support

Power Quality

The power quality delivered by SOFC systems is comparable to best-in-class UPS systems deployed in data centers and other mission critical facilities. They are able to utilize state-of-the-art PWM (pulse-width modulation) inverter technology for conversion of fuel cell DC power to 480V AC power. The waveform of the current supplied to the customer is generated by a sophisticated multi-level current-source inverter control scheme. The high inverter switching frequency supplemented by a high-performance filter in each inverter module means that harmonics are virtually eliminated.

These systems are designed to meet or exceed power quality requirements of standards relevant to distributed power generation and distribution such as:

- UL-1741
- IEEE-1547
- IEEE-519
- Other utility grid interconnection requirements in the USA and other countries around the world.

Density

Fuel cells provide significant power generation in a small footprint which makes it an ideal power solution for smarter space utilization. Unlike large, multi-megawatt generating combustion engines, SOFC systems can be deployed in increments as small as 200 kW, enabling power sources to be distributed and land to be used for critical business applications. Due to this minimal onsite footprint compared to other generation technologies, facilities can utilize available land for higher value uses – such as data centers expanding their facility to accommodate more server racks, or a hospital adding a new wing to accommodate more patients.

Redundancy

The modular, redundant architecture makes it such that a solid oxide fuel cell system can continue powering facilities while operation and maintenance activities occur on individual modules. Each power module independently connects and feeds power to a DC bus. When a power module needs maintenance, that module will be safely ramped down and shut off while the remaining modules continue producing consistent electrical output. The power module will then be repaired or replaced, and then ramped up to full power, ensuring consistent output from the fuel cells without disruption to a customer's business operations.

Reliability

SOFC systems deliver 24 x 7 x 365 baseload power, with mission critical reliability and grid independent capabilities. They can be easily configured to eliminate the need for traditional backup power equipment such as diesel generators, batteries and uninterruptible power systems.

What's more, by generating power at the point of consumption, the system is able to avoid the vulnerabilities and line losses of conventional transmission and distribution. It draws from two continuously operating independent sources – existing gas infrastructure, which is a redundant underground mesh network, and the utility grid. The probability of both failing simultaneously is extremely low. Modularity and combined with this fault-tolerant design means that SOFC systems are able to reliably operate at very high availability around the clock.

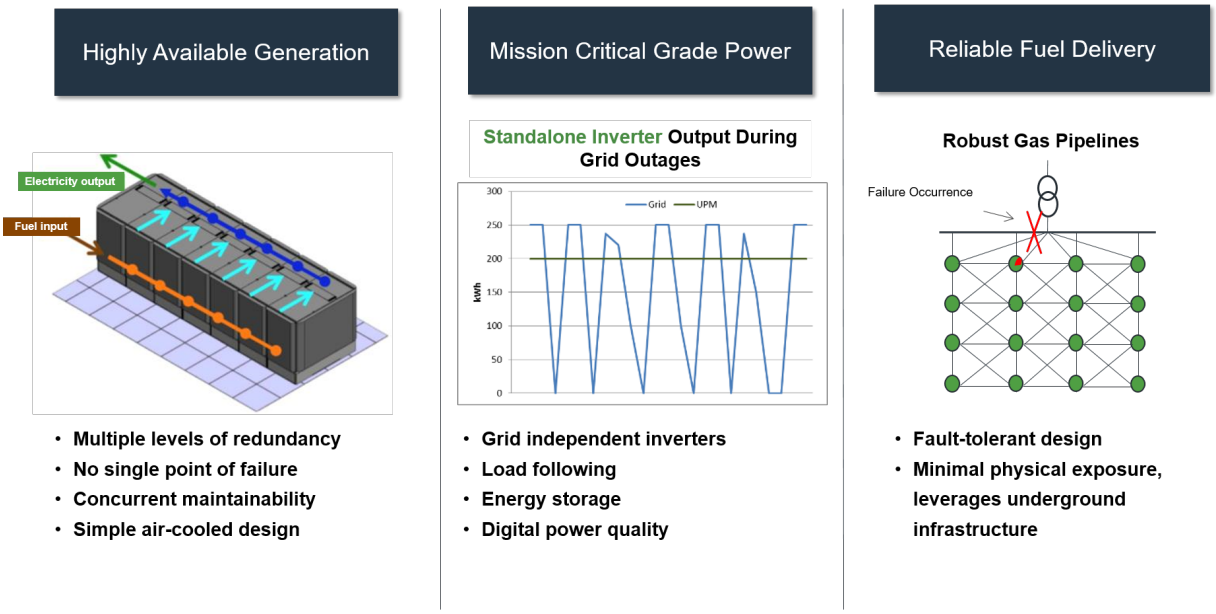


Figure 6: Key differentiators of SOFC for mission critical support

Resiliency

Fuel cell microgrids can provide critical resilience from power instability, driven increasingly by climate-related extreme weather events. The following case study highlights a technology manufacturing company who deployed Bloom Energy's AlwaysON solution to secure their operations and mitigate the negative impacts associated with escalating climate risks.

Case Study: Spotlight Save

A large technology company manufactures electronic test and measurement equipment and software that supports the larger digital ecosystem of e-mobility, network monitoring, 5G, LTE, and IOT. Their business needs are closely tied to the reliability of critical systems in operation at their HQ campus, located in a high-risk wildfire zone in Northern California.

The numerous unplanned outages they were experiencing throughout the year were costly and damaging to highly sensitive manufacturing equipment. With backup infrastructure reaching the end of its useful life, they were seeking new solution with mission critical capability that could power the whole campus independent of the grid.

In the wake of a growing number of outages and elevated risk to its operations, the company implemented an AlwaysON Microgrid, which enabled 2.8 MW of mission critical systems to operate independently when disruption to the electric grid occurs.

In the summer of 2019, California utilities began implementing transmission-level Public Safety Power Shutoffs (PSPS) to mitigate wildfire risk. Faced with one instance that left millions of customers without grid power for multiple days, Bloom's AlwaysON microgrid kept their campus online and operational throughout, reinforcing the technology's proven resiliency in the field.

9. Conclusion

In the last two decades, diverse groups across the energy sector has been working hard to create and commercialize innovative alternatives to centralized energy generation and delivery. Widespread decarbonization is occurring quickly through policy changes, technology innovation, and the willingness to adapt. However, it's becoming increasingly clear that alongside decarbonization sits an equally critical resiliency challenge. Hardening of our modern energy system will be a major focus for utilities going forward but businesses can support these efforts by helping break off smaller challenges through with distributed generation.

Fuel cell microgrids have changed the way businesses look at their critical power infrastructure, serving as a viable alternative source of primary power. The versatility of solid oxide technology operating at the core of a microgrid creates distinct advantages that will enable applications across the entire energy value chain for years to come. Its core efficiency advantages and ability to scale within a rapidly evolving energy landscape makes it the ideal distributed generation solution – one that delivers the greatest value now and provides the clear path needed to propel us all towards a better energy future.

Abbreviations

DER	Distributed Energy Resource
IEEE	Institute of Electrical and Electronics Engineers
IEEE-1547	IEEE standard for interconnecting distributed resources with electric power systems - provides requirements relevant to the performance, operation, testing, safety considerations, and maintenance of the interconnection.
IEEE-519	IEEE recommended practice and requirements for harmonic control in electric power systems
PSPS	Public Safety Power Shutoffs
PWM	Pulse-width modulation
SOFC	Solid oxide fuel cell
UL-1741	UL safety standard for inverters, converters, controllers and interconnection system equipment for use with distributed energy resources
UPM	Uninterruptible Power Module *(inherent to fuel cell system mission critical architecture)*
UPS	Uninterruptible Power System

Modernizing Cox Communication's Access and Aggregation Network Infrastructure for Remote PHY Deployment

A Technical Paper prepared for SCTE by

Deependra Malla
Lead Network Design Engineer
Cox Communications Inc.
6305 Peachtree Dunwoody Road
Atlanta, GA 30328
602-694-4429
deependra.malla@cox.com

1. Introduction

Like other Multi System Operators (MSOs) and Internet Service Provider (ISPs), Cox Communications needed to transform its access and aggregation network infrastructure to meet increased bandwidth demands and support next generations of multi-services. The evolution of Internet of Things (IOTs), increased usage of social media and video sharing platforms have drastically increased traffic volume putting lot of stress in the existing legacy outside plants and related access network infrastructures. The legacy cable plant relies on RF technologies and analog optics that are difficult to scale, expensive to maintain and not suited to meet demands for new and emerging digital technologies. This required Cox to modernize its access network infrastructures by implementing Distributed Access Architecture (DAA) and going fiber deep to expand customer serviceability, increase bandwidth, and improve efficiency and performance of the access network. To enable the deployment of Distributed Access Architecture and related new technologies, Cox introduced a new packet switched access aggregation network called Converged Interconnect Networks (CIN) based on IPv6 routed technology to connect its headend with new access network.

Cox Communication has one of the largest deployments of Converged Interconnect Networks, Distributed Access Architecture and Remote PHY in North America that covers almost 40% of Cox's footprint. Cox's vision to design a standardized, elastically scalable, and future proof IPv6 routed network has enabled Cox to meet its immediate and future evolution of modern access network paving its path for Next Generation of DOCSIS technology.

2. Legacy Access Networks in Cox and its Limitation

Most Multiple Services Operators (MSOs) use linear point-to-point optical fiber on their access fiber network between headend and primary optical node. However, Cox communication access fiber network, as shown in figure 1, is unique in that it uses a single SMF fiber in a diverse ring topology to add a level of protection from the fiber cuts. The access network fiber topology uses a passive optical architecture to select the shortest path as a primary path between headend and primary optical node. The primary optical node can be located anywhere along the fiber route, resulting in a short path and long path that must be managed optically. The access fiber path can range in overall distance up to 60KM and meets at the primary optical node into an optical bypass switch. The optical bypass switch is responsible for selecting the primary and backup path and provide an optical failover associated with loss of light on the primary path during a fiber cut event. A typical primary optical node in Cox HFC network services 500 household passed (HHP).

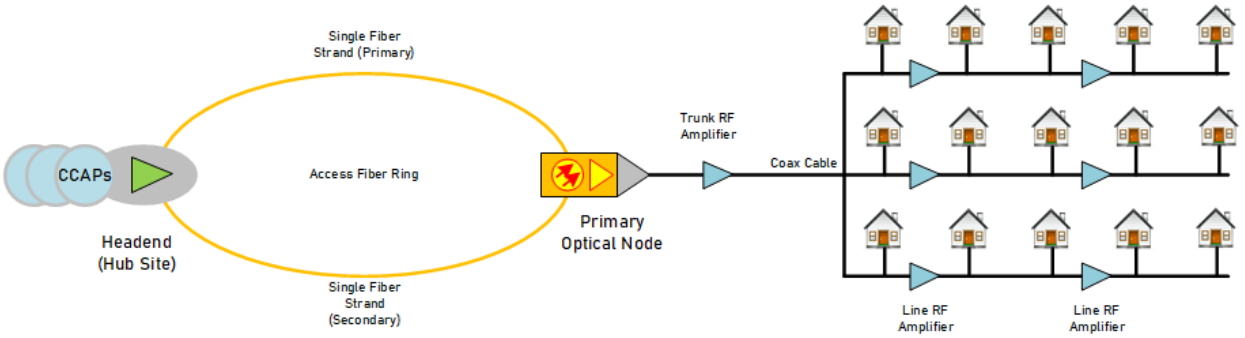


Figure 1 – Cox Communication Hybrid Fiber Coaxial (HFC) Network Architecture

Figure 2 shows the current access network infrastructure of Cox which includes several analog technologies embedded deep into the cable network. The Legacy cable networks have served customers well for a long time. However, with the increase bandwidth demand and new service requirements, these legacy networks

are stretched to their capacity limits. Also, the analog nature of the infrastructure possesses several limitations to the expansion of the network and provisioning of new services. Today's legacy cable plants are prone to signal noise and signal quality degradation because of lots of RF amplifiers between customers and headend equipment. The analog technology also has distance limitation and capacity constraints. We are required to install new CMTS and related auxiliary equipment to provide services to communities that fall just beyond the boundary of the existing headend. Such analog infrastructure is also expensive to maintain and operate and are susceptible to environmental factors.

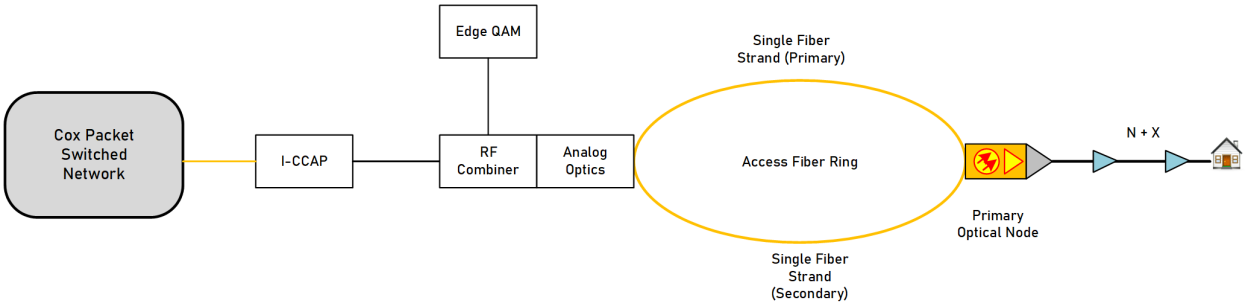


Figure 2 – Cox Communication Legacy Cable Network Infrastructure

As the demand for new and emerging digital technologies grew, the legacy analog architecture was not suited to meet the future requirement of digital communication services. To leverage and benefit from future access technologies like DOCSIS 4.0, R-OLT and beyond, it was imperative for Cox to modernize its analog access infrastructure that would meet customer demands and business requirements of current as well as future services. That is why Cox decided to implement the new access architecture to unlock new possibilities of future network capabilities. However, in doing so Cox needed to carefully integrate the current fiber and network assets to the new access network to maximize the utilization of existing infrastructure and reduce overall capital expenses.

3. Access and Aggregation Network Evolution in Cox

When we discuss the access network evolution in Cox network, it is relevant to know about the overall network architecture of Cox and various segments of Cox network. Like other ISP, Cox network is also a layered network as shown in Figure 3. Cox core backbone is a mesh of backbone routers that connect all its markets, data centers and peering centers. Each market constitutes a network called Regional Data Center (RDC) network that serves as an aggregation layer for that market's metro network. The metro network in each market is a collection of hubs/headends network that is based on hub and spoke architecture. Each headend network is also a hub and spoke network with a pair of hub aggregation routers that aggregates several access, business, and service routers.

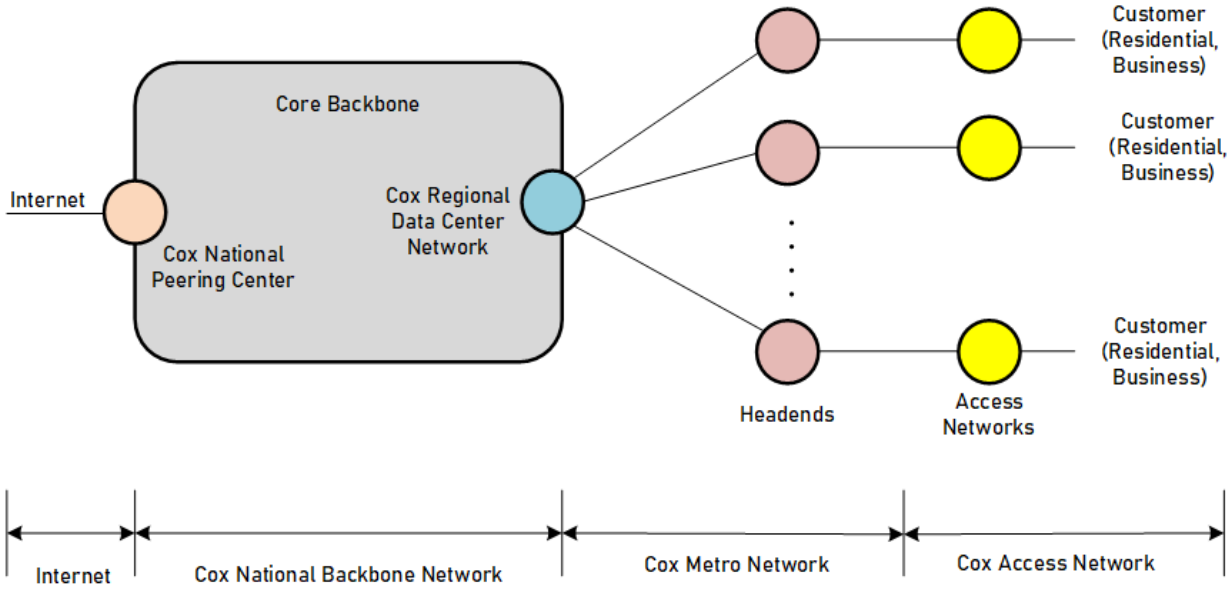


Figure 3 – Cox Communication High Level Network Architecture

As Cox faced continued pressure of increased bandwidth demand and scalability challenges in its legacy access network, it needed to evolve and transform its analog access network to modern digital access network. Distributed Access Architecture (DAA), a fiber deep technology developed by CableLabs was adopted to replace the current access network in Cox.

3.1. Distributed Access Architecture (DAA)

Cox has significant number of fibers and coax network infrastructure in its serving areas. It was important to utilize those existing fibers and network assets to save capital expenditure while continuing to provide higher capacity and reliable services to customers. CableLab's DAA technology was the right access architecture that would fully utilize Cox's existing access network infrastructure and transform it to modern digital access network.

DAA technology allows cable operators to disaggregate traditional I-CCAP into several key network components and functions and move them closer to the subscribers. It helps in reducing power and space requirement in already constrained headend and improve signal quality from customers to headend. The digital transformation also enables operators to automate and virtualize various aspects of new access network infrastructure.

DAA can be broadly classified into two technological variants:

a. **Remote PHY Architecture:** This architecture relocates PHY component of Integrated CCAP closer to the subscriber. A Remote PHY Devices (RPDs) replaces the existing fiber node or a primary optical node in Cox case. Given the maturity of RPHY specifications from CableLabs and availability of RPHY devices from various vendors, Cox has chosen RPHY technology as its DAA architecture. RPDs in Cox are deployed in N+0 as well as N+X model. Currently, Cox has successfully deployed 11,000+ RPDs covering 3.27 million HHP and 2.13 million subscribers.

b. **Remote MacPHY Architecture:** Another variant of DAA is Remote MAC PHY technology where the PHY as well as MAC domains are relocated close to the subscriber. The new access and aggregation network need to support seamless transition from RPD to RMD solution where the control plane and data plane are truly separated. These RMD devices eventually will replace the primary optical node and co-exists along with RPDs for sometime.

Deployment of large-scale Distributed Access Architecture in Cox had to overcome several challenges. Migrating from legacy access architecture to DAA required significant amount of planning, investment, and coordination among several departments in Cox. Deployment of DAA in Cox network was a paradigm shift from design, implementation, operation, and support perspective. The introduction of Layer 3 IP network in the form of Converged Interconnect Network between Digital CCAP and RPD has changed the operational and support modality of Cox metro and access network.

3.2. High – Level Cox Distributed Access Architecture (DAA) Topology

Figure 4 shows the high-level architecture of Cox DAA network deployment. As mentioned in section 1 of this paper, Cox access fiber network is unique in that it is a diverse ring topology from headend to primary optical node instead of linear point to point fiber. To preserve the ring topology of the access fiber, Cox designed and deployed Optical Communication Module Link (OCML) Extender as a DWDM component that would transport multi-wavelength optical signal over the existing ring fiber infrastructure. OCML amplifies and multiplexes 20 unique 10G DWDM wavelengths onto a single fiber. It also demultiplexes all DWDM wavelengths in the reverse direction.

Similarly, Mux/Demux Module (MDM) is a passive device that is located at the point where the legacy primary optical node exists today. MDM demultiplex all the unique DWDM wavelengths coming from OCML and separate them out to ports where RPDs are connected. MDM also multiplexes all the DWDM wavelengths coming from several RPDs onto a single fiber towards the OCML.

Cox utilizes ITU channel #18 through channel #37 for all downstream wavelength mapping whereas ITU channels from channel #42 to channel #61 are utilized for all upstream wavelength mapping. The concept of OCML and MDM has helped Cox to rapidly deploy DAA over existing access fiber infrastructure while preserving it unique ring topology. The use of standard ITU complaint wavelength in OCML/MDM also allows Cox to deploy 100G wavelength solutions in future.

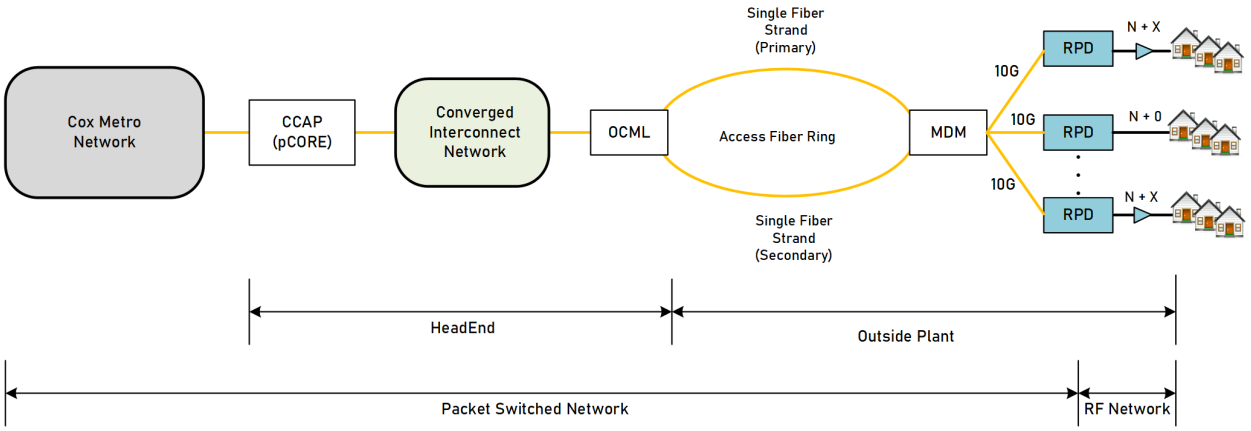


Figure 4 – Cox Communication High Level DAA Network Topology

3.3. Cox Converged Interconnect Network (CIN) Architecture

To enable the deployment of Distributed Access Architecture, Cox introduced a new packet switch network called Converged Interconnect Network (CIN) that connects Cox's headend with new digital access network. Cox's CIN infrastructure is Layer 3 based and is built on native IPv6 only that leverages loop-free communication using industry-standard routing protocols like IS-IS and BGP in control-plane and uses Zero Touch Provisioning (ZTP) to deploy massive number of aggregation switches needed to support the growing CIN deployment. Cox's CIN network is true IPv6 fabric that provides non-blocking, highly elastic, extensible network, and support any-to-any connectivity between RPDs and CCAP cores. This enables Cox to quickly move RPDs between CCAP cores without any physical work needed, and also enables Cox to deploy an "Extended CIN" technology for space-constrained head-ends. The CIN architecture solution deployed at Cox is also able meet the future requirement of regionalized controller based FMA architecture that is separate from the data plane for DOCSIS technology.

CIN Design Principles

As CIN was adopted to interconnect RPDs and various CCAP cores at headend, it blurred the line between several technologies and departments inside Cox. To maintain the CIN network, the collaboration between various teams and departments was integral. Although the CIN network primarily connects RPDs with digital CCAP's cores, it is capable of supporting future access network termination need like RMDs and R-OLTs. Following are few of several design principles Cox adhered to while designing the CIN architecture.

- Effectively and efficiently utilize existing network assets and architectures
- Standard deployment using one common architecture for different size hub sites/head ends
- Provide any – to – any reachability between RPDs and CCAP cores in the market
- Routed L3 only network based on industry standard and proven protocols like BGP, IS-IS and PIM
- IPv6 only network to conserve scarce IPv4 network and avoid complex architecture
- Efficiently utilize existing access fiber topology
- Automation friendly and ZTP capable

CIN Spine-and-Leaf Network Architecture

The implementation of Converged Interconnect Network required an extensive network architecture analysis, testing and meticulous planning. Several design architectures were considered for Converged Interconnect Network. Cox's desire to create Next Generation Access Aggregation Network that is elastic, resilient, fast, and easy to manage led to the implementation of already proven spine – and – leaf architecture in its CIN layer that is popular in Data Center Network today. By using the spine and leaf architecture in CIN, Cox was able to create a highly scalable and hierarchical network that can aggregate extremely large volume of physical connections at every headend.

Several possible combinations of spine – and – leaf architectures were considered for CIN in Cox metro network. Based on Cox's current metro design architecture and design requirement for new access aggregation network, Cox deployed independent RPD aggregation and DPIC aggregation at leaf layer and utilized existing hub aggregation routers as spine layer. Leaf routers that aggregate RPDs are called RPD Aggregation (RPA) routers and routers that aggregate Digital PIC interfaces of digital CCAP are called DPIC Aggregation (DPA) routers. Both RPAs and DPAs are connected to spine layer HUB routers. The overall CIN network as deployed in Cox metro network is show in figure 5.

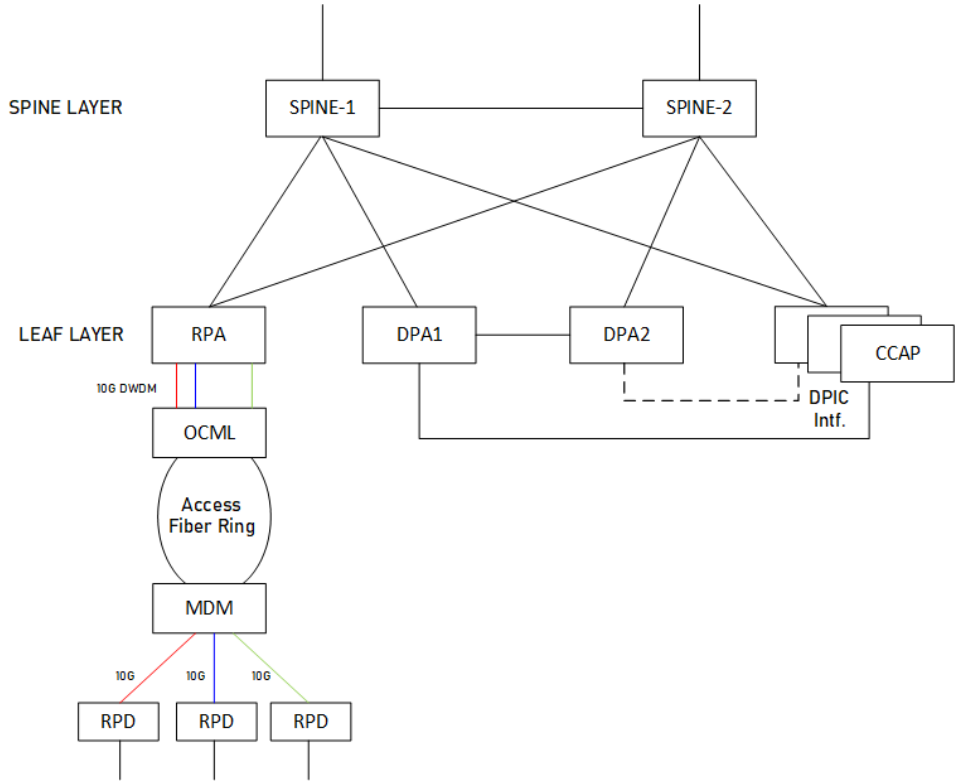


Figure 5 – Cox Communication CIN Topology

The new Converged Interconnect Network consists of following three segments:

- **Leaf Layer – DPIC Aggregation (DPA) Routers:** DPA routers aggregate DPIC interfaces from digital CCAP and are deployed in pair to provide High Availability feature.
- **Leaf Layer – RPA Aggregation (RPA) Routers:** RPA routers aggregate RPDs from outside plants and supports 10G DWDM Bidi optics
- **Spine Layer – Hub Aggregation (HUB) Routers:** A pair of P routers that aggregates RPAs, DPAs and CCAPs.

CIN Routing Design

The routing architecture of new access aggregation network needed to seamlessly integrate with existing metro routing architecture. To achieve all the design and business requirement of new access network, Cox deploys following routing protocols and services in CIN network.

- **Routed IPv6 only Layer 3 network:** To eliminate the dependency on IPv4 address and simplify the network architecture, Cox uses IPv6 only routed network in CIN. The IPv6 routed network allows for globally unique addresses thereby eliminating the need for complex network architecture.

- **IS – IS as an IGP:** Intermediate System – to – Intermediate System (IS – IS) Multi Topology in level – 1 is used as Interior Gateway Protocol in CIN to announce infrastructure prefixes like loopback and point – to – point interfaces of CIN devices.

- **MP – BGP:** Multi-Protocol BGP (MP – BGP) with unicast and multicast address families is used to advertise non-infrastructures prefixes. Spine routers are route reflectors to all leaf routers. BGP traffic engineering, like manipulation of local – preferences and route – aggregation, are used to enable deterministic routing and predictable traffic flows in CIN.

- **PIM:** As 99% of traffic in CIN is IPv6 multicast, PIM SSM is deployed as multicast routing protocol. MDLv2 is enabled on RPD Aggregation routers so that multicast listeners are automatically discovered on all MLD enabled interfaces.

- **Quality of Service:** To prioritize various types of traffic (control, user data, voice, video), class of service is configured with proper classification, mapping, and queuing.

- **802.1x:** To authenticate RPDs, RPD aggregation routers are configured with 802.1x authentication.

- **Other protocols and service**s: As a standard practice, all CIN layer routers are configured with other necessary protocols and services like LLDP, DHCPv6 relay, TACACS, NTP, SNMP etc.

3.4. Cox Extended CIN (E-CIN) Architecture

One of the main design features of CIN was to accommodate any – to – any connectivity between RPDs and CCAP cores in the market. The architecture, called Extended CIN, allows RPDs to connect to any CCAP cores that are not co-located at same headend. In E-CIN architecture, CCAP cores are located on a separate headend (called host site) than the RPD aggregation routers (called the E-CIN site). The high-level architecture for E – CIN is show in figure 6.

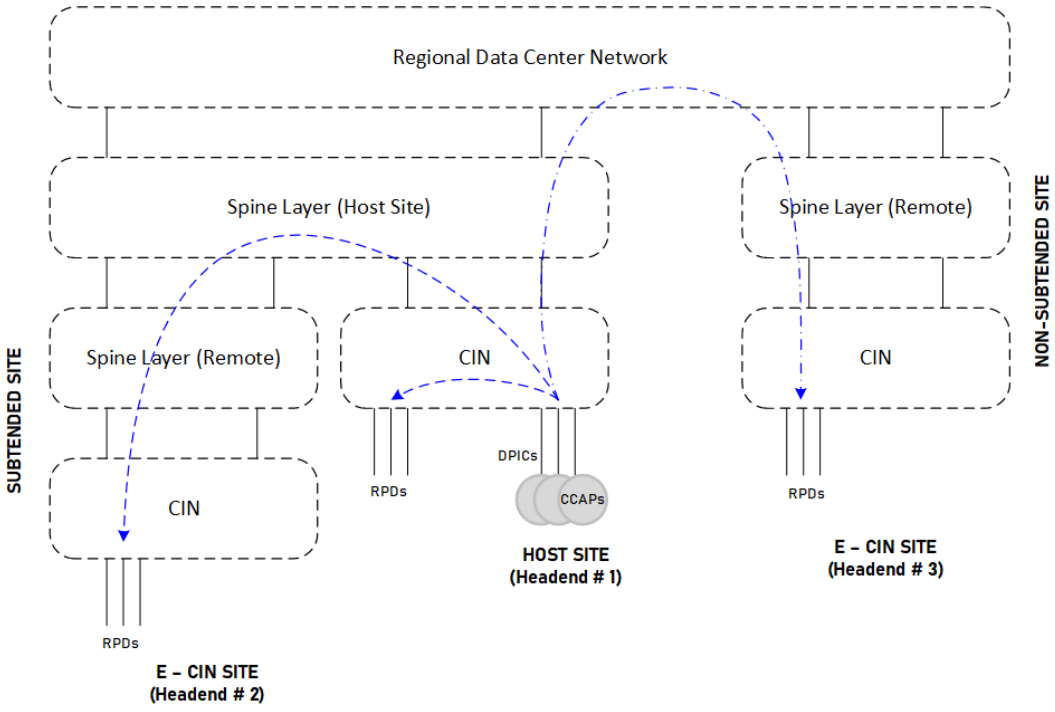


Figure 6 – Cox Communication High-level E – CIN Topology

In the E – CIN architecture of figure 6, host site's (Headend #1) CCAP cores serve headend # 2 (Subtended Headend to Headend #1) and headend # 3 (Non-subtended headend). Extended CIN, in some cases increases optical distance and introduces more hops in the routing path between host site and E – CIN sites which might add challenges to sensitive control plane and real time traffic. To address such issues, Cox has implemented some design features and specified the optical fiber distance for which E – CIN architecture can be implemented without sacrificing any traffic quality. Cox also implements IGP traffic engineering to keep E – CIN site traffic in optically shortest path. Also, proactive bandwidth capacity planning and augmentation ensures that there is enough capacity between host site and extended CIN sites all the time.

3.5. PTP Architecture

For proper DOCSIS operation, the timing synchronization between CCAP and RPDs is extremely important. To maintain such timing synchronization, Cox utilizes IEEE 1588/PTP protocols in hierarchical architecture. A pair of Grandmaster clock at the RDC level are master for all boundary clocks deployed at each headend. These boundary clocks provide timing for CCAP cores and RPD devices. Given the sensitive nature of timing synchronization between RPDs and CCAP, PTP traffic is symmetrically routed in CIN and metro network to eliminate packet delay variation. It is also treated as priority traffic in network to minimize queuing delay.

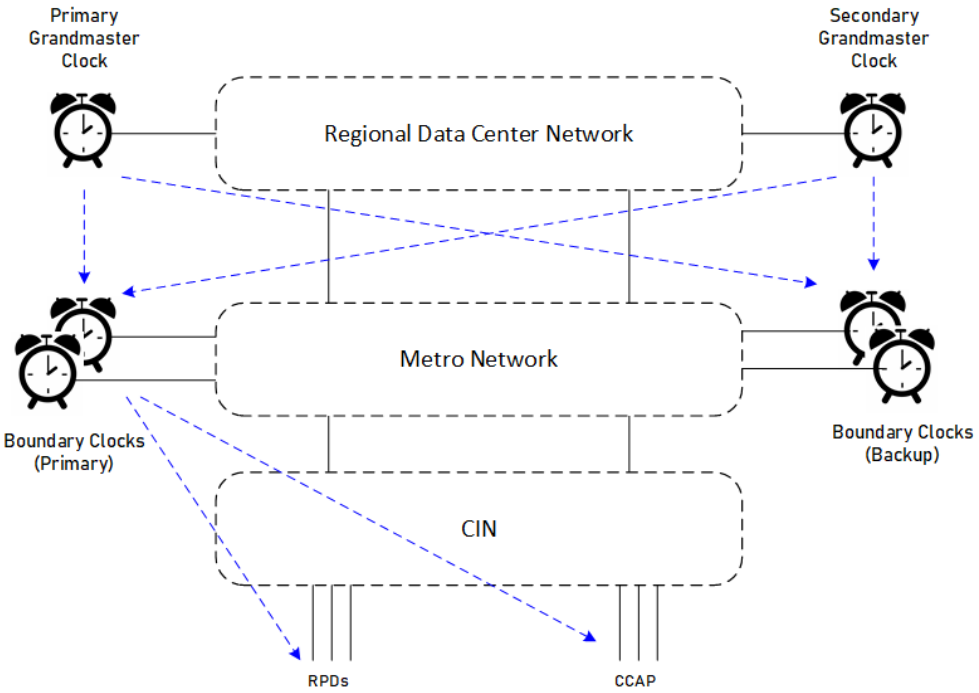


Figure 7 – Cox Communication High-level PTP Architecture

4. Access Aggregation Network Standardization and Automation

For the large-scale deployment like access aggregation network/CIN, it is extremely important to standardize the network deployment to minimize misconfigurations and enforce predictable network performance. A CIN playbook was created to document all protocols, services and configuration standards related to CIN deployment. To make the network automation easy, configurations on routers are done in such a way that they are user friendly and easily parse by any choice of programming languages. Configuration compliance are strictly enforced by compliance tools.

Network modeling and scale analysis showed that Cox will be deploying huge numbers of RPD and DPIC aggregation routers. As manual deployment of these devices is operationally expensive and consume lots of time and resources, Zero Touch Provisioning (ZTP) has been utilized to turn up these devices. ZTP engine is responsible to create configurations and apply them to all corresponding devices during new device turn up.

5. Conclusion

The tremendous growth in demand for bandwidth is pushing the legacy analog infrastructure to its capacity limits opening the opportunity to deploy modern network infrastructure in the access side of the network. Building upon this momentum of increased demand, Cox is deploying a new DAA and CIN architecture to push fiber deep into its serving areas. Cox started with Remote Phy (RPHY) as its preferred DAA technology. The foundation laid during the modernization of access and access aggregation network has far reaching impacts in Cox network. The access agnostic nature of new aggregation network allows Cox to deploy several IP and non-IP services over the same converged network. Pushing ubiquitous digital transmission technology as close to the subscriber as possible has helped Cox in delivering better services with improved quality to its valued customers. The real time performance monitoring of network assets deep into the network has also helped Cox to proactively address impending issues before it gets service impacting.

As Cox embarks on a path to the next generation of DOCSIS technologies like RMD/FMA, R-OLTs and business services, the new access and access aggregation network will offer the new and improved digital capabilities and versatilities as demanded by new and emerging technologies.

Abbreviations

DAA	Distributed Access Architecture
CIN	Converged Interconnect Network
CCAP	Converged Cable Access Platform
OCML	Optical Communication Module Link Extender
MDM	Mux/Demux Module
DWDM	Dense Wave Division Multiplexing
MP – BGP	Multi-Protocol Boarder Gateway Protocol
IS-IS MT	Intermediate System to Intermediate System Multi Topology
PIM	Protocol Independent Multicast
PTP	Precision Time Protocol
RPD	Remote PHY Device
RMD	Remote MAC PHY Device
FDX	Full Duplex
ESD	Extended Spectrum DOCSIS
DPIC	Digital Physical Interface Card

Bibliography & References

DWDM Access for Remote PHY Networks Integrated Optical Communications Module (OCML), Harj Ghuman; 2017 SCTE-ISBE and CTA

Monitoring and Troubleshooting at Scale with Advanced Analytics

A Technical Paper prepared for SCTE by

Nitin Kumar
Principal Architect
Harmonic, Inc
2590 Orchard Parkway, San Jose, CA 95131
+1 408 542 2559
Nitin.Kumar@harmonicinc.com

Amir Leventer
Senior Director, System Architecture and Networking
Harmonic, Inc
2590 Orchard Parkway, San Jose, CA 95131
+1 408 542 2559
Amir.Leventer@harmonicinc.com

Asaf Matatyaou
Vice President, Solutions and Product Management, Cable Access
Harmonic, Inc
2590 Orchard Parkway, San Jose, CA 95131
+1 408 542 2559
Asaf.Matatyaou@harmonicinc.com

1. Introduction

Access systems such as Cable Modem Termination Systems (CMTS) have traditionally been monitored using tools that leverage CLI and SNMP interfaces. These same interfaces are also used to gather live information while troubleshooting issues in the field. The data exposed by these interfaces are limited by the standard set of commands and MIBs supported by the system. With the move to a distributed access architecture (DAA), deployments can support much higher scale, and their software-centric designs make available richer data useful for vendors in operating, debugging, and optimizing the systems, but access to this increased amount of data is bottlenecked by the limited performance and limited extensibility of CLI and SNMP interfaces. Newer system designs support streaming logs and telemetry to overcome these limitations, and this paper looks at how these features not only overcome the limitations of traditional interfaces but also enable more efficient monitoring and troubleshooting workflows. We will show the building blocks of a system that implements streaming logs and telemetry. The performance and security improvements offered by such a system will be noted. We will illustrate some monitoring features and describe something we call "time-travel debugging" — using the streamed data for easier troubleshooting. Finally, we will give an overview of using the data for advanced analytics and machine learning, specifically for faster root cause analysis of field issues.

2. Monitoring and Troubleshooting at Scale with Advanced Analytics

Monitoring and troubleshooting traditional CMTS deployments can be a labor-intensive activity. Deployment health is monitored via SNMP or CLI scripts using metrics such as modem registration status or call volumes to customer support. Basic troubleshooting of common issues can be done by operations or vendor support personnel logging into each platform remotely to run monitoring commands and look at locally stored logs. Complicated issues may require direct involvement of development engineers to analyze logs and perform further troubleshooting steps. Some issues with significant service impact may occur only intermittently and may not be easily reproducible. Fault monitoring and resolution is mostly a manual process.

Compared with traditional architectures, the DAA allows deployments to be launched quickly, and operators can provide increased bandwidth to more subscribers. The troubleshooting load will grow with the scale of deployment, with a corresponding increase in operational overhead.

Monitoring and troubleshooting workflows can be augmented with tools that enable operators to handle the increased scale while keeping operational overhead under control. There are non-obvious but nevertheless important implications to vendors and operators, which will be explored in subsequent sections.

3. Troubleshooting

Let's look at some of the challenges with troubleshooting production deployments at scale and explore some of the solutions possible with new tools and workflows.

3.1. Challenges

As deployments scale to handle increases in bandwidth and number of subscribers, there's a corresponding increase in issues that occur and need to be debugged. Issues having significant operational impact can be quite complex to root-cause. Here is a list of potential issues:

- Transient cable plant problems

- Vendor-specific interpretation of DOCSIS specifications
- Non-compliant cable modems (sometimes with different firmware versions exhibiting different behavior on the same model)
- Incorrect handling of new DOCSIS features by older equipment
- Intermittently malfunctioning equipment
- Complicated interactions between specific configurations and plant population
- Combination of multiple issues

3.1.1. Limited Storage

With limited onboard storage on the CMTS, only a reduced number of data samples are stored for a short time span, severely hampering analytics and troubleshooting efforts that require a longer history of data. The logs pertaining to the issue may no longer be available, and troubleshooting efforts can be delayed until the issue occurs again, or effort is made to actively reproduce the issue. Because storage capacity is limited, only limited logging is enabled by default, and this means that logs are not as information-rich as they could be for some root-cause analysis.

3.1.2. Availability of Expertise

When logs are available, there may be only a limited number of engineers with the knowledge required to interpret the logs and correlate them with the observed issues. Thus, troubleshooting can be limited by availability of skilled personnel to handle the increased workload.

3.1.3. Security Considerations

Scale operations involve software running across many IP hosts, and this necessitates maintaining a list of IPs for support personnel to log in to. Login accounts need to be set up and/or passwords shared for debugging efforts. Managing access to these accounts and hosts need to be done carefully to avoid security issues. For example, the integrity of a deployment can be compromised by a single compromised account, or by a compromised host accessing a jump host.

3.1.4. Complexity of Issues

A DAA deployment has many hardware and software components. Problems can occur in any of these components, and some issues will be due to complex interactions between these components.

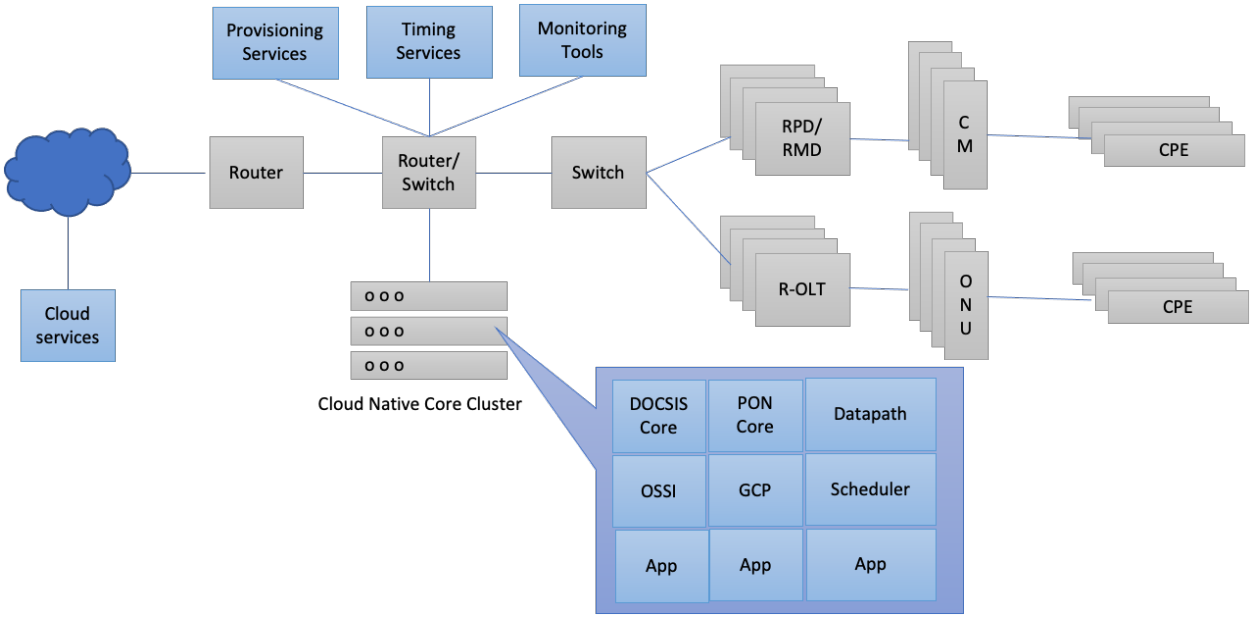


Figure 1 - A deployment has many possible fault sources

Logs and monitoring data from multiple sources must be analyzed to pinpoint the cause of a failure. Traditional troubleshooting processes can be overwhelmed by the volume of logs and metrics generated by the many software components of a modern cloud native solution. For instance, logs from multiple sources might need to be examined to troubleshoot a modem bandwidth problem. Generating a timeline of events across multiple system components will be challenging without new tools.

3.1.5. Legacy Monitoring Tools

Traditionally, the monitoring interface of a CMTS is defined by the DOCSIS OSSI standard, and tools such as CLI and SNMP are used to query the system to gather performance monitoring metrics. The OSSI standard does not specify data objects for all metrics that can be exposed by modern software designs, and CLI and SNMP tools do not scale well to handle vast amounts of fast-updating data. With increasing scale, using CLI for manual real-time analysis becomes impractical. A data pipeline constrained by these legacy tools limits the richness of data available for analysis and increases latency in detecting critical changes in key metrics.

3.2. Cloud Native Solutions

Cloud native, software-based solutions are now the popular field-proven choice for cost-effective, high-scale CMTS deployments. Service is provided by loosely coupled microservices deployed in containers running in a compute cluster. Each microservice exposes its own metrics and logs that provide highly granular insights into their operation to aid with advanced monitoring and troubleshooting. A vast amount of data is generated by the software components, and new tools are needed to process and consume the data.

3.2.1. Streaming Logs

Logs are streamed out to an external (cloud-based) service such as Elasticsearch. Analysis tools automatically monitor logs for known patterns and can generate alerts on critical events. Dashboards summarize log statistics and visualize events in time-series graphs, making it easier for operations

personnel to monitor and troubleshoot systems. Search capabilities allow complex log patterns to be easily searched across multiple log sources.

3.2.2. Streaming Telemetry

Each software component exposes performance and state metrics that are continuously streamed to external services for monitoring, analysis, and visualization. Threshold-based alerting indicate when metric values are outside expected bounds, allowing operations personnel to investigate potential issues before they lead to failure. Dashboards help visualize the state of the system, and visual indicators draw attention to potentially problematic states.

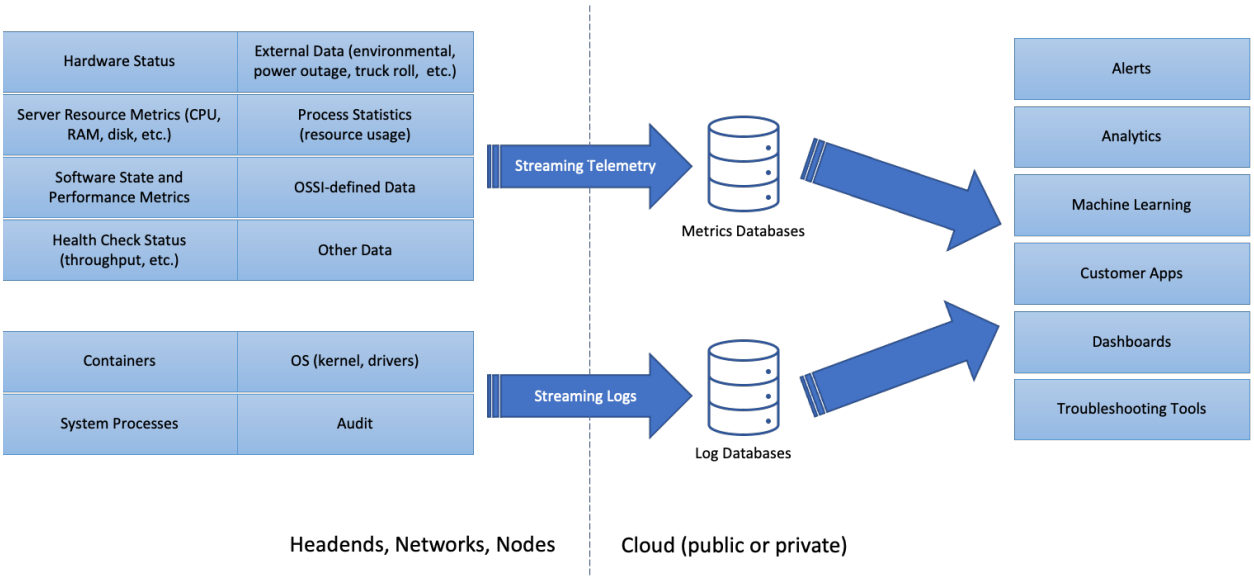


Figure 2 - Logs and metrics are streamed out and used for various applications

Streaming logs and telemetry mitigate the challenges with troubleshooting mentioned previously. The next sections will discuss how.

3.2.3. Scalability

The scalability of compute and storage resources in the cloud allows vast amounts of logs and metrics to be stored and analyzed. Since the data is continuously streamed off the system, local storage no longer limits availability of the data. Also, DAA deployments consist of many software components running across many servers, and tracking which server runs what software component while also figuring out which components are contributing to a field issue quickly becomes a problem if troubleshooting requires direct access to the server. With logs and metrics streamed out to accessible and external cloud-based services, such considerations are no longer an issue.

3.2.4. Automation and Analysis Tools

Logs and metrics are automatically analyzed for indicators of failure or violations of SLA guarantees. Alerts are automatically generated to draw attention of operations personnel to problematic areas. Data is visualized in dashboards and reports generated to ease understanding of system state without manually digging through logs or running debug commands. Search interfaces and interactive data exploration tools allow filtering out data irrelevant to a troubleshooting activity and drilling down to additional details in

areas of interest. The data pipeline is continuously augmented with new analysis algorithms, dashboards, alerts, and other features based on experience and input from operations and domain experts so that troubleshooting activities are continuously being simplified. Since most of this runs in the cloud, adding new features to the pipeline do not generally require changes on core servers that might introduce risks such as increased CPU usage and core software bugs.

3.2.5. Security

Access to logs and metrics no longer require logging into a deployment instance since all debug and system state information are streamed out and made available in secure and globally accessible external systems. There may be many deployments streaming data to a few external entities. Managing access to these external entities is much simpler than managing direct access to deployments. Since logs and metrics are pushed out to the external systems, there are no incoming connections into the deployments, further enhancing security and simplifying firewall rules.

3.3. "Time Travel" Debugging

"Time Travel" here refers to being able to look at the state of the system at a point in time in the past while troubleshooting a problem.

In many cases, the failure that needs to be diagnosed occurred in the recent past, and operations personnel might have already executed recovery procedures to restore service before support engineers have had a chance to access the system to gather information needed to root cause the failure. Without streaming logs and telemetry, most of the information needed to diagnose problems would be stored on the system itself. As previously mentioned, the amount of data stored is limited because of onboard storage capacity limitations. By the time someone is able to access the system, the logs relevant to the failure might have already been flushed from the system. Further, default logging levels might not have all the information needed to debug the problem. Since recovery procedures were executed, the system is no longer in a problematic state, and state information critical to determining the root cause of the failure has been lost.

Such situations lead to prolonged troubleshooting and root cause analysis efforts. Lack of system state and logs relevant to the failure would require waiting for the failure to occur again and being able to capture logs and system state leading up to the failure. Recording system state leading up to the failure can be challenging if that state is not captured in logs or by OSSI tools such as SNMP. Initial analysis might determine that additional logs or debug features need to be turned on, requiring further reproduction of the failure. Additionally, if live debugging is needed (requiring access to the system in the failed state), there could be prolonged service outages while the failure is being analyzed. Debugging under time pressures can result in erroneous analysis.

Streaming logs and telemetry mitigate these issues. Onboard capacity constraints are no longer a limiting factor, so logs are available in an external log database for extended time periods, and richer information can be logged by default (although there are cost implications; see later section). This means that the logs pertaining to a failure can be looked at, long after the failure occurred. Streaming telemetry makes periodic (many times a minute) state snapshots available in external databases, enabling the state of the system at any point in time to be analyzed. Operations personnel can execute recovery procedures without worrying about losing critical historical debug information.

In a sense, operations and support personnel and development engineers can "travel through time" to observe the state of the system at any point in the past. This can be done simply by using database query interfaces to search for information pertaining to the system at any time. Tools can show how various

state metrics change over time leading up to a failure. This means that failures can be analyzed without having to wait for multiple reproductions of the issue.

The state information streamed out is comprehensive and not constrained by the limited time resolution and information richness of CLI or SNMP data models, so almost any information needed to analyze the system at any point in time is always available. During root-cause analysis, various subsystems might need to be analyzed as the analysis progresses. Various domain experts will need to be brought in at different points of time. Since the streamed-out state is rich with information across all subsystems, there is seldom a situation where additional reproduction efforts are needed because information in some specific domains are lacking. Separate subsystems can be analyzed independently — each stage of the analysis can use the tools to explore system state across various subsystems over various periods of time. The result is faster root-cause analysis and minimal downtime.

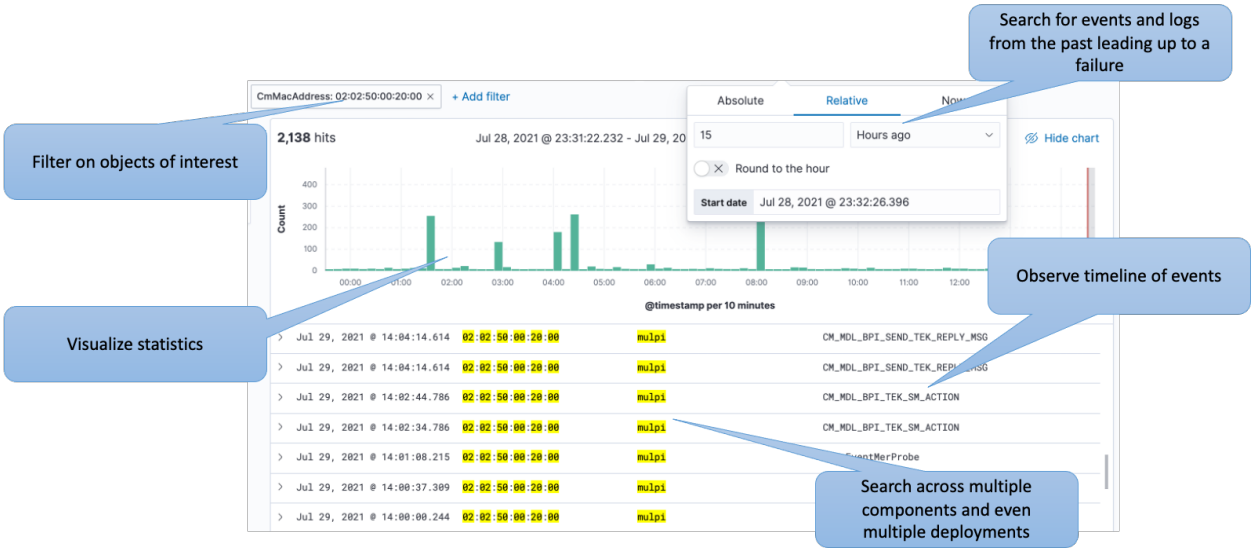


Figure 3 – Search and analysis of logs using tools such as Elasticsearch and Kibana

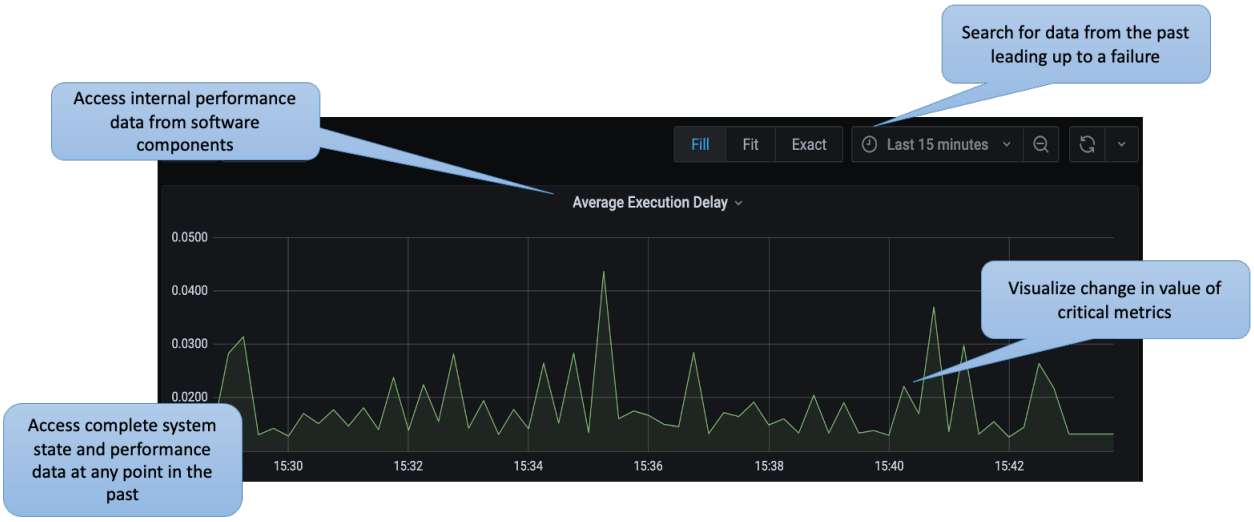


Figure 4 - Data visualized and analyzed using tools such as Prometheus and Grafana

Looking into data from the past can help determine the reason of a current system fault. Current symptoms and logs and telemetry data can be compared with analysis from previous investigations and corresponding logs and telemetry signatures. For help with future investigations, a database of failure events and their corresponding logs and telemetry snapshots may be created. Frequency of each event can be tracked to generate correlations between the events and certain logs and metrics, and this can be used to isolate root causes to specific system components.

4. Advanced Analytics

The data available in the cloud can be used for other data-driven apps such as billing and analytics. Data from the past can be used to train predictive algorithms and fed into machine learning pipelines. Closed-loop controls can be implemented where external applications analyze the streamed data and call back into system configuration interfaces to make performance-optimizing adjustments. In this section, we will take a brief look at some of these applications.

4.1. Auto Observability

Observability of systems must scale with the number of deployments. Manually monitoring systems and taking manual corrective actions impede the ability to expand service footprint and operate at scale while maintaining SLAs. A robust data analytics pipeline can help with this.

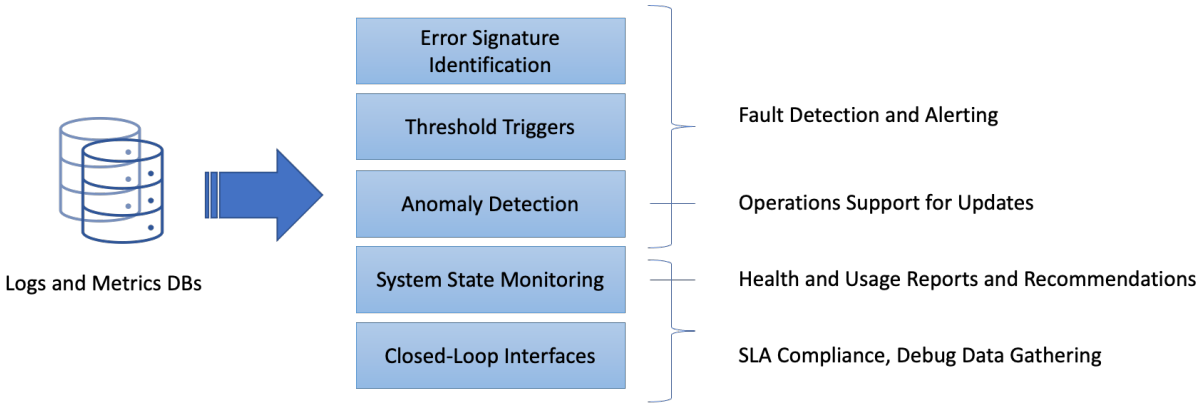


Figure 5 - Observability tasks are automated using data processing pipelines

4.1.1. Identify Failures

Comprehensive system state information is available in streamed telemetry and logs. Failures can be identified using basic data analytics (such as by identifying log signatures and metrics thresholds corresponding to failures) or by advanced approaches such as machine learning.

4.1.2. Rapidly Root-Cause Issues

At scale, failures must be analyzed quickly before they propagate and affect a large population of users. Failures detected during upgrades or operational updates must be diagnosed immediately to prevent stalling of planned operational procedures and rollouts. Potential issues introduced by an update can be identified by anomaly detection algorithms analyzing streamed logs and telemetry.

4.1.3. Proactive Debug Data Collection

Some investigations require more debug data than what's provided by logs and telemetry (such data might incur too much overhead to be streamed in logs and telemetry). For instance, an issue with specific kinds of data traffic might require analyzing a packet capture. Such data may be collected automatically when a potential issue is identified. This can be done by triggering advanced data collection when certain thresholds are exceeded. For example, a set threshold of subscriber-reported traffic problems can automatically trigger a traffic recording process for more advanced investigation.

4.1.4. Maintain SLAs

Performance metrics must be monitored across all deployments to ensure service quality. Monitoring tools must analyze streamed data to identify, report, and potentially correct problems as soon as they occur. Reports can identify subscribers who may benefit from moving to a different service tier. Closed-loop pipelines can observe streamed data and call into system configuration interfaces to take corrective action or adjust configuration to maintain SLA targets. For example, OFDM RxMER data from cable modems can be monitored to dynamically create and apply OFDM profiles optimized for current plant conditions. Having such closed-loop pipelines run in the cloud enables the associated algorithms to be easily updated and the performance of the algorithms themselves be monitored.

4.1.5. Characterize System Health

Data provided by streaming telemetry and logs must be analyzed to characterize system state and usage over time. The knowledge of system behavior provided by such analysis will aid in optimizing existing deployments and forecasting and planning for future needs. Automated reports can give an overview of system usage, and analysis pipelines can provide recommendations on managing deployments (such as recommending node splits based on bandwidth data or enabling features such as dynamic load balancing and equalization).

4.2. Multi-Variable KPI

Key Performance Indicators (KPIs) tracking customer satisfaction can be optimized for using data analysis and machine learning. As an example, bandwidth utilization could be a KPI, and optimizing it can involve tuning multiple parameters. RF parameters (modulation profile), subscriber density, cable modem configuration, and connectivity bandwidth, are just some of the things that contribute to optimal bandwidth usage. However, these parameters can also have adverse influence on some other metrics. Let's look at a couple of these parameters.

4.2.1. Modulation Profile

More robust modulation profiles may be configured when RF plant conditions deteriorate. Streaming telemetry includes RF metrics such as counts of uncorrectable and corrected FEC codewords, and this can be monitored to determine an appropriate modulation profile. A single-variable optimization would be to maintain the counts below a certain threshold. However, the more robust modulation profiles usually reduce available bandwidth as well. If the optimization is based on just optimizing the FEC-related counts, bandwidth utilization is adversely affected.

4.2.2. Subscriber Density

More per-user bandwidth can be made available by reducing the number of users sharing the RF spectrum. This involves a node-split. Bandwidth utilization is reported via telemetry, and peak bandwidth

can be used as a threshold to decide on splitting a node. Additional nodes imply additional cost; therefore, optimizing on just the peak bandwidth can adversely affect cost.

Thus, monitoring and optimizing a single parameter can adversely affect other KPIs. What's needed is optimization across multiple variables: configuration must be dynamically adjusted to maximize utility over multiple metrics such as FEC, bandwidth utilization, and cost. One method of achieving this is by using a machine learning pipeline based on reinforcement learning.

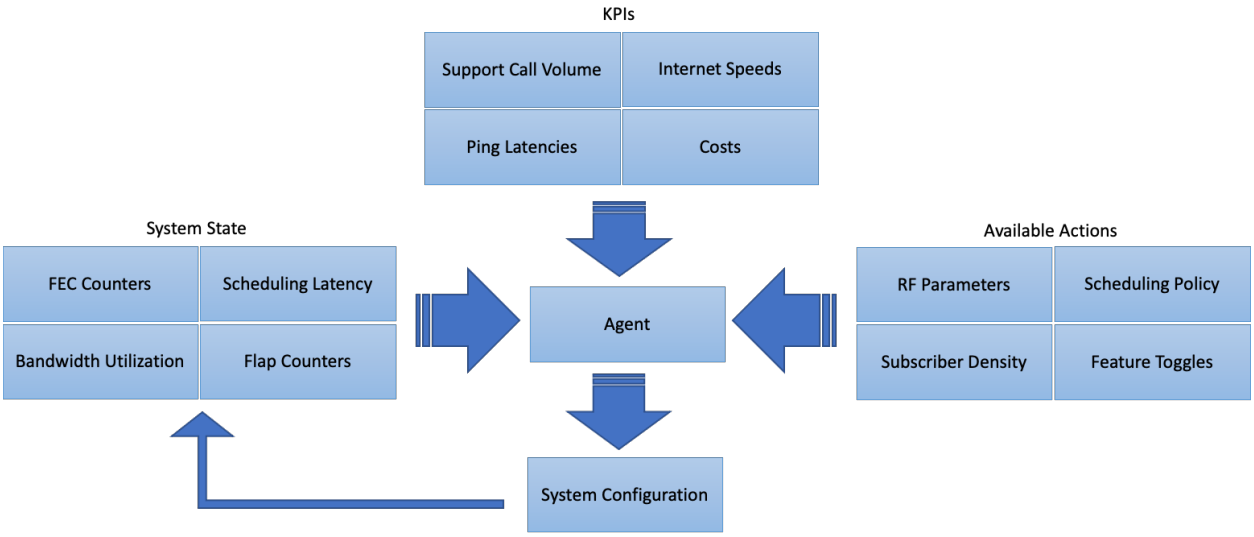


Figure 6 - Machine learning pipelines optimize key performance indicators

4.3. Fault Isolation

A DAA deployment contains many hardware elements (RPD, core servers, RF plant components, network switches and routers, etc.) and software components (RPD software, core software, software on network elements, OSSI tools, etc.). Each of these components are further composed of many subsystems, and many components are duplicated for redundancy. The overall system can be quite complex, with many potential points of failure. When a problem occurs, it can be challenging to pinpoint the source.

As an example, consider a case where some cable modems go offline. The root cause could be plant RF issues, a power outage event, a transient failure on a node or RPD or RF chassis, a network issue, a software crash, a hardware resource issue, or any number of other possibilities. A first step in analyzing the failure would be to identify the point in the topology where the failure originated. If the root cause is a power failure, then the metrics for multiple components would show a problem depending on the affected outage area. Identifying the cause of failure is not easily done just by looking at system dashboards. However, an analytics pipeline that is fed all the relevant data, including data from power supplies and third-party data sources such as utility companies, can correlate failures to the observed events, and combined with topology information, indicate the origin of the failure.

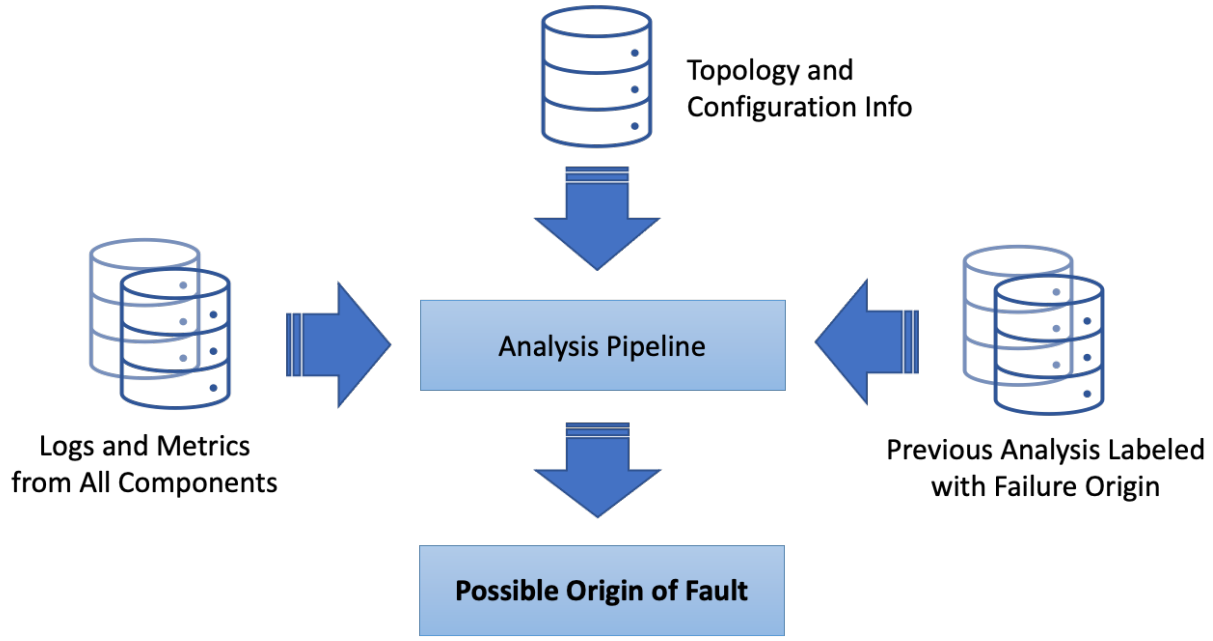


Figure 7 - Logs and metrics signatures indicate fault origin

Past analysis efforts can be used to automate future identification of root causes. Streamed data is continuously analyzed for patterns and anomalies, and problems identified in the past can be used to label the data and used to train machine learning algorithms. The algorithms can then automatically identify failures and indicate the root cause in future occurrences.

5. Other Considerations

Data analysis and machine learning enable tools that provide powerful new capabilities. There are some potential issues that need to be considered when deploying these tools at scale. This section will examine a couple of them.

5.1. Data Privacy

Logs and telemetry data can contain privacy-sensitive data. Information such as cable modem registration attempts, subscriber data usage, subscriber data activity, IP and MAC addresses, and other identifiers can potentially be combined with external data sources to gather information about subscribers. It is best to not collect such data in logs and telemetry. Identifiers must be sanitized, and any data collected must be anonymous. Data collection practices must be guided by both moral and legal (such as GDPR) considerations.

5.2. Infrastructure Cost

The most convenient and scalable method for running the components of the analytics pipeline discussed in this paper involves use of public cloud infrastructure. The associated cost can be significant if not managed well. It can be challenging to assign a dollar amount to the value provided by the pipeline so that a budget can be allocated for the public cloud services. For instance, the value is close to nil for the troubleshooting use case when everything works and there are no problems to be debugged. However, the services provided by the pipeline can be invaluable when a complex issue occurs that would have

otherwise required the operator's or vendor's most experienced engineers to spend significant chunks of time analyzing and root causing the issue. The cost analysis is further complicated by the fact that this time spent troubleshooting takes away valuable time from those engineers.

The features provided by the pipeline, such as smart alerting and usage reports with data analysis, can significantly improve the operator's customer satisfaction metrics. The value provided by such features also needs to be considered when assessing the cost.

Regardless of the challenges assigning a value to the services, there are some knobs that may be used to control the costs.

5.2.1. Compute vs. Storage

The most expensive cost component is cloud compute resources. The data streamed to the cloud needs to be processed for analysis and display. For instance, compute resources are needed to index logs to enable fast searches. On the other hand, storing the data is relatively inexpensive. The pipeline can be tweaked to store raw data and use compute resources only on demand. The main drawback of this strategy is increased latency in accessing the data.

5.2.2. On-Prem vs. Public Cloud

Running the services on-premises instead of using a public cloud provider can reduce operational expenditure (depending on how the cost associated with operating the on-prem infrastructure is accounted for), with the caveat that additional overhead is incurred maintaining the on-prem hardware and software. Service expense is traded for capital expenditure. Drawbacks of this approach include additional difficulty in aggregating and analyzing data across multiple deployments, and more challenging upgrade processes and infrastructure scalability. In general, on-prem infrastructure might be a good option if all use cases are well-defined and already implemented and scale of operations is known in advance. Machine learning and advanced data applications are still evolving and use of public cloud infrastructure can provide flexibility in accommodating the changing use cases.

5.2.3. Data Retention

Storage costs can be managed by limiting the amount of data that is retained. For example, data older than a certain age can be discarded. Some architectures support a warm or cold storage option where older data is stored in cheaper (but higher latency) storage mediums. Discarding older data has the disadvantage that longer-term seasonal patterns may be missed by data analysis tools; this may be mitigated by feature generation capturing traits of the data and storing that longer term.

6. Conclusion

In this paper, we looked at how modern data analysis and machine learning tools help with operating critical infrastructure at scale. We examined the challenges posed by legacy tools and systems and how those challenges are mitigated by software-centric solutions that support streaming logs and telemetry. We explored how data and tools can be used as a framework for powerful troubleshooting capabilities that simplify analysis of failures. We looked at some of the applications made possible by using automated data analysis and machine learning algorithms. We also looked at some potential issues that need to be considered when deploying these applications at scale.

The applications and use cases presented in this paper are just a sample of what's made possible by rich data sets and advanced data analytics. As deployments scale to support more subscribers and bandwidth,

such applications will be an essential component of successful operations and critical to ensuring customer satisfaction.

Abbreviations

CLI	command line interface
CM	cable modem
CMTS	cable modem termination system
DAA	distributed access architecture
DOCSIS	Data-Over-Cable Service Interface Specifications
FEC	forward error correction
GDPR	General Data Protection Regulation
IP	Internet Protocol
KPI	key performance indicator
MAC	media access control
MIB	management information base
OSSI	operations support system interface
OFDM	orthogonal frequency division multiplexing
RF	radio frequency
RPD	remote PHY device
RxMER	receive modulation error ratio
SLA	service-level agreement
SNMP	simple network management protocol

Bibliography & References

Deep Learning: A Visual Approach, Andrew Glassner; No Starch Press

Elasticsearch: The Definitive Guide, Clinton Gormley & Zachary Tong; O'Reilly Media

Practical Time Series Analysis, Aileen Nielsen; O'Reilly Media

Cloud Native Definition, Cloud Native Computing Foundation

DOCSIS Specifications, CableLabs

Navigating the Transition to a Post-Quantum World

A Technical Paper prepared for SCTE by

Chujiao Ma
Senior Security R&D Engineer
Comcast Cable Communications, LLC
Philadelphia, PA, USA
Chujiao_ma@comcast.com

Vaibhav Garg
Sr. Director Cybersecurity Research & Public Policy
Comcast Cable
Blacksburg VA
Vaibhav_garg@comcast.com

1. Introduction

Quantum computing is an emerging technology that can dramatically change the security landscape. Unlike traditional computation that processes information in binary bits, 0 or 1, the information in quantum computing is stored in a particle in a quantum state called a "qubit." Qubits exist in a superposition, which means they can be 0 or 1 or everything in between, until measured. This allows quantum computers to simultaneously perform computations for a range of inputs. In practice, this reduces the computational complexity of certain algorithms from exponential to polynomial time -- from $O(e^n)$ to $O(n)$. This means that to solve certain classes of problems that would take classical computers hundreds of years, quantum computers may only take days or even hours.

For example, Grover's algorithm provides a quadratic speedup on unstructured search problems [1], whereas Shor's algorithm can be used to factor the product of two large prime numbers in polynomial time [2]. This has an impact on the security of current crypto systems. Grover's algorithm halves the security of current symmetric keys and Shor renders all public key cryptosystems insecure [3]. These classical public key algorithms are used ubiquitously in security protocols for digital signatures, authentications, key transport and authorization. These algorithms secure cyber infrastructure, from software distributions to virtual private networks. Thus, the construction of a large enough quantum computer will require a transition to quantum-safe alternatives to ensure the continued security of these systems.

It is unclear whether such computing capacity will be available in the near future -- but it's a matter of when, not if. Many experts posit that there is a substantial probability of this happening in the next 20 years [4]. This may create the impression that the threats associated with quantum computing have a long-time horizon. However, it is important to consider the challenges in crypto transitions. For example, the transition from Secure Hash Algorithm 1 (SHA1) to Secure Hash Algorithm 2 (SHA2) took over 10 years and cost organizations $5M on average [5]. In that light, a transition across all cryptography can seem overwhelming. However, there is much cause to keep calm and carry on.

In the U.S., the National Institute of Standards and Technology (NIST) has a process underway to determine a list of post-quantum cryptography (PQC) algorithms to replace current public-key cryptography [6]. Open Quantum Safe has open-source implementations of many PQCs for benchmarking and exploration [7]. Additional support is available from European Union projects PQCrypto and SAFEcrypto, as well as CREST Crypto-Math project in Japan [41]. Cloud computing providers, such as Amazon Web Services (AWS), already provide the option of using hybrid cryptography, which combines PQC within a classical cryptography wrapper [8].

There are three components of a quantum transition. First is the choice of crypto algorithms themselves. Second is the implementations of these algorithms in core technologies such as FPGAs. The third component is the evolution of supporting infrastructure to support PQC. For best results, these three components get combined in a broader crypto agility strategy. This sets the stage for a smoother transition, within a time period commensurate with an organization's risk tolerance.

In this paper we help you navigate a transition to the post-quantum world. We begin by providing an overview of post-quantum cryptography, including the various standardization efforts. Next, we introduce the different implementations that currently exist for incorporating PQC algorithms into your infrastructure. Then we discuss the state of complementary solutions -- specifically certificates, protocols, and cloud computing. We also describe crypto agility frameworks that can be used to develop a transition strategy and identify potential gaps. Finally, we close with a roadmap to help you move forward efficiently according to your organizational needs.

2. Post-Quantum Cryptography

All current public-key crypto-systems assume that certain mathematical problems are hard to solve, i.e. the time to solve them on classical computers increases exponentially in proportion to the size of the input. For example, RSA assumes that factoring the product of two very large prime numbers is difficult to do with current computing technology. Thus, the security of RSA is predicated on this assumption staying true. A large enough quantum computer may render these assumptions false [3]. For example, Shor's algorithm can factor RSA keys in polynomial time [2]. Consequently, any transition to a quantum safe future requires new classes of algorithms with hardness assumptions that will not be impinged upon by quantum computers [13].

The impact of quantum computing will be different based on the type of algorithms. According to NIST [41], larger key sizes for symmetric key and larger output for hash functions may be needed to ensure security in the post-quantum world, while public key cryptography such as RSA (Rivest Shamir Adelman), ECDSA (Elliptic Curve Digital Signal Algorithm), ECDH (Elliptic Curve Diffie-Hellman) and DSA will no longer be secure. Thus, the development of new classes of crypto algorithms focuses on public key cryptography used for key exchange and digital signature schemes. In this section we provide an overview of these efforts.

2.1. Quantum-Safe Algorithms

Quantum-resistant cryptography is primarily based on one of six different mathematical problems. Each problem has distinct hardness assumptions as well as pros and cons in terms of performance. These are listed below:

1) **Lattice-based cryptography** is based on the hardness of well-studied lattice problems in the construction itself or in the security proof. Two popular sub-categories of it are NTRU signature and ring-LWE (Learning With Errors.) These algorithms are simple, efficient, and parallelizable. However, they have larger public key sizes than RSA. Additionally, it is difficult to give precise estimates of the security using known cryptanalysis techniques [14].
2) **Code-based cryptography** relies on error-correcting codes. Examples include McEliece encryption algorithm and CFS (Courtois-Finiasz-Sendrier) signatures. They have large key sizes and attempts to reduce them so far all resulted in compromised security. There has been more
3) with implementing it for encryption than for signatures [4].
4) **Hash-based signatures** are digital signatures constructed using hash functions such as Merkle signature scheme and XMSS (Extended Merkle Signature Schemes.) The security of hash functions is well studied. However, corresponding schemes can only produce a limited number of signatures, and many require a secure record of the exact number of previously signed messages. Together with the much larger signature, the drawbacks make it tricky to implement for large-scale environments [43].
5) **Multivariate cryptography** is based on the difficulty of solving systems of multivariate polynomial equations over a single finite field. The multivariate encryption schemes are not very efficient, due to large public keys and long decryption times. However, they are more successful for building signature schemes because they provide some of the shortest signatures among the post-quantum algorithms [4].
6) **Isogeny** uses mathematics of super-singular elliptic curves and super-singular isogeny graphs to create a Diffie-Hellman-like key exchange. This mathematical problem is the most recent basis for any post-quantum candidates and is therefore less studied. However, it has one of the smallest key sizes [48].

7) **Zero knowledge proof (ZKP)** proves validity without revealing underlying information. It is currently only used by one PQC, Picnic, where it was made non-interactive and turned into a signature scheme using the traditional Fiat-Shamir transform [15].

The mathematical structure for different PQC algorithms varies widely. A detailed discussion is beyond the scope of this paper and is available elsewhere [16].

2.2. NIST PQC Standardization

In 2017, NIST started a post-quantum cryptography standardization effort to select algorithms that will supplement or replace existing public key cryptography. There were 82 submissions received in the 1st round, and 69 accepted, with a focus on the security analysis. Round 2 was started in 2019 with 26 candidate algorithms, and a focus on the hardware and software performance as well as security. To keep the diversity but reduce the numbers, NIST encouraged mergers of similar submissions. In July 2020, NIST announced the candidates for the third round, which included 7 primary and 8 alternate candidates. At the time of this writing (summer 2021), the finalists are still being reviewed for standardization at what is the conclusion of the third round. Algorithms with structured lattice schemes appear to be the most promising general-purpose algorithms for public key encryption and digital signature schemes. Several of the alternate candidates have worse performance than the finalists but might be selected for standardization if there's high confidence in their security. Others have acceptable performance but require additional analysis to inspire sufficient confidence in their security [6]. NIST will select which alternates to keep studying in a 4th round and expect the finalized standard to be ready around 2024 [40].

NIST's standardization effort focuses on two categories of PQC algorithms: 1) public key encryption/key establishment and 2) digital signatures. The current candidates all offer a range of security based on the parameter set, that range from two to eighteen. The key sizes and ciphertext sizes differ based on the parameter set selected, as shown in Table I and Table II. The * denotes an alternate candidate.

Table 1. NIST Finalists: Public-key Encryption

Name	Type	Public Key (bytes)	Private Key (bytes)	Ciphertext Size (bytes)
Classic McEliece [9]	Code-based	261120 - 1357824	6492 - 14120	128 - 240
Crystals-Kyber [10]	Lattice	800 - 1568	1632 - 3168	768 - 1568
NTRU [11]	Lattice	699 - 1230	935 - 1590	699 - 1230
Saber [11]	Lattice	672 - 1312	1568 - 3040	736 - 1472
*BIKE [11]	Code-based	2542 - 6206	3110 - 13236	2542 - 6206
*FrodoKEM [11]	Lattice	9616 - 21520	19888 - 43088	9729 - 21632
*HQC [11]	Code-based	2249 - 7245	2289 - 7285	4481 - 14469
*NTRU Prime [11]	Lattice	897 - 1322	1125 - 1999	1025 - 1184
*SIKE [11]	Isogeny	197 - 564	28 - 644	197 - 596

Table 2. NIST Finalists: Digital Signature Algorithms

Name	Type	Public Key	Private Key	Signature
Crystals-Dilithium [11]	Lattice	1312 - 2592	2544 - 4880	2420 - 4595
Falcon [11]	Lattice	897 - 1793	1281 - 2305	690 - 1330
Rainbow [11]	Multivariate	60192 - 1930600	64 - 1408736	66 - 212
*GeMSS [12]	Multivariate	352000 - 10400000	13100 - 12300	240000 - 600000
*Picnic [11]	ZKP	33 - 65	49 - 97	14612 - 209510
*SPHINCS+ [11]	Hash based	32-64	64-128	8080 - 49216

The appropriate PQC algorithm may depend both on the security and the asset constraints. Consider the public key encryption candidates. Lattice-based algorithms such as NTRU and Saber have a much smaller public/private key size than the code-based Classic McEliece. However, Classic McEliece has a smaller ciphertext size than the lattice-based algorithms. Similarly for digital signature algorithms multivariate-based Rainbow has a much larger key size, but much smaller signature size, than the lattice-based Falcon and Crystals-Dilithium.

PQC generally has larger key sizes, but some algorithms at lower security levels have a comparable size to classical algorithms at a higher security level. The five security levels are denoted as:

- Level 1: At least as hard to break as AES-128 using exhaustive key search.
- Level 2: At least as hard to break as SHA-256 using collision search.
- Level 3: At least as hard to break as AES-192 using exhaustive key search
- Level 4: At least as hard to break as SHA-384 using collision search.
- Level 5: At least as hard to break as AES-256 using exhaustive key search.

The focus of NIST competition is on security level 1, 2, and 3. Take the key exchange algorithms, for example, as shown in Figure 1. SIKE, SABER, Crystal-Kyber and NTRU all have comparable key and ciphertext sizes with RSA/DH. For the digital signature candidates, shown in Figure 2, Falcon and Crystals-Dilithium have similar signature and public key sizes as RSA.

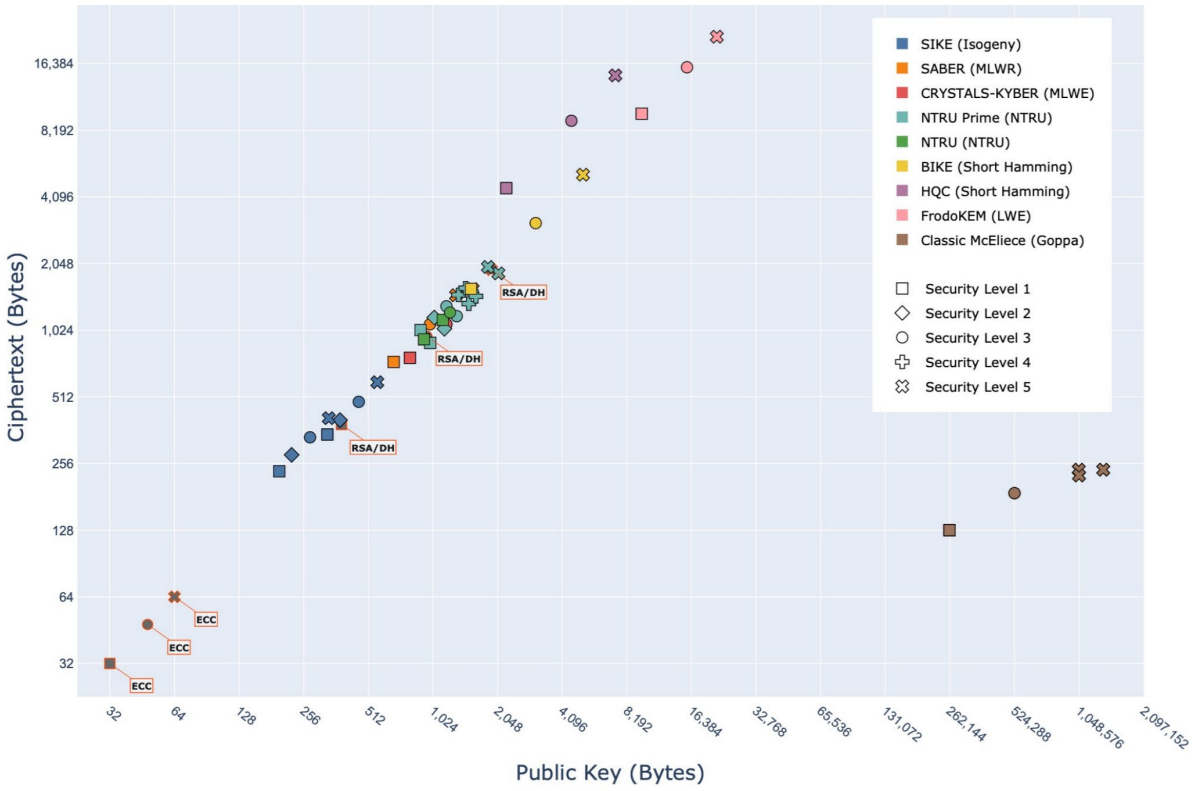


Figure 1. Ciphertext and public key sizes for NIST 3rd round key exchange + classical PKE [20]

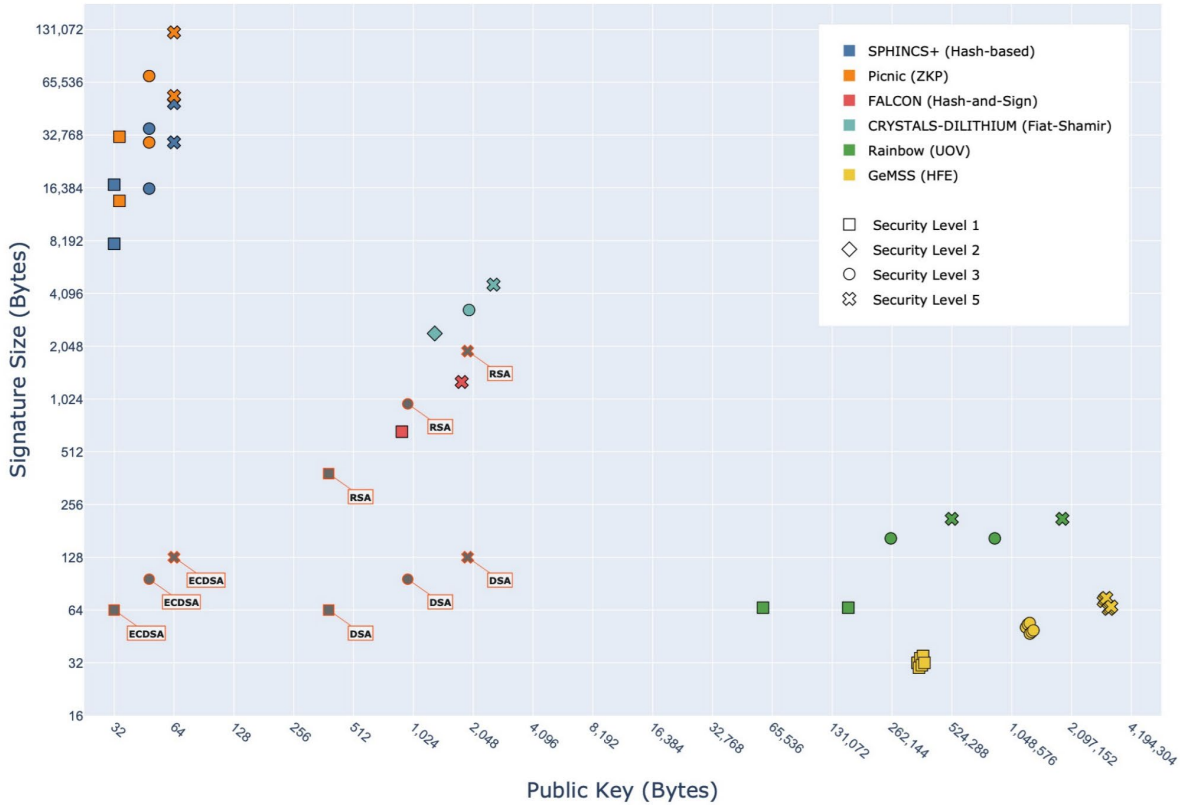


Figure 2. Signature and public key sizes for NIST 3rd round and classical digital signature [20]

2.3. CACR Competition

Aside from NIST, China also held their own post-quantum cryptography competition. The Chinese Association for Cryptologic Research (CACR) issued a notice of algorithm competition in August of 2018. The competition focused on functionalities, security, and performance. While NIST separated the competition into two tracks (digital signature and public key crypto/key encapsulation), China separated the competition into three tracks: digital signature, public key cryptography and key exchange. The first round received 36 submissions, with the majority being lattice-based algorithms. Unlike NIST, which included multiple rounds, the CACR competition only had one round and concluded in December 2019. The results of round 1 contained 14 finalists with 1st, 2nd, and 3rd place winners across the three categories [21]. The list of finalists is provided in Table III and the details are from the Announcement of the Results of the National Cryptographic Algorithm Design Competition Algorithm Selection, a technical report published by the CACR in 2020 [22].

Table 3. CACR Competition Finalists

Rank	Name	Category	Type
1st place	Aigis-sig	Signatures	Lattice
1st place	LAC.PKE	KEM	Lattice
1st place	Aigis-enc	KEM	Lattice
2nd place	LAC.KEX	KEX	Lattice
2nd place	SIAKE	KEX	Isogeny
2nd place	SCloud	KEM	Lattice
2nd place	AKCN	KEM	Lattice
3rd place	OKCN (SKCN-MLWE)	KEX	Lattice
3rd place	Fatseal	Signature	Lattice
3rd place	Mulan	Signatures	Lattice
3rd place	AKCN-E8	KEM	Lattice
3rd place	TALE	KEM	Lattice
3rd place	PKP-DSS	Signature	PKP
3rd place	Piglet-1	KEM	Code-based

2.4. Europe

While Europe did not hold its own post-quantum competition, there have been multiple initiatives exploring the current solutions. The European Technology and Standards Institute (ETSI) is a recognized regional standards body dealing with telecommunications, broadcasting, and other networks and services. It recognized that cryptanalysis and the standardization of algorithms require significant time and effort for their security to be trusted by governments and industry. Thus, ETSI is taking a proactive approach and formed the Cyber Quantum Safe Cryptography (QSC) Working Group to assess and make recommendations for quantum-safe cryptographic primitives, protocols, and implementation considerations. The focus is on practical implementation with consideration of performance, capabilities, benchmarking, architecture, and protocols, instead of development of the cryptographic primitives [14]. The European Union Agency for Cybersecurity (ENISA) also published a study that provides an overview of NIST finalists as well as proposals that system owners can implement now in order to protect the confidentiality of their data against a quantum-capable attacker [4].

3. PQC Implementations

The performance of PQC algorithms will depend on their implementation and deployment environments. Most of the published benchmarks are based on algorithms from round 2 of NIST's competition. They include benchmarking of seven lattice-based algorithms [17]; benchmarking of algorithms in hardware [18]; benchmarking using FPGA [20] and benchmarking in TLS [19]. Because these algorithms are currently evolving, their actual performance may differ from published benchmarks, which is also informed by implementation platforms. The most comprehensive benchmarking of the round 3 algorithms, as of this publication, is from the Open Quantum Source (OQS) profiling project and includes runtime, memory use and performance on x86 [44]. The code for OQS profiling is open source and geared toward collecting information across the algorithms at different levels of the software and network stack. It does not provide testing at the raw algorithm level, which can be done using SUPERCOP, a

toolkit that measures crypto primitives according to length of the key/message and time to generate, encrypt or authenticate [45]. Because each PQC algorithm has its own limitations and might be appropriate for distinct platform or assets, it's best to test and explore the implementation challenges on your target device.

3.1. Libraries

There are several libraries that implement PQC for distinct systems and corresponding requirements. One of the earliest is libpqcrypto, a cryptographic software library produced by the PQCRYPTO project that includes software for 77 cryptographic systems. It includes AES-256 and Salsa20 for symmetric/secret-key cryptography, McBits for public-key cryptography and SPHINCS+ for signatures. PQCRYPTO can be used with OpenSSH and OpenIKED. The project includes a benchmarking and testing framework as well as libraries for ARM Cortex-M4 and FPGAs. However, the library is research-oriented and not ready for use in production environments [23].

Another library that is more widely used is liboqs from Open Quantum Safe (OQS) [7]. It is an open-source library written in C and has well developed support for many platform and languages. The library can run on Linux, Mac and Windows. It supports x86 and ARM architectures, as well as compilers from Clang, GCC (GNU Compiler Collection) and Microsoft. It also contains language wrappers for C++, Go, Java, .Net, Python and Rust. OQS provides code for integration into TLS, SSH, x.509, CMS and S/MIME via OpenSSL and OpenSSH. The library has been used by many external projects, including Microsoft Post-Quantum Cryptography VPN, Mullvad VPN, Thales eSecurity Go wrapper, Liesware Coherence Cryptographic Server, IBM Cloud and others [39]. However, it is built as a library for testing and not commercial deployment, so some components are more mature than others. Some algorithms implemented have large stack usage and may cause failures when run on threads or in constrained environments [11].

In addition to general libraries, Cloudflare released the source code of CIRCL, a cryptographic library written in Go, in 2019. It contains a package that combines an implementation of Diffie-Hellman with SIKE (Super-singular Isogeny Key Encapsulation), allowing developers to experiment with post-quantum key exchange schemes for TLS 1.3. They are currently looking to add lattice-based algorithms such as NTRU and Crystals-Kyber, and post-quantum signature algorithms to CIRCL [24].

Aside from open-source libraries, there are commercial alternatives. ISARA's Radiate Quantum Safe Toolkit supports system Android, iOS, Linux, macOS, Windows 10 and FreeBSD Crypto library and integration tools. It also includes a hybrid mode that has PQC but is backward compatible and maintains the current security measures [25]. Then there's also PQshield, which helps customer transition to quantum-secure standards with a post-quantum cryptography library, hardware for embedded devices, an SDK for mobile and server, as well as a solution for messaging platforms and apps [26].

Other efforts in post-quantum cryptography are happening within the United Arab Emirates. The Technology Innovation Institute (TII), part of Abu Dhabi's Advanced Technology Research Council (ATRC), made available its first PQC software library for the nation in early 2021. Not much is known about the algorithms used, but the library is written in C and supports a wide variety of architectures and OS. A hardware (FPGA-based) implementation has also been developed. The first release of the library has already been integrated in several secure communication products [27].

3.2. Hybrid

There are currently two options to implement post-quantum cryptography, either by replacing current public key algorithms such as RSA and ECDHE, or as part of hybrid cryptosystems. Post-quantum cryptography

has not yet been tested with a real quantum computer, so there is a risk of security or implementation flaws. Thus, many current solutions are exploring the option of hybrid cryptography.

Hybrid cryptosystems combine two or more different cryptographic techniques to perform the same function. There are three different types of hybrids: classical/quantum-safe hybrids, which secure against classical attacks at least; quantum-safe/quantum-safe hybrids, which are good for future attacks, but the security of quantum-safe algorithms remain uncertain; and classical/QKD hybrids, which combine the security of classical algorithms with the only one known quantum-safe method. Hybrid cryptosystems will in theory remain secure if at least one of the underlying cryptographic schemes remains unbroken. However, they can be slower, have a larger footprint for key storage, and be less efficient [8].

4. PQC Infrastructure

The first step to a PQC transition is to identify which algorithm is appropriate for your asset and explore the implementations of these algorithms in core technology. However, because these algorithms are drastically different from classical algorithms, from mathematical foundation to key sizes to performance overhead, there may be changes needed in the infrastructure to support both them and the transition to them. The infrastructures most affected are those that use public-key cryptography, such as certificates and PKI, protocols and cloud solutions. In this section we'll explore some of the changes needed in the infrastructures when transitioning to PQC.

4.1. Certificates and PKIs

The most common certificate sizes on the internet today vary between 500-1500 bytes. Most of the proposed post-quantum schemes have public key and signature sizes of 10-200 kilobytes (kb), which is significantly bigger and can pose challenges for the infrastructures that would use them in X.509 certificates. These challenges include transmission overhead, IP fragmentation and wasted bandwidth for connections. To support new proposed post-quantum signature schemes in X.509, new algorithm identifiers that correspond to certain post-quantum signature scheme parameters and structures will need to be defined [28]. While the x.509 data format allows for long public keys and signatures, some applications may put size limits on the x.509 fields. Also, the cost and performance overhead will differ depending on the implementation and usage of the certificates. Some devices or system may or may not be fully upgradeable due to software or hardware limitations.

The transition to PQC can be a huge undertaking that takes a long time. To ensure a smooth transition, there will be a need for a certificate that can work with both PQC-enabled systems and non-upgraded systems, or hybrid certificate. A hybrid certificate is an X.509 certificate with additional quantum safe components, so you only need to support one certificate instead of two no matter the system. The hybrid certificate would contain extra X.509 certificate fields for quantum-safe keys and signatures as well as encoding for a quantum safe algorithm. NIST has updated the guidance on transition in SP800-56C Rev. 2 to permit the use of hybrid mode. In hybrid mode, an unapproved (i.e. PQC) algorithm can be combined with a NIST-approved algorithm and still receive FIPS validation [42].

One way to implement a hybrid certificate is to offer the option of choosing to use the classical or post-quantum algorithm. This allows the relying parties that cannot update their cryptographic suites to be in the same infrastructure as other relying parties that use a stronger validation algorithm. Another way is to implement the certificate such that it encrypts/decrypts using both classical and post-quantum algorithms. Instead of choosing one or the other, the server/client now needs to use both. This way the certificate is

protected against classical and quantum attacks. However, this does present an implementation challenge, because you need to upgrade the PKI system, and maybe the servers and the clients as well. The signing and validation might also need to be upgraded [29]. There are currently multiple collaborative efforts on new digital certificate formats that can work with both classic and post-quantum algorithms from ISARA, Cisco, CableLabs, DigiCert and Entrust. For more details, refer to [46] for a method of embedding alternative sets of cryptographic materials into digital certificates as well as how creation, verification, signing and revocation would work in such cases.

4.2. Protocols

Once successfully adopted, security protocols tend to be long lived in products and networks. Thus, protocols typically allow for some elements of flexibility in changes to the key sizes and cryptographic parameters in case of algorithm degradation. However, protection against quantum attacks may require more drastic changes, where the cryptographic primitives may need to be replaced entirely or protocol-level changes may be needed. This can be an easy or difficult process depending on how crypto-agile the protocols are. An overview is provided here, and more details can be found in the white paper from ETSI [14].

Internet Key Exchange (IKEv2) is a protocol used mainly for setting up VPNs, using three exchanges to set up a security association. First, a common key is derived using the Diffie-Hellman key agreement algorithm. Second, the key is authenticated using certified digital signatures or pre-shared authentication key. Thirdly, the key agreement is conducted again to generate new ephemeral keys for the IP packet. The protocol standard is rigid and only offers a small set of cryptographic algorithms. Making IKE quantum-safe will require replacing the algorithms used in all three exchanges.

The Transport Layer Security (TLS) protocol, previously SSL, establishes a protected tunnel between a client and server for transmission of application data. It starts with a handshake sub-protocol that authenticates server and client, then establishes shared secret keys for transmission of application data. The shared secret keys are then used in the record layer subprotocol to encrypt and authenticate application data. The handshake uses public key cryptography and will have to be replaced with quantum-safe alternatives. The subsequent record sub-protocol uses symmetric key cryptography and just needs to increase the key sizes. The design of TLS is largely independent of cryptographic algorithms and allows the parties to negotiate the cipher suites to be used. While quantum-safe algorithms with large public keys or signatures may require additional changes to the standard, there are currently libraries available to help test the post-quantum algorithm implementations and identify implementation challenges.

Secure/Multipurpose Internet Mail Extension (S/MIME) is used to securely send email messages. It allows email to remain encrypted during the entire path from sender to recipient, preserving end-to-end confidentiality and data integrity. Content encryption in S/MIME relies upon symmetric ciphers and is believed to be quantum-safe. However, the digital signatures for authentication and integrity use DSA or RSA, which will need to be replaced with quantum-safe alternatives. S/MIME does support extended key size and encryption methods, so it is possible to upgrade signature and key-establishment algorithms without replacing the entire protocol.

Secure Shell (SSH) is used to encrypt information sent over an insecure network and allows remote login, file transfer or operations without compromising data integrity or confidentiality. The SSH protocol involves three major sub-protocols: 1) The transport layer protocol that creates a secure channel and runs over top of TCP/IP; 2) The user authentication protocol that authenticates the client to the server; and 3) The connection protocol that takes the encrypted tunnel generated by transport layer and multiplexes it into

several channels for login, proxy forwarding and accessing secure subsystems etc. The SSH protocol includes a high level of cryptographic agility and allows servers and clients to negotiate the algorithms, so the addition of quantum-safe controls should not require significant changes to the base SSH protocol.

4.3. Cloud Solutions

At its most basic level, quantum-safe solutions involve using quantum-safe cryptography, supported by certificates and protocols that accept the quantum-safe option. At a higher level, quantum-safe cloud computing means quantum-safe server, endpoint, and network infrastructure. The Cloud Security Alliance has published a note on cryptanalytic and mathematical research that builds meaningful confidence in the algorithms' security [30]. It's not an analysis on implementation, performance or application to protocol. However, many companies have already taken steps to explore performances and integrate some post-quantum cryptography into their offerings.

Google took a first step towards post-quantum cryptography by researching and prototyping lattice-based public-key cryptography. In 2016, Google launched an experiment to incorporate the lattice-based algorithm into its Chrome browser in developer mode. It is implemented for OpenSSL and designed to provide post-quantum security for TLS [31]. They also explored the performance of Apache HTTP server using post-quantum key exchange algorithms BCNS, NewHope, NTRU and Frodo (only NTRU and FrodoKEM remained as finalists in NIST's competition.) By looking at throughput, connection time and handshake size, they have concluded that the additional overhead in serving typical webpages (between 10KB and 100KB) with a post-quantum cipher suite will only decrease server throughput by less than a factor of two [32].

IBM researchers developed lattice cryptography suites which include the NIST finalists crystals-Kyber and Crystals-Dilithium, and is working with open-source community to develop open standards implementations as part of Open Quantum Safe. Currently, Kyber has been integrated as part of IBM Key Protect for IBM Cloud, a full-service encryption solution that leverages cloud-based hardware security modules. The algorithm performance may be affected by network profile, CPU speed and API call rates. The quantum-safe TLS is currently supported through the Key Protect software development kit and available both in hybrid mode and quantum-safe mode. It is currently only available in Linux, but support for additional operating systems is anticipated [33].

AWS started incorporating PQC since round 2 and now supports post-quantum TLS in AWS KMS. It supports ECDHE with BIKE and ECDHE with SIKE [8]. For the two hybrid algorithms tested, ECDHE with BIKE have a larger size than ECDHE with SIKE. However, ECDHE with SIKE is slower than ECDHE with BIKE. Which algorithm to use would depend on the constraints of the asset, and whether memory or computational speed is more of a priority.

Microsoft is working with academia and industry on four candidates for cryptography systems: Rooke, SIKE, Picnic and qTESLA. Each algorithm may be appropriate for different scenarios where different trade-offs regarding performance and key size are preferred. In addition to working with Open Quantum Safe to develop a post-quantum branch of TLS and SSH, Microsoft also worked on a fork of OpenVPN integrated with post-quantum cryptography to enable testing and experimentation [34].

5. Crypto Agility

According to NIST, "continued progress in the development of quantum computing foreshadows a particularly disruptive cryptographic transition." Once quantum computers and exploitation of such attacks becomes practical, protecting stored keys and data will require re-encrypting them with a quantum-resistant

algorithm and deleting or physically securing backups. The integrity and sources of information will become unreliable unless they are processed or encapsulated with quantum-resistant mechanisms. In the best case, 5-15 or more years will elapse after the publication of the standards before a full implementation of those standards is complete. Without proper planning, it may take decades to replace most of the vulnerable public-key systems currently in use. Thus, NIST encourages enterprises to identify where and for what it is employing public-key cryptography and all the use characteristics, as well as developing a playbook for crypto agility [13].

While there are many libraries and solutions available to help with the quantum transition, many information systems are not designed to encourage support of rapid adaptations of new cryptographic primitives and algorithms. Cryptographic algorithms cannot be replaced until all components of a system are prepared to process the replacement. This may require not only the replacement of cryptographic algorithms, but also updates to the protocols, hardware, dependent operating systems and procedures, especially in the case of post-quantum cryptography where the parameter and structures can be very different. In addition to replacing algorithms, other non-security issues, such as adoption rates, backward compatibility and performance must also be considered. The Crypto Agility Risk Assessment Framework (CARAF) can be used to combine all the factors into a broader crypto agility strategy that allows for a smoother transition within a time period commensurate with an organization's risk tolerance [36]. The risk framework consists of five phases:

1. **Identify threat** – the threat in this case is security risk attributable to quantum computing, and more specifically the challenges and risk of migrating to PQC. The assets that will be phased out before quantum computers become practical, or NIST publishes a definitive standard for PQC, can be eliminated from the risk assessment. Meanwhile, NIST has published transition guidance on the recommended algorithms and key lengths [35].
2. **Inventory of assets** - will give an overview of how crypto-agile assets are and help in identifying which assets should be prioritized in any migration. In this case, we need to identify where, how and for what public-key cryptography is employed, as well as use characteristics. NIST has published a draft on how to identify and what information should be recorded for post-quantum migration based on 5 scenarios: FIPS-140 validated hardware and software modules, cryptographic libraries, cryptographic applications, embedded code in computing platforms and communication protocols [47].
3. **Estimate risk** – given the lack of information on attack vectors, the risk estimation for crypto agility is based on a combination of timeline, to realize the threat, and cost to mitigate.
4. **Secure asset** - based on an evaluation of the risk and the resources available, the organization may prioritize the assets and decide to accept the risk, mitigate it, or phase out the asset. While it is important to assess how crypto-agile your assets are when assessing the risk, it's also important to keep crypto agility in mind when choosing a new solution to implement, especially since PQC has not been tested against quantum computers and may still undergo changes in the future. Some of the properties to keep in mind for the new solutions are extensibility, removability, interoperability, compatibility, flexibility, and updatability [37].
5. **Create roadmap** - the solutions will likely go through a few iterations and will affect different assets differently. In addition to performing benchmark testing to evaluate the performance and impact on your assets, it's important to make sure that all teams and vendors are on the same page and take future changes into account in policies and guidelines. What the roadmap should consists of depends on the mitigation methods and the organization. Security standards, best practice documentations, installation and administration documentations may all need to be changed or replaced.

The transition to PQC and making assets crypto-agile requires cooperation and collaboration between different teams and vendors. It can be overwhelming at first, but the more elements of crypto agility are implemented, the easier the next step will be. By preparing now for the upcoming transition, we can ensure a more orderly, less costly, and minimally disruptive changeover [38]. Because many of the PQC solutions have not been rigorously tested yet, they may have systemic weakness and go through several iterations before they become secure. Crypto agility will provide a practical framework to address updates to crypto threats in a quick and efficient manner.

6. Conclusion

Transitioning to post-quantum algorithms is a big undertaking. Different algorithms have different key lengths, performance, and operational constraints. There is no one size fits all solution. Benchmarking of the algorithm and crypto agility assessment of the target assets will help determine what algorithm is appropriate, as well as the potential overhead. Even if your asset implements post-quantum cryptography, there will be backward compatibility problems if others don't. Thus, it is important to plan and create a roadmap for your PQC transition.

The first step in transitioning to a post-quantum world is assessing which assets are capable of post-quantum transition using a crypto agility risk assessment. The more elements of crypto agility your assets implement, the easier the transition will be. Based on the risk assessment, action plan options range from phasing out the asset before quantum computing becomes available, accepting the risk, or securing it.

The true security of post-quantum cryptography won't be testable or tested until a practical quantum computer becomes available. For that reason, and in the meantime, a more proactive approach is to focus on the implementation and performance of hybrid cryptography.

Picking the appropriate quantum algorithms is a decision that is tied to security requirements and asset/system constraints. As with all new technological environments, testing how post-quantum algorithms work with your assets will be informative; different implementations may provide distinct benefits. Also important: Transitioning to support a quantum-based crypto environment, and the identifying which phases happen when, based on existing or desired risk tolerance levels.

Abbreviations

AES	Advanced Encryption Standard
API	Application Program Interface
ATRC	Abu Dhabi's Advanced Technology Research Council
AWS	Amazon Web Service
CACR	Chinese Association for Cryptologic Research
CARAF	Crypto Agility Risk Assessment Framework
CPU	Central Processing Unit
DSA	Digital Signature Algorithm
ECDH	Elliptic-Curve Diffie–Hellman
ECDSA	Elliptic Curve Digital Signature Algorithm
ENISA	European Union Agency for Cybersecurity
ETSI	European Technology and Standards Institute

HTTP	Hypertext Transfer Protocol
IBM	International Business Machines
IKE	Internet Key Exchange
IP	Internet Protocol
KMS	Key Management Service
NIST	National Institute of Standards and Technology
OQS	Open Quantum Safe
PKI	Public Key Infrastructure
PQC	Post-Quantum Cryptography
QKD	Quantum Key Distribution
QSC	Cyber Quantum Safe Cryptography
RSA	Rivest–Shamir–Adleman
SSH	Secure Shell
S/MIME	Secure/Multipurpose Internet Mail Extension
TCP	Transmission Control Protocol
TII	Technology Innovation Institute
TLS	Transport Layer Security
VPN	Virtual Private Network
XMSS	eXtended Merkle Signature Scheme
ZKP	Zero Knowledge Proof

Bibliography & References

[1] Post Quantum Cryptography Team. A Quantum World and How NIST is Preparing for Future Crypto, 2014.
[2] Peter Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. SIAM J. Comput., 26(5):1484– 1509, October 1997.
[3] Tyson Macaulay. Comments in Response to NIST Cyber Security Framework Draft 1.1 and NIST Roadmap for Improving Critical Infrastructure Cybersecurity Draft 1.1. Technical report, National Institute of Standards & Technology, 2018.
[4] Ward Beullens, Jan-Pieter D'Anvers, Andreas Hulsing, Tanja Lange, Lorenz Panny, Cyprien de Saint Guilhem, and Nigel P. Smart. Post-Quantum Cryptography: Current State and Quantum Mitigation. Technical report, ENISA.
[5] Chris Hickman. CSO Insights: 4 Reasons Why You Can't Ignore Crypto Agility, 2019.
[6] PQC Standardization Process: Third Round Candidate Announcement. Technical report, National Institute of Standards & Technology, 2020.
[7] "liboqs," Open Quantum Safe. 2021. https://openquantumsafe.org/liboqs/.
[8] Andrew Hopkins. Post-Quantum TLS Now Supported in AWS KMS. https://aws.amazon.com/blogs/security/post-quantum-tls-now-supported-in-aws-kms/.
[9] Classic McEliece: Conservative Code-Based Cryptography. Technical report, 2020. https://classic.mceliece.org/nist/mceliece-20201010.pdf.
[10] "Kyber," CRYSTALS: Cryptographic Suite for Algebraic Lattices. 2020. https://pq-crystals.org/kyber/
[11] "Algorithms in liboqs," Open Quantum Safe. 2021. https://openquantumsafe.org/liboqs/algorithms/.

[12] GeMSS: A Great Multivariate Short Signature. 2020. https://www-polsys.lip6.fr/Links/NIST/GeMSS_specification.pdf
[13] William Barker, William Polk, and Murugiah Souppaya. Getting Ready for Post-Quantum Cryptography: Exploring Challenges Associated with Adopting and Using Post-Quantum Cryptographic Algorithms, 2021.
[14] Quantum Safe Cryptography and Security: An Introduction, Benefits, Enablers and Challenges. Technical report, ETSI.
[15] Daniel Kales and Greg Zaverucha. Improving the Performance of the Picnic Signature Scheme. 2020.
[16] Olaf Grote, Andreas Ahrens, and Cesar Benavente-Peces. Paradigm of Post-Quantum Cryptography and Crypto-agility: Strategy Approach of Quantum-safe Techniques. 2019.
[17] Farnoud Farahmand, Viet Ba Dang, Michal Andrzejczak, and Kris Gaj. Implementing and Benchmarking Seven Round 2 Lattice-Based Key Encapsulation Mechanisms Using a Software/Hardware Codesign Approach. In NIST Second PQC Standardization Conference. National Institute of Standards & Technology, 2019.
[18] Kris Gaj. Challenges and Rewards of Implementing and Benchmarking Post-Quantum Cryptography in Hardware. In Proceedings of the 2018 on Great Lakes Symposium on VLSI, pages 359–364, 2018.
[19] Christian Paquin, Douglas Stebila, and Goutam Tamvada. Benchmarking Post-Quantum Cryptography in TLS. 2019.
[20] Kris Gaj. Implementation and Benchmarking of Round 2 Candidates in the NIST Post-Quantum Cryptography Standardization Process Using FPGAs. 2020.
[21] The Latest Progress of PQC Competition in China. Technical report, 7th ETSI QSC/IQC Workshop, 2019.
[22] Announcement of the Results of the National Cryptographic Algorithm Design Competition Algorithm Selection. Technical report, CACR, 2020.
[23] PQCrypto Usage and Deployment. IANIX. 2021. https://ianix.com/pqcrypto/pqcrypto-deployment.html
[24] Kris Kwiatkowski and Armando Faz-Hernandez. Introducing CIRCL: An Advanced Cryptographic Library, 2019.
[25] ISARA Radiate. https://www.isara.com/products/isara-radiate.htm
[26] Post-Quantum Cryptography Hardware, Firmware, SDK, Toolkits - PQShield. https://pqshield.com/
[27] Nitin Dahad. UAE Building First Quantum Computing and Cryptography Library. EE Times Asia, 2021.
[28] Panos Kampanakis, Peter Panburana, Ellie Daw, and Daniel Van Geest. The Viability of Post-Quantum X.509 Certificates. Technical report, 2018.
[29] Massimiliano Pala. Docsis pki: A Proposal for a Next-Generation Quantum-Resistant Infrastructure. SCTE ISBE, 2020.
[30] Confidence in Post Quantum Algorithms. Cloud Security Alliance. 2021. https://cloudsecurityalliance.org/artifacts/confidence-in-post-quantum-algorithms/
[31] Jeremy Kirk. Google Tests Post-Quantum Crypto, 2016. https://www.bankinfosecurity.com/google-adds-quantum-computing-armor-to-chrome-a-9253
[32] Joppe Bos, Craig Costello, Leo Ducas, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Ananth Raghunathan, and Douglas Stebila. Frodo: Take Off the Ring! Practical, Quantum-Secure Key Exchange From LWE. ACM Conference on Computer and Communications Security, 2016.
[33] Introduction to Quantum-Safe Cryptography in TLS for IBM Key Protect, 2021. https://www.ibm.com/cloud/blog/introducing-quantum-safe-crypto-tls-for-ibm-key-protect
[34] Post-Quantum Cryptography. Microsoft. 2021. https://www.microsoft.com/en-us/research/project/post-quantum-cryptography/
[35] Elaine Barker and Allen Roginsky. Transitioning the Use of Cryptographic Algorithms and Key Lengths. Technical report, National Institute of Standards & Technology, 2019.

[36] Chujiao Ma, Luis Colon, Joe Dera, Bahman Rashidi, and Vaibhav Garg. CARAF: Crypto Agility Risk Assessment Framework. Journal of Cybersecurity, 7(1), 05 2021. tyab013.
[37] Hassane Aissaoui Mehrez and Othmane EL OMRI. The Crypto Agility Properties. The 12th International Multi-Conference on Society, Cybernetics and Informatics, 2018.
[38] Matt Campagna, Brian LaMacchia, and David Ott. Post Quantum Cryptography: Readiness Challenges and the Approaching Storm. Computing Community Consortium, 2020
[39] "External Users of OQS," Open Quantum Safe. 2021. https://openquantumsafe.org/applications/external.html.
[40] Dustin Moody, "NIST PQC Standardization," 2021. https://www.nccoe.nist.gov/sites/default/files/3-PQC%20NCCoE.pdf .
[41] Lily Chen, Stephen Jordan, Yi-Kai Liu, Dustin Moody, Rene Peralta, Ray Perlner and Daniel Smith-Tone. Report on Post-Quantum Cryptography. Technical report, NIST, 2016.
[42] Elaine Baker, Lily Chen and Richard Davis "Recommendation for Key-Derivation Methods in Key-Establishment Schemes," NIST, 2020.
[43] Andreas Hülsing, Stefan-Lukas Gazdag, Denis Butin and Johannes Buchmann. " Hash-based Signatures: An Outline for a New Standard," NIST, 2015.
[44] "OQS Algorithm Performance Visualizations," Open Quantum Safe. 2021. https://openquantumsafe.org/benchmarking/.
[45] "eBACS: ECRYPT Benchmarking of Cryptographic Systems." https://bench.cr.yp.to/supercop.html.
[46] Alexander Truskovsky, Daniel Van Geest, Scott Fluhrer, Panos Kampanakis, Mike Ounsworth and Serge Mister, "Multiple Public-Key Algorithm X.509 Certificates," Internet Engineering Task Force, 2018. https://datatracker.ietf.org/doc/html/draft-truskovsky-lamps-pq-hybrid-x509-01.
[47] William Barker and Murugiah Souppaya, "Migration to Post-Quantum Cryptography," NIST, 2021. https://www.nccoe.nist.gov/sites/default/files/library/project-descriptions/pqc-migration-project-description-draft.pdf.
[48] Cong Peng, Jianhua Chen, Sherali Zeadally and Debiao He, "Isogeny-Based Cryptography: A Promising Post-Quantum Technique" in IT Professional, vol. 21, no. 06, pp. 27-32, 2019.

Network in A Box with Open Source EPC/HSS and 'Zero Touch' Control

A Technical Paper prepared for SCTE by

Joerg Ahrweiler
Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+15613065021
Joerg.Ahrweiler@charter.com

Hany Heikal
Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+1720-595-4645
Hany.heikal@charter.com

Hossam Hmimy
Senior Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+17204049716
Hossam.hmimy@charter.com

1. Introduction

The Network In A Box (NIB) definition started with a 4G network that is operating the core network and base-station in a single box that is portable and self-organizing, which provides seamless connectivity to a group of mobile users, offering services such as internet connectivity and closed group communication (Push To Talk/Video/Text). The NIB setup expands cellular network coverage in various environments and different use cases, such as terrestrial disaster relief, private networks in-flight, at sea and in other scenarios and environments where an ad-hoc cellular network is required. The NIB concept is a miniature evolution of the traditional cell on the wheel (COW) concept that has been widely used for cellular mobile voice communication.

In this paper, the NIB cellular network architecture in a CBRS (Citizen Broadband Radio System) environment is described, and the background of combining an 'off the shelf' 4G eNB with a software-based open-source core network and a 'Zero Touch' system bring-up and control SW implementation will be presented. Finally, the testbed implementation of the 4G core network and eNB will be installed in a bare-metal environment with MEC (Mobile Edge Compute) demonstration application environment. The paper will conclude with lessons learned, and outlook to migrate to a 5G and container-based NIB.

2. Cellular Network CBRS Architecture

The network architecture in Figure 1 shows the traditional 4G network architecture with a CBRS environment. The RAN (Radio Access Network) and CN (Core Network) architecture in the NIB setup is very much like the traditional LTE network in commercial solutions; main difference is the small-scale nature of the setup. Since the use of the frequency band B48 (3550 MHz – 3700 MHz) is ideal for the NIB approach – shared spectrum/no spectrum needs to be owned – the connectivity of the NIB setup to a SAS (Spectrum Access System) is mandatory.

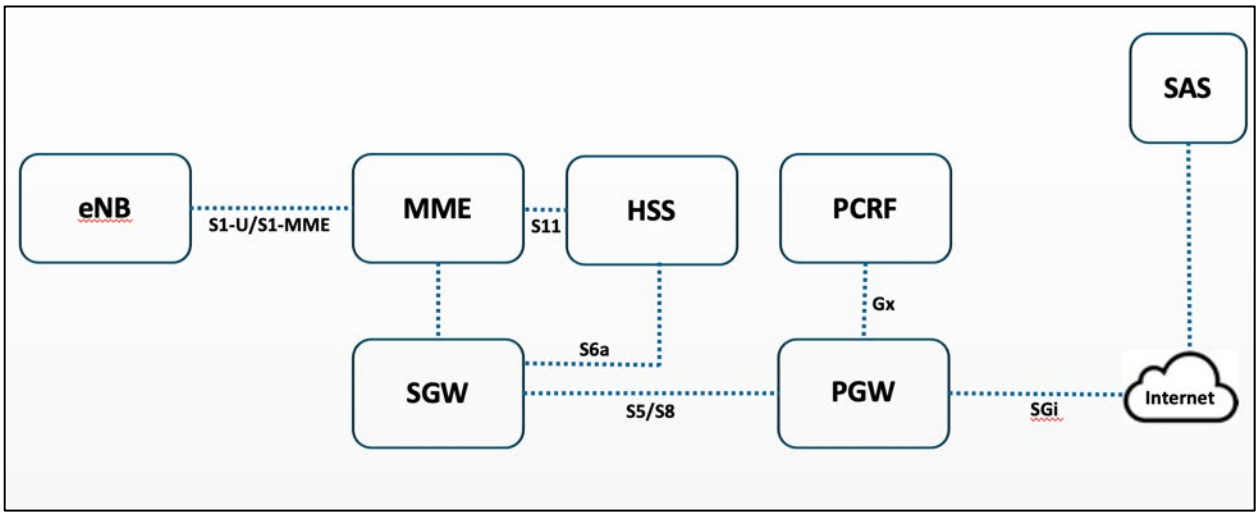


Figure 1 - Network architecture

2.1. Radio Access Network (eNB)

For the eNodeB, a commercial small cell Pico eNB is used. The output power was limited to 14 dBm/25mW per antenna port but can be increased to 24 dBm/250mW if required for respective use case.

Table 1 – LTE eNB Equipment Specifications

Specification	Value
Product	LTE Pico eNB
Band Support	B48 (3.55 – 3.7 GHz)
Carrier Aggregation	Up to 3 CA (only 1 Carrier used in NIB setup)
MIMO	2x2
Frame Configuration	TDD Frame Configuration 2 Special Sub Frame 7
IBW/OBW	150 MHz / 60 MHz
Output Power	2 x 10 W
Antenna	Built-in average 0 dBi
Modulation QAM DL/UL	256 / 64
BF Capability	No
CBRS Classification	CBSD CAT B

2.2. Core Network (EPC/HSS/PCRF)

The Core Network functions are established by utilizing an open-source application suite running on a bare-metal industrial small scale PC with a Linux-based operating system.

There are nowadays many options of open-source EPC/HSS solutions, see below for a list (not conclusive) of most popular open source core network:

- Open5Gs - Formerly NextEPC
- OpenAI Core Network - Related to / branched from OMEC
- Magma - Based on OMEC, with a focus on Fixed Wireless more than mobile
- OMEC – Open Evolved Mobile Core
- OpenMME – MME
- OpenCORD
- srsEPC

3. End To End NIB integration and setup

To be able to easily deploy the NIB, all the components are installed in a Pelican case for easy transportation. Additionally an embedded screen and wireless keyboard/mouse are included in the setup.

Figure 2 - Fully-integrated NIB setup

Figure 2 is showing the industrial type PC, running the EPC/HSS and O&M functionalities, on the left side in the case. The right component in the case is the commercial Pico eNB.

3.1. Physical connectivity

The eNodeB is physically interconnected to the industrial PC via 2 * RJ45 Ethernet cables. One connection is for the S1-MME&S1-MME LTE connectivity, the other one is for O&M administration purposes. The O&M link is utilized to automatically control the bring-up and maintain the system. A third Ethernet interface on the industrial PC is for the Backhaul connectivity to the Internet (user traffic and SAS connectivity). The same function can also be established via an internal WiFi or (Commercial) Cellular module. In the future, NIB versions and other BH connectivity options will be explored, e.g. MANET in unlicensed spectrum.

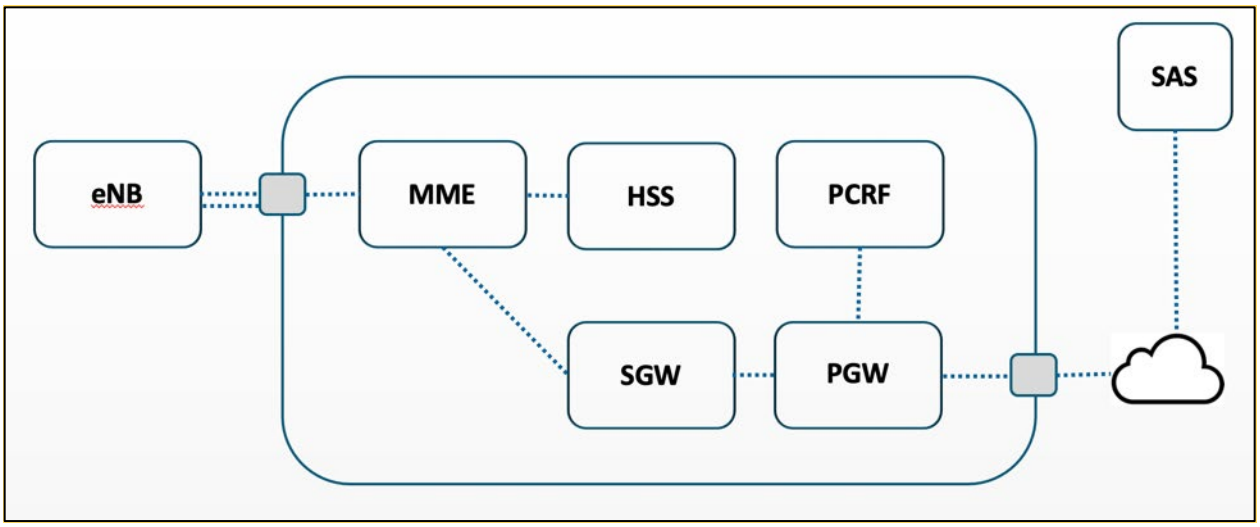


Figure 3 - Logical and physical connectivity

3.2. Automated system bring-up and device monitoring – "Zero Touch"

3.2.1. Introduction

For the purpose of a 'zero touch' and controlled system boot-up, a software application suite was developed to overcome any possible grace conditions and system malfunctions while the NIB setup is coming in service. The application starts automatically after power-up of the NIB and observing the physical and functional status of all components. If a delay or intervention during the boot-up process is necessary, the application will do so. Once the system in completely in service and the Radio is On-Air, the user will be notified and the power-on of the end-user devices can begin.

The requirement was to automate all manual processes to start the private wireless network, known as Network in a Box. Therefore, the start mechanism has to follow the new concept of Zero Touch.

In order to implement a Zero-Touch start mechanism for the Network in a Box, a three-tier software solution has to be developed.

The most common software architecture for typical client-server applications is three-tier architecture, which divides applications into three logical and physical computer tiers, as shown in Figure 4.

The presentation tier, or user interface; the application tier, where data is processed; and the data tier, where the data associated with the application is stored and managed, is a well-established software application architecture that organizes applications into three logical and physical computing tiers.

For many years, the three-tier design was the standard for client-server applications. Most three-tier programs are now candidates for modernization and cloud migration using cloud-native technologies like containers and microservices.

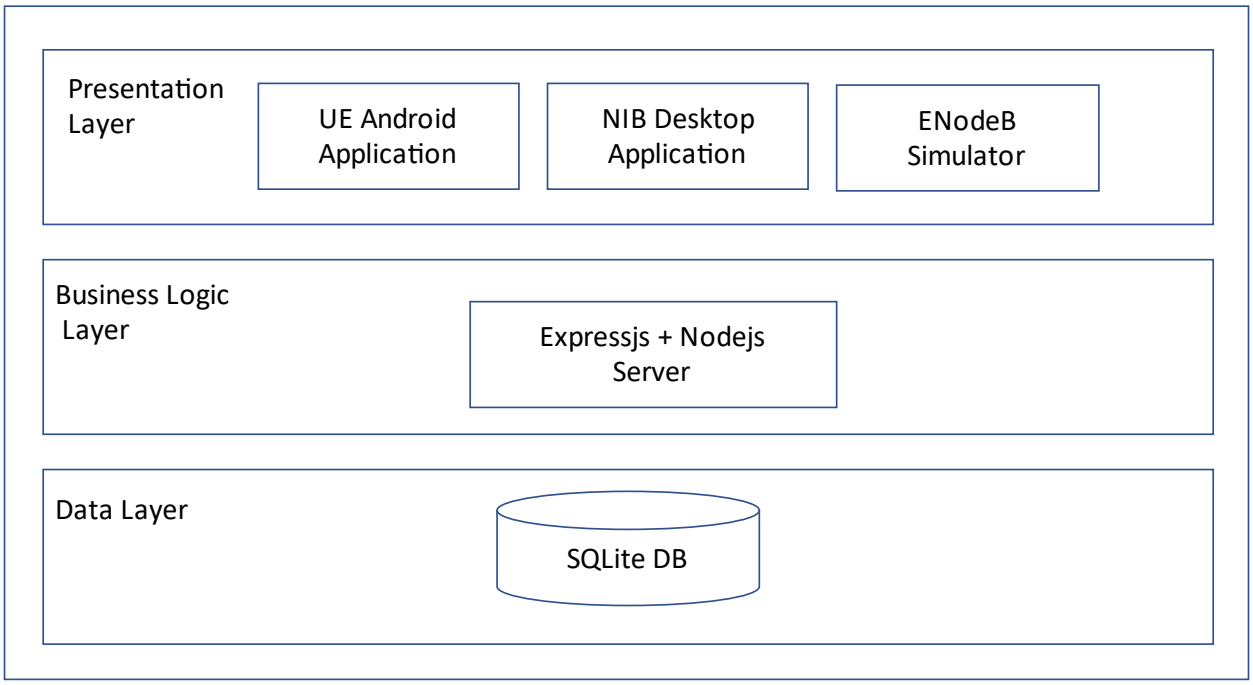


Figure 4 - NIB 'Zero Touch' portal architecture

3.2.2. Presentation Layer

The idea here is to have two applications, the first one is a desktop application, and the second one is an android application.

The desktop application runs on the ubuntu server, and it starts automatically once the server starts. The desktop application automatically starts the private LTE components (MME, SGW, PGW, PCRF, and HSS). It checks the log file for all of these components to ensure that each component starts correctly. It displays a block diagram to represent the elements of the network and it uses the color scheme to represent the status of each element as shown in the screenshots below.

The software used to build this layer are Reactjs and Electronjs and can be located in referenced links 1 and 2.

The desktop application has to communicate with the operating system and it executes a number of the commands in the command-level. In order to do that, the JavaScript child_process package had to be used.

The child_process module creates new child processes of our primary Node.js process. We can execute shell commands with these child processes.

Using external processes can improve the performance of an application if used correctly. For example, if a feature of a Node.js application is CPU intensive, as Node.js is single-threaded, it would block the other tasks from executing while it is running.

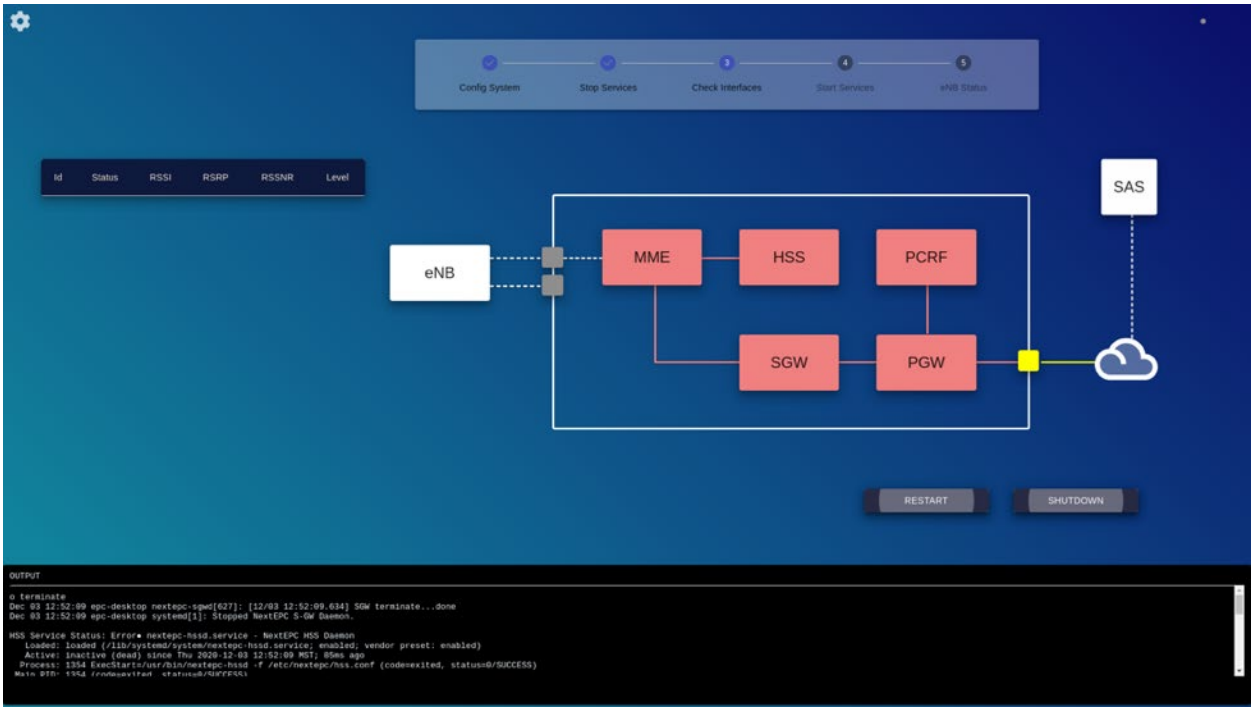


Figure 5 - Initial status after power-up - next step wait for physical interfaces to come in service

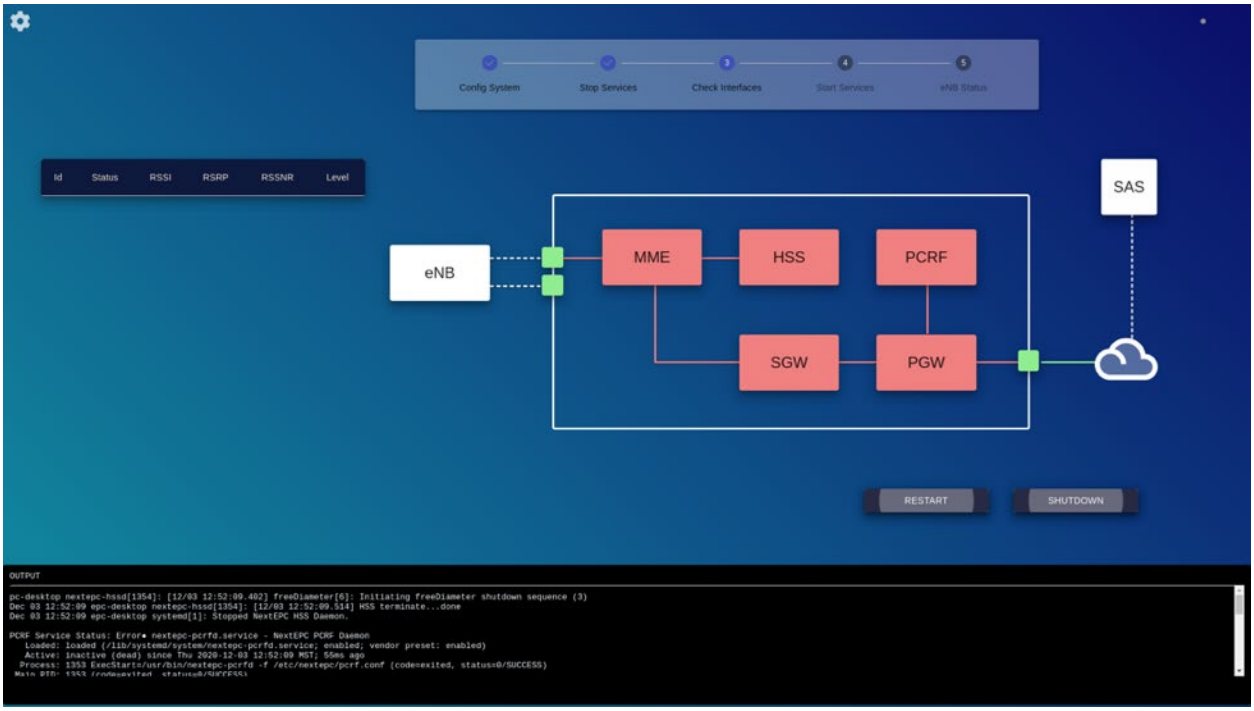


Figure 6 - All Ethernet interfaces are in service – next step start of EPC/HSS services

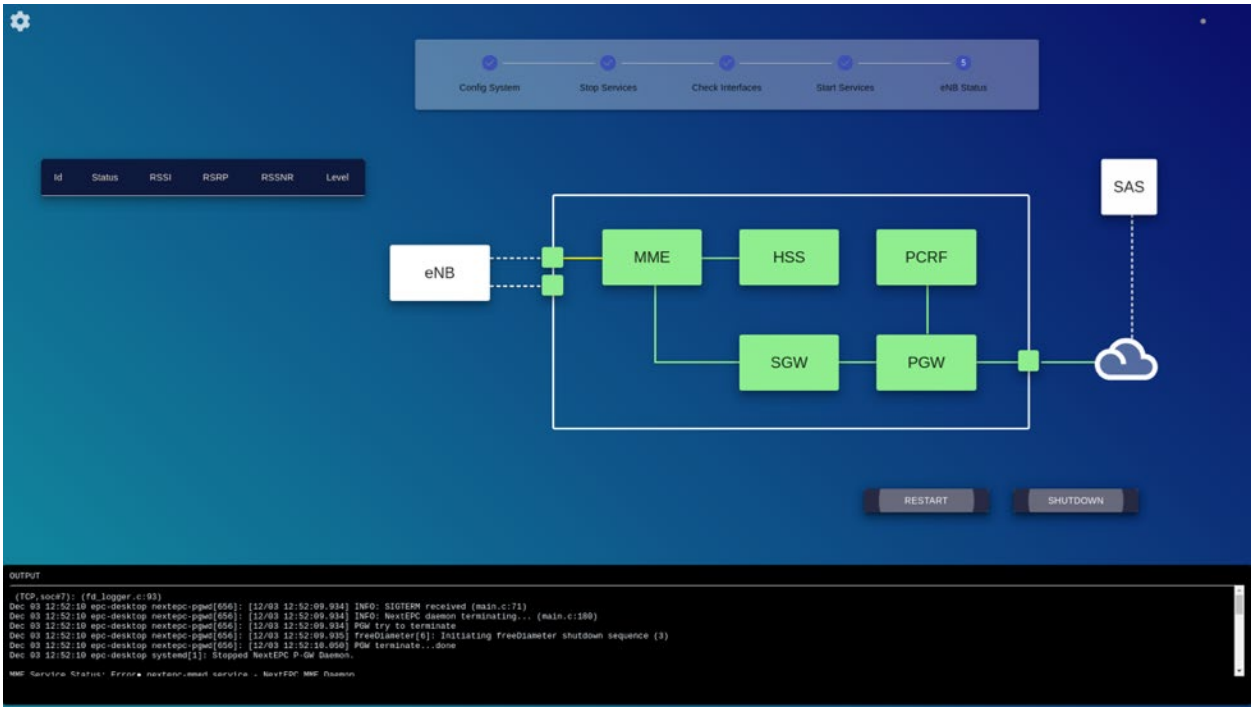


Figure 7 - EPC/HSS services started – next step activation of S1 link

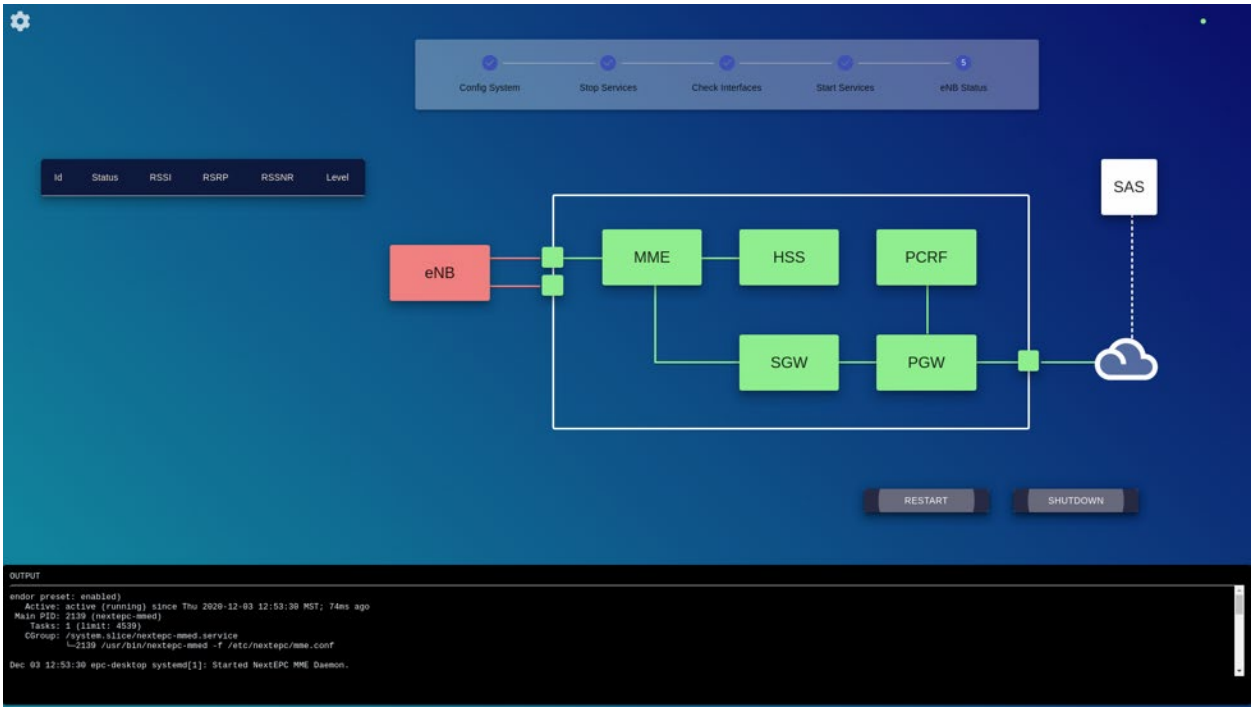


Figure 8 - S1 link is activated – next step wait for positive SAS communication and eNB to come On-Air

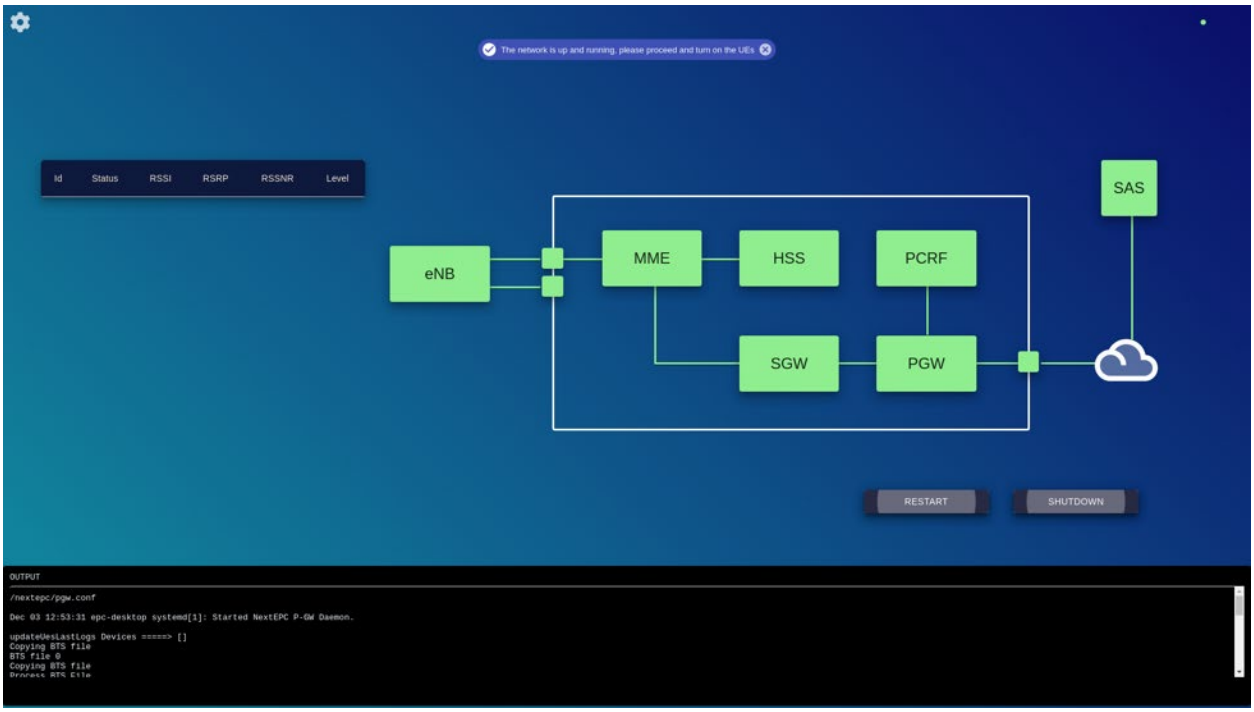


Figure 9 - The system is fully in service and eNB On-Air – user is notified that devices can be powered-on

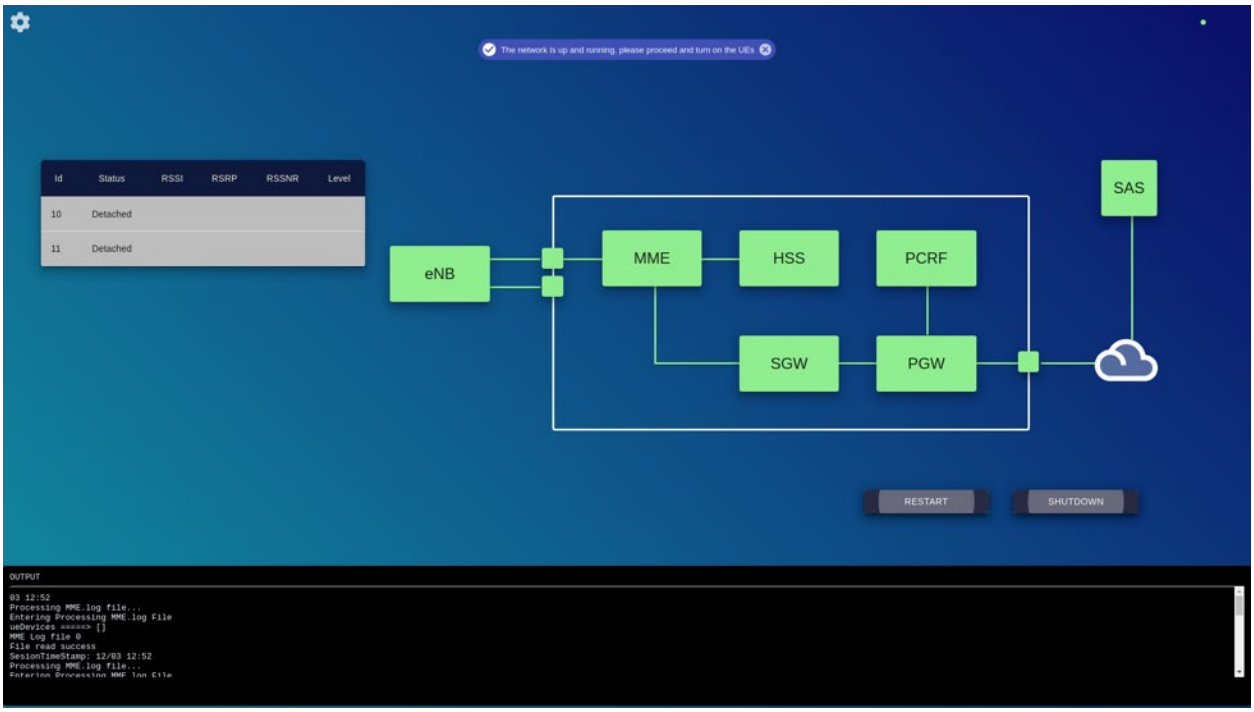


Figure 10 - After devices are powered-on, the device table is activated and device status monitored

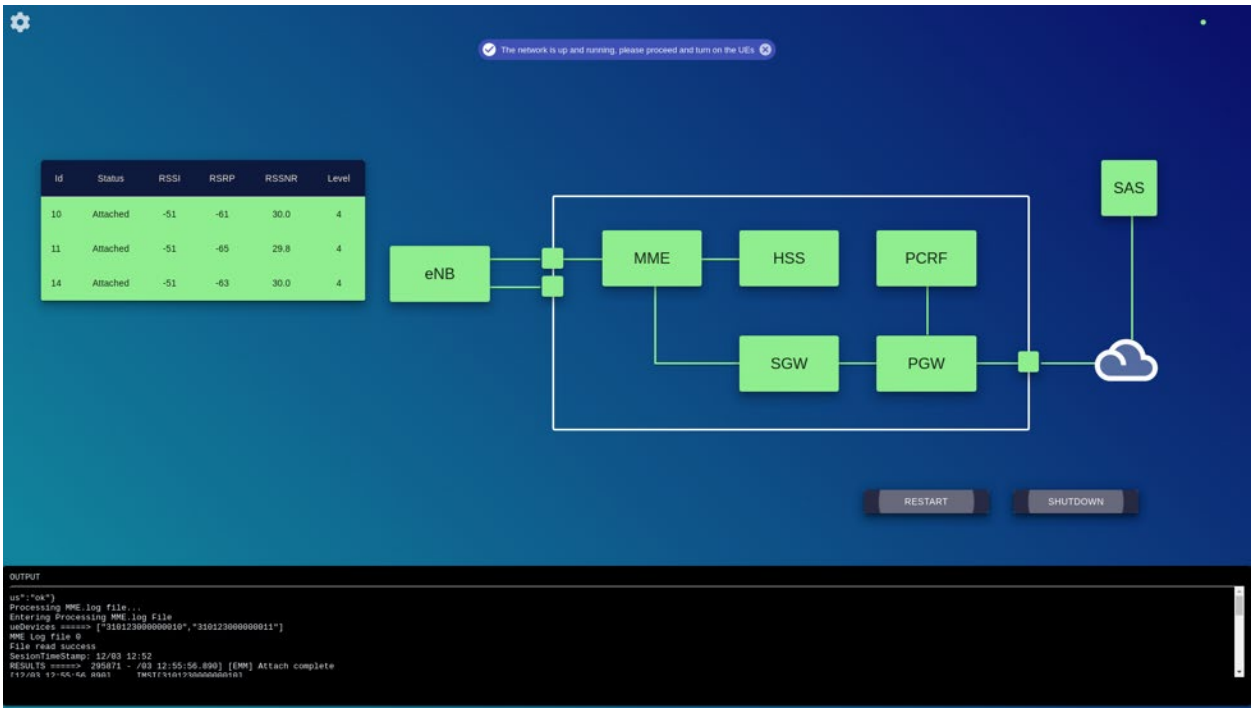


Figure 11 - The devices are successfully attached to the NIB network and the RF conditions reported and presented in the NIB portal

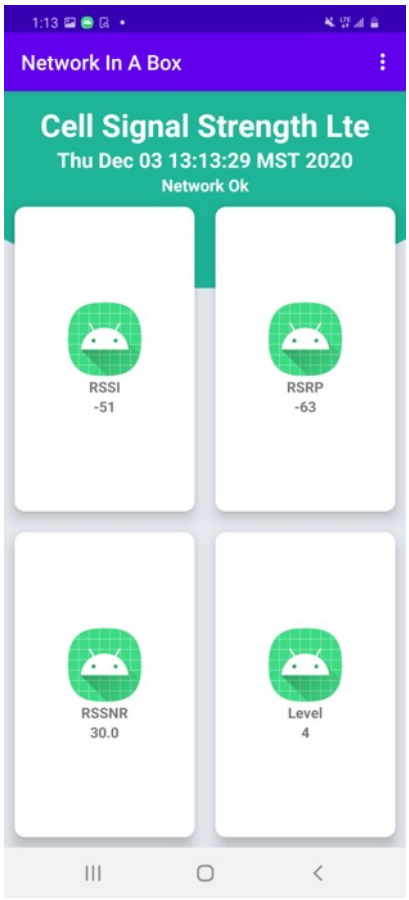


Figure 12 - Corresponding device NIB application

3.2.3. Business Logic

Three software components had to be used in this layer. The software list is ExpressJS, NodeJS and Puppetee,r and can be located into the following referenced links 3-5.

The ExpressJS was used to build a Rest API interface to serve all the HTTP requests from the Desktop Application.

Express is a Node.js web application framework that offers a comprehensive range of functionality for both web and mobile apps.

Express adds a thin layer of basic web application functionality without obscuring the Node.js capabilities users already know and appreciate.

Node.js is a scalable network application builder that uses an asynchronous event-driven JavaScript engine. Many connections can be handled at the same time in the "hello world" example below. The callback is invoked with each connection, but if there is no work to be done, Node.js will sleep.

Puppeteer is a Node library that provides a high-level API to control headless Chrome or Chromium over the DevTools Protocol. It can also be configured to use full (non-headless) Chrome or Chromium.

A headless browser is a great tool for automated testing and server environments where one does not need a visible UI shell. For example, one can run some tests against an actual web page, create a PDF of it, or inspect how the browser renders an URL.

The Chrome DevTools Protocol allows for tools to instrument, inspect, debug and profile Chromium.

3.2.4. Data Layer

The database which was used to store the data is SQLite and it can be found in the referenced link 6.

The SQLite file format is stable, cross-platform, and backwards compatible
The following information is stored in the database:

UeID: User Entity Identification
RSSI: Received Signal Strength Indicator
RSRP: Reference Signal Received Power
RSSNR: Reference Signal Signal to Noise Ratio
CQI: Channel Quality Indicator
level: Battery Level
timeStamp: Time Stamp

3.2.5. Development Strategy

It was not possible to have an environment for each developer who designs with the application.

The strategy was to create a simulator for the ENodeB as shown in Figure 13.
The ENodeB simulator was built using the ReactJS and NodeJS in the referenced links 1-2.

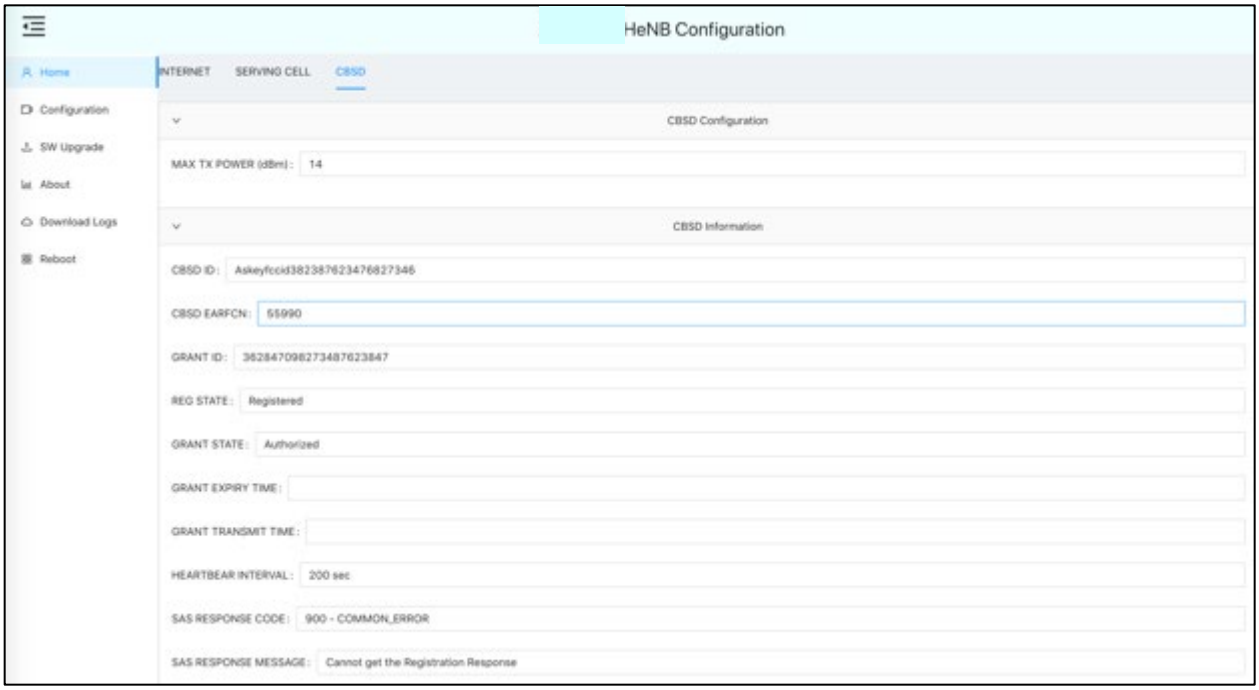


Figure 13 - eNodeB Simulator

3.2.6. Configuration Management

A Configuration had to be added for the user to be able to choose between which ENodeB is connected to the PLTE Core Network, as shown in Figure 14.

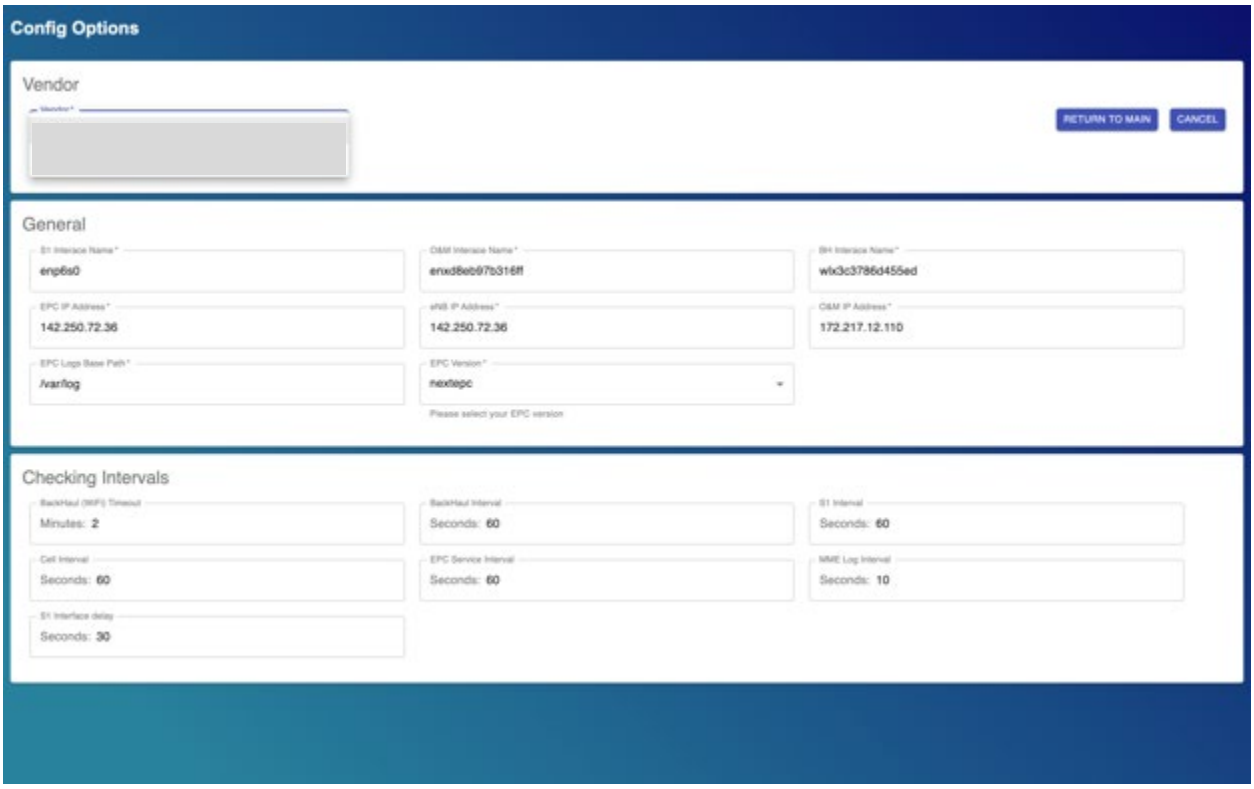


Figure 14 - Desktop Configuration

4. Mobile Edge Compute Use Case Examples

Beside generic end-user internet access, different options for MEC-based use cases were explored. The following lists two examples:

Closed Group Communication

Closed Group Communication aka 'Push To Talk/Video/Text' is an ideal communication method in an enterprise private wireless environment, e.g. factory, hotel etc... End users can be easily communicating with their respective groups (all or some at once or one-to-one), or a local PTx network dispatcher, by a push of a button. In our NIB solution, we installed a demo version the PTx server and dispatch applications in the MEC and respective PTx client application on the end-user devices.

VR (Virtual Reality) low latency applications

VR Workplace: Teleport to virtual office floor, Interact virtually with lab - machines and sensors

VR smart light control: Able the city workers to access Smart light for control and data retrieval using VR

VR Healthcare: Control instruments remotely, read patient charts

VR Industrial: Troubleshoot and analyze machine performance

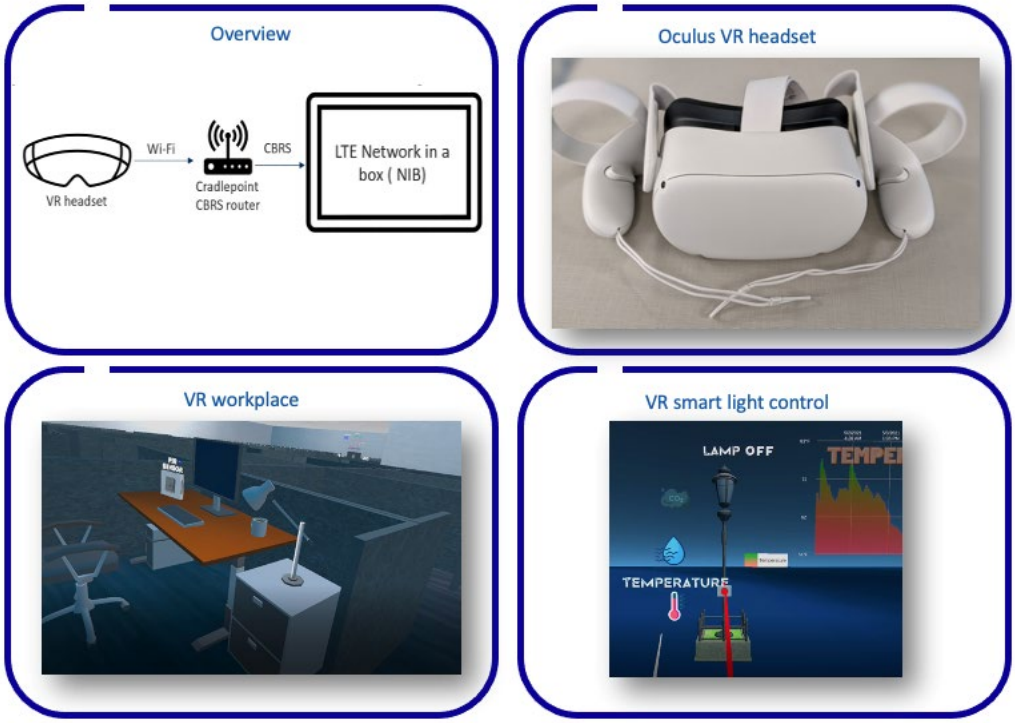


Figure 15 - VR over NIB with low latency

5. Conclusions

Network In A Box has a well-deserved reputation in the cellular industry. With the addition of 'zero touch' control and MEC in the same compute platform, the use for any users and many applications domains is opened up.

- Research and further development of NIB will continue which includes:
  - Evolve to a 5G system, e.g. 5G n48 Pico eNB and a 5G SA Core Network
  - Expand on the list of supported RAN vendors in the NIB 'zero touch' portal
  - Migrate the CN to a virtualized/containerized implementation

Abbreviations

CBRS	Citizen Broadband Radio System
CN	Core Network
COW	cell on the wheel
MEC	Mobile Edge Compute
NIB	Network In A Box
RAN	Radio Access Network
SAS	Spectrum Access System
VR	virtual reality

Bibliography & References

1- https://reactjs.org/

2- https://www.electronjs.org/

3- https://expressjs.com/

4- https://nodejs.org/en/

5- https://developers.google.com/web/tools/puppeteer

6- https://www.sqlite.org/index.html

New Service Paradigm With 5G Private Network

A Technical Paper prepared for SCTE by

Curt Wong
Sr. Dir – Wireless Standards, R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
425-395-4379
Curt.Wong@Charter.com

Yildirim Sahin
Director – Wireless Standards, R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
720-536-9394
yildirim.sahin@charter.com

Deh-Min Richard Wu
Director – Wireless Standards, R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
256-763-1202
deh-minrichard.wu@charter.com

Umamaheswar Achari Kakinada
Director – Wireless Standards, R&D
Charter Communications, Inc
6360 S Fiddlers Green, Greenwood Village, CO 80111
847-544-6560
Achari.Kakinada@charter.com

1. Introduction

Traditionally, a universal subscriber identity module (USIM) and roaming agreement between the serving network (visited public land mobile network (PLMN)) and home network (home PLMN) is needed in order for the user equipment (UE) to gain access to normal services (i.e., data/voice) when roaming. This type of "control" has been a long tradition used within the mobile network operators (MNO) to control which roaming network that their subscribers are authorized to gain services using 3GPP radio access technologies (RAT). With the ongoing development of 5G private network in the 3GPP ecosystem, a new service paradigm with 5G standalone architecture (5G SA) and 3GPP RAT is being created. USIM (or eSIM) are no longer the only storage mechanism for credentials when using 3GPP RAT. Onboarding can be performed locally for non-initialized UE to access the network. UE credential storage at the network side can be separated from the network that is providing the access to the users. And overall, the 5G private network can take advantage of the native support for edge computing to allow service hosting environment to be locally deployed for services with ultra-low latency and better quality of experience (QoE).

This paper first describes in a high level on how a UE obtains services with the current 4G cellular networks to contrast with the new capabilities offered with 5G private networks (also known as Stand-alone non-public network (SNPN)[1] [2]). These new capabilities defined by 3GPP 5G system architecture standards for SNPN set a course for a new service paradigm possibility for 5G private networks.

2. Services Provided By The Current Cellular Networks

Nowadays, a typical service provided by the cellular networks is mainly a best effort bit pipe connecting the UE to the external data networks (i.e., the Internet). When the bit pipe is characterized as best effort or non-guaranteed bit rate, it means that everyone shares the available bandwidth within the network and in some cases the user may exhibit congestion where the QoE became unsatisfactory.

Some services like real-time voice or public safety related services can be tailored by the network by setting the quality of service (QoS) parameters for that the bit pipe. This requires the network to be aware of at least two properties: 1) who is using the network, and 2) the type of QoS for which the user is eligible.

[1] SNPN is defined by 3GPP. Another type of 5G private network called public network integrated non-public network which is not described in this paper.

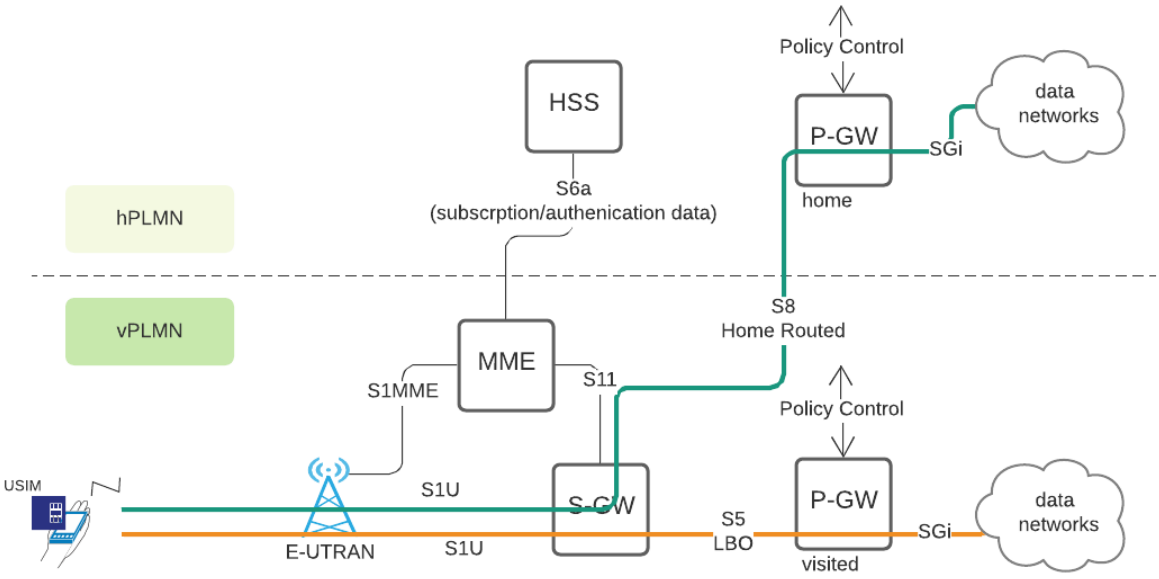


Figure 1 – Simplied EPS Architecture With Roaming Interfaces Illustrated

To determine who is using the network, a USIM with mutual authentications is used between the UE and the vPLMN. This requires the vPLMN to obtain the authentication vectors from the user's home (i.e., hPLMN) in order to perform mutual authentications with the USIM. With this requirement, a user must somehow obtain a USIM from a hPLMN before a cellular service can be rendered. In addition, a business relation (i.e., roaming relation) between vPLMN and hPLMN is needed in order to fetch the authentication vectors from hPLMN.

In the roaming example shown in figure 1, the UE has two protocol data network (PDN) connectivity services established from the vPLMN. The user's subscription in the home subscriber server (HSS) indicates that S8 home routed for internet traffic is used and local breakout (LBO) is used for e.g., voice media with IP-multimedia subsystem (IMS). Effectively, this means one user plane tunnel is breaking out locally at the vPLMN while the other is routed back to hPLMN.

The type of PDN connectivity service allowed (and their associated QoS profile) is based on user subscription data from hPLMN (but can be modified by the vPLMN if dynamic policy control is used).

The following list shows the QoS parameters related to each PDN connection:

- QoS class identifier (QCI) – a scalar that points to a set bearer level packet forwarding treatment (delay budget, packet loss, scheduling priority, etc.).
- Allocation and retention priority (ARP) – for admission related criteria during congestion with priority level (scalar), the pre-emption capability (flag) and the pre-emption vulnerability (flag).
- Guaranteed bit rate (GBR) (as oppose to the best effort).
- Maximum bit rate (MBR) – the upper limit of the bit rate that is expected to be provided by a bearer.
- Access Point Name Aggregate Maximum Bit Rate (APN-AMBR) – to limit the aggregate bit rate that can be expected to be provided across all non-GBR bearers and across all PDN connections of the APN.

- UE Aggregate Maximum Bit Rate (UE-AMBR) – to limit the aggregate bit rate that can be expected to be provided across all non-GBR bearers of a UE.

3. Why 5G Private Network Is Different?

As described in the previous section, a user who is able to obtain connectivity service in a vPLMN must first have a prior relationship with a hPLMN and that both vPLMN and hPLMN must have roaming relationship with each other in order to allow the vPLMN to authenticate the user and to obtain the subscription information for QoS settings and for user plane routing between the UE and the networks.

For 5G private networks – namely SNPN, it is neither necessary to use USIM to gain access to local access network using 3GPP RAT nor that a roaming interface is required from the SNPN in order to reach hPLMN.

3GPP SNPN [2] offers the following new capabilities which are different from technologies used for prior cellular services described in the early section.

- Subscriber identifier (SUPI) is no longer restricted to IMSI. Network access identifier (NAI) using the NAI RFC 7542 based user identification (e.g., lux_luther@phantom.zone) can be defined and used.
- Mutual authentication is no longer restricted with the shared secret stored in the USIM. Extensible authentication protocol (EAP) with TLS/TTLS can be used (e.g., certificate, username/password) as well. The credentials can be stored in the device itself as well.
- SNPN is identified with the addition of NID (Network identifier) – i.e., PLMN ID + NID. PLMN ID is not required to be unique. It the traditional cellular network, PLMN ID (i.e., MCC and MNC pair) is assigned by either ITU or regional organization like ATIS IOC (IMSI Oversight Council) and is globally unique. For SNPN using MCC with 999 [1] any MNC and also any NID can be used. Operator can also choose to use its own dedicated PLMN ID plus any self-managed NID value as deployment choice. In addition, 3GPP allows the option of using a globally unique NID value independent of the PLMN ID.
- NG-RAN may also broadcast a human-readable network name to ease the user to select SNPN. This is mainly used for manual selection based on network name awareness of the user.
- Onboarding service allows a UE without any prior relationship (i.e., no credential to access the network) to obtain the necessary authorization and credentials on demand from a local SNPN.
- Authentication vectors can be stored externally from SNPN (e.g., in an external AAA server, other UDM from 3rd party) to allow neutral hosts type of deployment.
- Accessing to hPLMN service via SNPN is also possible with untrusted access procedure over 3GPP RAT.

Please note that real time voice communication and emergency session can also be supported with SNPN toward the local PSTN.

The following figure illustrates an example of a simplified SNPN architecture.

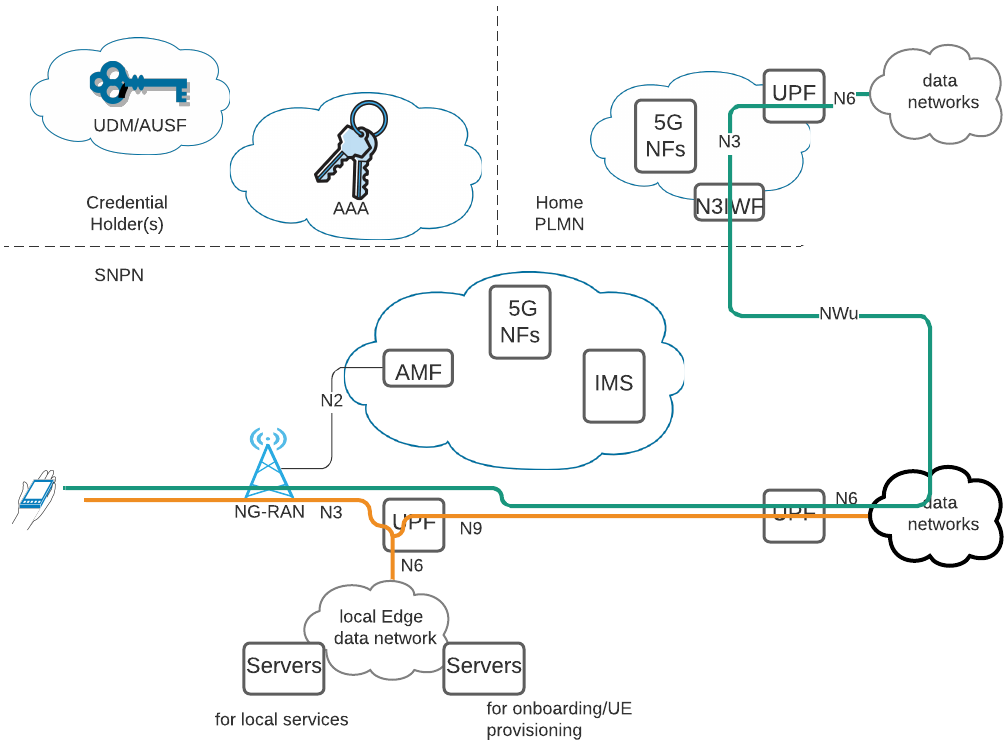


Figure 2 – Simplied SNPN Architecture With User Plane Routing For Different Services

The following section describes some of the key capabilities [2] [3] with more details using SNPN that is differed from PLMN.

3.1. UE Onboarding

Onboarding is the service to allow a non-provisioned UE to obtain SNPN credentials in order to get connectivity service from the SNPN.

The following figure shows a simplified view on how a non-initialized UE is provisioned with SNPN credentials.

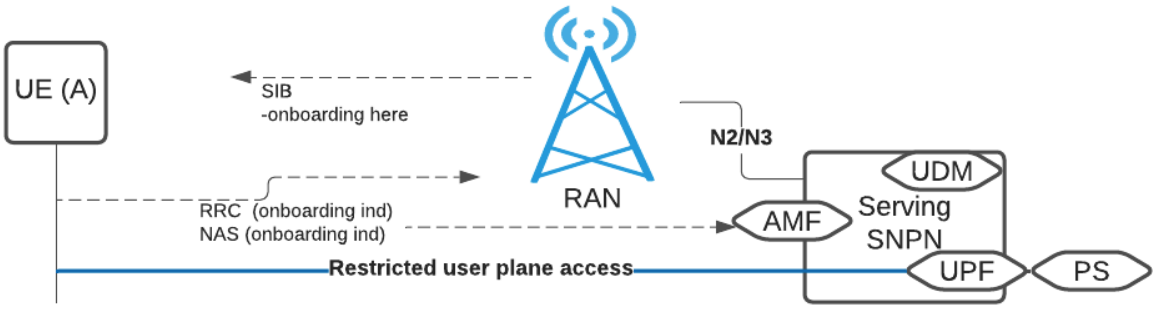


Figure 3 – Onboarding Process For A Non-initialized UE.

The SNPN which allows onboarding process will broadcast an indication at the cell level that onboarding is supported. UE, which wants to be provisioned with SNPN credentials for connectivity service from the SNPN, will signal to the network with onboarding indication in both radio resource control signaling (RRC) and non-access stratum signaling (NAS). This allows the RAN to forward this request from the UE to a dedicated onboarding access and mobility management function (AMF). AMF uses the NAS level indication to select a session management function (SMF) used for remote provisioning and to setup a restricted user data plane between the UE and the provisioning server (PS). The SMF may send the PS fully qualified domain name (FQDN) to the UE as part of protocol configuration options (PCO) in the PDU session establishment response. After PDU session is set up, the application in the UE is then interacting with the PS to obtain the SNPN credentials. The PS may request certain information from the UE/user (e.g., name, type of connectivity services requested, credit card, etc.) prior to sending the SNPN credentials to the UE. The PS may also provision the UDM with corresponding SNPN subscription data if it has not been prefilled already. The applications in the UE for handling provisioning task is out of scope of 3GPP standards and is left for OEM implementation.

Once the UE has received the SNPN credentials from the PS, it will restart and use the downloaded SNPN credential to access the network.

3.2. Access to SNPN With Credentials Owned By An Entity Separate From The SNPN

Credentials holder is defined in 3GPP [2] as an entity which authenticates and authorizes access to an SNPN separate from the credentials holder. It means that the serving SNPN does not store the credentials that can be used to authenticate/authorize the UE. This is also commonly known as neutral host offering as it allows the SNPN to provide connectivity service to the users using credentials from 3rd party.

The following simplified figure shows how a serving SNPN interacts with credential holders.

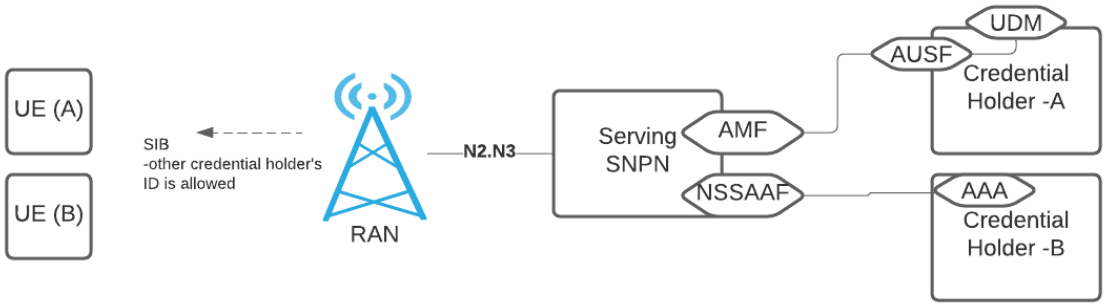


Figure 4 – SNPN Interacting With Credential Holders

The following figure shows two 5G architecture views from 3GPP TS 23.501 [2] with credentials holder and SNPN.

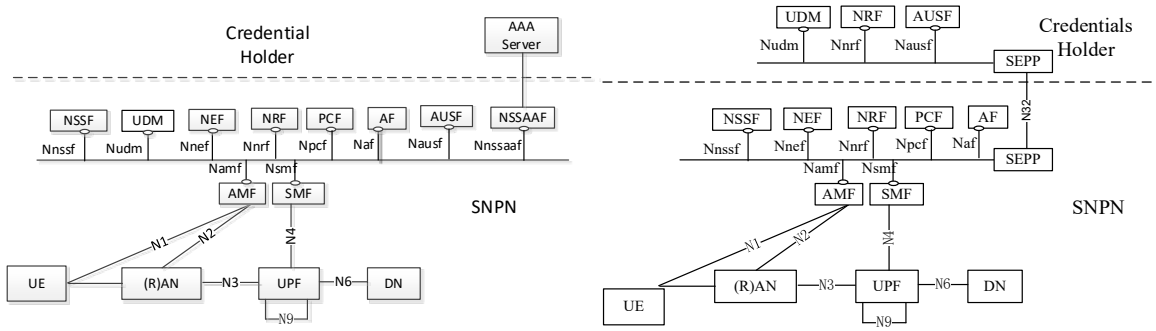


Figure 5 – SNPN and Credential Holder From 3GPP TS 23.501

Please note that from 3GPP's point of view, the architecture for SNPN and credentials holder is always depicted as a non-roaming reference architecture even though a roaming reference point (i.e., N32) is used. In this sense, the 3GPP does not support roaming for UE between SNPN and MNO.

When accessing a SNPN with credentials from 3rd party (credential holder), the UE must first check the system information block (SIB) broadcast message from the RAN to determine which 3rd party (credential holder) is supported by this SNPN. This could be in the form of a list of PLMN ID + NID or group ID (GIN) for network selection. GIN represents a group of 3rd parties using a common Network ID to minimize the broadcast list in the SIB. If the UE has been configured with a credential from one of those 3rd parties that is shown in the SIB then the UE may proceed to register to the SNPN using the 3rd party credential.

RAN may also broadcast an additional indication in the SIB to indicate that UE with any 3rd party (credential holder) information can try to access this SNPN. This type of uncontrolled access may be useful for general public usage (e.g., public access at the park, public library, etc.).

Credential holder can be from another SNPN or PLMN or from enterprise domain in the case of AAA.

If the credential information is stored in 3rd party UDM, AMF in SNPN forwards the EAP message to the AUSF of the 3rd party based on PLMN ID + NID received from the UE.

If the credential information is stored in an AAA Server in the credential holder, the authentication function (AUSF) in SNPN selects a network slice-specific and SNPN authentication and authorization function (NSSAAF) to handle the related EAP messages from the UE. The NSSAAF selects AAA server based on the domain name to the realm part of the SUPI, relays EAP messages between AUSF and AAA server (or AAA proxy) and performs related protocol conversion. The AAA server acts as the EAP server for the purpose of primary authentication. UDM is still used for storing subscription information and to decide that the primary authentication is performed by AAA or AUSF.

3.3. Accessing PLMN Services From SNPN

Interesting enough that even when 3GPP does not support roaming between SNPN and PLMN, the specification does allow a UE with dual credentials (e.g., similar to dual SIM dual standby (DSDS) or dual SIM dual active (DSDA)) to access both SNPN and PLMN services at the same time. This is particularly useful when a user wants to separate their usage between enterprise and personal domains or when the user is in a remote SNPN environments like caves or shielded factory where PLMN coverage is not available or accessible.

3GPP defines the following terms to cover this usage:

- Overlay network: When UE is accessing SNPN service via NWu using user plane established in PLMN, SNPN is the overlay network. When UE is accessing PLMN services via NWu using user plane established in SNPN, PLMN is the overlay network.
- Underlay network: When UE is accessing SNPN service via NWu using user plane established in PLMN, PLMN is the underlay network. When UE is accessing PLMN services via NWu using user plane established in SNPN, SNPN is the underlay network.

The following figure illustrate this architecture setup using hPLMN as overlay as an example. Please note that the other direction is also valid (i.e., accessing SNPN via PLMN's 3GPP RAT) but is not used here for illustration.

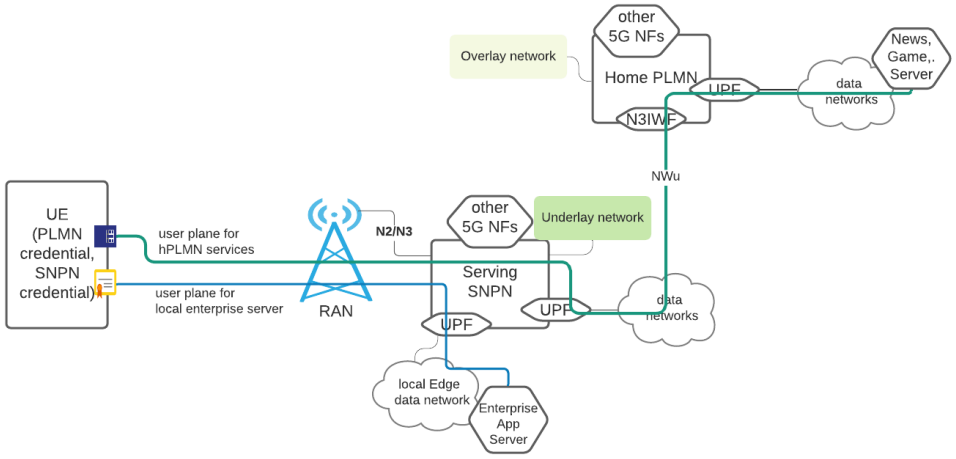


Figure 6 – Accessing HPLMN Services Via SNPN

The UE has both the SNPN credential and the PLMN credential. The enabler here is that the UE is able to use the untrusted non-3GPP access procedure over the 3GPP RAT of the SNPN in order to create a NWu

tunnel toward the hPLMN. The SNPN credential is needed in order to get access to the data network via SNPN. The PLMN credential is needed in order to access the hPLMN via the NWu tunnel.

On the SNPN side, UE maintains at least one PDU session in the network that can reach the N3IWF of the hPLMN, and UE keeps its state in CM-CONNECTED state as if the UE is using WiFi access.

From hPLMN's perspective, UE is accessing the hPLMN as if the UE is using any WiFi access even though the UE is actually using 3GPP RAT from SNPN.

When UE is moving out of the SNPN coverage and into the 3GPP radio coverage of the hPLMN, the user data session can also be transferred from SNPN to hPLMN using non-3GPP to 3GPP handover procedures as described in 3GPP specifications. In other words, when UE has detected that the radio coverage of the PLMN is sufficient, it can initiate a registration procedure to the hPLMN with an indication that an existing user plane session from NWu needs to be moved over to this 3GPP radio interface. In the figure below, the data carried via green path will be moved to blue path after this handover procedure is completed.

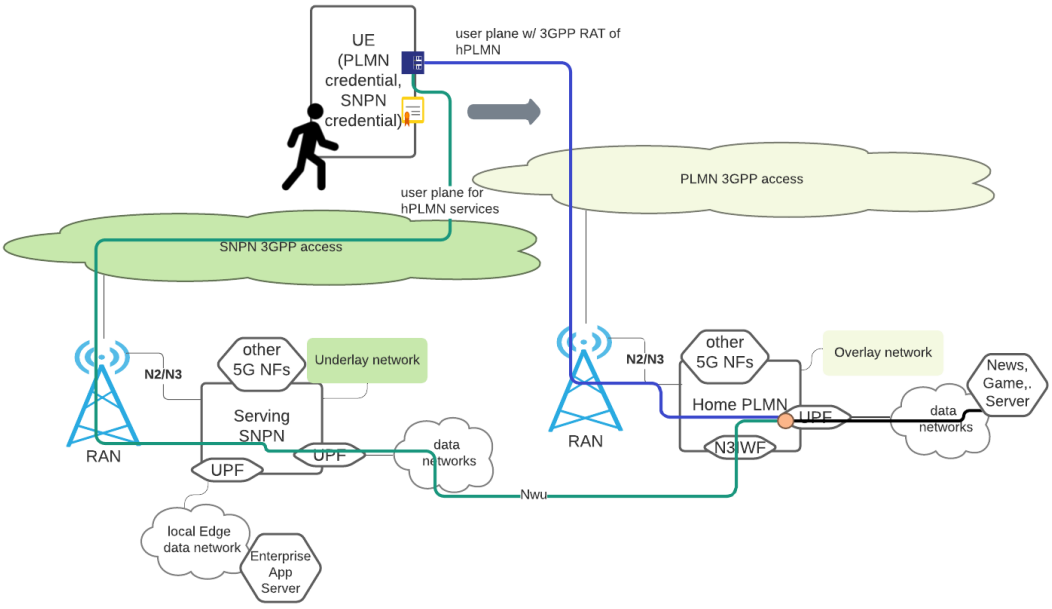


Figure 7 – Mobility Into PLMN's 3GPP RAT Coverage From SNPN

With the network based non-3GPP to 3GPP handover procedure, the data path is switched from one tunnel to another (i.e., from blue to green in the above figure). This means that even if the UE is able to maintain Rx/Tx connections to both 3GPP RATs at the same time, the UE must decide which tunnel is used for data transfer.

If the UE is bouncing between both networks then the tunnel switching will also happen each time the UE is switching between the networks. This may induce excessive network signaling procedure with the hPLMN. To minimize the signaling induced by tunnel switching with HPLMN, and to allow mobility procedure between SNPN and PLMN, a 3GPP feature called access traffic steering, switching, splitting (ATSSS) can be used. ATSSS allows the UE and the network to transfer user plane data via one or more data paths using MPTCP and/or ATSSS low-layer (ATSSS-LL). MPTCP is a protocol defined by IETF

and has been widely used in the industry already for reliable delivery via different data paths for TCP traffic. For UDP, 3GPP develops its own ATSSS-LL as IETF MP-QUIC is not yet finalized. Overall, ATSSS defines the following different steering modes to steer the traffics between two paths:

- Active-standby: One path is denoted as primary access and is always used even when the other path is available.
- Smallest delay: The path with the smallest round-trip time (RTT) is used.
- Load-balancing: Traffic is split across different paths.
- Priority-based: The primary path is used until it reaches certain congestion level in which case the other path is used.

4. Services That Can Be Offered With 5G Private Networks

3GPP in Release 18 has been studying new use cases for 5G networks providing access to localized services and identifying relevant requirements. The study takes a number of relevant aspects into consideration as depicted in a business model in Figure 8 [5]:

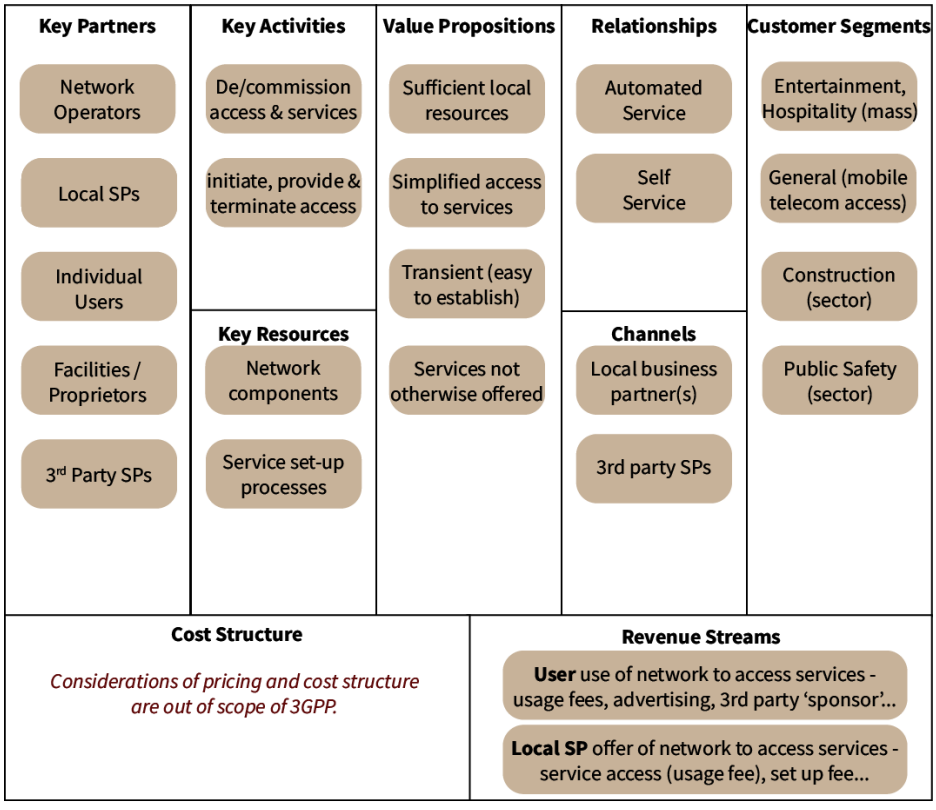


Figure 8 – Business Model For Providing Access To Localized Services

- *Key partners* may include network operators (MNOs, private network operators, MSOs, etc.), local service providers, individuals (users), owners of facilities or proprietors of businesses in which the local access network will be available, and 3rd party service providers. These stakeholders need to work together to provide local access to services.
- *Key activities* may involve how to commission/decommission access to local network and a localized services offered by any of the key partners using the local network. Since the access is local and may

be bounded in time and space, the effort to commission/decommission accesses and services needs to be aimed to be requiring short lead-time and low complexity, etc. From a user perspective, the user needs to become aware of access and local services to choose and access them. The process for the user and his/her equipment to gain access to the network, to use and terminate access and services needs to be efficient, simple and convenient. The offered resources and services cannot be accessed by any other way.

- *Value proposition* perspective providing access to local services can result in some distinct opportunities for users and service providers.
    - The access can be provided that is sufficient in areas that otherwise would lack them, for example, on a fairground established far from other infrastructure.
    - The access to local services can be simpler than access would be without this service. For example, obtaining network access may result in associated local service configuration and effortless presentation to the user.
    - The access and local services operation can be established as needed, without the need for long term business relationships such as facilities, permanently installed equipment, etc.
- In order to establish the set of business *relationships* and arrangements from the perspective of all involved key partners that could be very complex, business processes need to be automated, or at the very least available for self-service.
- *Customer segments* for localized services can be broad. The ones listed in the figure are just some sample ones.
- Two kinds of key *resources* are needed to provide access to local services:
    - Network components: The network operator needs to have or be able to make use of other operators' network infrastructure, both for mobile access and for configuration of services, authorization and other aspects.
    - Service set-up processes: A set of service set-up processes need to be in place such that the service providers (local service providers, 3rd party service providers and the network operator provided services) are able to arrange for their services to be offered via the local access.
- These local service access can be promoted and arranged through different *channels*. Principally the local service operators, such as an entertainment venue, can provide and promote information to potential users so that they can seek to access the local services. Third party service provider, such as a sports club, can also inform, motivate and prepare their users to expect local access to services in a particular place and at a particular time.
- *Cost* and *revenue* due to offering and usage of localized services can vary such that the key partners need to make necessary business agreements to get their proper share. The cost of using some localized services to users may be free-of-charge or usage-based. Some offered localized services may be sponsored, for example, by any of the key partners, advertisers, or 3rd party sponsors.

In the light of the aspects listed above, as an example for providing access to localized services, an SNPN deployed at a sporting venue may offer various services to their customers visiting the venue for a game: The offered services may be like the Internet access, real-time video or AR/3D services from various camera angles of the field (e.g., the perspective of the referee or a favorite player) that are only available to visitors at the venue. Any services in the venue may be offered by one or more key partners such as the sporting venue owner, infrastructure network provider, broadcasting agency like a TV station, 3rd party sponsor or one of the athletic clubs, etc. The visitors can be made aware of the offered services and how to get onboarded to the local access network and the localized services before or when they come the venue by one or more of the key partners. The offered services can be broadcasted by the local access network and compatible UEs can show the available services to the users; and guide the interested users for easy onboarding and then start using the services. When new services or an update to the previously

offered services become available, the visitors can be made aware via various means such as updating service broadcast at access network or via a venue application, etc. While some or all services offered might be free-of-charge or usage-based to the users. The service provider may generate advertising revenue by offering advertising services to 3rd party companies while users rendering localized services.

5. Standards Development For 5G Private Networks

Standardization work for 5G private networks (i.e., SNPN) was started from 3GPP Rel 16 using 5G standalone architecture as the baseline. The development has continued in Rel 17 [3] with more capabilities and features. Currently, the industries are discussing further possible enhancements in Rel 18 [4] [5] [6].

Rel 16 was frozen at June 2020. Rel 17 is currently under development and is scheduled to be frozen by March 2022. R18 has already been started with new requirements gathering by 3GPP SA1 since the end of 2020 and most likely that the freeze date for Rel 18 will be by the end of 2023 or early 2024.

The following figure shows some of the key features related to 5G private networks per 3GPP release cycle under the umbrella of the system architecture group (3GPP SA). Rel 18 content is only at the initial phase of the development; hence, it is a best guess from the authors of this paper.

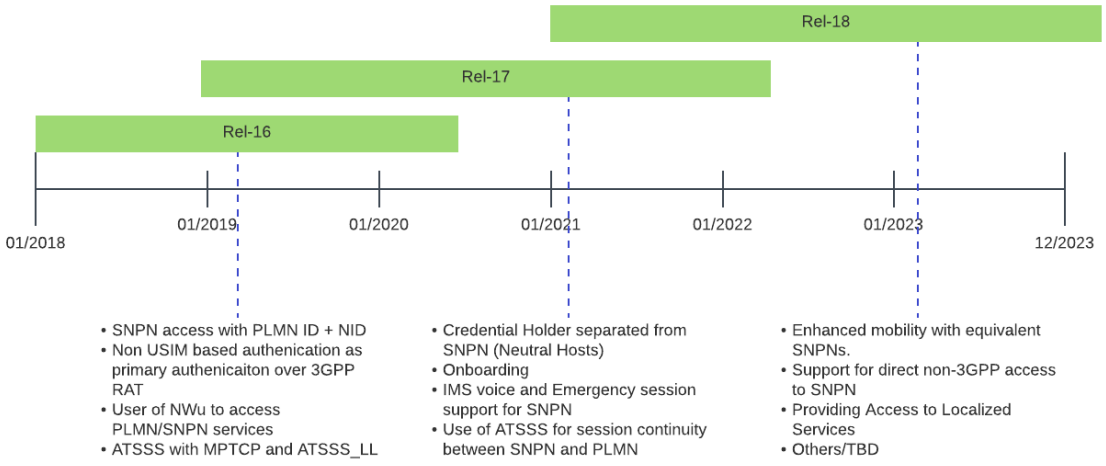


Figure 9 – SNPN Related Feature Per 3GPP Release

In 3GPP Release 18, in addition to providing access to localized services studied in [5], other potential Rel 18 features are aiming for better mobility experience. Enhanced mobility with equivalent SNPNs, allows a UE with subscription for one of the SNPNs has access to its all equivalent SNPN(s), similar to how equivalent (h)PLMNs work today. 5G system architecture can already support non 3GPP access like Wi-Fi or fixed broadband access. The SNPN feature has so far been focusing on using 3GPP RAT due to time pressure, however documenting the support for SNPN for non-3GPP access will make the standards specifications to be more holistic.

6. Conclusion

The new capabilities available for 5G private networks (i.e., SNPN) open a new realm of service possibilities that was previously not available with 4G systems based on evolved packet core (EPC). The on-demand credential provisioning with onboarding feature, the adoption of using network access identifier (NAI) for user identity and EAP with TLS/TTLS, flexible network ID assignment (PLMN ID + NID), and neutral host offering are some of the features that make 5G private network unique from any of the previous 3GPP systems, and this allows much wider applicability of using 3GPP eco-system with 3GPP RAT and non-3GPP RATs (WiFi, fixed-broadband access) for new service creation.

Abbreviations

3GPP	3rd Generation Partnership Project
AAA	authentication, authorization, accounting
AMF	access and mobility management function
ATIS	Alliance for Telecommunication Industry Solutions
ATSSS	access traffic steering, switching and splitting
ATSSS-LL	ATSSS low-layer
AUSF	authentication server function
CH	credentials holder
DSDA	dual SIM dual active
DSDS	dual SIM dual standby
EAP	extensible authentication protocol
eSIM	embedded SIM
EPC	evolved packet core
FQDN	fully qualified domain name
GIN	group ID for network selection
hPLMN	home PLMN
HSS	home subscriber server
IETF	Internet Engineering Task Force
IMS	IP-multimedia subsystem
LBO	local breakout
MCC+MNC	mobile country code + mobile network code
MNO	mobile network operator
MSO	multi system operator
MP-QUIC	multi-path quick UDP internet connections
MPTCP	multi-path transport control protocol
N3IWF	non-3GPP interworking function
NAI	network access identifier
NAS	non-access stratum
NF	network function
NG-RAN	next generation radio access network
NID	network identifier
NSSAAF	network slice-specific and SNPN authentication and authorization function
NWu	reference point for untrusted non-3GPP access network to 5GC
PCO	protocol configuration options
PDN	packet data network
PDU	protocol data unit
PLMN	public land mobile network
PS	provisioning server
QoE	quality of experience
QoS	quality of service
RAN	radio access network
RAT	radio access technology
RRC	radio resource control

RTT	round trip time
SA	Stand Alone
SIB	system information block
SMF	session management function
SNPN	stand-alone non-public network
TCP	transport control protocol
TLS	transport layer security
TTLS	tunneled transport layer security
UDM	unified data management
UDP	user datagram protocol
UE	user equipment
UPF	user plane function
USIM	universal subscriber identity module
vPLMN	visited PLMN
GBR	Guaranteed bit rate
MBR	Maximum bit rate
ARP	Allocation and retention priority
APN-AMBR	Access Point Name Aggregate Maximum Bit Rate
QCI	QoS class identifier
UE-AMBR	UE Aggregate Maximum Bit Rate
SUPI	Subscriber identifier
ITU	International Telecommunication Union
OEM	Original Equipment Manufacturer (to describe smartphone vendor)
TX	Transmit
RX	Receive

Bibliography & References

[1] International Telecommunication Union (ITU), Standardization Bureau (TSB): Operational Bulletin No. 1156; http://handle.itu.int/11.1002/pub/810cad63-en (retrieved October 5, 2018)

[2] 3GPP TS 23.501 V16.8 and V17.1, System Architecture for the 5G System (5GS); Stage 2

[3] 3GPP TR 23.700-07 V17.0, Study on Enhanced Support of Non-Public Networks (NPN)

[4] 3GPP Study Item proposal for Rel-18, Study on Enhanced Support of Non-Public Networks; Phase 2 (S2-2104337)

[5] 3GPP TR 22.844 V18.0, Study on 5G Networks Providing Access to Localized Services; Stage 1

[6] 3GPP Work Item proposal for Rel-18, 5G Networks Providing Access to Localized Services (SP-210588)

Node Health Within Cox ACOE's Service Health Framework

Improve the Health and Quality of HFC Services through Predictive Analytics

A Technical Paper prepared for SCTE by

Shane Yates
Executive Director
Cox Communications
6305 Peachtree Dunwoody rd.
Atlanta, GA 30328
888-566-7751
Shane.Yates@Cox.com

Brian Stublen
Director
Cox Communications
6305 Peachtree Dunwoody rd.
Atlanta, GA 30328
888-566-7751
Brian.Stublen@Cox.com

Alexis Hwang
Manager
Cox Communications
6305 Peachtree Dunwoody rd.
Atlanta, GA 30328
888-566-7751
Alexis.Hwang@Cox.com

1. Introduction

The Service Health framework is a suite of analytic models designed to predict and prevent customer impacting issues, while improving overall service quality. While this framework includes a multitude of Network, Customer, Premise, etc. focused components, this project and abstract are focused on Node Health and improving the effectiveness of the workforce that operates the HFC network, the Field Network Technician (FNT).

The FNT group is responsible for responding to all network impacting events, restoring customers' services, and performing proactive network maintenance to prevent issues before they occur. Previously, this work was created through a combination of rules-based telemetry processing and manually generated events. While this approach provided insight into clear rules violations regarding customer service availability, performance, and utilization, it was both limited in its ability to find novel patterns impacting customers and in its effectiveness with prioritizing work based on the potential impact to customers.

However, through the Node Health Framework, Cox can analyze thousands of combinations and permutations of RF patterns that may impact customer's services now and into the future. Additionally, this framework uses the predicted impact to prioritize work, both daily at a macro-level and hourly at a micro-level. As a result, FNTs have increased the volume of proactive network maintenance, prevented thousands of customer impacting events, reduced transactions to record low levels (note: Cox has attributed a 10%-15% reduction in transactions from the Node Health initiatives), and radically improved customer service quality.

2. History

Network Health is an initiative with roots going back to 2015 within Cox Communications. This started as a KPI-driven initiative to measure the health of the HFC network. The ultimate output of the early initiative was a measurement system called OSP7 (pictured below).

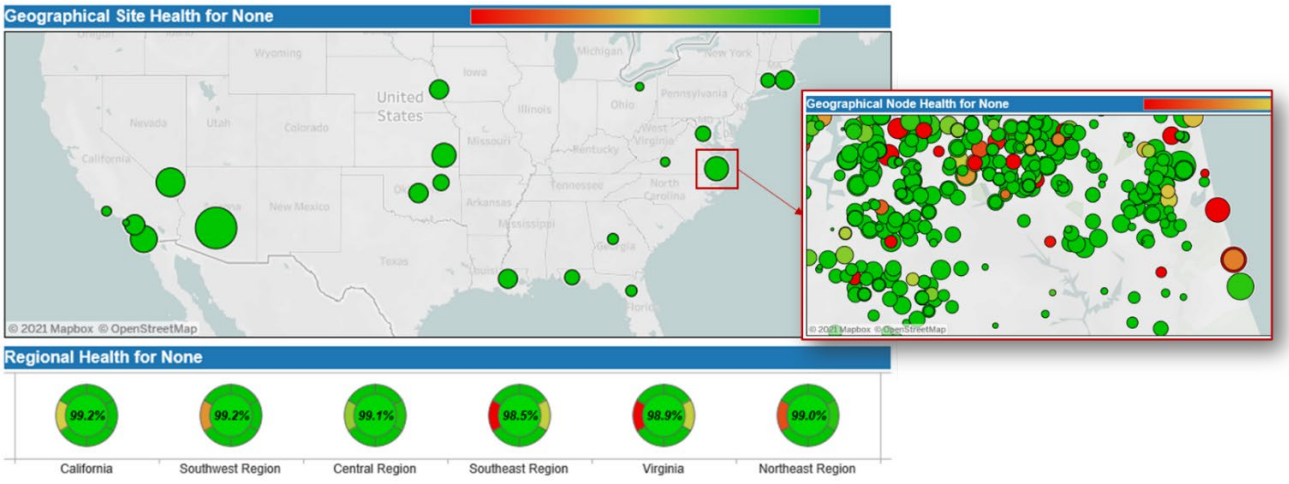


Figure 1 – OSP7 Measurement System

This measurement system began as a preliminary set of KPIs that measured a wide array of metrics, 35 to be specific, which ranged from operational performance to technician throughput/effectiveness.

However, the analytics team believed there was room for improvement, so they put these KPIs through the two-factor test below:

1. Does this metric correlate with operational processes and practices? Or to put this more simply, can the Field Network Technician (FNT) organization successfully influence the metric either positively or negatively?
2. Does the metric correlate with business and financial objectives?

As a result of this two-factor test they created the OSP7 (or Outside Plant 7), which contained the 7 metrics represented below that correlated best with both FNT labor processes and business/financial objectives. The table below represents the first draft of the OSP7 program:

Table 1 – OSP7 Program Metrics

	Metric	Description	Spec
1	**DS RX (Downstream Receive)**	Downstream Power-level at the Customer modem within Docsis Spec	-12dB to +15
2	**US TX (Upstream Transmit)**	Upstream Power-level at the Customer modem within Docsis Spec	+30dB to +52
3	**DS SNR (Downstream Signal to Noise)**	Downstream SNR ration at the Customer modem within Docsis Spec	>+32 dB
4	**US SNR (Upstream Signal to Noise)**	Upstream SNR ration at the Customer modem within Docsis Spec	>+30 dB
5	**DS FEC (Downstream Forward Error Correction)**	Downstream packet loss at the Customer modem	<1% packet loss
6	**US FEC (Upstream Forward Error Correction)**	Upstream packet loss at the Customer modem	<1% packet loss
7	**Aggregate Health**	The combined score between all metrics	n/a

Within this program, the magic was not derived from the KPIs themselves, but rather from the user's ability to navigate the map (pictured above) and quickly identify problematic areas that required more focus from the FNT team. The FNT was now able to make decisions based on analytical modeling (i.e. what is healthy in green and unhealthy in red) provided to them using simple visualization principles. As a result, Cox experienced a substantial improvement in overall network performance (~1% improvement to OSP7 over 2 years associated with this work) and a corresponding drop in transactions (~5% reduction to Tech Support calls and UHT truck rolls). This reduction was calculated using several controlled experiments, where we measured improvements at the node, headend, and market/site level vs. control groups to validate and quantify the effect with precision.

3. Service Health Framework

As the organization matured, there was a clear need to evolve the space of analytics. As a result, Cox announced the formation of the Analytics Center of Excellence (ACOE) in 2017, which centralized all analytics teams across Cox's Technology and Operations organizations. With the formation of this new team, the ACOE launched new advanced analytical capabilities focused on the health and quality of the customers' services. This led to the creation of the Service Health Framework.

The Service Health Framework is based on three principles:

1. Most customer-impacting events can be predicted
2. Customer-impacting events are complex but unique and identifiable
3. We can shift from the customer as a diagnostic to prediction as a diagnostic

Additionally, there are four key focus areas that shape the framework:

1. Deep understanding of data that describes quality and use of service by our customers
2. Ability to predict future service impacts and mitigate or prevent them
3. Machines executing the right actions to take, at scale
4. Continuous learning capabilities to improve over time

From the intersection of these principles and focus areas, the ACOE designed the following structure to define the Service Health framework:

- **Node Health**: Predict, isolate, and prevent any customer experience issues at the node level
- **Location Health**: Predict home/business wiring, device, Wi-Fi customer experience issues
- **Usage Health**: Understand customer experience impact based on individual customer usage and behavior
- **Interaction Health**: Link human conversation to Service Health root cause

4. Node Health

As the Service Health framework launched, the previous foundations of network health, outlined in the history above, were used to quickly advance the overall framework. In particular, the lessons learned from using the OSP7 methodology/dashboard on how to direct labor, became an essential component for how Cox's ACOE developed the Node Health component of the Service Health framework. To help frame this up, the ACOE developed three pillars as the core to the Node Health concept:

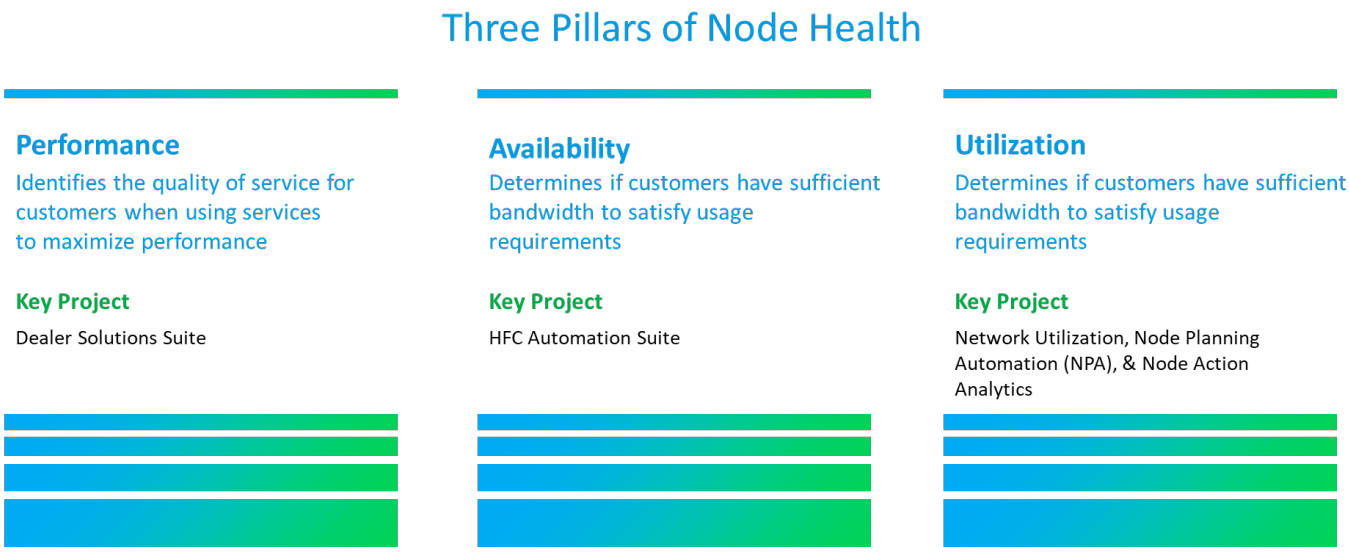


Figure 2 – Three Pillars of Node Health

These three pillars are the backbone of the Node Health framework and contain a suite of projects that seek to predict problems before customers recognize issues and prevent unnecessary transactions. Projects within these pillars span from predictive analytics, proactive customer resolution, and real-time root cause analysis. To dig a little deeper, there are six overarching projects that will deliver the core of the value within the Node Health framework.

These projects are described in the diagram below:

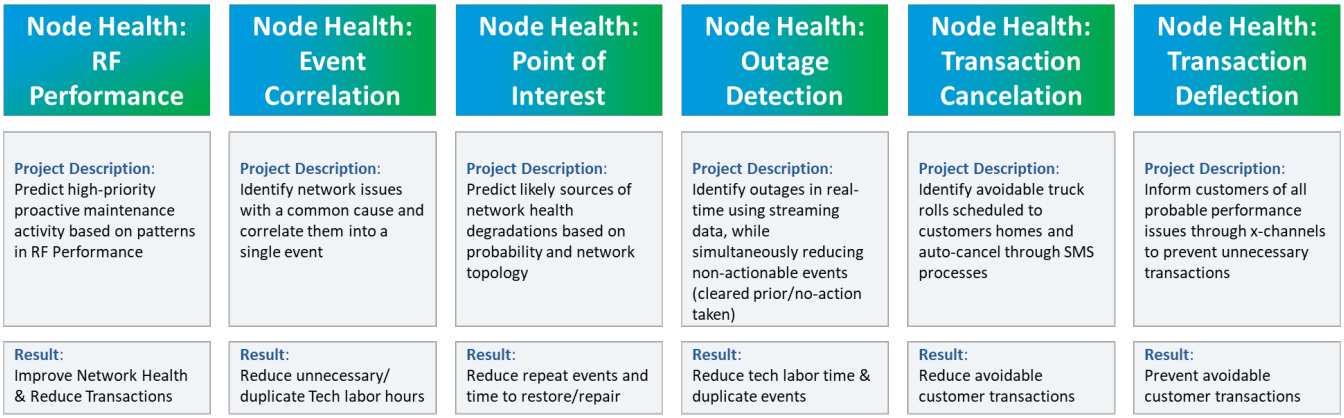


Figure 3 – Node Health Projects

The first program, titled "Node Health: RF Performance", is a proactive network maintenance model that takes in data from a multitude of sources and makes a prediction about the most impactful work that a technician can perform. Specifically, the model makes a prediction about the impact/value of performing work in every HFC node by creating patters based on historical node performance and the response/improvement those patterns recognized following proactive maintenance. Following this

prediction, the model then stack ranks those nodes based on the calculated value and routes the work according to priority to ensure the most impactful work is driven to the top priority. To make these predictions, we use the following suite of models listed below. One component left out of the diagram below is the continuous learning component of the model, which continues to learn how RF impairment patterns are impacted by proactive network maintenance activity to retrain and optimize the model.:

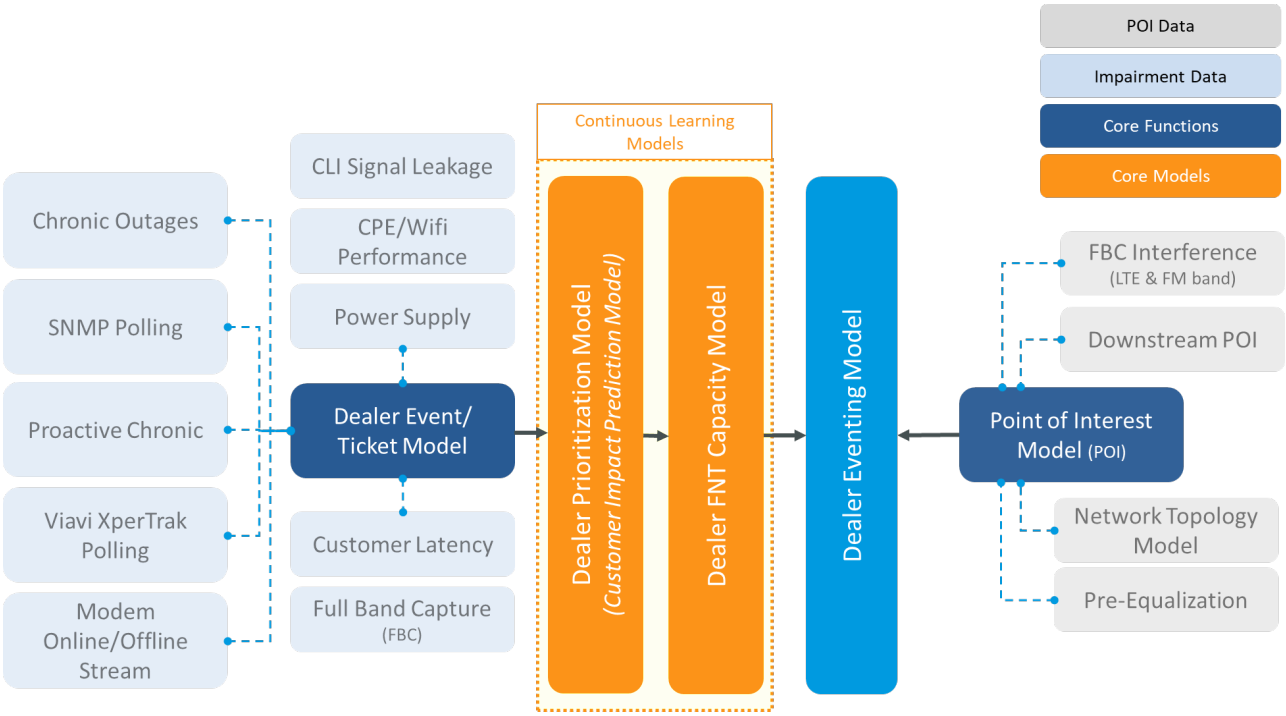


Figure 4 – Node Health Models

5. How Does Node Health RF Work?

5.1. Overview

Node Health RF Performance is a proactive platform for identifying areas of the network with recurring, intermittent and service affecting issues in order to optimize the use of Field Network Technicians' (FNT) time and labor before customers recognize an issue. With the introduction of Node Health RF Performance, Cox moves away from leveraging our customers as a primary diagnostic to leveraging machine learning models and their predictions as a primary diagnostic.

5.2. BAU Process (Previous Approach)

Before Node Health, our journey started with a customer calling in an issue. The customer in this case is uninformed throughout the entire process. Painfully, our technicians and care agents also have no visibility into the status of the issue. As a result of this reactive process, most operators have a difficult time enabling technicians or leveraging enhanced tooling to monitor and respond to the network events. Furthermore, managing customer interactions during the process is quite challenging as none of the various channels (web, app, phone, etc.) have visibility to these impairments. Below is a diagram outlining a recent example:

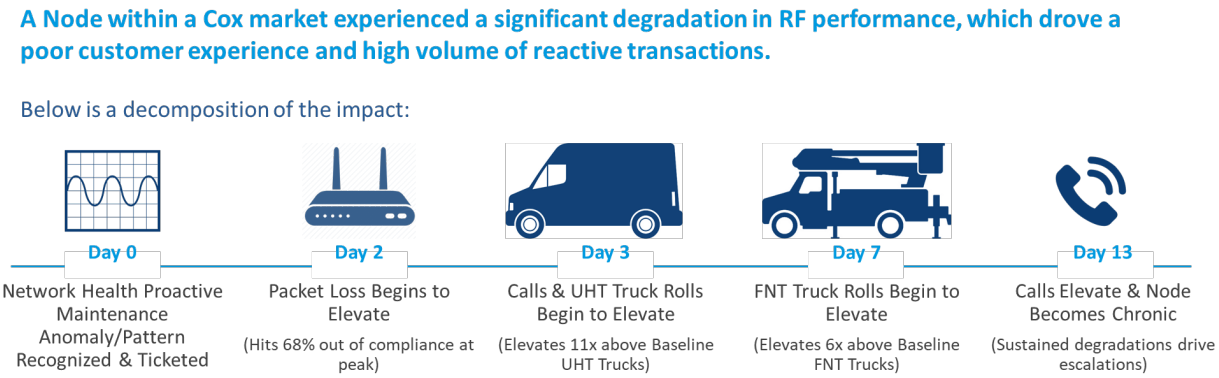


Figure 5 – Decomposition of the Impact of a Degradation RF performance Node

5.3. Problem Statement

Data shows a strong correlation between the amount of labor devoted to proactive maintenance event and the overall improvement in the health of the network (see diagram below, which depicts the correlation between proactive maintenance and network health). When network health is low, we see more outages, which leads to additional reactive work. However, the key challenge is how to prioritize this maintenance activity in a way that it can prevent the most transactions and can best improve customer experience. There are more proactive impairments to troubleshoot than there are technician labor hours in each week/month. As a result, much of the existing reactive work, both multi-customer outages and single-customer escalations, could be prevented if Cox successfully predicts which ***X*** impairments were the most critical and demanded work within ***Y*** days/hours.

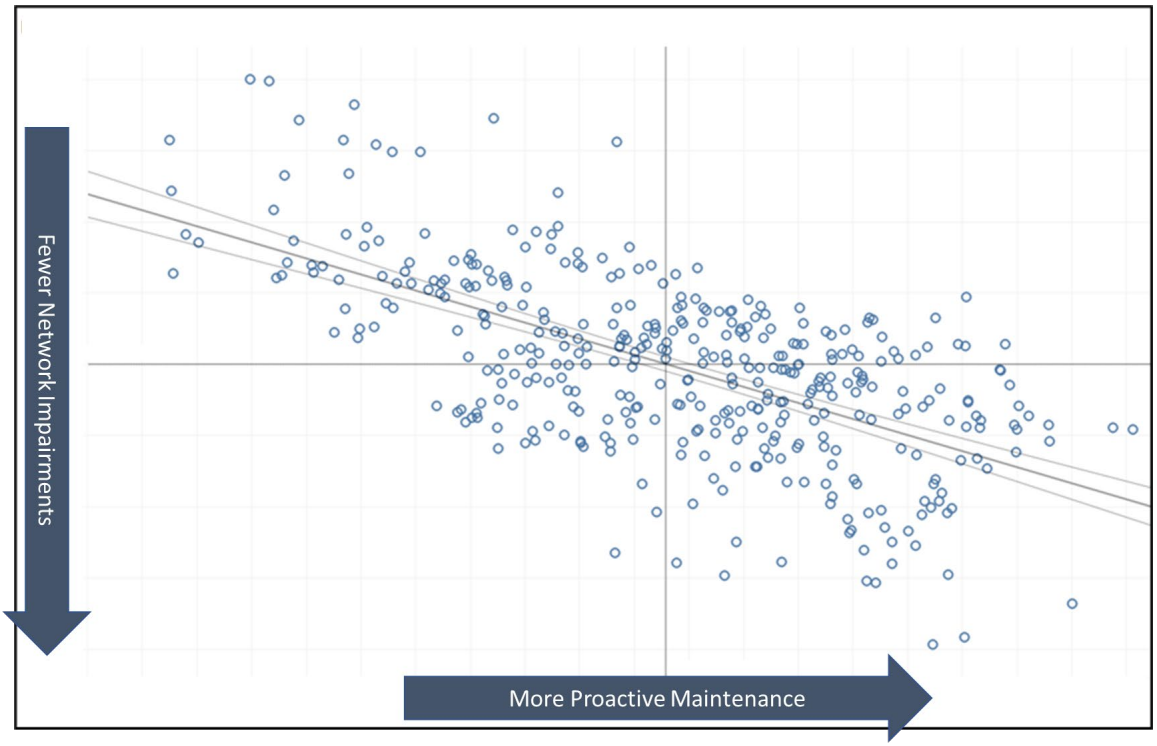


Figure 6 – Correlation Study betweeen Proactive Maintenance and Node Health

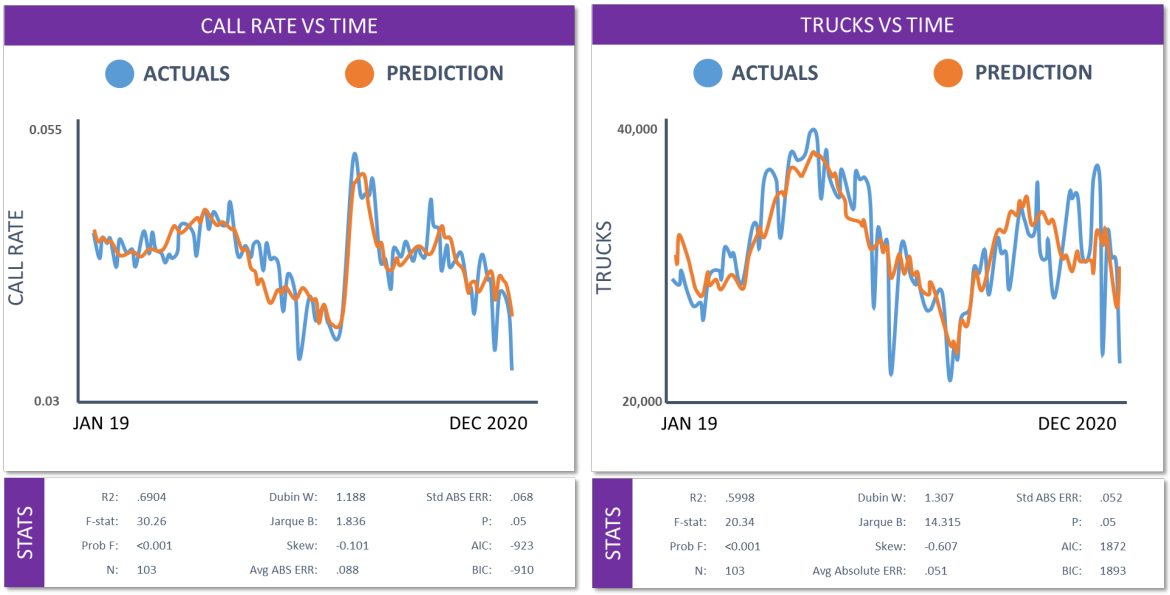


Figure 7 – Correlation with Node Health and Transactions

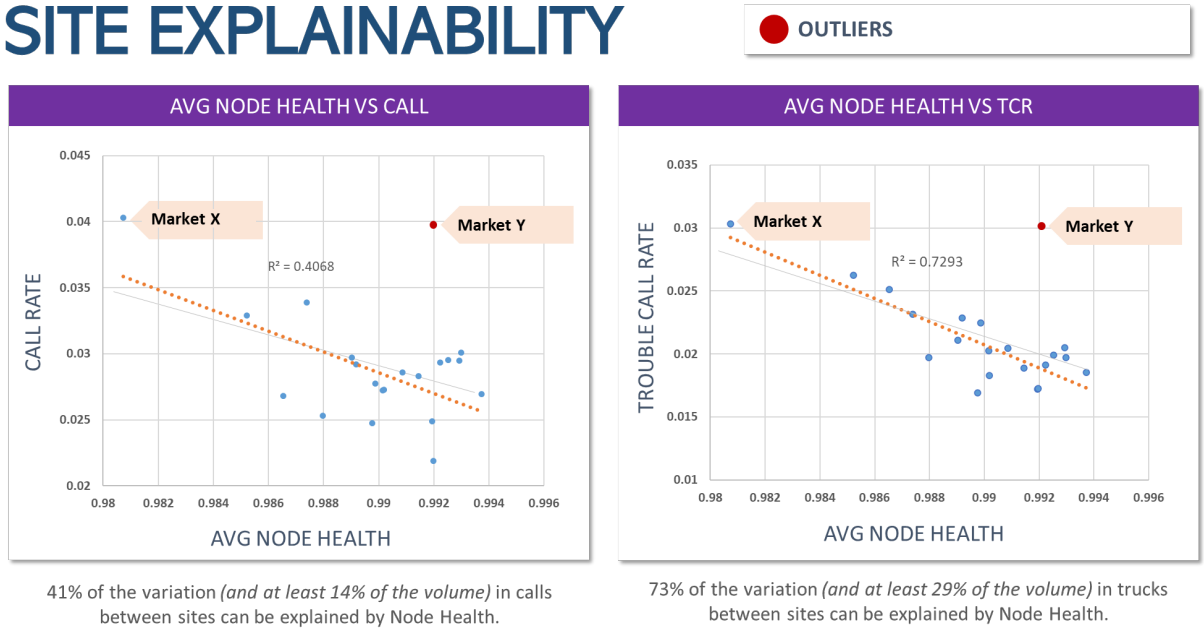


Figure 8 – Relationship between Node Health and market-level transaction level/variation

5.4. Solution Proposal

Cox uses a multitude of network and customer premise equipment (CPE) telemetry when identifying an issue. Specifically, they use the combination of sources highlighted in the diagram above, which retrieves the RF (Radio Frequency Signal data) and impairment metrics for multiple customer and network devices. By leveraging RF signal data and impairment data, devices with statistically similar patterns can be grouped together to act as an indicator of an issue on the network. Also, by leveraging real-time signal level and the frequency of occurrence of rapidly changing patterns, we can predict a time of day when the issue is present, which allows a technician to isolate the issue and find regular/irregular patterns in signal.

As a part of the capacity model, Cox also predicts the daily volume of reactive tickets to identify the volume of available labor hours/technicians grouped by their skills and schedules. Specifically, this model allows the Field Network Technician (FNT) team to receive a volume of proactive network maintenance work that meet the available capacity.

Finally, Node Health RF Performance ensemble models proactively creates a prioritized list of events/tickets that identify nodes with recurring/intermittent issue or continuous issue. This list of events/tickets is then routed based on priority and available technician capacity, which is performed daily (note: this process can run every 15 minutes but is currently running daily in accordance with existing processes). As a result, we can recast the diagram above with the following remediation and response:

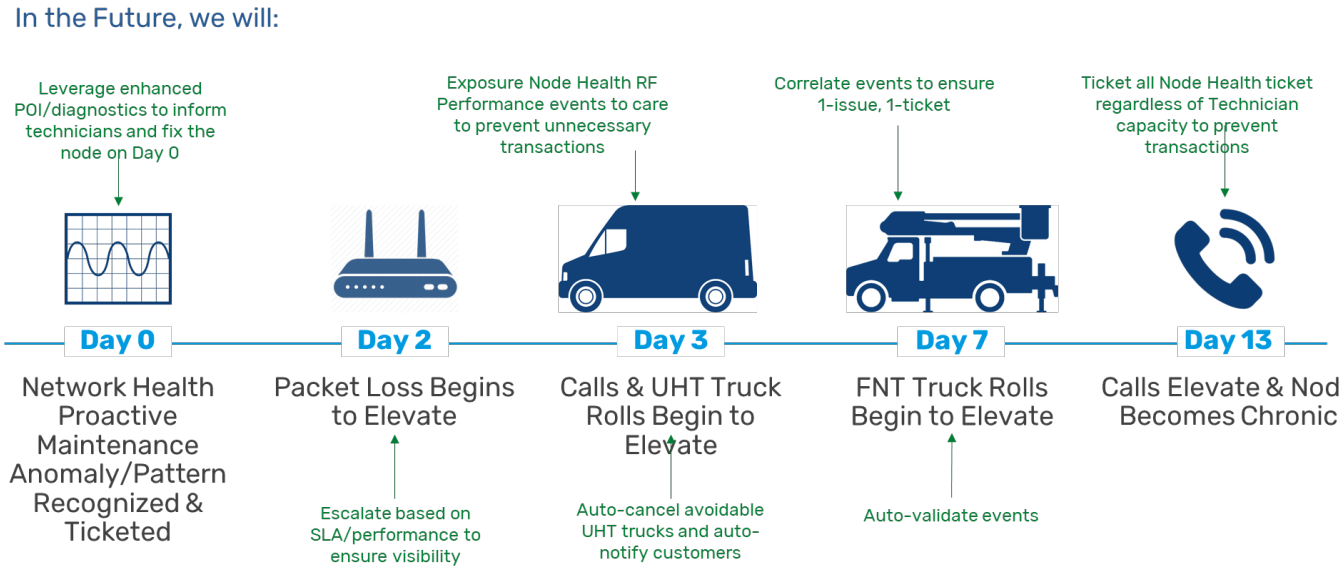


Figure 9 – Model of improved customer interactions post-Node Health models

5.5. Conclusion and Business Impact

The Node Health RF Performance application has been a significant driver in reducing the number of service-affecting events, multi-customer disruptions, and single-customer escalations. This has reduced unnecessary transactions, prevented customer service disruptions, and improved customer experience. Additionally, these capabilities have not only made technicians more efficient and effective, but they have also enabled next-generation communication with customers, allowing Cox to inform on potential issues and preemptively perform maintenance on parts of the network that are predicted to degrade. As a result of these ensemble models, Cox has experienced a ***5%*** reduction to Tech Support calls and in-home technician Service Call Truck Rolls.

6. Putting the Model into Action: AWS Cloud Architecture

The ACOE team used Amazon Web Services (AWS) as the primary cloud vendor for hosting this new technology. The following architecture diagram shows, generally, the services used by the Node Health RF Performance program.

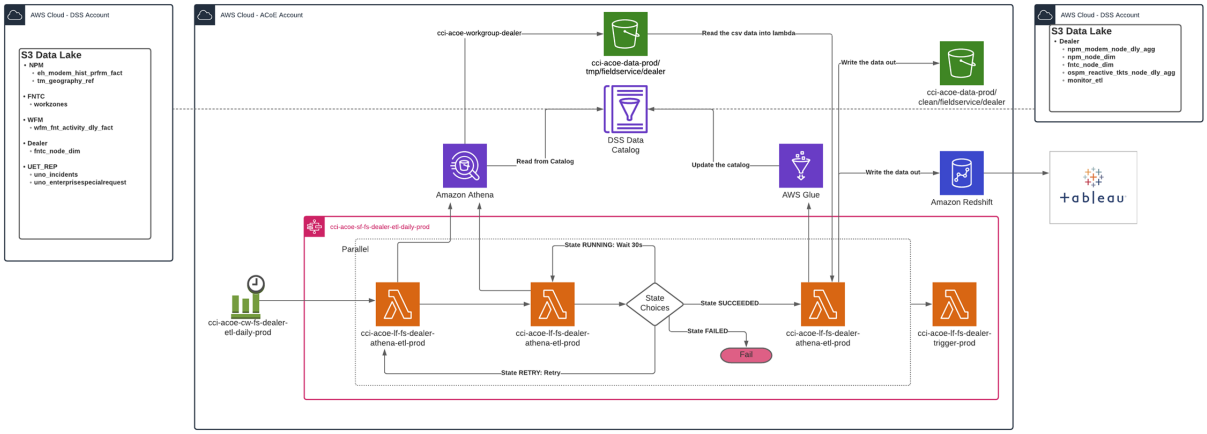


Figure 10 – Cloud-based Node Health architecture

When selecting to use a cloud partner, the ACOE had five primary considerations:

1. Scale: How do we scale our infrastructure easily when new projects/programs are approved?
2. Rapid Development: How do we get from ideation to production more quickly?
3. Next-Generation Capabilities: What platform gives Cox's ACOE capabilities that are multiple generations beyond existing tools?
4. Talent: What platform allows Cox's AOCE team to recruit and retain the best talent in the marketplace?
5. Future-Proof: What platform hardens our environment to ensure it is not obsolete in 2-5 years?

Moreover, the above architecture, which leverages Lamda, Athena, S3, Redshift, and Glue were selected for the following reasons:

- The overall cost of this pattern was the most cost effective model at the time
    - We don't need always-on database service since Node Health runs in batches periodically during the day rather than all of the time
    - The primary data sources were already in an S3 data lake
    - The tables are partitioned well such that we don't incur unnecessary cost when using the Athena service which charges by the GB scanned
- This pattern does not require ongoing maintenance of a database cluster
- The skills and experience required to create and manage Lambda functions are better aligned with the skills of our Data Scientists and Analysts over management of an EC2 instance.
- The Dealer application ETLs run (at the most frequent) once per hour and we did not want to incur the added time and money cost of turning up and down an EMR cluster (or EC2 Instance) or leaving one active
- Athena is a very fast service for its cost

7. Further Discussion

While this white paper focused on components of Node Health within the Cox ACOE Service Health framework, in future sessions we can focus on additional components of both Node Health and Service

Health. Areas of further discussion on Node Health that may be of interest and that are only briefly touched on in this whitepaper include:

- **Advanced event correlation**: How to improve grouping and correlation of related events based on predictive models and real-time root-cause analysis (RCA)
- **Point of interest (POI)**: How to predict potential drivers of problems within the node to reduce technician time to restore services and prevent repeat/unnecessary transactions
- **Outage detection**: How to reduce non-actionable events (i.e. often coded off as no-problem found/cleared prior), improve detection of true outages, and reduce the time to restore customers services using predictive analytics
- **Transaction cancelation**: How to cancel and prevent in-flight transactions using multiple models making predictions about customers' services and creating a dynamic interaction with the customer through SMS messaging
- **Transaction deflection**: How to improve predictions about service issues impacting customers and enrich the interaction through multiple channels to inform and educate the customer regarding restore/repair

These additional five projects augment the Node Health RF Performance project and add additional features that further improve the health of the network, reduce operating expense, and improve customer experience. Notably, these additional programs will reduce Tech Support calls and Service Call Truck Rolls (i.e. in-home technicians) by an incremental ***5%-10%***.

Abbreviations

RX	the receive power level to customers' modems
TX	the transmit power level from customers' modems
FEC	forward error correction
SNR	signal to noise ratio
US	upstream network performance
DS	downstream network performance
SCTE	Society of Cable Telecommunications Engineers
AWS	Amazon Web Services
POI	point of interest
ACOE	Analytics Center of Excellence
GB	gigabit
S3	AWS Simple Storage Services
EMR	AWS Elastic MapReduce
ETL	extract transform load (data transformation technique)
FNT	field network technician
CPE	customer premise equipment
RF	radio frequency
HFC	hybrid fiber-coax
OSP7	outside plant seven metrics system
UHT	universal home technician
KPI	Key Performance Indicator

OFDMA Predistortion Coefficient and OFDM Channel Estimation Decoding and Analysis

Remove the Linear Delay, Examine the Group Delay

A Technical Paper prepared for SCTE by

Tom Williams
Distinguished Technologist
CableLabs
858 Coal Creek Circle, Louisville, CO 80027
303.661.9100
t.williams@cablelabs.com

Alberto Campos
Fellow
CableLabs
858 Coal Creek Circle, Louisville, CO 80027
303.661.9100
a.campos@cablelabs.com

Lin Cheng, CableLabs

James Lin, CableLabs

Jason Rupe, CableLabs

Jay Zhu, CableLabs

1. Introduction

Upstream DOCSIS® carriers use predistortion data to cancel linear distortion. Equalization is a required step in a digital transmission system to remove linear distortion. In continuous DOCSIS downstream transmissions the equalization is done in the cable modem (CM) receiver, but in DOCSIS upstream the burst transmissions are pre-distorted by the CM so that they arrive at the cable modem termination system (CMTS) receiver with linear distortion removed. The predistortion is adjusted during a periodic process called ranging, which occurs about every 20 to 30 seconds.

An analogy for predistortion is corrective eyeglasses, which custom pre-distort an image so that it arrives on a wearer's retina in focus. An optometrist can measure lenses and determine what visual impairments are being corrected, such as myopia or astigmatism. Likewise, upstream predistortion can be analyzed to determine what linear distortion impairments are being cancelled, such as echoes, group delay, tilt, suck-outs, filters, etc.

DOCSIS 3.0 single carrier upstream equalization data has been mined for over 10 years to provide troubleshooting information about plant impairments. Now both DOCSIS 3.1 upstream orthogonal frequency division multiple access (OFDMA) multicarrier equalization and downstream orthogonal frequency division multiplexing (OFDM) multicarrier channel estimation data are available from CMs using management information base (MIB) objects. OFDMA is being used mostly in mid- and high-split cable plant because of spectrum availability. OFDM and OFDMA multicarrier data should be more valuable than single frequency carrier data because of a much wider potential bandwidth relative to narrowband DOCSIS 3.0 single carrier coefficients. This wider bandwidth provides high resolution time responses, enabling cable problems to be identified with high accuracy.

Two application categories have been used to date: examination of a single response for impairments and comparing multiple responses to form groups with common impairments.

OFDM and OFDMA data obtained to date needs to be pre-processed before analysis due to the addition of delay and phase from the CM-CMTS interaction during ranging. This paper discusses the processing steps, provides field and lab data, and discusses new applications which are now feasible.

Just as full band capture (FBC) came to be described as a "free spectrum analyzer in every CM," OFDM and OFMDA coefficients can be viewed as a "free pair of network analyzers in every CM." The OFDM & OFDMA equalization data are compelling because there are many millions of deployed DOCSIS 3.1 CMs, and the data have both high time resolution and large dynamic range.

2. Background

As mentioned above, predistortion analysis has been used operationally on cable upstream single-carrier transmissions for a while, and readers are directed to the bibliography for publications on this common operational practice, done under the CableLabs® umbrella of proactive network maintenance (PNM). DOCSIS version 1.1 upstream used an 8-tap finite impulse response (FIR) filter which was upgraded to a 24-tap FIR filter for DOCSIS upstream versions 2.0 and 3.0.

There is also a possibility of reading coefficients for downstream single carrier signals which are modulated with 64 or 256 QAM. Decision feedback equalization (DFE) is used for downstream equalization. Unfortunately, this practice has not been highly successful due to the large number of configurations of feed forward and feedback taps, plus inconsistent coefficient reporting. This was caused in part by an incomplete and confusing MIB description.

A common limitation for both upstream and downstream single carrier signals has been the relatively narrow bandwidth, which is 6.0 MHz in the USA for downstream and frequently 6.4 MHz in the upstream. The narrow analysis bandwidth results in a limited time accuracy resolution. The limited time accuracy implies limited distance accuracy, resulting in an imprecise location of plant damage. This limitation makes locating a point of damage on cable plant harder to identify, which can be particularly annoying when locating damage in buried cable.

In DOCSIS 3.1 specifications, both OFDM and OFDMA were standardized. Pre-distortion for upstream OFDMA carriers was continued, although made optional. On both modulation types, MIBs were specified to reveal the frequency domain (FD) complex coefficients associated with each of the potentially thousands of active subcarriers. A MIB object request is sent to the CM, and the CM performs a trivial file transfer protocol (TFTP) transfer of a file containing the requested in-phase (I) and quadrature (Q) coefficients. Note that the upstream coefficients returned are for the corrections to remove impairments, not for the impairments themselves. The two responses are inverses of each other. In the downstream, the CM can report channel estimation coefficients for OFDM, which are also complex coefficients.

With both upstream and downstream single carrier (SC) signals, the returned complex coefficients are in the time domain (TD). Alternately, the returned coefficients in the multi-carrier (MC) systems, both OFDM and OFDMA, are FD coefficients. The discrete fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) may be used to convert between time and frequency domains.

3. Past and New Applications for Equalization Data

In the past, two main application categories were developed from analysis of the SC coefficients. The first was a stand-alone path analysis. At a CM this analysis revealed echo tunnels created by two or more impedance mismatches along with other problems such as excessive (GD) group delay, suck-outs, peaking, and tilt. The second application was a matching path analysis, indicating that a group of CMs share a common impairment. This information, along with a strand map of the plant, reveals likely locations of plant damage. Figure 1 illustrates upstream single carrier groupings done within a node, revealing homes with common impairments, which are color coded on the map and spectral plots.

Figure 1 - Upstream SC groupings showing common plant impairments.

Applications for equalization data previously mentioned are single response analysis for impairments and matching common impaired responses to assist in the location of line damage. This second application can be expanded to a third application where matching nonimpaired (normal) responses are compared to determine if subscribers are neighbors sharing a common cable line. This can be done to validate records, or as a source of information for locating CMs in the network when location data are not available.

Another application for equalization data is for plant stability analysis. Intermittent connections will cause re-adaption of coefficients to the new echoes. Connections may change infrequently, such as daily in the case of thermal-caused intermittent problems, or several times per second if intermittent connections are experiencing wind or traffic vibrations. Comparing responses with earlier stored coefficients reveals the presence of intermittent connections. If the changes are large, or frequent, the problem may be severe.

Equalization can also be affected by interference, such as band-limited ingress LTE signals or wider bandwidth interference such as common path distortion (CPD) interference or passive inter-modulation (PIM). An examination of the "smoothness" of an equalization solution can reveal noise causing a poor equalization setting. These data can be complementary to MER-per-subcarrier data available on both OFDM and OFDMA. Figure 2 shows a lab equalization response with a poor equalization solution in an LTE band.

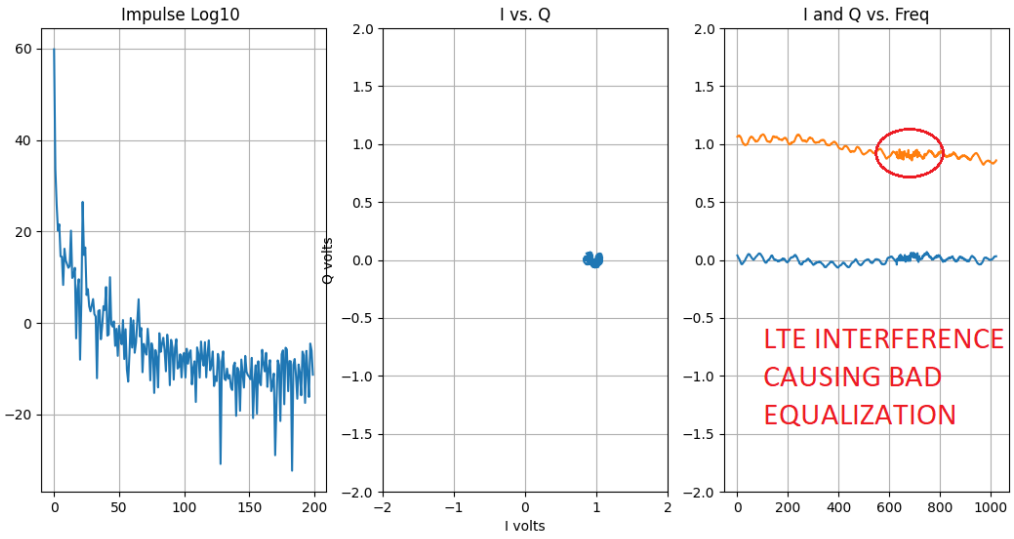


Figure 2 - LTE interference causing poor equalization shows up in the right pre-equalization plot.

Another application is determining if a selected cyclic prefix (CP) is sufficiently long to accommodate the longest significant echoes that are being corrected. Figure 3 shows an example of the strength of echoes vs. delay time. The longer a signal travels in a cable, the greater the attenuation. Knowing the cable type allows this attenuation vs. time curve to be corrected down to near the noise floor. Another application is determining if the programmed cyclic prefix length is long enough for the longest expected echo, which is shown in green in Figure 3. Any echo longer than the CP must be sufficiently weak to cause no harm.

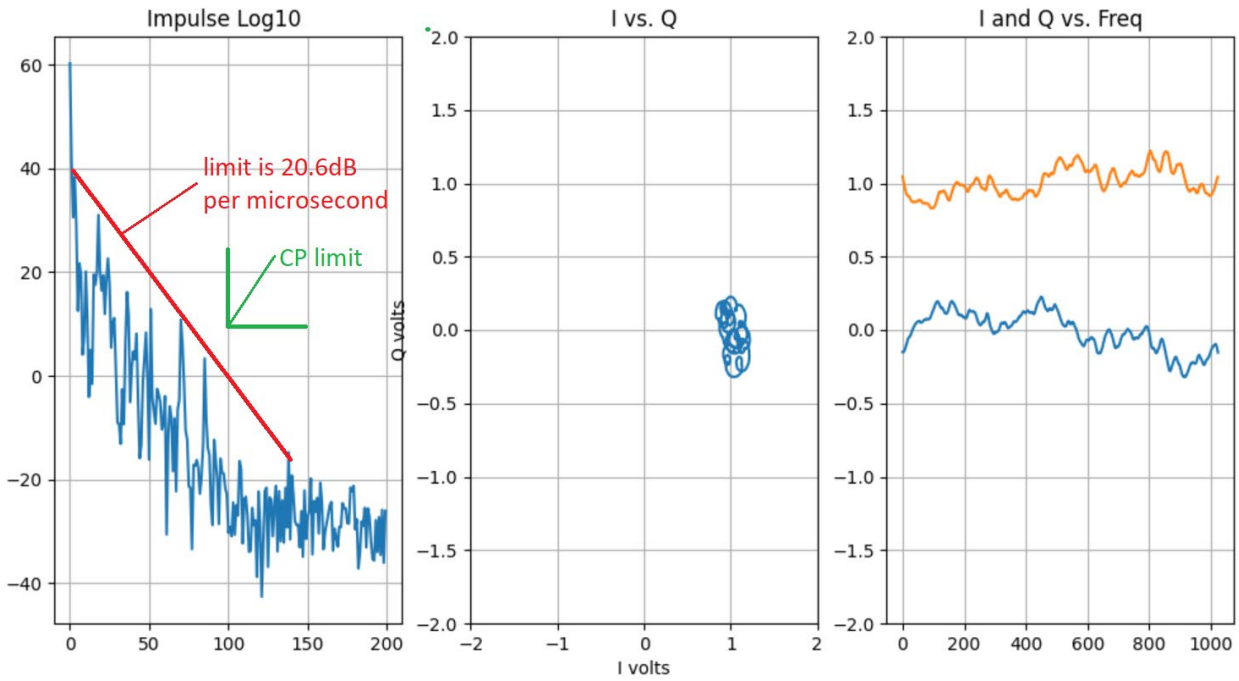


Figure 3 - See the impulse response on the left with the added red limit of 20.6 dB per microsecond, showing no unexpected echo tunnels as peaks above the line. Also no echoes appear over the green corner threshold, indicating sufficient CP is being allocated.

4. Pre-Processing OFDM and OFDMA Coefficients

Presented next are the processing steps for correcting raw (direct from CM) phase so we can make use of it for the described use cases. The basic idea to correct a frequency response is to remove added uniform time delay, which manifests itself in the frequency domain as phase rotating linearly with frequency. We assume that there is a direct current (DC) component to the frequency response, and it can be used as a single "pilot" subcarrier. In this method, there is a coarse delay removal step done in the frequency domain followed by a phase rotation done in the time domain. The phase rotation is done to make the DC term real-only, so the imaginary component of the DC term is forced to zero. In other words, the DC "pilot" is calibrated. With this assumption, the remaining signal reflects the plant condition, which allows us to process these data for our described use cases.

At a high level, the steps are as follows.

1. Import the complex FD data.
2. Estimate linear delay ($d\varphi/d\omega$) – calculate ω using arc tangent to get phase angle.
3. Coarse delay removal – Flatten the angle vs. frequency plot for the middle portion of the band. Using only the middle portion of the band prevents group delay (GD) from effecting the calculation.
4. Convert corrected response to time domain using IFFT, and measure phase on the DC term.
5. Correct phase on all time samples so Im[0]=0 by a time-domain rotation
6. View response in FD by a FFT calculation.
7. Optionally normalize coefficients. For example, make the total power in all coefficients equal to unity, or make the DC term 1.0.

8. (Optional) compare with another response using FD division

The next few figures show the transformation of the processing steps. Appendix 1 contains C++ computer code to remove the undesired (random, linear) time delay and Python code to plot the results as a set of 3 graphs per CM.

As mentioned above, I and Q coefficients for both upstream OFDMA and downstream OFDM arrive raw from a CM with a random time delay inserted that must be removed before processing the coefficients for estimating the channel-plant. In the TD, a delay manifests itself as the main tap being moved from its desired position, which is ideally the 0th or DC term. In the FD, the time delay appears as a constant rotation of phase angle with frequency. Figure 4 illustrates a raw downstream FD polar response with rotation, and Figure 5 illustrates that same response as I and Q plots vs. frequency.

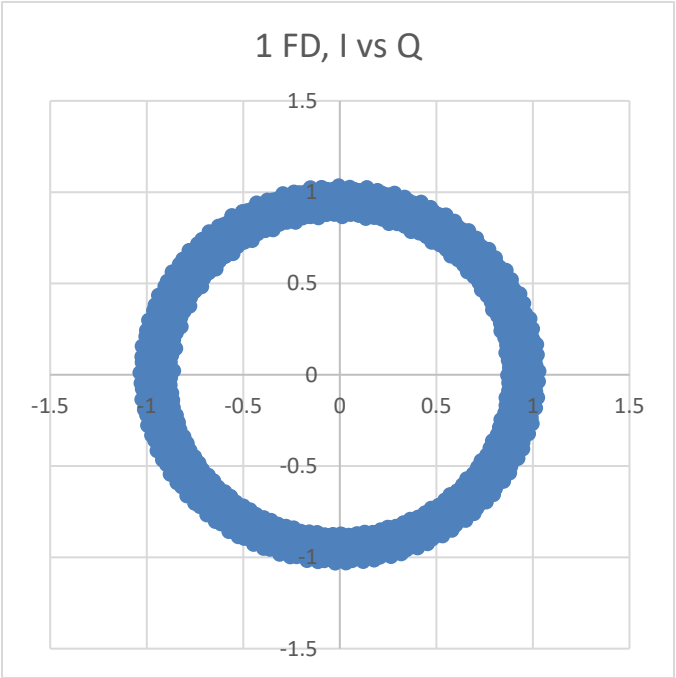


Figure 4 - Raw downstream equalization coefficients, no echo or other impairment.

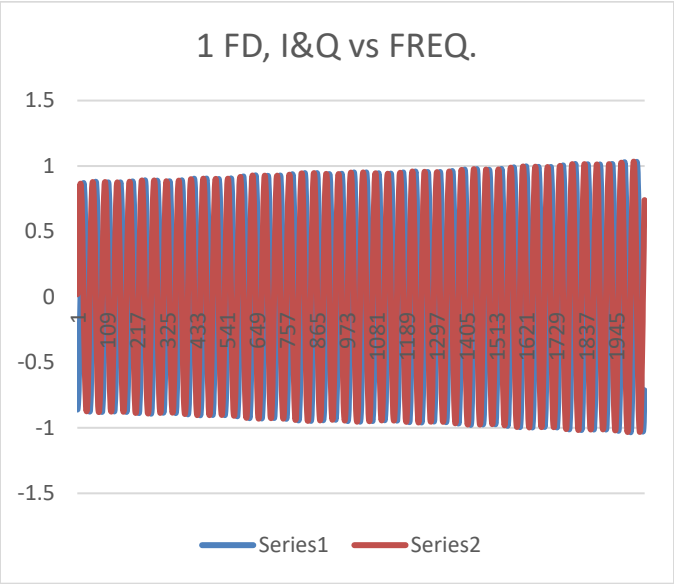


Figure 5 - Raw downstream equalization coefficients versus frequency, no echo or other impairment.

To heuristically explain and demonstrate the removal of delay in the frequency domain, an Excel spreadsheet program was made with two slide controls. It is available from CableLabs. The first slide bars adds or subtracts linear delay, and the second slide bar adjusts the phase angle. The pre-equalization data in the example was obtained from upstream lab testing.
The slide bars operate in the frequency domain. One slide bar changes the rate of change of phase vs. frequency:

$$\Theta = \Theta_0 + k_1\omega$$

where Θ is the new phase angle in radians for every frequency, Θ_0 is the starting phase, and ω in the angular rate of change in radians per second. The value of k_1 is determined by a first slide bar position. Magnitude values for each frequency are not altered; only the phase angle is adjusted.

The second slide bar rotates the response phase equally for all frequencies:

$$\Theta = \Theta_0 + k_2$$

The value of k_2 is determined by a second slide bar position.

See Figure 6 through Figure 12. Fig. 6 is a raw I-Q upstream response in FD polar form, having approximately 450 degrees of unwanted rotation before delay removal. Figure 7 is a plot of angle versus frequency associated with Figure 6. The angle in radians is obtained by taking the arctangent of the Q coefficient value divided by the I coefficient value. The object of the slide processing is to remove delay by adjusting the phase response to be flat in the middle of the spectrum, and then adjust phase angle of the DC term in the time domain to be real-only with an imaginary value of zero. Figure 8 is the angle versus frequency plot with delay (rotation in FD) removed. Observe that the middle of the band has been flattened, but there is still movement at the low and high frequencies. This is caused by group delay associated with upstream high-pass filtering on the low end, and upstream low-pass filtering on the high end.

Figure 9 is an I-Q polar plot showing the removal of rotation. Observe that the plot still has some residual rotation due to group delay, which is not uniform with frequency.

Figure 10 is a plot of group delay vs. frequency. Group delay is the derivative of the slope of phase vs. frequency.

$$GD = -\frac{d\varphi}{d\omega}$$

If the frequency steps are 50 kHz, a group delay value of 0.035 from the graph would be 111ns = 0.035/(2*π*50,000).

Figure 11 is corrected real and imaginary responses vs. frequency. Finally, Figure 12 shows the TD impulse response which is obtained by taking the IFFT of the coefficients in Figure 10. Phase angle has been adjusted to produce a maximum DC real value and zero DC imaginary value.

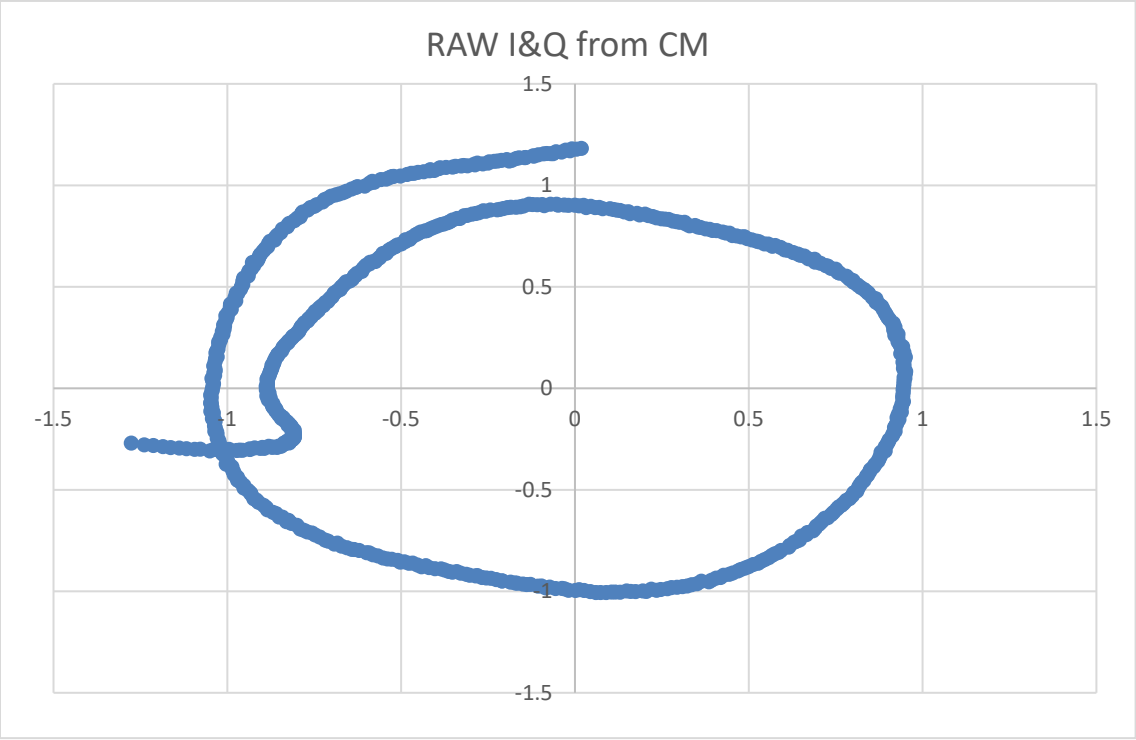


Figure 6 - In-phase (I) versus quadrature (Q) coefficients as supplied by a CM, contaminated with a time delay, which manifests itself as a uniform rate of phase rotation with frequency.

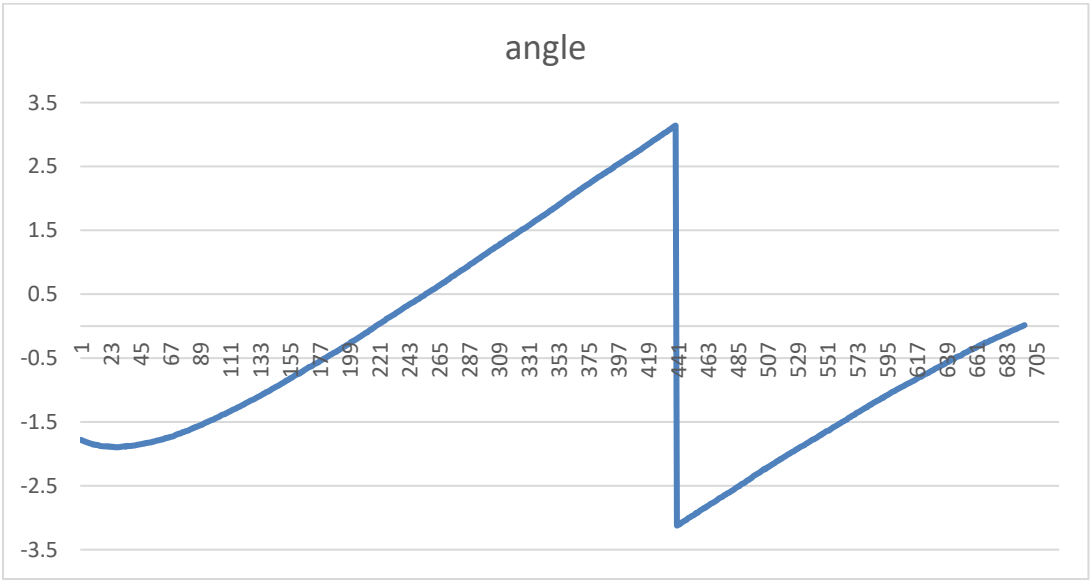


Figure 7 - Angle plot of the data from Figure 6. The rate of change of angle (in radians) versus frequency is delay.

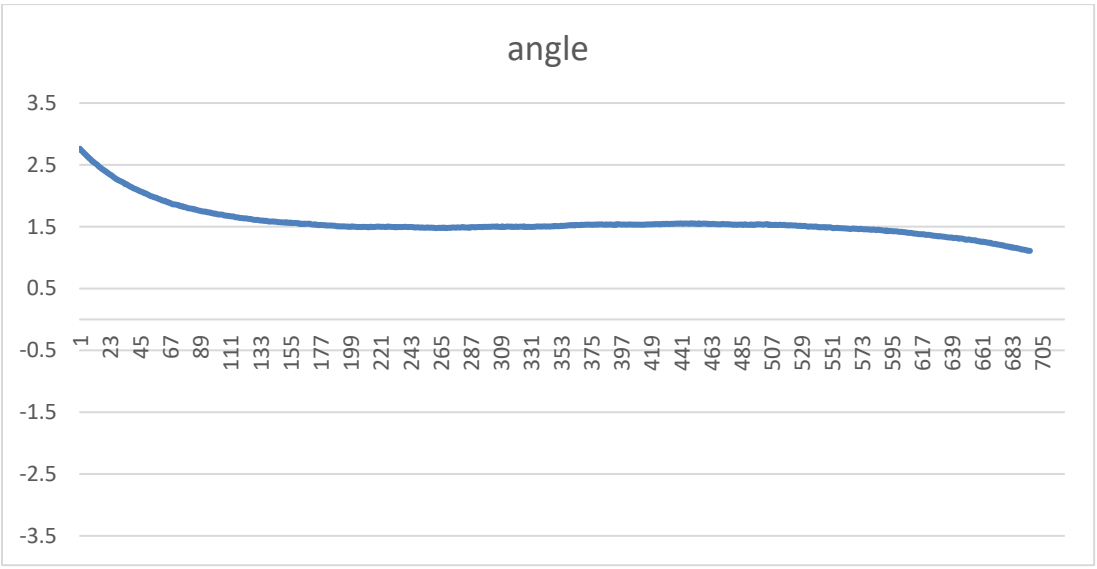


Figure 8 - Phase angle of I and Q coefficients after delay removal. Phase is adjusted to be flat in spectral band without excessive group delay.

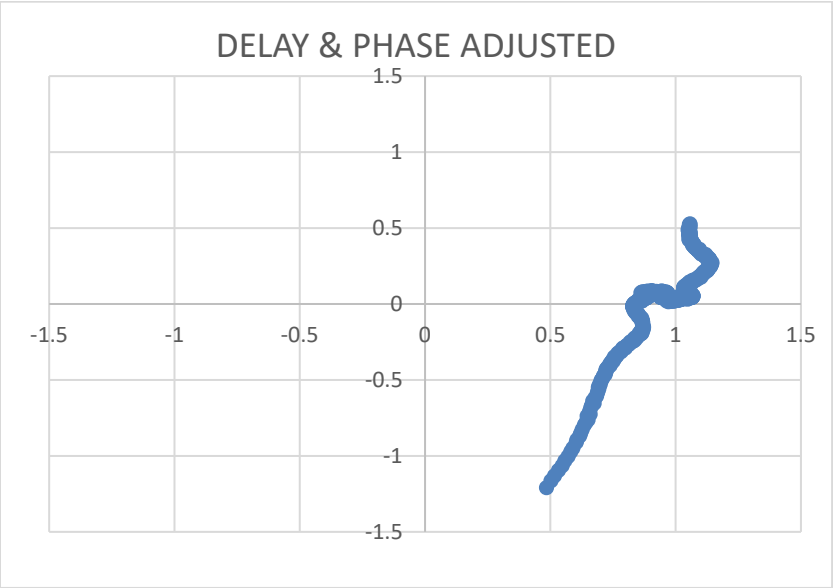


Figure 9 - I and Q coefficients as a polar plot, after delay removal.

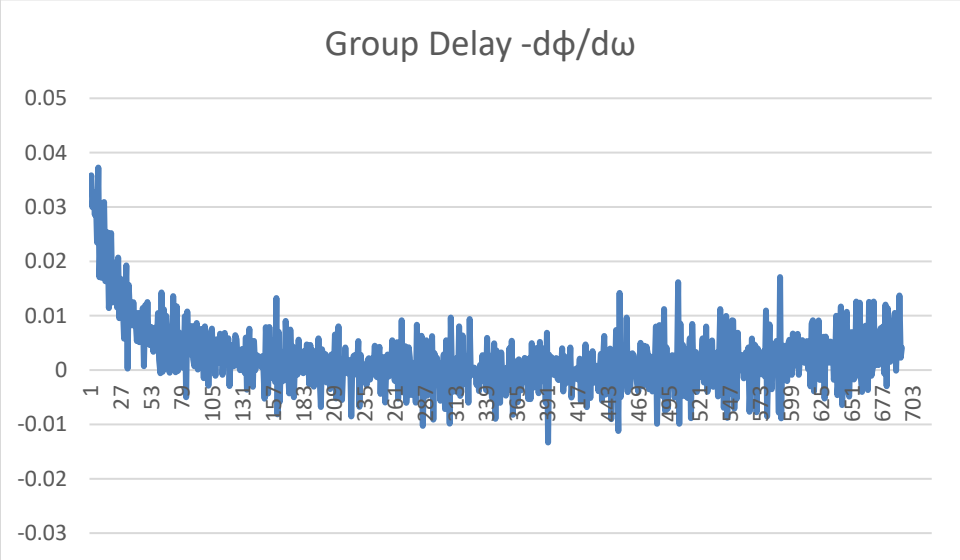


Figure 10 - Group delay, created by taking the derivative of the phase response with respect to frequency. Group delay is a "bending" of the phase response, and is caused by a high-pass filter at 5 MHz.

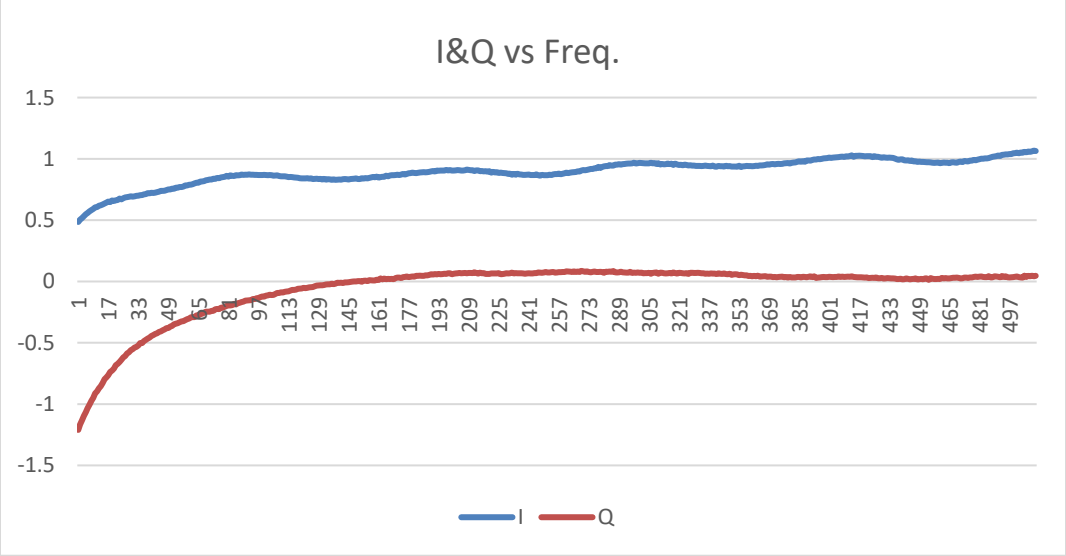


Figure 11 - De-rotated I and Q values.

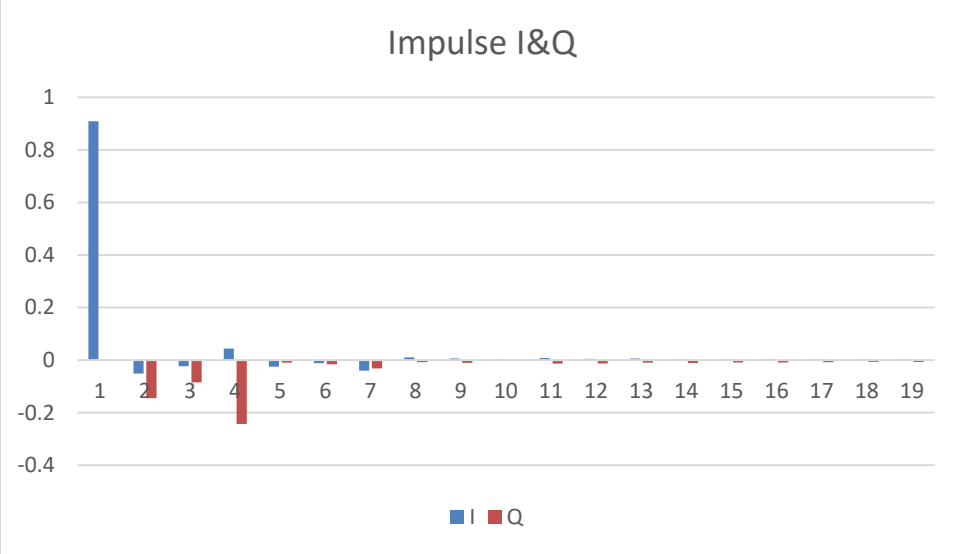


Figure 12 - Impulse (time domain) response that was obtained by taking the IFFT of the re-rotated response. It shows a maximum DC real value (I) and zero DC imaginary (Q) value.

Because of group delay at the low end of the frequency band, only the middle portion of the band was flattened. As mentioned above, computer code has been written to automate delay removal and angle adjustment with a target of maximum energy in the real DC coefficient and zero energy in the DC imaginary coefficient. The code for executing that process is in Appendix 1 of this paper.

A question arises: why is there a time delay (same as FD rotation) addition added to the frequency domain coefficients? A best guess is that delay is an adjustment method that allows an optimal selection of time domain samples on which to perform an FFT. Symbols selected for FFT processing come from an OFDM frame preceded by a CP (or cyclic extension, or guard interval). Lab vector network analyzers (VNA) also add delay to responses as a result delay from connecting cables.

5. Downstream Lab Results

In the lab, tests were performed to compare transmission (S21) on a Keysight VNA with a CM's downstream equalization coefficients after passing through a created echo. 1800 complex coefficients were analyzed with a subcarrier spacing of 50 kHz, starting at 612 MHz.

A plot of raw VNA downstream I and Q values with an echo and accompanying delay is illustrated in Figure 13. The same delay-corrected data are shown in Figure 14, which illustrates a downstream single-recursion echo created with a pair of splitters and a long 100' cable with a 12dB attenuator, and a short 1' cable with no attenuator (see Appendix 2 on types of echoes). These data are shown as I and Q coefficients plotted versus frequency. On the left of Figure 14 is a response from the VNA and on the right is a response from a CM. Figure 15 shows a time plot showing the magnitude impulse response after both responses were transformed with IFFTs. Note that the delayed echo was about 16.5 dB due to an additional 4.5 dB of cable loss.

An interesting VNA test result is that the coefficients were produced by the instrument with an accompanying electrical delay, which needed to be removed for analysis and comparison. The delay in S21 in Figure 13 was caused by cable length.

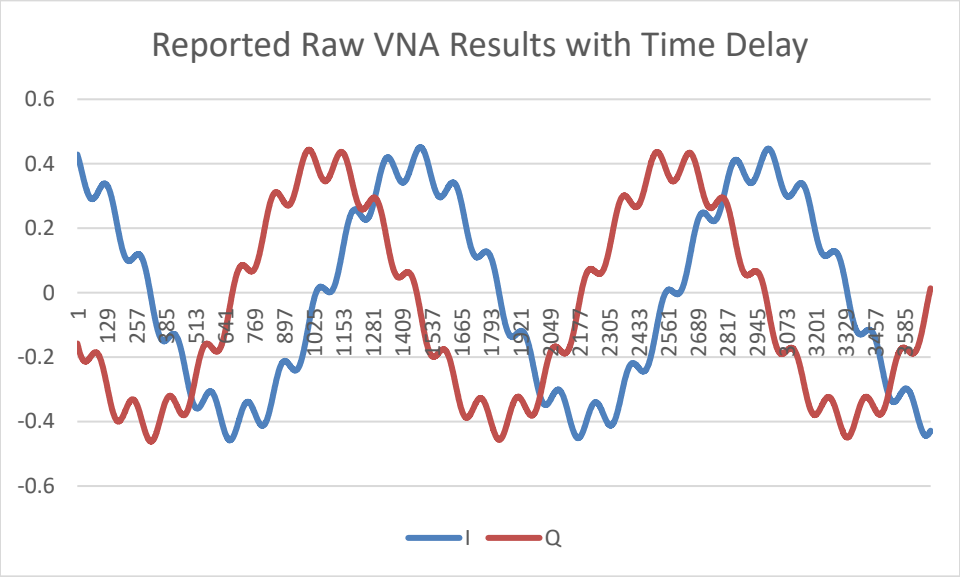


Figure 13 - Raw VNA data showing echo (fine ripple) and time delay (large coarse ripple).

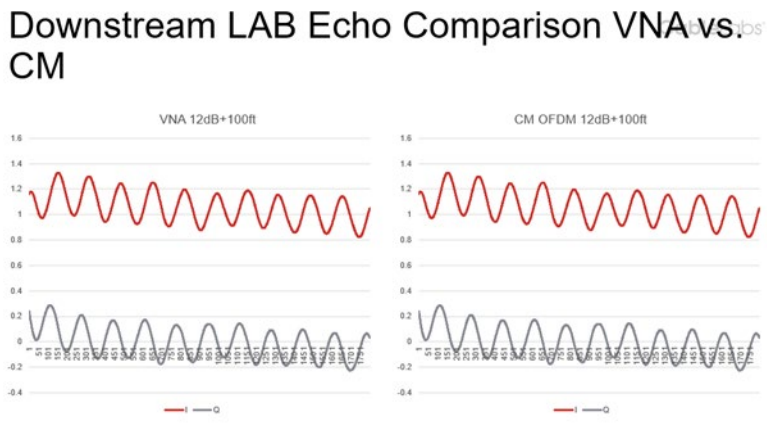


Figure 14 - A downstream single-recursion echo created with a pair of splitters and a long 100' cable with a 12dB attenuator, and a short 1' cable with no attenuator, as I and Q coefficients plotted versus frequency.

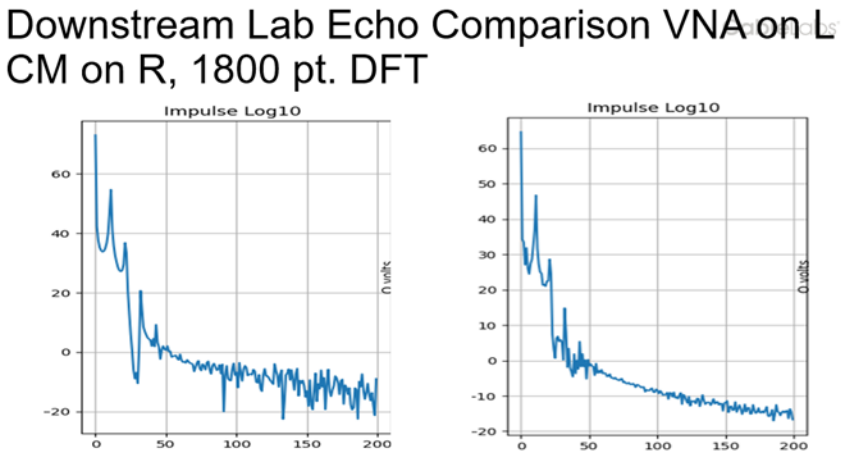


Figure 15 - Time plots showing the magnitude impulse responses of the VNA (left) and CM (right) after both responses were transformed with IFFTs.

Another experiment that was done in the lab was to measure reflection (S11) using a CM and then compare test results with those obtained by the VNA. See Figure 16 for test plant wiring. It illustrates a fiber node feeding signals to a string of taps. The output cable of a fiber node was tapped with a high impedance -20dB probe, and a cable modem (CM METER) connected to the probe was allowed to range and register on a downstream OFDM channel. The standing wave of the OFDM received signal was captured in the complex frequency domain as coefficients, and then inverse Fourier transformed. The red lines show one path the reflections can take to make a standing wave. The green graph in Figure 16 shows an expected impulse response with tap reflections. There should have been an impulse response returned from each tap, and possibly an indication of plant damage, if there was any. Figure 17 shows the OFDM's transformed coefficients (impulse response) from the lab string of taps, and Figure 18 shows the impulse response from a Keysight vector network analyzer. Other than the VNA having more dynamic range than a CM, the results are comparable.

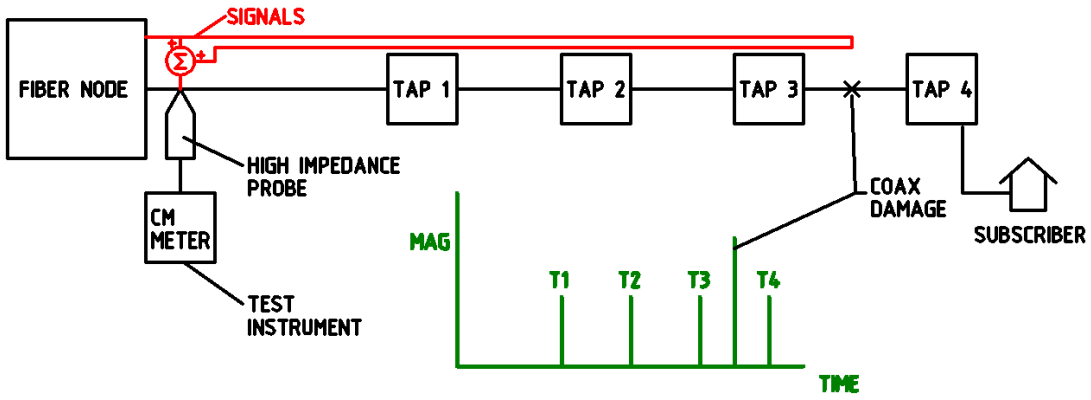


Figure 16 - Test wiring for the second experiment.

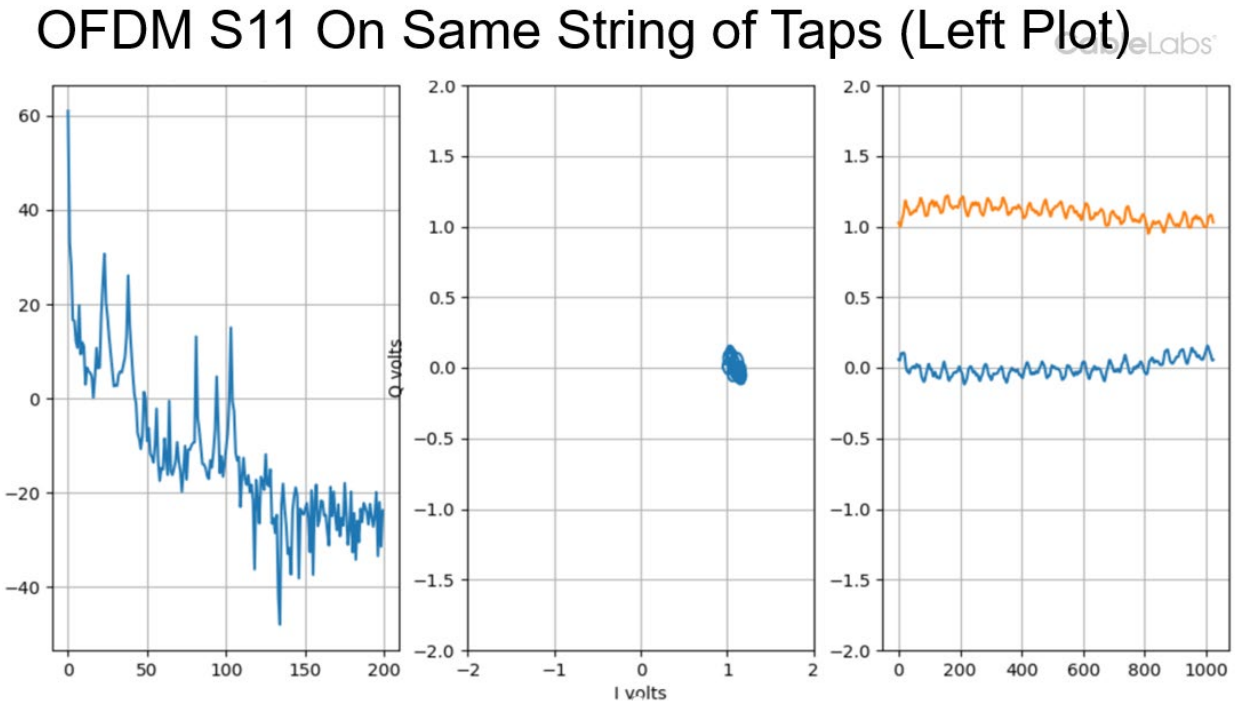


Figure 17 - Impulse response from the lab string of taps in Figure 16, obtained using a CM.

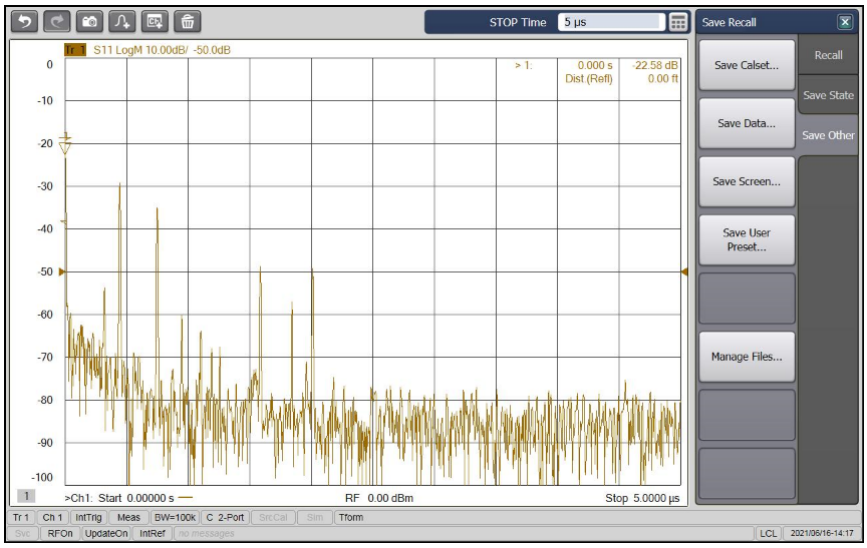


Figure 18 - Impulse response from the lab string of taps in Figure 16, obtained using a Keysight VNA.

Generally, the downstream responses contain much more uniform delay relative to upstream responses.

Lab testing has revealed that if a CM is rebooted, the delay value in the coefficients will probably change to a new value. Likewise, a different CM will likely obtain new delay values when it comes online.

Another observation that has been made is that if there is an exclusion band programmed into the OFDM(A) spectrum, the delay values "jump" over the vacant band. That is, different delay values are applied on occupied bands on either side of an exclusion band. Authors are not sure why. But this does complicate the correction work.

A lab confirmation was that, while OFDMA coefficients are commonly reported as the applied correction (demonstrated in the next section), CMs report OFDM coefficients as the channel's actual response, as expected from the specifications. The channel's correction and its response are frequency domain reciprocals of each other.

The horizontal axes of spectral data are index units, which for these tests were 50 kHz frequency steps.

On the horizontal axes of impulse responses, the units are also indices, which can be converted to time. For example, assume the downstream FFT size is 4096 and symbol rate is 204.8 Mega symbols per second. That results in a symbol period of 4.88 ns per symbol. OFDM frame time is 20 µsec. But because 4096 coefficients are not available, the FFT transform size used is reduced to 1024 (time and frequency symbols). This increases the time per symbol by a factor of 4. Therefore, there are 19.53 ns per index number.

6. Upstream Lab Results

The downstream comparison tests were repeated on the upstream OFDMA channel, which ran 10-85 MHz and used 50kHz channel spacing. This gave a 1500 point discrete Fourier transform. The network analyzer results on the echo are shown in Figure 19, and the CM results are shown in Figure 20. Note

that the VNA is showing an echo, but the CM coefficients are showing the channel's correction, which is the inverse, and infinitely recursive as described in Appendix 2. In the frequency domain, if the ripple goes up for the VNA at some frequency, it is reported as going down from the CM to make a correction. In the frequency domain the inverse solution is infinitely recursive, but in the time domain the measurement is a single recursion.

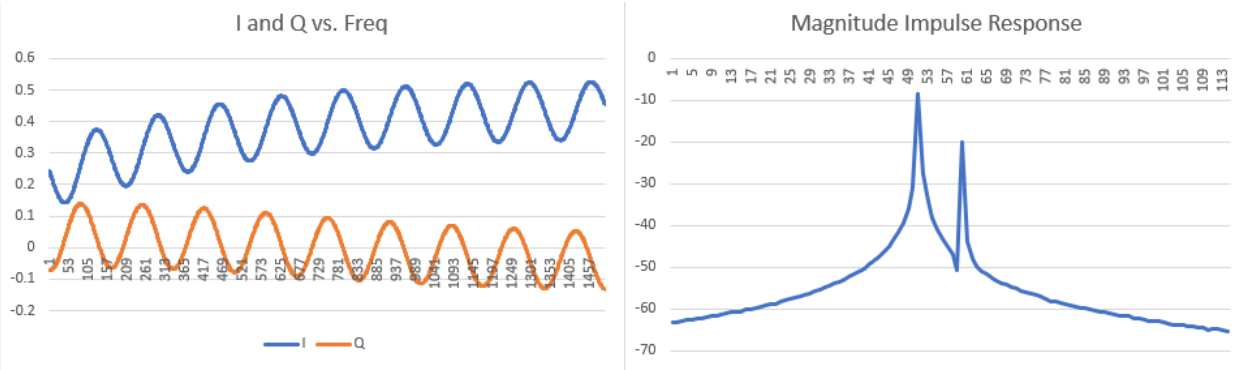


Figure 19 - Network analyzer upstream results in time and frequency.

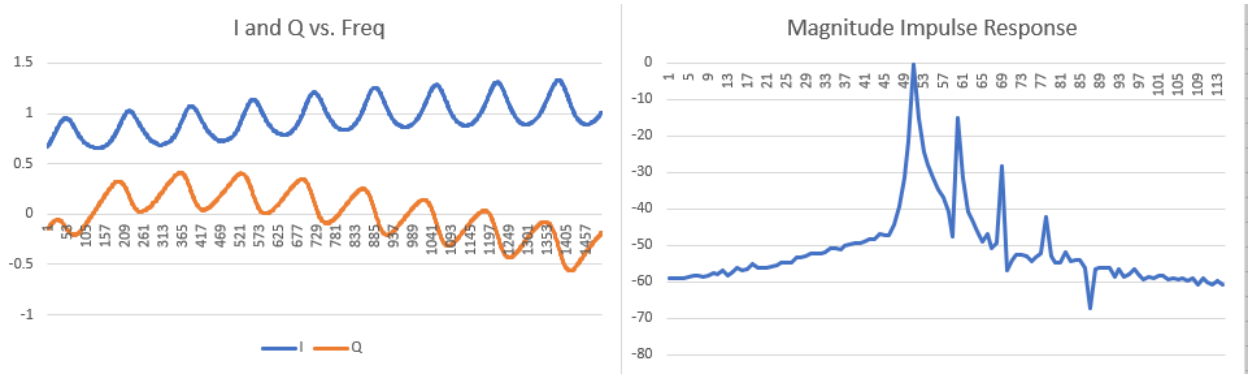


Figure 20 - Cable Modem upstream results in time and frequency. Notice the results are the inverse of the network analyzer results because the CM is returning the correction, not the channel response.

7. Field Data

A browsing tool was created in Python to display equalization coefficients. It is included in Appendix 1. The data are preprocessed for removal of the added delay. Two example plots are illustrated in Figure 21 and Figure 22. The center plot is FD linear I vs. Q. The right plot is FD linear I and Q vs frequency. The left plot is TD magnitude log impulse response. The time scale is IFFT index numbers and is 19.53125 ns per point. On the impulse response, echoes appear to the right of the main tap (delayed) and group delay effects will appear to the left of the main tap.

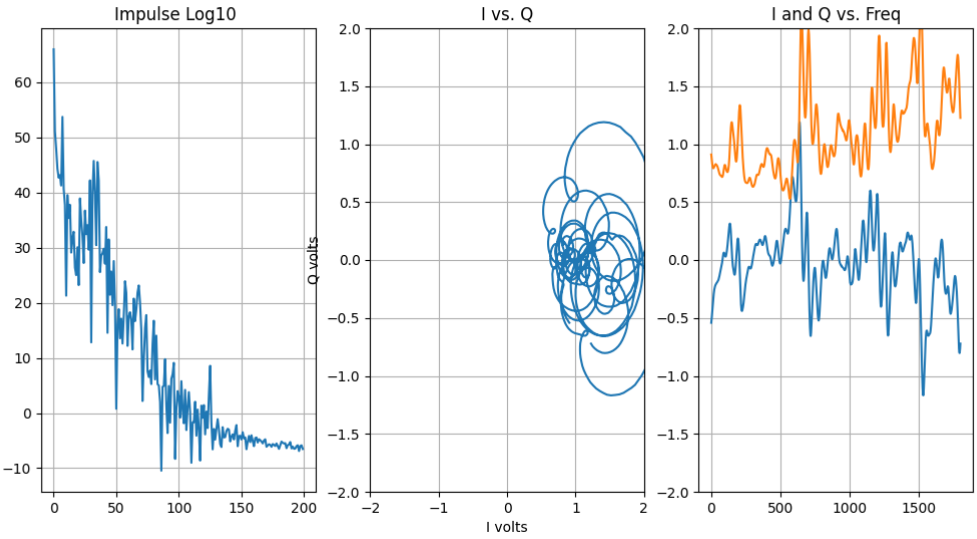


Figure 21 – Downstream example shows an unknown big problem, maybe water in coax.

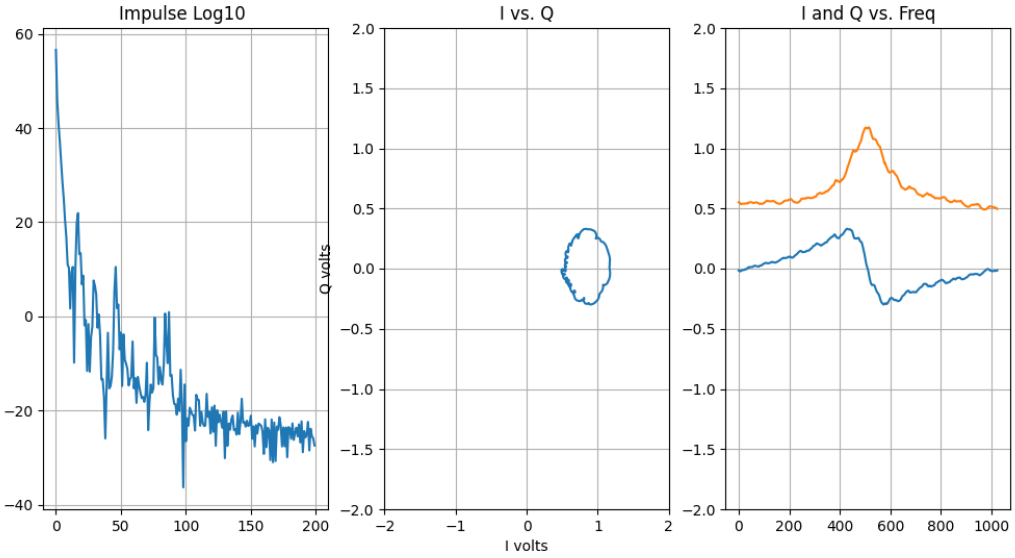


Figure 22 – Second downstream response showing resonant peaking response.

The most common impairment observed was a standing wave created by an echo tunnel. Taps, amplifiers, and other network components generally have a return loss around 16-18dB. This implies, if the coaxial cable was lossless, the strongest observed echo would be 32dB from two reflections between any two components. Of course, cable is not lossless, but has predictable loss as a function of diameter and frequency. For every nanosecond a reflected signal travels in the coax predicted attenuation should occur given by:

dB/ns = A dB/meter * VoP * 0.3meters/ns

Example: 0.5 cable at 750MHz has 7.09dB per 100m, or 0.079dB/1m

The speed of light is 3e8m/sec, and the velocity of propagation (VoP) is 87%.

So, every nanosecond an echo is attenuated by 0.0206dB, or 20.6dB per microsecond.

This allows a mask to be applied to an impulse response, as illustrated in Fig. xx. Any impulse crossing a threshold signifies an out of spec. component. This simple test allows cable damage producing less than a 16dB discrete reflection to pass undetected unless echo timing is considered.

Figure 23 is a similar plot to Figure 3. It shows a CM downstream response with an added threshold red line of some number of dB of attenuation per microsecond slope. Because the cable diameter (loss) was not known, the line's correct slope is a guess. This line can be used as a limit for finding any echo that peaks above the line, indicating of a potential plant problem. A problem is shown where the echo limit is crossed. The far-right FD plot indicates echoes (ripple) that align with the peaks in the left plot. The green CP limit line shows that the selected CP length is sufficient.

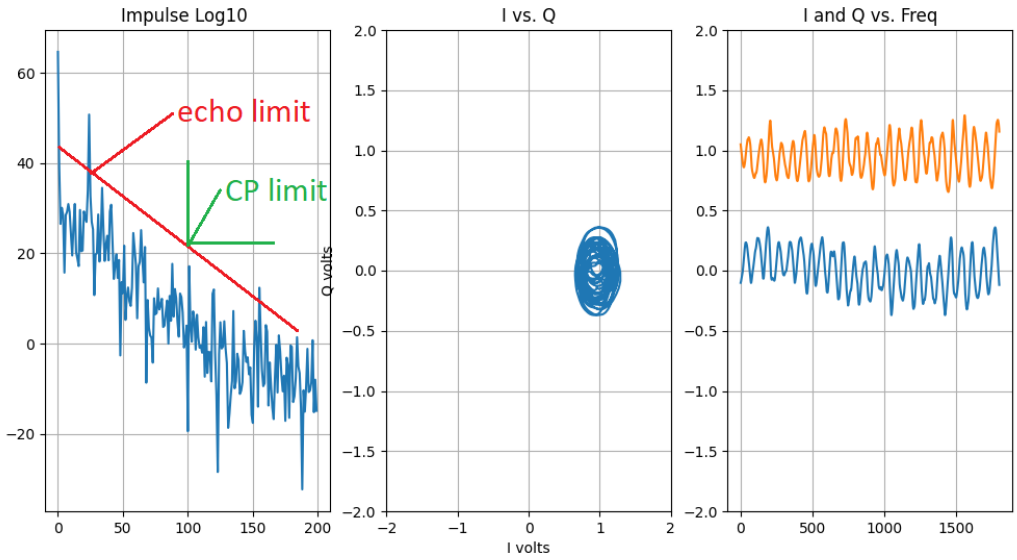


Figure 23 – Adding a line that reflects the CP, so that any peak above the line indicates a signal problem.

Appendix 4 holds a gallery of interesting downstream and upstream data plots from field data with comments in the figure captions.

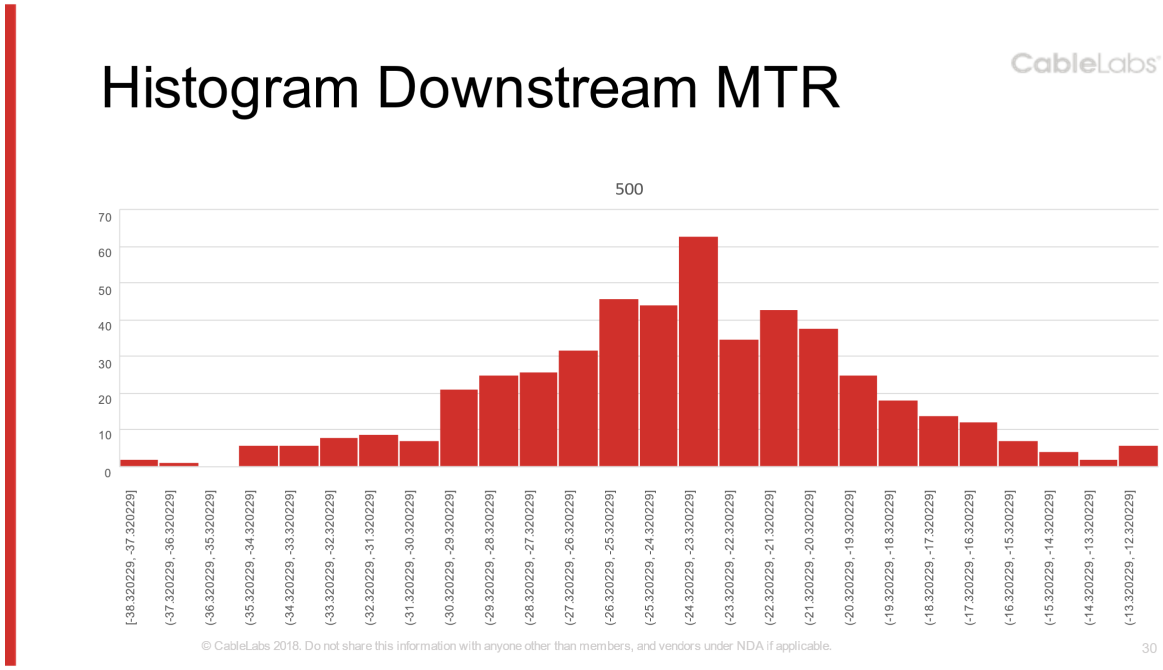


Figure 24 - A histogram of DS main tap ratio values from about 500 CMs from the field.

Fig. 24 is a histogram of downstream main tap ratio (MTR) values for a field population of about 500 CMs. The values range between 12dB and 37dB with a peak around 24db. MTR is the measure of the energy in the main tap to the energy in all other taps combined. If residual delay is not removed, the MTR values will be lowered.

8. Response Matching Algorithms

For many years PNM engineers have been discovering which field responses are similar by using frequency domain division, followed by an IFFT. Similar responses may indicate that modems share a common plant impairment, as illustrated in Figure 1. This makes for easier and quicker plant troubleshooting when a strand map is available. More recently, unimpaired normal responses are being matched, allowing discovery of which common cable lines are used, indicating CMs are in the same neighborhood. More details on this in Appendix 3.

In this matching process, a set of coefficients from one cable modem is divided into another CM's coefficients at each subcarrier frequency using complex frequency domain division. The result is a set of quotients which will be all 1.0 at 0 degrees, if responses are identical. This set of quotients is processed with an IFFT to make a time response. A temporal power measurement can be made of the DC term (coefficient zero) relative to the power in all other coefficients combined. A large ratio of main tap to all other taps combined indicates a good complex response match. This is illustrated in Figure 25, where two I-Q polar responses look very similar. Their quotient is illustrated in Figure 26 as both I and Q vs. frequency and I vs. Q. When this spectral response is transformed with an IFFT, almost all of the transform's power will be in the temporal DC term.

To form clusters, processing is generally only done on CM data from inside a fiber node. Each CM's response is used as a reference and compared with each other CM response, and all matches are noted and formed into groups.

In the past this has been done on 24 coefficient single carrier upstream responses. Now it can be done on both upstream OFDMA and downstream OFDM responses with hundreds of coefficients. Trials on downstream field data show it works well. Furthermore, the effects of different house wiring can be masked by eliminating (zeroing) low value time domain coefficients when performing the power calculation.

Using complex coefficient matching generally produces noticeably better results than using magnitude-only coefficient response matching.

Figure 25 - Two CM's responses, one used in a numerator and the other (reference) used in the denominator.

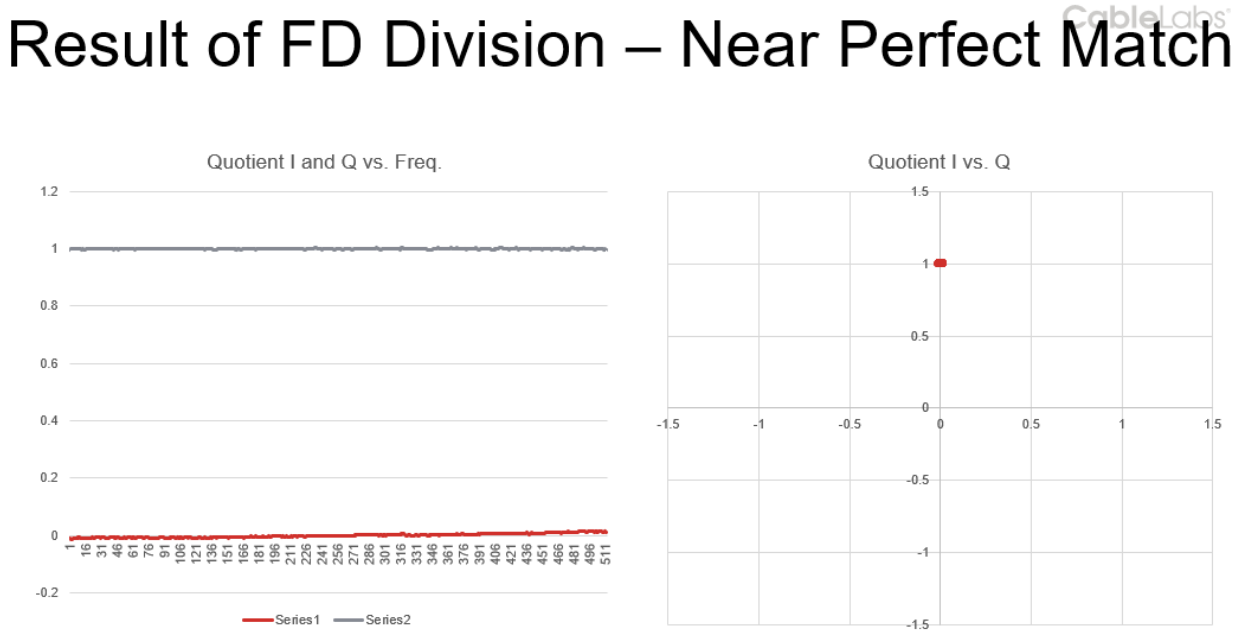


Figure 26 - Two nearly identical responses produce quotient coefficients of 1.0 at an angle of 0 degrees. Nearly all the energy is in the DC term.

9. New Applications for Wideband Equalization Data

Existing applications of impaired response matching and individual response matching work better with multicarrier than single carrier because the equalization data has more dynamic range, in part due to using elevated pilots. The wider bandwidth of the signals also provided for more accurate time resolution.

New possible applications are:

1. Wideband downstream and upstream response matching.
2. Investigation of correct sizing of OFDM(A) CP length.
3. Intermittent detection from unstable equalization responses.
4. Interference or ingress causing band-limited wrong equalization values.
5. In-house wiring problems, such as broken coaxial shields, identified by energy in low index coefficients.
6. Water detection in coaxial cable.
7. Locating subscribers connected to the same coaxial line by matching responses or identifying amplifiers in a cascade.
8. Field test gear, including time domain reflectometer (TDR) functionality.

Experience from using PNM objects has taught us that operational value is gained by learning what physical problems are manifested in observed channel responses. Unfortunately, we have not been provided feedback from our operators on what problems have created the unusual responses, partially due to time and distance, and partially due to the global pandemic.

An operational disservice occurs when time and resources are wasted on chasing non-problems. An impairment that is time-variable in nature is more suspicious than one that is static. Likewise, an impairment that increases with time, such as water slowly entering a cable, should be flagged for inspection and possible repair.

DOCSIS equalization is powerful, and many impairments can be tolerated. For example, if a tap's return loss drops from a specified 18dB to 15dB, the equalizers can normally fix that issue.

10. Conclusions

OFDM and OFDMA equalization coefficients both contain useful complex frequency response data, but the data need to be pre-processed to remove added random, linear delay which is phase rotation. Code to do this processing is in the CableLabs PNM GIT repository C3, and a verbose version is provided in the first appendix of this paper. The first appendix also has Python code for plotting results. This equalization analysis should work for water wave detection and have field test equipment applications. The wide measurement bandwidth obtainable from OFDM and OFDMA provides detail in the time domain, which allows more precise distance measurements. The impulse response can show if your CP (or Guard Interval) is the correct length or if echoes are above expected limits.

While spectrum data, which is magnitude only, has been proven to be very valuable for impairment detection, the complex data from pre-equalization and channel estimation data may be more useful for determining more precisely the location and cause of the issue, better informing maintenance decisions.

Many thanks to the members who contributed data for this analysis.

Abbreviations

CM	Cable modem
CMTS	Cable modem termination system
CP	Cyclic prefix
CPD	Common path distortion
DC	Direct Current
DFE	Decision feedback equalization
FBC	Full band capture
FD	Frequency domain
FFT	Fast Fourier transform
FIR	Finite impulse response
GD	Group delay
I	In-phase
IFFT	Inverse Fast Fourier Transform
MC	Multi-carrier
MIB	Management information base
MTR	Main tap ratio
OFDM	Orthogonal frequency division multiple
OFDMA	Orthogonal frequency division multiple access
PIM	Passive inter-modulation
PNM	Proactive network maintenance
Q	Quadrature
SC	Single carrier
SCTE	Society of Cable Telecommunications Engineers
TD	Time Domain
TFTP	Inverse Fast Fourier Transform
TDR	Time Domain Reflectometer
SCTE	Society of Cable Telecommunications Engineers
VNA	Vector network analyzer

Bibliography & References

Alan V. Oppenheim, Alan S. Willsky, S. Hamid Nawab, *Signals and Systems, second edition,* Prentice Hall, 1996. (See page 53 for a discussion on differences between linear and nonlinear.)

Alberto Campos, Bruce Currivan, Charles Moore, and Tom Williams titled “Upstream Cable Echoes Come in Two Flavors,” CED Magazine, 2010.

Data Over CableService Interface Specification Proactive Network Maintenance “Primer for PNM Best Practices in HFC Networks (DOCSIS® 3.1),” CM-GL-PNM-3.1-V02-210114 (Cable Television Laboratories).

Data-Over-Cable Service Interface Specifications DOCSIS® 3.1 Cable Modem Operations Support System Interface Specification CM-SP-CM-OSSIv3.1-I16-190917 (Cable Television Laboratories).

DOCSIS® Best Practices and Guidelines PNM Best Practices: HFC Networks (DOCSIS 3.0) CM-GL-PNMP-V03-160725 (Cable Television Laboratories).

Data-Over-Cable Service Interface Specifications DOCSIS® 3.1 Physical Layer Specification CM-SP-PHYv3.1-I17-190917 (Cable Television Laboratories).

Data-Over-Cable Service Interface Specifications DOCSIS® 3.1 CCAP™ Operations Support System Interface Specification CM-SP-CCAP-OSSIv3.1-I18-200610 (Cable Television Laboratories).

Appendix

1. Removing the delay in the TD, which is rotation in FD.

The computer code explained in Section 4 of the paper is provided in this Appendix 1. First, we provide C++ code for delay removal from raw responses, followed by Python browsing code which plots responses in 3 graphs.

1.1. C++ code

The following code is meant to be copied and pasted into your software editor for direct use.

```
//PaperCodeR3.cpp a program to process OFDM equalization coefficients

//R3 added choosing a freq band where to flatten angle, s1 to s2

//R2added normalization, group delay

//copyright Cable Television Laboratories Inc. 2021

/*

Legal notice

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated

 documentation (the "Software") the rights to use, copy, modify, merge, publish, distribute, sublicense,

 and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject

to the following conditions:

The above copyright notice and this permission notice shall apply to such person and  be included in all copies

 or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED

 TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT.

 IN NO EVENT SHALL CABLELABS, ITS MEMBERS, OR ITS SUBSIDIARIES BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,

 WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM OR RLEATED TO THE USE OF THE SOFTWARE  OR

 OTHER EXERCISE OF THE RIGHTS AS LICENSED HEREIN.

*/

#include <iostream>

#include <math.h>

#include <stdio.h>
```

```
#include <stdlib.h>
#include <malloc.h>
#include <time.h>
#include <string.h>
#include <unistd.h>
#ifdef _WIN32
#include <conio.h>
#endif
typedef struct {float real, imag;} COMPLEX;
extern void correct(COMPLEX *,int,int,int,int,int );
extern void idft(COMPLEX *, COMPLEX *, int);
extern void dft(COMPLEX *, COMPLEX *, int);
void FDinvert(COMPLEX *,int);
float normalizePower(COMPLEX *,int);
float PI = 3.141592653, MTR;
int fftSize , cnt, good;//int fftSize = 404, cnt, good;
//e.g. from command line: main.exe 1 1880 1 1880
int main(int argc, char *argv[] )//int main()
{
        int nrIQ,nrLines=1,nrSamples=1880,s1=0,s2=1880;
                fftSize = nrSamples;
        if((int)argc >1){
        nrLines = atoi(argv[1]);//command line argument
                printf("argc = %d\t %s\t%s\n",argc, argv[1],argv[2]);
                nrSamples = atoi(argv[2]);
                nrLines = atoi(argv[1]);//command line argument
                s1=atoi(argv[3]);
                s2=atoi(argv[4]);
                printf("you entered %d arguments\t fftSize=%d\tnrLines=%d\n",argc-1,fftSize,nrLines );
        }
   printf("running batchOFDM\n");
```

```
COMPLEX *w;
w=(COMPLEX*) calloc(4096, sizeof(COMPLEX));
if(!w) { printf("\n Unable to allocate input memory.\n"); printf("\x7"); exit(1); }
        FILE *input, *output,*outputR;
int i,j,offset=0,lineLength=404;
char *pch;
char delim[] = "    ";//delimiter is a tab or a space
char data[80000] = {0};
        float A[4096],B[4096];
if( (input = fopen("input.txt", "r") ) == NULL)
{   //read data
                printf("could not open file.\n");
                system("pause");
   free(w);
                exit(0);
        }
if( (output = fopen("output.txt", "w") ) == NULL)
{   //write results to a file
                printf("could not open file.\n");
                system("pause");
   free(w);
   fclose(input);
                exit(0);
        }
if( (outputR = fopen("outputR.txt", "w") ) == NULL)
{   //write results to a file
                printf("could not open file.\n");
                system("pause");
   free(w);
   fclose(output);
   fclose(input);
```

```
                exit(0);
        }
for(cnt=0;cnt<nrLines  ;cnt++)
{
    //if less than cnt lines, data will be empty (except for newline),
    // strtok will return NULL, and atof will crash
    fgets(data,80000,input);
    pch = strtok(data,delim);
    if (pch == NULL)
    {
        //printf("No data\n");
        continue;
    }
    A[0]=atof(pch);
    pch = strtok(NULL,delim);
    if (pch == NULL)
    {
        //printf("No data\n");
        continue;
    }
    B[0]=atof(pch);
    i=1;
    while( pch !=NULL)
    {
        pch = strtok(NULL,delim);
        if (pch == NULL)
            break;
        A[i]=atof(pch);
        pch = strtok(NULL,delim);
        if (pch == NULL)
            break;
```

```
    B[i]=atof(pch);

    i++;

}

lineLength = i;//how many numbers were read in

for(i=0;i<fftSize;i++)

{//all lines set to this length

    if(i >= offset && i<fftSize*2+offset)

    {

        w[i-offset].real=A[i];

        w[i-offset].imag=B[i];

    }

}

//data looks like I value followed by Q value followed by next I value followed by next Q value

//eg: -1.343994141        -0.490112305        1.233276367        -0.723999023        -0.2265625        1.406616211
        -0.934570313        -1.072509766        1.419677734        -0.033813477

//offset is used to pick a start point not at the start of the read file.  This can reduce effects of GD. But you must make sure

//you have enough I-Q samples on the line or an error will occur

fprintf(output,"file = %d\tLine Length=%d\tRead Offest=%d\n",cnt,lineLength, offset);//SC=subcarriers

printf("\nfile = %d\tLine Length=%d SC\tRead Offest=%d SC\n",cnt,lineLength, offset);

for(i=0;i<fftSize;i++)

{//input data

        fprintf(output,"F%d\t%f\t%f\n",i,w[i].real,w[i].imag);

}

fprintf(output,"\n");

                s1=100;s2=1800;//set these for widest possible band without GD from band edge filters// probably want to do
some limit checking here

correct(w,fftSize,s1,s2,cnt,0);//this function is used to process both pre and post equalziation data

//do a second time to remove residual rotation

correct(w,fftSize,s1,s2,cnt,1);

                //void correct(COMPLEX *w,int np, int s1, int s2,int fn,int iteration)

//FDinvert(w,fftSize);  //optional step to look at what a VNA would see, not the inverse

normalizePower(w,fftSize);

for(i=0;i<fftSize;i++)
```

```
        {
            fprintf(outputR,"%f,",w[i].real);
            if(w[i].imag >= 0)
            {
                fprintf(outputR,"+%f,",w[i].imag);
            }
            else
            {
                fprintf(outputR,"%f,",w[i].imag);
            }
        }
                fprintf(outputR,"\n");
            }
    fclose(outputR);
    fclose(output);
    fclose(input);
    return(0);
}//end of MAIN////////////////////////////////////////////////////////////////////////////////////////////
float normalizePower(COMPLEX *w,int np){
            int i;
            float a,pow=0;
            for(i=0;i<np;i++){
                        pow += w[i].real*w[i].real +  w[i].imag*w[i].imag;
            }
            pow = pow/(float)fftSize;
printf("power = %f\n",pow/(float)fftSize);
            for(i=0;i<np;i++){
                        w[i].real = w[i].real/sqrt(pow);
                        w[i].imag = w[i].imag/sqrt(pow);
            }
            return pow;
```

```
}


void FDinvert(COMPLEX *w,int np)
{//take FD reciprocal
        int i;
        float a,b,DENOM;
        for(i=0;i<np;i++)
   {
                a=w[i].real;
                b=w[i].imag;
                DENOM = a*a + b*b;
                w[i].real = a/DENOM;
                w[i].imag = -b/DENOM;
        }
}
void correct(COMPLEX *w,int np, int s1, int s2,int fn,int iteration)
{
        FILE *output;
        int i,sp=1;//0 sp supresses fprintf, 1 prints inline
        float th1=0,mag=0,angle[4096],dphi=0,theta,delta=0,pow=0,DCpow=0,oldang,gd;
        double ang;
        char readme[16];
   sprintf(readme,"output%d.txt",0  );//to supress making a file for every response
        COMPLEX *s;
        s=(COMPLEX*) calloc(4096 , sizeof(COMPLEX));
        if(!s)
   {
      printf("\n Unable to allocate input memory.\n");
      printf("\x7");
      // add cleanup here
      exit(1);
```

```
}

        COMPLEX *r;
        r=(COMPLEX*) calloc(4096 , sizeof(COMPLEX));
        if(!r)
{
   printf("\n Unable to allocate input memory.\n");
   printf("\x7");
   // add cleanup here
   exit(1);
}

        COMPLEX *t;
        t=(COMPLEX*) calloc(4096 , sizeof(COMPLEX));
        if(!t)
{
   printf("\n Unable to allocate input memory.\n");
   printf("\x7");
   // add cleanup here
   exit(1);
}

        COMPLEX *u;
        u=(COMPLEX*) calloc(4096, sizeof(COMPLEX));
        if(!u)
{
   printf("\n Unable to allocate input memory.\n");
   printf("\x7");
   // add cleanup here
   exit(1);
}
```

```
        COMPLEX *ws;
        ws=(COMPLEX*) calloc(4096 , sizeof(COMPLEX));
        if(!ws)
{
   printf("\n Unable to allocate input memory.\n");
   printf("\x7");
   // add cleanup here
   exit(1);
}

        COMPLEX *x;
        x=(COMPLEX*) calloc(4096, sizeof(COMPLEX));
        if(!x)
{
   printf("\n Unable to allocate input memory.\n");
   printf("\x7");
   // add cleanup here
   exit(1);
}

        COMPLEX *y;
        y=(COMPLEX*) calloc(4096, sizeof(COMPLEX));
        if(!y)
{
   printf("\n Unable to allocate input memory.\n");
   printf("\x7");
   // add cleanup here
   exit(1);
}
```

```
        if( (output = fopen(readme, "w") ) == NULL)
{
            printf("could not open file.\n");
   system("pause");
   // add cleanup here
   exit(0);
}

        printf("\n");
//middle frequency is np/2
th1 = atan2(w[np/2].imag,w[np/2].real);
for(i=0;i<fftSize;i++)
{
   ws[i].real = w[i].real; //copy and store  the input
   ws[i].imag = w[i].imag;
}
printf("delay correction pass #%d\tFFT Size=%d\tAnalysis between FFT sample %d and sample %d\n",iteration, fftSize,s1,s2);
//printf("freq=%d\txR=%f\txI=%f\tang=%f\n", np ,w[np/2].real,w[np/2].imag,th1,th2);

if (sp==1)
   fprintf(output,"P1_1 input FD data: I, Q, angle[]\tLine # = %d\n", fn );
// /* begin rotation removal function
for(i=0;i<np;i++)
{
   angle[i]=atan2(ws[i].imag,ws[i].real);
   ang=atan2(ws[i].imag,ws[i].real);
   if(sp==1)fprintf(output, "~%d\t%f\t%f\t%f\n",i,ws[i].real,ws[i].imag,angle[i]); //raw data  Dan bug fix 2
}
float phi = 0;
```

```
        for(i=s1 ;i<s2;i++)
 {
   dphi = angle[i]-angle[i-1];
   phi += dphi;
   if(dphi > PI) {phi -= 2*PI;}//  printf("going clockwise");}
   if(dphi < -PI) {phi += 2*PI;}//printf("going counter clockwise");}
}
printf("fftSize = %d\t s1=%d\ts2=%d\n", fftSize, s1,s2);
delta = -phi/(float)(s2-s1);//delay estimate
printf("Radians rotation =%f\tSlope=%f radians per subcarrier\n",phi,delta );//reduce the delta by number of -pi to pi jumps
//end rotation removal function
if(sp==1)
   fprintf(output,"2 rotation removed in FD\n");
for(i=0;i<np;i++)
{
   ang = -(float)i*delta;//(float)np;
   u[i].real = ws[i].real*cos(-ang) - ws[i].imag*sin(-ang);//w is impaired
   u[i].imag = ws[i].real*sin(-ang) + ws[i].imag*cos(-ang);
   ang=atan2(u[i].imag,u[i].real);
   if(sp==1)
      fprintf(output, "2C%d\t%f\t%f\t%f\n",i,u[i].real,u[i].imag,ang);
}
//now split fd data into sidebands for using ifft
for(i=0;i<np/2;i++)
{//upper sideband
   s[i].real = u[i+np/2].real;
   s[i].imag = u[i+np/2].imag;
}
for(i=0;i<np/2;i++)
{//lower sideband
   s[fftSize-np/2+i].real = u[i].real;
```

```
    s[fftSize-np/2+i].imag = u[i].imag;
}
if(sp==1)fprintf(output,"3 sidebands swapped\n");
for(i=0;i<fftSize;i++)
{
    t[i].real = s[i].real;
    t[i].imag = s[i].imag;
}
for(i=0;i<fftSize;i++)
{
mag = sqrt(t[i].real*t[i].real + t[i].imag*t[i].imag) ;
if(sp==1) fprintf(output,"3Sf%d\t%f\t%f\t%f\n",i,t[i].real,t[i].imag,mag);
}
idft(s,r,fftSize);//tom
for(i=0;i<fftSize;i++)
{
    s[i].real=r[i].real;
    s[i].imag=r[i].imag;
}
if(sp==1)
    fprintf(output,"4 TD with angle err.\n");
for(i=0;i<fftSize;i++)
{
    mag=sqrt(s[i].real*s[i].real + s[i].imag*s[i].imag);
    if(sp==1)
        fprintf(output,"4St%d\t%f\t%f\t%f\n",i,s[i].real,s[i].imag,mag);
}
for(i=0;i<32;i++)
{
    mag=sqrt(s[i].real*s[i].real + s[i].imag*s[i].imag);
    if(sp==1)
```

```
        fprintf(output,"4St%d\t%f\t%f\t%f\n",i,s[i].real,s[i].imag,mag);
}
theta = atan2(s[0].imag,s[0].real);//rotate the DC term
printf("DC term rotation = %f radians\n",theta);
ang = theta;
if(sp==1)fprintf(output,"5 TD angle = 0 deg\n");
for(i=0;i<fftSize;i++)
{
    x[i].real = s[i].real*cos(-ang) - s[i].imag*sin(-ang);//w is unimpaired
    x[i].imag = s[i].real*sin(-ang) + s[i].imag*cos(-ang);
    mag=sqrt(x[i].real*x[i].real + x[i].imag*x[i].imag);
    if(sp==1)
        fprintf(output, "5@CT%d\t%f\t%f\t%f\t%f\n",i,x[i].real,x[i].imag,mag,20*log10(mag));
}
for(i=0;i<64;i++)
{
    mag=sqrt(x[i].real*x[i].real + x[i].imag*x[i].imag);
    if(sp==1)
        fprintf(output, "5@CT%d\t%f\t%f\t%f\t%f\n",i,x[i].real,x[i].imag,mag,20*log10(mag));
}
for(i=0;i<fftSize;i++)
{
    y[i].real = x[i].real;
    y[i].imag = x[i].imag;
}
dft(y,r,fftSize);
for(i=0;i<fftSize;i++)
{
    y[i].real=r[i].real;
    y[i].imag=r[i].imag;
}
```

```
//y has the correct answer
for(i=0;i<fftSize/2;i++)
{//swap sidebands back
   w[i].real = y[i+fftSize/2].real;
   w[i].imag = y[i+fftSize/2].imag;
}
for(i=0;i<fftSize/2;i++)
{
   w[i+fftSize/2].real = y[i].real;
   w[i+fftSize/2].imag = y[i].imag;
}
if(sp==1)fprintf(output,"6 FD corrected\n");
for(i=fftSize/2;i<fftSize;i++)
{
   mag=sqrt(y[i].real*y[i].real + y[i].imag*y[i].imag);
                        ang = atan2(y[i].imag,y[i].real);
                        gd = -(ang-oldang)/(6.28*50000); //assuming a 50kHz subcarrier spacing
   if(sp==1)
   fprintf(output, "6CF%d\t%f\t%f\t%f\t%f\t%f\t%e\n",i,y[i].real,y[i].imag,mag,20*log10(mag),ang,gd);
                                    oldang = ang;
}


for(i=0;i<fftSize/2  ;i++)
{
  mag=sqrt(y[i].real*y[i].real + y[i].imag*y[i].imag);
                        ang = atan2(y[i].imag,y[i].real);
                        gd = -(ang-oldang)/(6.28*50000); //assuming a 50kHz subcarrier spacing
   if(sp==1)
   fprintf(output, "6CF%d\t%f\t%f\t%f\t%f\t%f\t%e\n",i,y[i].real,y[i].imag,mag,20*log10(mag),ang,gd);
                        oldang = ang;
          }
```

```
    pow=0; //initialize
    for(i=0;i<fftSize;i++)
    {
        pow += x[i].real*x[i].real + x[i].imag*x[i].imag;
    }
    DCpow =  x[0].real*x[0].real + x[0].imag*x[0].imag;//imag component should be zero
    printf("total power = %f\tDCpower= %f\n",pow,DCpow);
    printf("power correction to unity power is %f\t%f dB\n",1/pow, 10*log10(1/pow) );
    float Vcorr = sqrt(1/pow);
    printf("voltage correction is %f\n", Vcorr );
    for(i=0;i<fftSize;i++)
    {
        x[i].real *= Vcorr;
        x[i].imag *= Vcorr;
    }
    DCpow =  x[0].real*x[0].real + x[0].imag*x[0].imag;//imag component should be zero
    MTR = 10*log10(1-DCpow)/1.0;
    printf("DC power  =%f\t Other power =%f\t MTR=%f\n",DCpow,1.0 - DCpow, MTR) ;
    fclose(output);
} //END Correct ////////////////////////////////////////////////////////

/*****************************************************************************
dft - Discrete Fourier Transform
This function performs a straight DFT of N points on an array of
complex numbers whose first member is pointed to by Datain.  The
output is placed in an array pointed to by Dataout.
*****************************************************************************/
void dft(COMPLEX *Datain, COMPLEX *Dataout, int N)
{
    int i,k,n,p;
    static int nstore = 0;      /* store N for future use */
    static COMPLEX *cf;         /* coefficient storage */
```

```
    COMPLEX *cfptr,*Dinptr;
    double arg;
/* Create the coefficients if N has changed */
    if(N != nstore) {
        if(nstore != 0) free((char *) cf);    /* free previous */
        cf = (COMPLEX  *) calloc(N, sizeof(COMPLEX));
        if (!cf) {
            printf("\nUnable to allocate memory for coefficients.\n");
            // add cleanup here
            exit(1);
        }
        arg = 8.0*atan(1.0)/N;
        for (i=0 ; i<N ; i++) {
            cf[i].real = (float)cos(arg*i);
            cf[i].imag = -(float)sin(arg*i);
        }
    }
/* Perform the DFT calculation */
    printf("\n");
    for (k=0 ; k<N ; k++) {
        Dinptr = Datain;
        Dataout->real = Dinptr->real;
        Dataout->imag = Dinptr->imag;
        Dinptr++;
        for (n=1; n<N; n++) {
        p = (int)((long)n*k % N);
            cfptr = cf + p;         /* pointer to cf modulo N */
            Dataout->real += Dinptr->real * cfptr->real
                        - Dinptr->imag * cfptr->imag;

            Dataout->imag += Dinptr->real * cfptr->imag
```

```
                        + Dinptr->imag * cfptr->real;
            Dinptr++;
        }
        if (k % 32 == 31) printf("*");
        Dataout++;          /* next output */
    }
    printf("\n");
}
/*************************************************************************
idft - Inverse Discrete Fourier Transform
This function performs an inverse DFT of N points on an array of
complex numbers whose first member is pointed to by Datain.  The
output is placed in an array pointed to by Dataout.
It returns nothing.
*************************************************************************/
void idft(COMPLEX *Datain, COMPLEX *Dataout, int N)
{
    int i,k,n,p;
    static int nstore = 0;      /* store N for future use */
    static COMPLEX *cf;         /* coefficient storage */
    COMPLEX *cfptr,*Dinptr;
    double arg;
/* Create the coefficients if N has changed */
    if(N != nstore) {
        if(nstore != 0) free((char *) cf);    /* free previous */
        cf = (COMPLEX  *) calloc(N, sizeof(COMPLEX));
        if (cf == 0) {
            printf("\nUnable to allocate memory for coefficients.\n");
            // add cleanup here
            exit(1);
        }
```

```
/* scale stored values by 1/N */
        arg = 8.0*atan(1.0)/N;
        for (i=0 ; i<N ; i++) {
           cf[i].real = (float)(cos(arg*i)/(double)N);
           cf[i].imag = (float)(sin(arg*i)/(double)N);
        }
    }
/* Perform the DFT calculation */
    printf("\n");
    for (k=0 ; k<N ; k++) {
        Dinptr = Datain;
        Dataout->real = Dinptr->real * cf[0].real;
        Dataout->imag = Dinptr->imag * cf[0].real;
        Dinptr++;
        for (n=1; n<N; n++) {
        p = (int)((long)n*k % N);
           cfptr = cf + p;        /* pointer to cf modulo N */
           Dataout->real += Dinptr->real * cfptr->real
                        - Dinptr->imag * cfptr->imag;
           Dataout->imag += Dinptr->real * cfptr->imag
                        + Dinptr->imag * cfptr->real;
           Dinptr++;
        }
        if (k % 32 == 31) printf("*");
        Dataout++;         /* next output */
    }
    printf("\n");
}
```

1.2. Python code

```
#makesGraphs from outputR.txt C++ program
#ver 15
import matplotlib.pyplot as plt
import numpy as np
```

```
from numpy.fft import fft, ifft
import math
import cmath
line =0
k=int(line)
s=int(0) #unit to display
dl=int(1600)#data length this value needs to be the number of complex I-Q
samples#fdl = float(dl)
with open('outputR_1600.txt') as f: #data file written by C++ code that
performed delay correction
    for line in f:
        data = line.split(",")
        re = []
        im = []
        fre = []
        fim = []
        y = []
        cx = []
        LX = []
        LY = []
        NLY = []
        TNLY = []

        for cnt in range(0,2*dl,2):
            re.append(data[cnt])
            im.append(data[cnt+1])

        for i in range(dl):
            v=re[i]+im[i]+"j"
            cx.append(v)

        for item in re:
            fre.append(float(item))
        for item in im:
            fim.append(float(item))

        for c in range(0,dl,1):
            x= complex(fre[c],fim[c])

        for item in range(dl):
            y.append(x)

        X=fft(cx)

        LX = np.array(X)

        for i in range(dl):
            LX[i] = LX[i].real*LX[i].real + LX[i].imag*LX[i].imag
            LX[i]=cmath.log10(LX[i]) #log values

        LY = np.append(LX,LX[0])
        NLY = LY[::-1]
        TNLY = NLY[:200]

        fig, (ax0,ax1, ax2,) = plt.subplots(nrows=1, ncols=3,)
```

```
ax0.plot(10*TNLY)# normally 20, not 10.  mag is squared
ax0.grid()
ax0.set_title('Impulse Log10 ')
k += 1
print("line ",k)
t = np.arange(0, dl, 1)

fig.set_figwidth(12)
fig.set_figheight(6)
ax1.plot(fre,fim) #X-Y plot
ax1.set_ylabel('Q volts')
ax1.set_xlabel('I volts')
ax1.set_title('I vs. Q')
ax1.grid(True)
ax1.set_xlim([-2,2])
ax1.set_ylim([-2,2])

ax2.plot(fim) # plot vs freq
ax2.plot(fre) # plot vs freq
ax2.grid(True)
ax2.set_title('I and Q vs. Freq')
ax2.set_ylim([-2,2])
plt.show()
```

2. Upstream Cable Echoes Come In Two Flavors

A first type of echo can be compared to a person standing within two walls of a canyon and yelling "hello". They hear their call repeated several times, each time getting weaker. The signals are bouncing off the 2 walls.

A second type of echo can be compared to a person standing in front of a big single wall and yelling "hello". They hear their call repeated only once.

Cable systems have both types, although the first type is by far more common. The first type, "infinitely recursive", is created by a pair of impedance discontinuities created on a cable line. It is of interest as the impedance mismatch may be caused by cable line damage.

The second type, "single recursive", may be caused by a signal finding two paths upstream. This condition was discovered from upstream transmissions that took the easy upstream path from a home through a high value tap port to the fiber node, and a slightly harder and longer path from a home, through the tap port to the tap's output port, and then off an impedance mismatch downstream to head upstream to the fiber node.

Figure A1 shows the first type of channel (top left) and how it is canceled with the top right impulse response. Essentially, each time an echo is canceled, a tap makes another smaller echo that also needs to be canceled until it is too weak to matter. The infinitely recursive solution impulse response is shown on the upper right.

Figure A2 shows the second type of channel (top left) and how it is canceled with the top right single recursion impulse response. Essentially, each time a multi-recursive echo is encountered, a second tap makes an inverse echo to cancel the next recursion. The solution impulse response is shown on the upper right.

Thus, the solution to an infinitely recursive echo is single recursion set of tap coefficients.

The solution to a single recursive echo is an infinitively recursive set of taps, up to the point when you run out of taps. Hopefully at that point the echo has died out.

This same principle works when observed in the frequency domain for both types of echoes because of time-frequency duality. In the frequency domain, when a frequency response ripple goes up, its inverse response ripple goes down.

OFDM/OFDMA does not employ time domain taps per se but multiplies each frequency domain subcarrier with a complex coefficient for correction. An IFFT can reveal the associated impulse response.

The Math

A single recursion echo can be modeled by: 1 + a where 1.0 is the main signal amplitude and 'a' is the echo's amplitude in linear terms. That is, for a -3dB echo a = 0.707. In a baseband channel 'a' is real, but in an RF channel 'a' may be complex. To be more precise,

$$a = Ae^{j2\pi fT}$$

where A is the amplitude of the echo, f is carrier frequency (MHz) and T is the delay of the echo (usec). However, the equations are easier to present if we simply use 'a'.

The equalizer solution is for a single recursion echo is:

$\frac{1}{1+a} = 1 - a + a^2 - a^3 + a^4 -$ So, the result is what was expected: infinitely recursive. The first term '1' represents the main tap of the equalizer. The second term '-a' acts to cancel the echo by subtracting it. However, in doing so, it causes another, smaller echo in the response. This smaller echo requires the third term to cancel it. This produces another, yet smaller echo, which requires the 4th term to cancel it, and so on until we reach the end of the equalizer delay line. After that, any remaining echo energy (hopefully very small) is not canceled and shows up as reduced RxMER (received modulation error ratio), essentially a noise floor, in the receiver.

If 'a' is a relatively big number such as 0.707 (-3 dB), and the delay T of the echo is several symbol periods, we might run out of taps in the pre-equalizer before we get an accurate solution. DOCSIS 2.0 and later pre-equalizers have 24 taps, with 7 taps normally assigned ahead of the main tap, leaving 16 taps to cancel the echo. For a small value of 'a' such as 0.1 (-20 dB), with a short echo delay, 16 taps is normally sufficient to cancel the echo with minimal residual energy.

Note that the recursions in the above equation have alternating signs. To check for this effect, we can examine the real and imaginary parts of the equalizer taps. If the response is alternating in sign, it is a hint that this type of solution may be present.

A multiple recursion echo may be modeled as:

$$1 + a + a^2 + a^3 + a^4 - .....$$

So its solution can be computed as:

$$\frac{1}{1+a+a^2+a^3+a^4-....} = 1-a$$

Which shows that an infinitely recursive echo can be canceled by a single recursion. So, you should be able to cancel a multiple recursion echo with an adaptive equalizer with only two taps – one for the main signal and one for the echo. To be precise, this example applies to an ideal case where the echo delay T equals a multiple of the symbol period, which for a 5.12 Msps DOCSIS upstream symbol rate is T_s = 195 ns. So, the pre-equalizer can exactly cancel the echo with a single tap if the single echo has delay T = 195 ns, 390 ns, or 585 ns, etc. If the echo lies between these multiples, the equalizer will activate additional taps to provide interpolation. In that case it will be more difficult to see the pattern of a single main recursion.

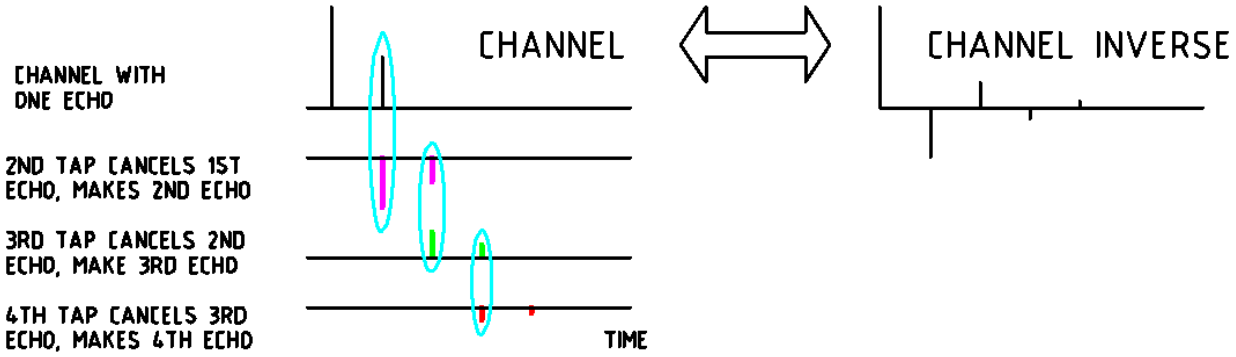


Figure 27 – A single recursive echo needs an infinite number of taps for cancelation. Plots are linear voltage vs time. Blue ovals show cancelation.

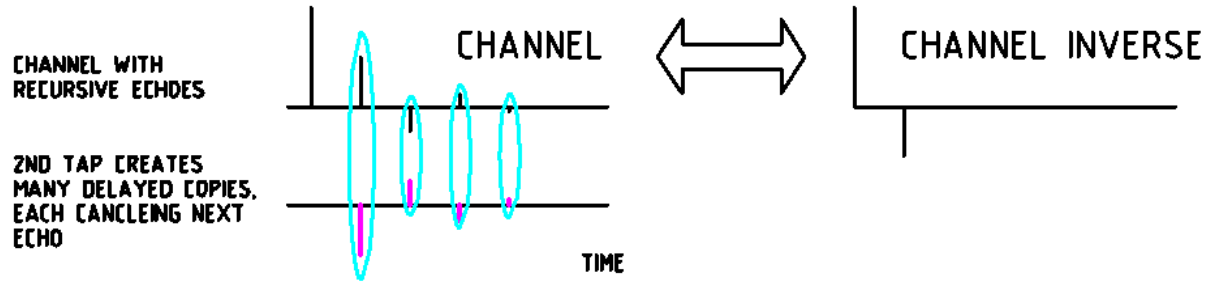


Figure 28 – An infinitely recursive echo is candeled with a two-tap equalizer. Each echo encountered cancels the next echo with a single tap.

3. Linear Distortion Analysis and Network Discovery

Since early days of PNM, we found that by analyzing the characteristics of equalization coefficients, we were able to identify characteristics and features of our HFC network. Many times, these characteristics are apparent when accompanied by a noticeable impairment or condition in the plant such as an impedance mismatch due to a loose connector or an amplifier malfunction. Prior to the DOCSIS 3.1 spec version, we relied on 6.4 MHz or narrower channels in the upstream leveraging up to 24 pre-equalization coefficients. In the downstream direction, although number of coefficients have not been specified, we have seen implementations using 32, 48 or higher number of equalization coefficients to compensate for distortion within the 6 MHz downstream channel. The upstream uses the ranging process to iteratively adjust the coefficients until they converge to a value that compensates channel distortion. In the downstream there is no handshaking with the CMTS so the CM has the sole responsibility in distortion compensation. We have introduced techniques to correlate common impairments within these narrow channels and figure out which CMs share the same impairment and infer the portion of the network that is common to them. One technique consisted of performing complex division of the FFTs of time domain equalization coefficients of the two CMs under comparison. High correlation results in the complex division approaching 1, while if uncorrelated values the division would be much different than 1.

The same approach is valid in DOCSIS 3.1 systems except that the coefficients are already in the frequency domain so one just has to divide the coefficients of CMs under comparison. The normalized equation is shown below.

$$DistortionMatchingMetric = \frac{\sum_{i=1}^{N} \left| \frac{Coef_i CM_k}{Coef_i CM_l} \right|}{N}$$

Where each of the i equalization coefficients of CM_k and CM_l, are complex divided with N being the total number of subcarriers.

Because in DOCSIS 3.1 systems, we are talking about thousands of coefficients rather than tens of coefficients, a much higher resolution is expected. The fact that you are leveraging a complex division process, the likelihood of false matches is much smaller compared to an amplitude only comparison operation. This greater sensitivity enables grouping of cable modems that have more subtle conditions in common, thereby enabling network discovery even on healthy portions of the network.

Another useful tool leveraging complex equalization coefficients analysis is group delay distortion. In DOCSIS 3.0 and earlier versions of the specifications, by analyzing the energy in the pre-main tap coefficients, we had a good indication of the amount of group delay distortion present, and we could infer whether the CMs where sitting behind 0, 1, 2 or higher numbers of amplifiers in cascade. One challenge was that very short micro-reflections would also cause pre-main tap energy to rise, so careful analysis is needed. Figure 32 shows the tap values from an example DOCSIS 3.0 CM.

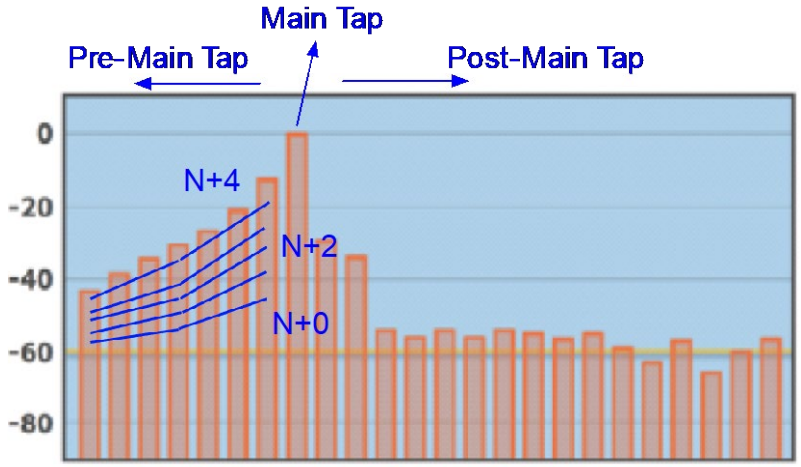


Figure 29 - DOCSIS 3.0 Pre-main tap coefficients and group delay.

In DOCSIS 3.1 systems, we have a view of not just 6 .4 MHz but we can cover the entire US and by obtaining group delay from phase response we can have an accurate view of how group delay changes at the band edges as the number of actives in cascade changes; see Figure 33.

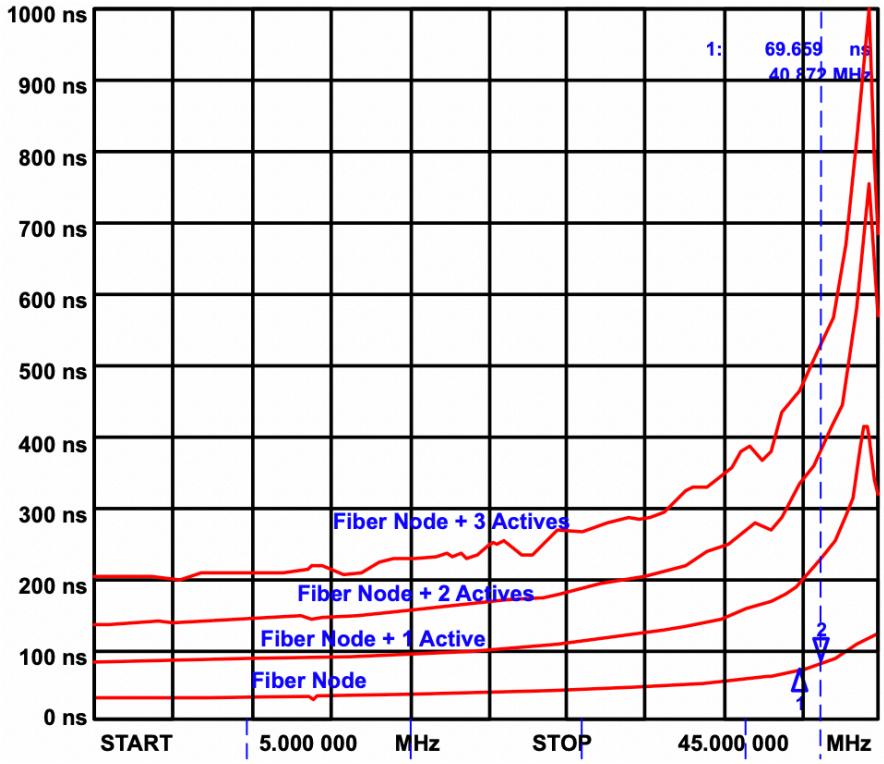


Figure 30 - Group delay and cascade value.

When reflections occurred at tap-coax interfaces, we have been able to deduce coaxial lengths between taps and amplifiers or between taps. This has been useful in determining plant characteristics. As DOCSIS 3.1 specifications came along, the channel became much wider, up to 192 MHz in the downstream and 96 MHz in the upstream. DOCSIS 3.1 specifications evolved from single carrier QAM to multiple subcarriers using either 50KHz or 25KHz, resulting respectively in up to 3800 or 7600 subcarriers in the downstream. In DOCSIS 3.1 specifications, the approach to distortion compensation is the same as with DOCSIS 3.0 and earlier specification versions, in the upstream pre-equalization leverages the ranging handshaking process between CMTS and CM and in the downstream only CM equalization takes place.

A change between DOCSIS 3.0 SC-QAM and DOCSIS 3.1 multicarrier OFDM/OFDMA is that while in DOCSIS 3.0 and earlier versions we had equalization coefficients represented in the time domain across several symbols, in DOCSIS 3.1 we are dealing with a frequency domain equalization representation with one coefficient per subcarrier. This difference is easily overcome through a translation from time to frequency domain using Fourier transformation.

A key difference is that instead of 24/32/48 coefficients in DOCSIS 3.0 and earlier specifications you have up to 3800 or 7600 coefficients in DOCSIS 3.1 specifications. The level of granularity that is possible is more than two orders of magnitude. The techniques discussed here enable the use of equalization coefficients' magnitude and phase for all processes.

Knowing magnitude and phase allows us to determine whether the reflection due to an open a short or a specific complex impedance. This becomes very useful to discriminate whether the distortion impacting CMs is the same or not and will play a key role in grouping devices with common features with high level of sensitivity. The goal is to correlate and group CMs not only when there is a noticeable event or change in channel conditions but also to detect subtle changes even when the network is not impaired, thereby making the HFC network discovery process not just impairment driven but feasible on a healthy network.

Regarding the wider bandwidth coverage, it is worth mentioning that the band-edges in the narrow SC-QAM channels introduce uncertainties as the nature of the FFT requires same values at both edges. The much wider DOCSIS 3.1 channel and frequency equalization operation circumvents that issue. The wider channel view allows us to better describe the group delay distortion in the upstream that is indicative of cascade value.

4. Several upstream (first operator, OFDMA Pre Equalization) and downstream (second operator, OFDM Channel Estimation) results in figures.

For this appendix section, we provide several field data examples in plots. The data come in as I-Q coefficients, one per subcarrier, and typically every 50kHz. The plotted spectral coefficients have linear delay removed before viewing. The FFT size is typically 1024 points for downstream, while the DFT size is 224 points for upstream. In each figure in this section, the center plot is polar I vs. Q, the right plot is I and Q vs. frequency, and the left plot is magnitude of time (impulse response) in dB vs. time. The horizontal scale is the index number. For downstream the time is 19.53ns. per index (4/204.8e6) with a 1024 sample FFT. The bandwidth is 50 kHz times the index, or about 11.2 MHz for upstream and 51.2 MHz for downstream. Note: not all the downstream data provided were plotted.

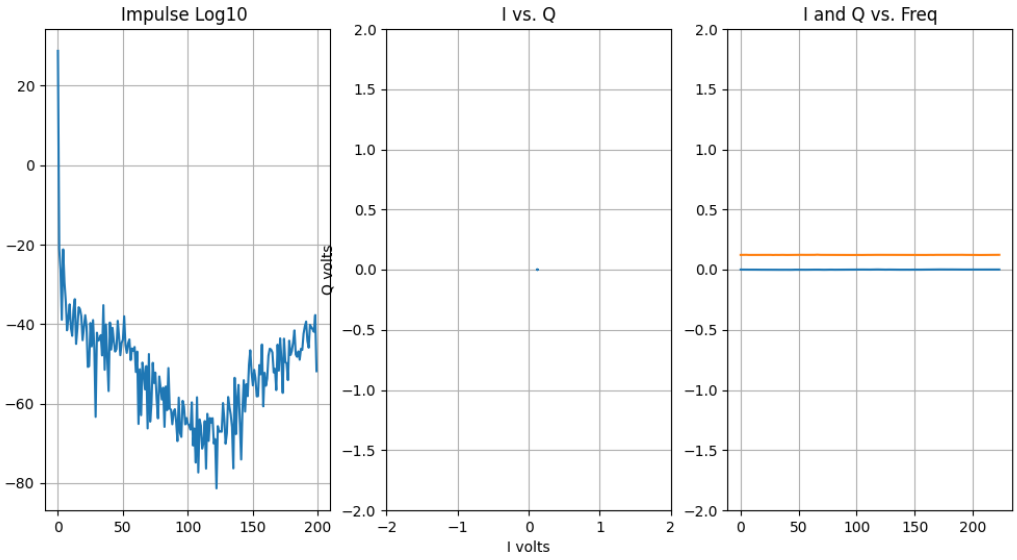


Figure 31 - Pre-equalization data that appears too clean to be cored.

Figure 34 above, shows an upstream response that is too good to be real. Half of the CMs in the data set we received have this response. We guess the result is possibly related to having pre-distortion turned off.

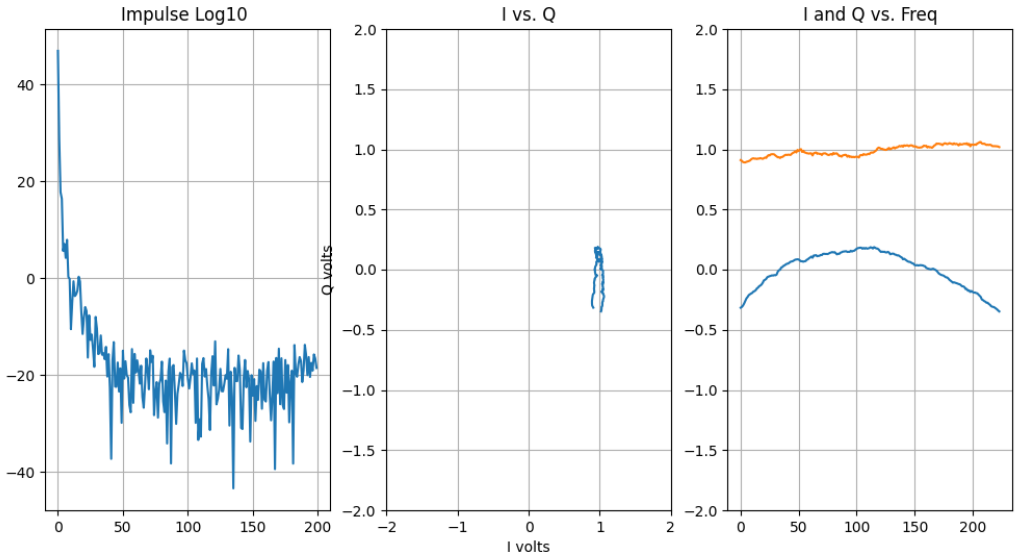


Figure 32 - A typically good pre-equalization result.

Figure 35 above shows a more believable upstream response. A typical upstream good response. Most CMs in our data set look similar to this one. Group delay variation is shown at the top and bottom of the band.

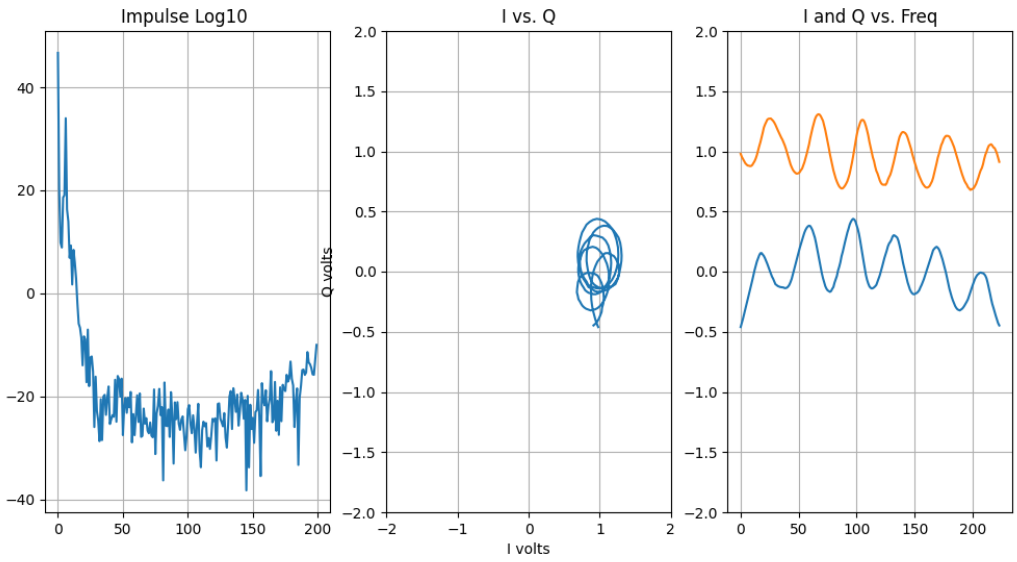


Figure 33 - Pre-equalization showing a standing wave.

Figure 36 above shows an upstream response with a standing wave caused by an infinite recursion echo, which we know because the solution appears as a single recursion.

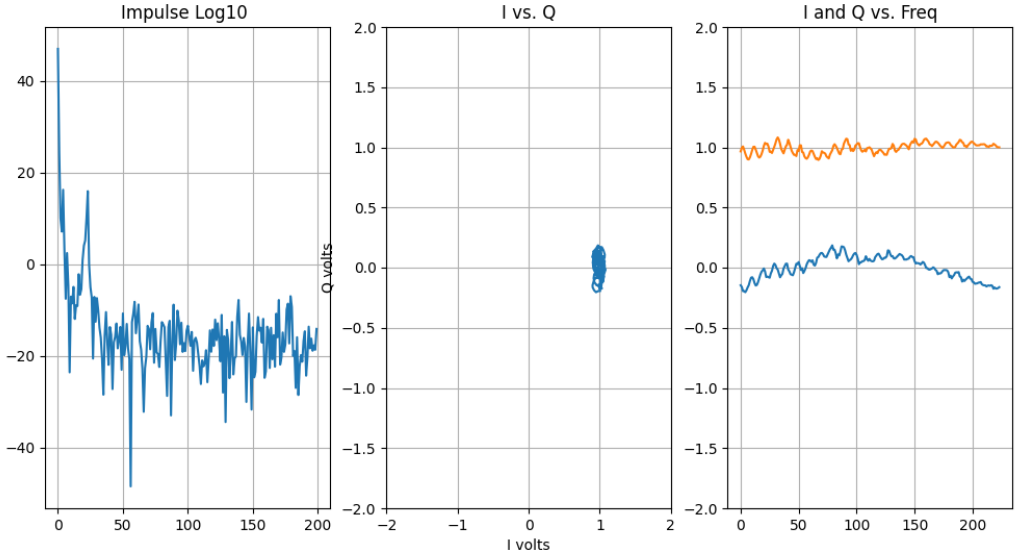


Figure 34 - Upstream pre-equalization data showing a relatively long echo.

Figure 37 above shows an example of an upstream response with a long echo. Note the tight ripple response in the far-right graph.

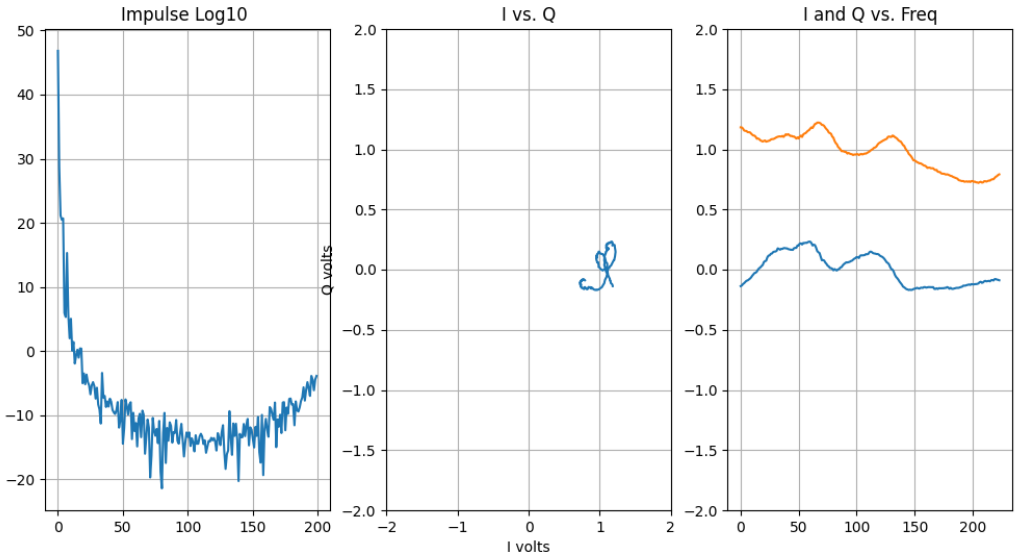


Figure 35 - Upstream pre-equalization, problem probably in house.

Figure 38 shows an upstream data plot of a close-in echo which is probably in the home.

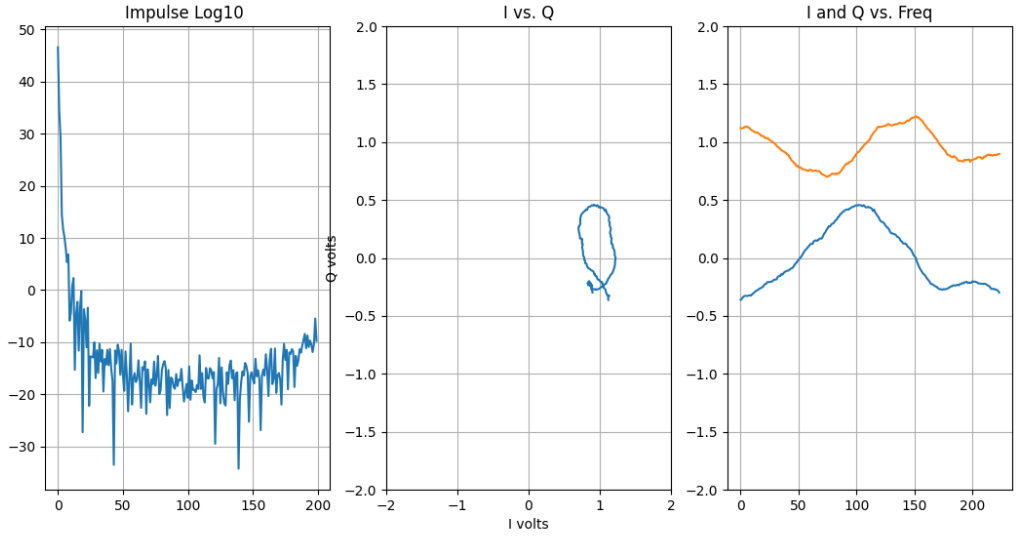


Figure 36 - Upstream pre-equalization data, in-home wiring, maybe open coax shield.

Figure 39 shows pre-equalization data from a CM where there is likely an open shield or other in-home wiring problem.

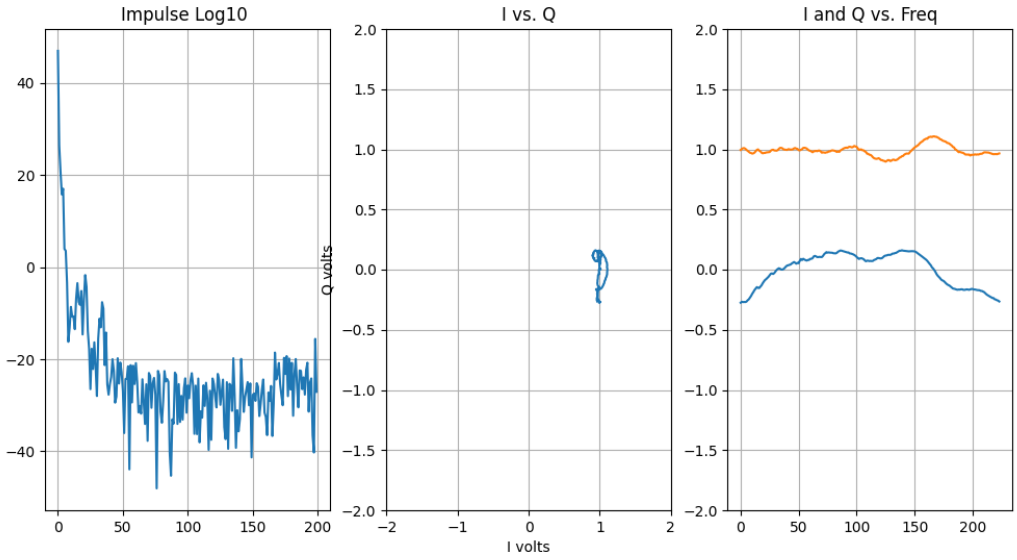


Figure 37 - Upstream pre-equalization, echo only in low frequency bands, so defect probably in home.

Figure 40 above shows an echo but only in the lower part of the band, which suggests an echo cavity, likely in the home.

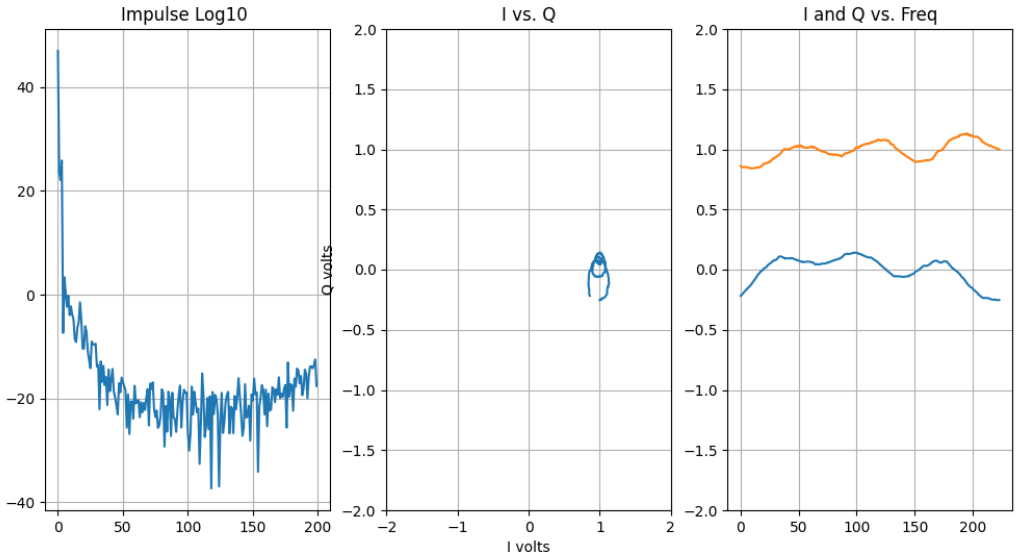


Figure 38 - Upstream pre-equalization data, standing wave.

Figure 41 above shows an example of a small standing wave. Echo is again probably in the house.

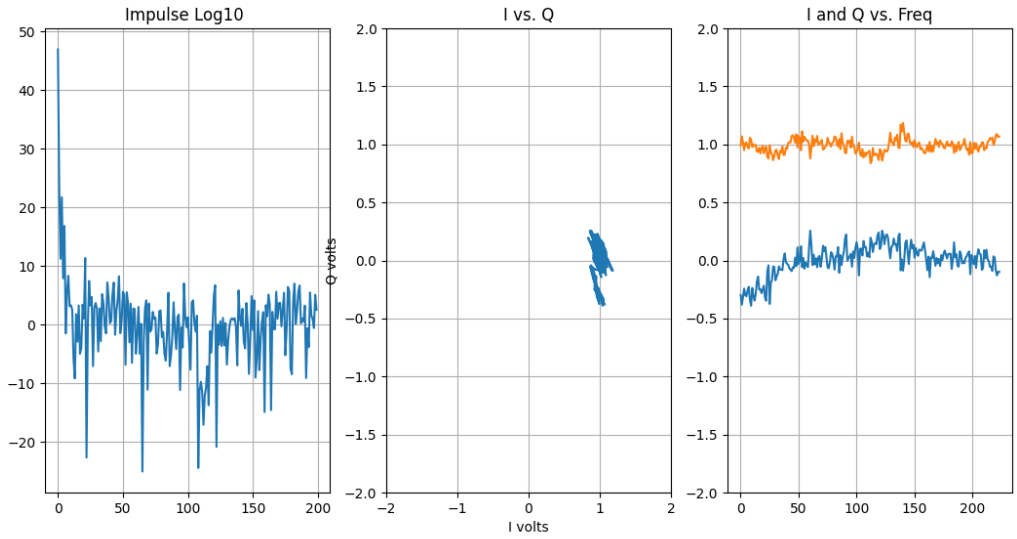


Figure 39 - Upstream pre-equalization, guessed to be water in coax.

Figure 42 above is rather rare in the data set we received. A guess is that it is water in the coax cable from the jagged frequency response. No field tests confirmed this conjecture.

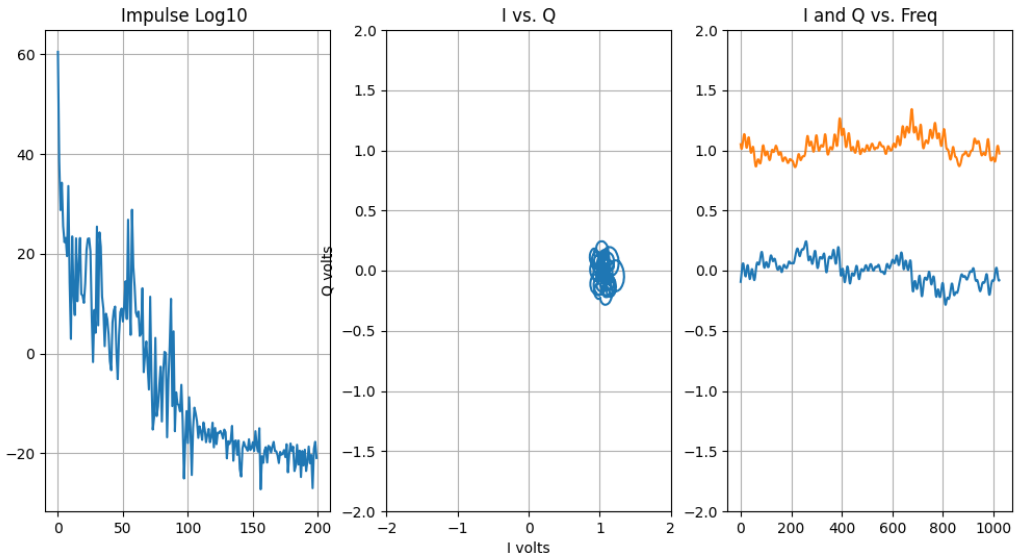


Figure 40 - Downstream, guessed to be water in coax.

Moving now to downstream data, in Figure 43 we see our first example which is guessed again be a case of water in the coax plant. Note the jagged plot on the right, and echo energy close in time on the far-left impulse response plot.

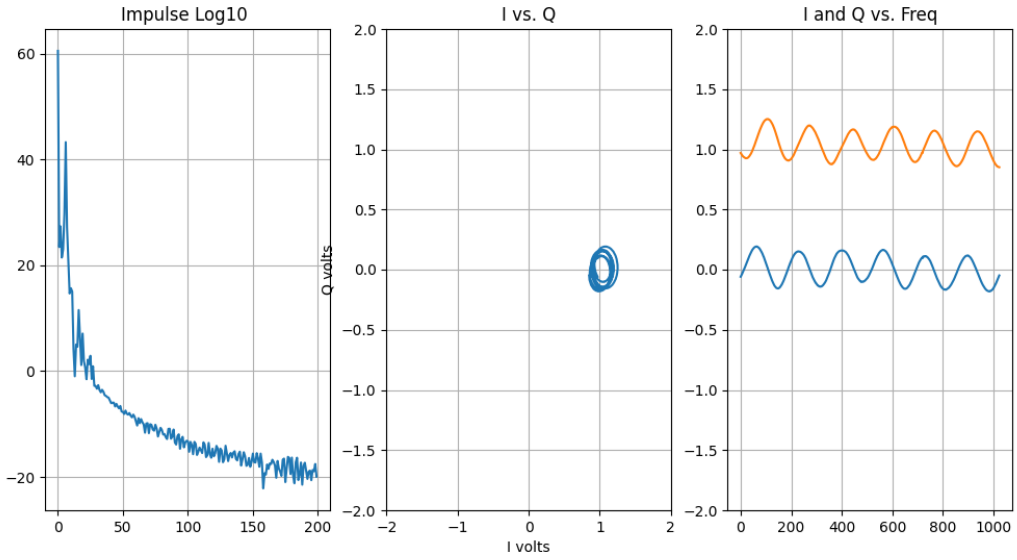


Figure 41 - Downstream lab test. CM is reporting channel response.

Figure 44 shows a confirming lab test, which we used to prove that the CM here was providing channel estimation data in the downstream, clearly showing the echo we inserted.

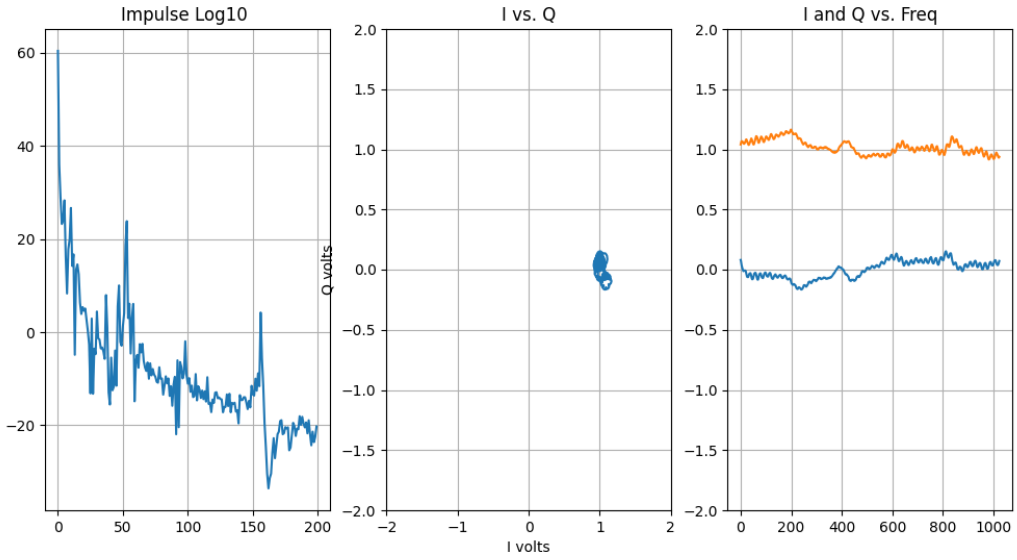


Figure 42 - Downstream, long echo tunnel so echo was weak. Source unknown.

Figure 45 is a channel estimation data plot of a long echo tunnel, which shows a weak response.

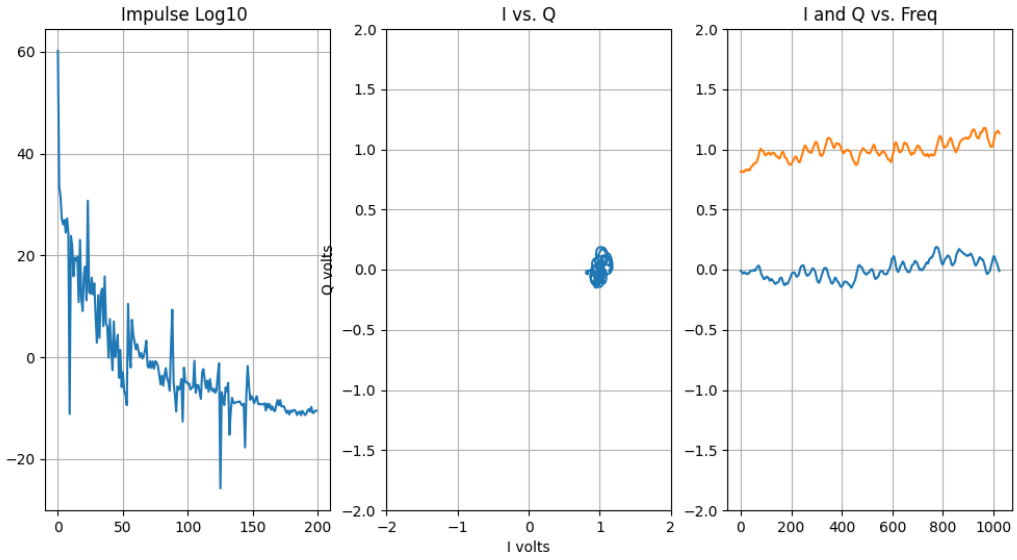


Figure 43 - Downstream, pair of long echoes.

Figure 46 shows a pair of long echoes in the channel estimation data.

On the Road to 10G - Converged Access Platform for HFC & Ultra Long (60+ km) NGPON2

A Technical Paper Prepared for SCTE by

Harj Ghuman
Principal
COX Communications
6305 Peachtree Dunwoody Road, Atlanta, GA, 30328
404-449-4711
Harj.ghuman@cox.com

1. Introduction

As the Cable industry investigates solutions for 10 Gigabits per second (Gbps) subscriber bandwidths and beyond, the convergence of the two main Multiple System Operator (MSO) network topologies in use today, Hybrid Fiber Coax (HFC) and Passive Optical Network (PON), becomes highly desirable. COX's current access platform, the Optical Communications Module Link Extender (OCML)[1], supports protected bi-directional transport of 10G Dense Wavelength Division Multiplexing (DWDM) optics over dual 5 to 60 km fiber links for Remote PHY Devices (RPDs), remote Optical Line Terminal (OLT) uplinks and business services. This paper describes COX's next generation converged access platform, the ROCML (Raman OCML), that is targeted for 10 Gbps subscriber bandwidth utilizing Data over Cable Service Interface Specification (DOCSIS 4.0) and/or long distance (60+ km) multi-wavelength 10 Gbps Next-Generation Passive Optical Network 2 (NGPON2). Extended Spectrum DOCSIS 4.0 (ESD) usage of a full 1.8GHz Radio Frequency (RF) spectrum requires 25 Gbps optical interfaces at the Remote PHY Device/Remote MAC-PHY Device (RPD/RMD). The ROCML's innovative combination of Raman, semiconductor and erbium doped fiber amplification (EDFA) transports both multi-wavelength NGPON2 and 25G DWDM optics over 60+ km optical links from the same integrated platform. It provides scalability to meet future bandwidth requirements such as multiwavelength 25G PON and Distributed Access Architectures (DAA) requiring >25G optical signals to RPDs/RMDs. The solution stays in line with COX's preferred methodology of keeping the outside plant passive. Hitherto, deploying PON over optical links >20 km has been a challenge requiring COX to deploy outside plant remote OLTs, either in cabinets or in clamshell enclosures.

The ROCML's long distance NGPON2 capability allows OLTs to remain in centralized indoor hub locations, creating a truly passive long distance PON, vastly improving reliability while reducing Op Ex and capital equipment costs. The ROCML's scalability to multi-wavelength 25G PON ensures a path to a future all Fiber to the home (FTTH network); it also allows for the transport of high capacity Coherent 100G - 400G optical signals, providing a powerful multi-purpose optical access platform for today's networks and tomorrow's requirements.

2. Data Projections

Figure 1 provides the twenty-year downstream data projections based on a 25 to 30% YoY CAGR. It shows the aggregate link usage for both DOCSIS 4.0 (D4.0) and 10G PON. Current COX RPD node size is approximately 350 homes passed, which with a 60% penetration rate gives approximately 200 subscribers per RPD. 10 Gigabit Symmetrical (XGS) PON deployments utilize 1:64 split (64 homes passed) per PON port, which using an average 60% penetration rates gives approximately 40 subscribers per link. For D4.0 to meet future data growth, node sizes can be reduced, or spectrum added such as in a 2x2 RPD, both requiring upgrades to the plant. 10G PON can meet projected data demand well into the future, with 25G PON meeting demand for the next twenty years. The maximum speed available to subscribers on either D4.0 or PON will be limited to half the link capacity, so with current D4.0 at 200 subscribers per link, maximum available speed will be 5 Gbps till 2027 and beyond that node sizes can be reduced, or spectrum added (2x2 RPD). On the other hand, 10G PON can support 5 Gbps till 2032 and beyond that a transition to 25G PON can be implemented to meet capacity requirements.

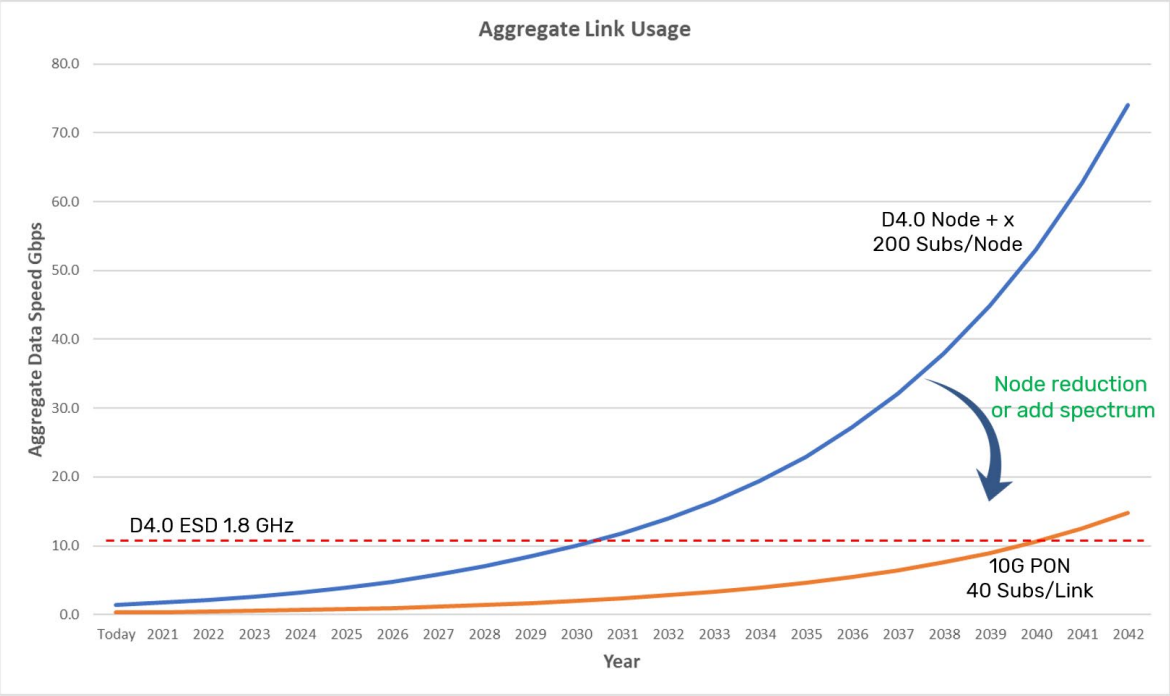


Figure 1 - Twenty Year Downstream Data Projections

3. ROCML Drivers

Some of the main drivers in developing the ROCML concept are given below:

- Long distance 60+ km 10G NGPON2, scalable to 25G PON
- 25 Gbps DWDM optical link for ESD 1.8GHz
- Optical transport over dual fiber links, switchable from primary to a backup fiber
- High system OSNR platform
- Coherent Capability 100G - 400G

The main challenge was in developing a solution which would provide long reach PON capability over 60 km and the inherent OSNR required to enable this. Since 25G DWDM was also an important requirement, dispersion management became an important factor given the inherent increased dispersion at high data rates. A move from 10 Gbps to 25 Gbps in Distributed Access Architectures (DAA) will be required in the next few years to meet projected subscriber demand. DOCSIS 4.0 ESD 1.8 GHz allows HFC networks to provide much higher bandwidths than current DOCSIS 3.1 which supports only 10 Gbps downstream and 1.5 Gbps upstream. ESD usage up to 1.8 GHz in DOCSIS 4.0 shows downstream rates exceeding 12 Gbps for a 1x1 node and 24 Gbps for a 2x2 node, hence a 25 Gbps optical input to RPDs will be needed. Next generation RMDs supporting ESD 1.8/3 GHz and 25 Gbps optical interfaces could meet projected bandwidth demand for the next decade. 25 Gbps DWDM direct detect Non-return-to-zero (NRZ) is planned to be used over 70 km links in the COX access network with the ROCML, with the assumption that RMDs will be able to support 25 Gbps optical interfaces. Figure 2 provides a breakdown of optical bandwidth requirements for DOCSIS 4.0 ESD 1.8 GHz.

Description	Data Gbps
Legacy 258 MHz – 1002 MHz	
2 OFDM Blocks @ 4096 QAM at 1.8Gbps	3.6
24 Ch DOCSIS 3.0	0.93
ESD 1002 MHz – 1794 MHz	
4 OFDM Blocks @ 1024 QAM	6
Total Data Payload	10.53
Total Data Payload (2x10.53)	21.06
Total Video Payload	1.4
IPv6 + L2TPv3 Overhead (5.3%)	1.19
TOTAL OPTICAL LOAD	**23.65**

Figure 2 - Optical Bandwidth Requirements For a 2x2 ESD RMD

4. Raman Optical Communications Module (ROCML) Overview

Figure 3 shows a block diagram of the ROCML which is intended to support ethernet bi-directional DWDM signals in the C-band over a single fiber and 10G PON transport utilizing the NGPON2 wavelength designation in a Point to Point (PtP) configuration. The ethernet transport can be 10G/25G NRZ direct detect signals or high capacity 100G-400G coherent transport. Dispersion compensation allows the 25G NRZ signals to be transported over longer distances. The ROCML provides switching capability to a secondary fiber in case of a fiber cut. Most COX access links are below 40 km, but a few links exist up to 65 km. Downstream amplification is provided by C and L-band EDFAs for the ethernet and PON signals respectively, while Raman pumps at the two output ports provide upstream amplification in the C band which is shared between the NGPON2 and ethernet signals. The Raman Mux DeMux (RMDM) is a passive integrated Mux/DeMux located in the outside plant and enables connectivity to either PON or remote MACPHY devices (RMDs). Eight PON ports allow 512 PON homes to be supported while twenty ethernet ports can be used to support RMDs or business services. The ROCML also incorporates Optical Time Domain Reflectometer (OTDR) ports for in-service (continuous) OTDR monitoring which enables instantaneous location of fiber cuts that drastically reducing time to repair.

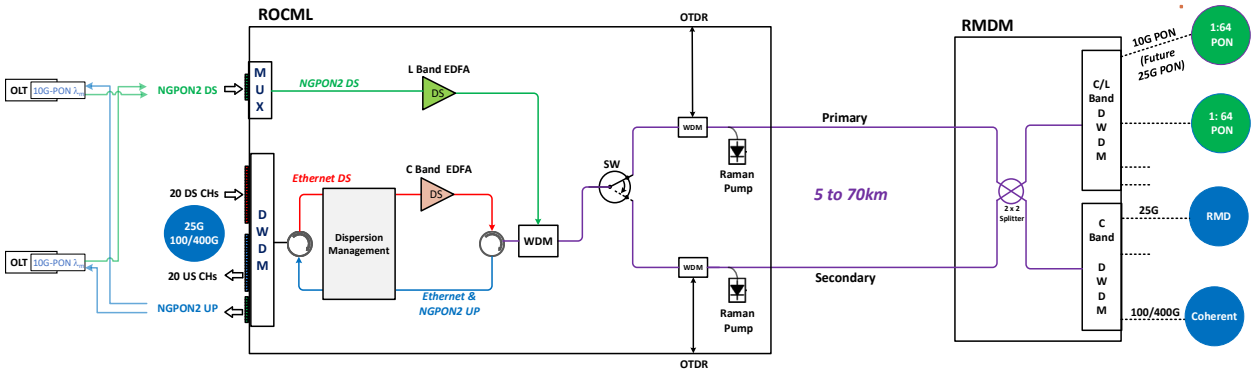


Figure 3 - ROCML Block Diagram

4.1. Optical Wavelengths

The ROCML - RMDM is designed for bi-directional transport over a single fiber and the wavelengths utilized are given in Figure 2 below. C-band DWDM is utilized for ethernet transport in both the downstream (DS) and upstream (UP) directions. For the NGPON2 signals, L-band is used in the DS and C-band for the UP signals.

ETHERNET					NGPON2				
	Downstream		Upstream			Downstream		Upstream	
Pair	ITU	Wavelength	ITU	Wavelength	Pair	ITU	Wavelength	ITU	Wavelength
1	14	1566.31	38	1546.92					
2	15	1565.5	39	1546.12					
3	16	1564.68	40	1545.32					
4	17	1563.86	41	1544.53					
5	18	1563.05	42	1543.73					
6	19	1562.23	43	1542.94					
7	20	1561.42	44	1542.14					
8	21	1560.61	45	1541.35					
9	22	1559.79	46	1540.56					
10	23	1558.98	47	1539.77					
11	24	1558.17	48	1538.98					
12	25	1557.36	49	1538.19					
13	26	1556.56	50	1537.4					
14	27	1555.75	51	1536.61					
15	28	1554.94	52	1535.82					
16	29	1554.13	53	1535.04					
17	30	1553.33	34	1550.12		NGPON2 L Band DS		NGPON2 C Band UP		NGPON2
18	31	1552.52	35	1549.32	1	71	1602.31	55	1533.47	UPSTREAM
19	32	1551.72	36	1548.52	2	72	1601.46	56	1532.68	1524-1544 nm
20	33	1550.92	37	1547.72	3	73	1600.6	57	1531.9	
					4	74	1599.75	58	1531.12	
					5	75	1598.89	59	1530.33	
					6	76	1598.04	60	1529.55	
					7	77	1597.19	61	1528.77	
					8	78	1596.34	62	1527.99	

Figure 2 - ROCML Optical Wavelengths

4.2. ROCML – RMDM Loss Budget

The ROCML - RMDM loss budget given in Figure 3 shows a 37 dB and 50 dB loss budget for the ethernet and PON signals respectively. The system goal was to provide minimum optical receive power levels in both downstream and upstream signals of -18dBm for the ethernet signals and -25dBm for the PON signals. 25G NRZ transport requires higher optical levels than 10G NRZ which can operate down to -22 dBm, hence the design ensures at least -18 dBm optical levels. 10G PON will work down to -28.5 dBm but since the ROCML will have to support 25G PON at a future date, it was decided to ensure at least -25dBm optical receive power levels for the PON signals.

PARAMETER	LOSS/GAIN			
	Ethernet		PON	
	DS C Band	UP C Band	DS - L Band	UP - C Band
Transmit Power (dBm)	1	1	7.2	6
ROCML	14.4	14.4	4	8.9
Fiber (65km)	14.3	14.3	15.6	14.3
RMDM	6.5	6.5	5.4	5.4
PON Splitter (1:64)			21	21
Margin	2	2	2	2
Total System Loss Budget (dB)	**37.2**	**37.2**	**48**	**51.6**
Gains (EDFA + Raman)	18	18	15	21
Optical Rx Power (dBm)	-18.2	-18.2	-25.8	-24.6

Figure 3 - ROCML RMDM Loss Budget

4.2.1. Optical Amplification

The ROCML uses conventional C and L-band EDFAs for the downstream ethernet and PON signals respectively. Upstream amplification is provided by Raman pumps located at both the primary and secondary output ports of the ROCML. The Raman pumps' power and wavelength are selected to provide sufficient amplification in the upstream C band used by the ethernet and PON signals. The Raman pumps also provide some gain in the downstream for both the PON and ethernet signals. The Raman pump powers/wavelengths are optimized to provide maximum gain for the PON signals and the Raman gain

profile for a 70 km link is given in Figure 4. This shows 21 dB upstream gain for the PON signals which is required to increase the PON loss budget from a typical 29 dB to 50 dB. An 18 dB gain for the upstream ethernet is sufficient to meet the loss budgets given in Figure 3. The Raman pumps also provide downstream gains of 7.5 dB and 5 dB for the ethernet and PON signals respectively.

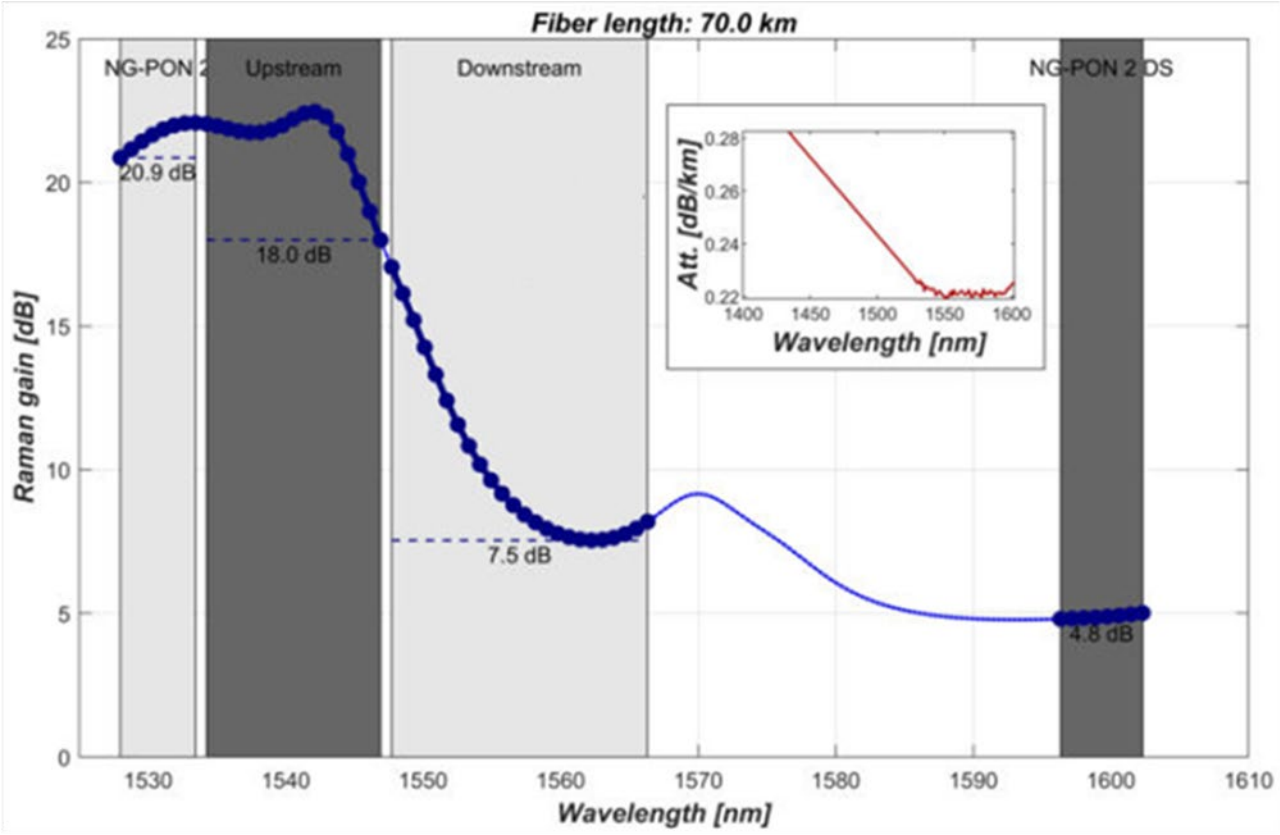


Figure 4 - Raman Gain Profile

4.3. Dispersion Management

Chromatic dispersion (CD) can cause laser pulses traveling down fibers to spread until they become unrecognizable as depicted in Figure 5. It occurs because different optical wavelengths/frequencies launched from the Tx light source travel at different velocities in a propagating dispersive medium, with the longer wavelengths traveling faster. Chromatic dispersion degrades the transmission quality and limits link distance and hence Dispersion Compensation Modules (DCMs) are added for Intensity Modulated Direct Detect (IM-DD) systems.

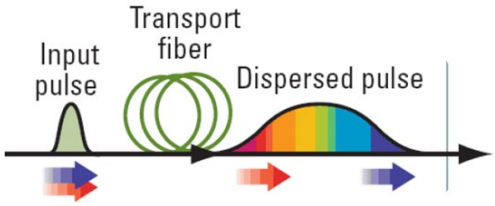


Figure 5 - Pulse Broadening Due to Dispersion

Maximum fiber link (L) due to chromatic dispersion CD is given by equation 1 below:

$Equation1:\ L = \frac{104{,}000}{CD*B^2}$

where

L: Distance in km
CD: Dispersion in ps/(nm*km)
B: Bit rate in Gbps

The maximum fiber links which can be transported for direct detect systems are tabulated in Figure 6 below using equation1. For 25G NRZ direct detect signals, links are limited to 10 km, so dispersion compensation modules (DCM) are employed in the ROCML to extend these links up to 70 km. The DCM can be made from Dispersion Compensating Fiber (DCF-DCM) or Fiber Bragg Grating (FBG-DCM). There are pros and cons with both dispersion compensation technologies. The DCF-DCM has a fixed amount of dispersion compensation but is bi-directional, so a single DCM can be used for DS and UP. The FBG-DCM can be tuned to the specific fiber length but are uni-directional and hence separate FBG-DCMs are needed for the DS and UP signals.

Data Rate Gbps	Maximum Fiber Link (km)
10	61.2
25	9.8
40	3.8
100	0.6

Figure 6 - Max link vs Data Rate

5. COX PON Architecture

COX current PON deployments utilize XGS-PON which is limited to fiber links of 20 km. XGS-PON utilizes L band downstream (1575-1580 nm) and O-band upstream (1260-1280 nm). The upstream wavelengths used have a minimum dispersion which is required since the Optical Network Unit (ONU) uses low-cost direct modulation and the resulting chirp combined with fiber dispersion fiber would cause performance degradation. The ROCML's innovative design manages fiber dispersion in the upstream, hence C band transmission as defined in the NGPON2 standard can be utilized over longer distances.

5.1. Current COX PON network

The current COX PON architecture (Figure 7) utilizes 10G ethernet uplinks from the current COX access platform OCML[1] to feed remote OLTs which are environmentally hardened, passively cooled, strand-mount devices. The OLTs can be mounted in a pedestal or on aerial strand and must be provided with plant power. A distribution split ratio of 1:64 is used which can be dropped to 1:32 to reduce contention.

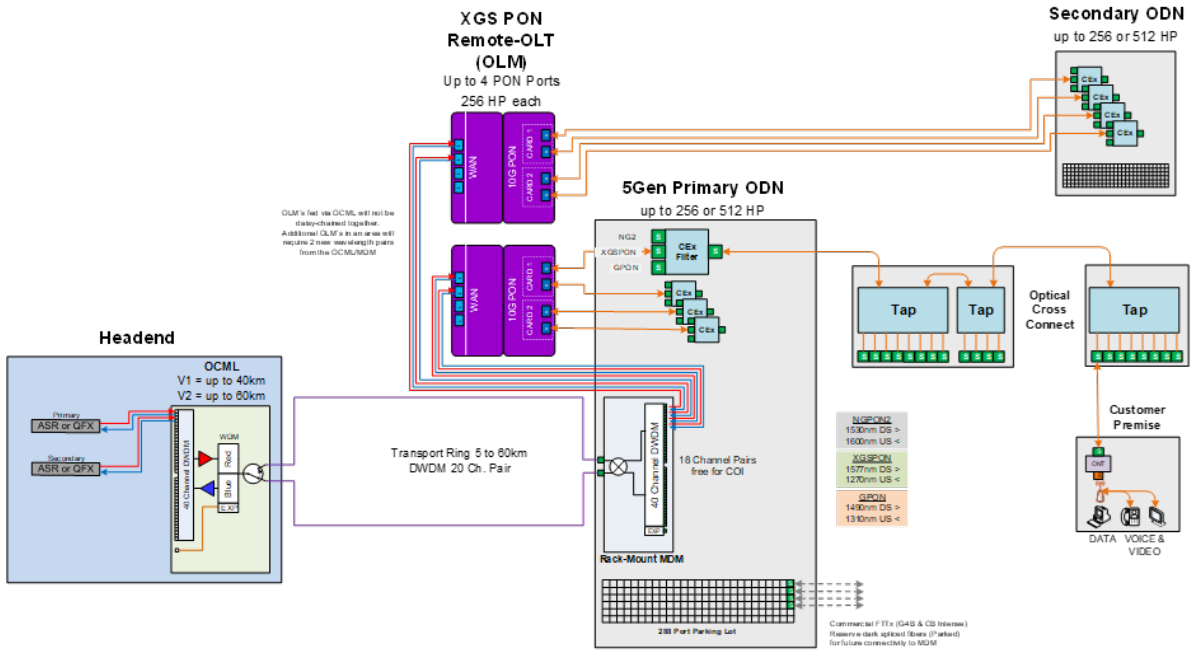


Figure 7 - COX XGS-PON Network

5.2. Proposed COX Next Gen PON network

The proposed PON network with the ROCML is shown in Figure 8, which shows centralized OLTs transported over a primary and back up secondary fiber which could be up to 70 km. Eight C/L band ports on the passive Mux/DeMux (RMDM) located in the outside plant can each support 64 subscribers for a total of 512. Larger split ratio of 1:128 can also be used over shorter links of 40 km (most COX networks) for a total of 1024 subscribers. This network creates a true passive optical network with all active components inside the headend, vastly improving reliability, network downtime and reducing OPEx.

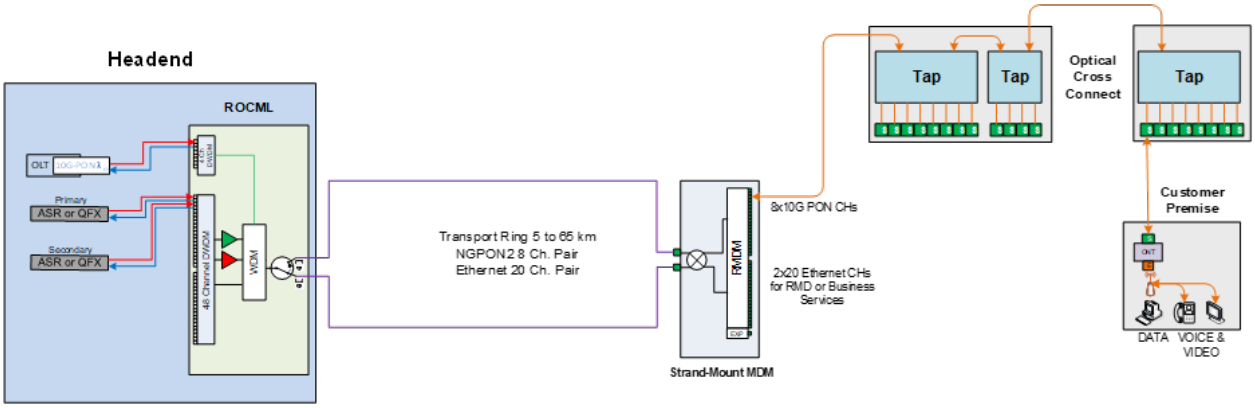


Figure 8 - ROCML 10G PON Network

6. 25G PON

25G PON is the next step in PON evolution driven by concrete demand and use cases, including 5th generation mobile network (5G), advanced enterprise applications, and wholesaling services. As shown in Figure 1, COX would need to move to 25 Gbps PON, possibly within this decade to meet projected data demand, especially if 10 Gbps were to be offered to subscribers. 25G PON Delivers a 250% increase in capacity over today's 10 Gbps XGS-PON and can deliver true 10 Gbps symmetrical services. It leverages mature, high-volume data center optical technology.

6.1. 25 Gigabit Symmetric Passive Optical Network (25GS-PON)

25-Gigabit-capable symmetric passive optical network (25GS-PON) is an optical access network for residential, business, mobile back/mid-haul, and other applications, operating over a point-to-multipoint infrastructure at 25 Gbps downstream and 10 or 25 Gbps upstream. The main features of the 25G PON network shown in Figure 9 are:

- 29 dB loss budget
- 20 km maximum fiber link
- O-band single wavelength
- NRZ transmission
- Low cost Directly Modulated Laser diode (DML) in ONUs
- No dispersion compensation
- No optical amplification

This solution would require remote OLTs as in the current COX XGS-PON deployments with high capacity ethernet uplinks.

Figure 9 - 25GS-PON Network

6.2. 25G PON Over ROCML

One of the drivers of the ROCML was to create a scalable long-distance 60+ km PON solution which could seamlessly migrate from 10G PON to 25G PON. To enable low-cost direct modulation NRZ to be utilized at the ONU, it was critical to be able to manage dispersion at the higher 25Gbps data rates and to have a high system OSNR. It is envisaged that the same NGPON2 wavelengths (L-band DS and C-band UP) could be utilized to transport multiwavelength 25G PON over the ROCML-RMDM system and extend the loss budget to 50 dB. The main features of the 25G PON "ROCML" network shown in Figure 10 are:

- 50 dB Loss budget
- 60+ km
- L/C-Band multiwavelength, 8 X 25G
- NRZ transmission
- Low cost DML lasers in ONUs
- Dispersion compensation
- Optical amplification (EDFAs and Raman)
- 512 subscribers over 60+ km, 1024 subscribers over 40 km
- Less hardware
- Increased reliability with no field actives
- Reduced OpEx

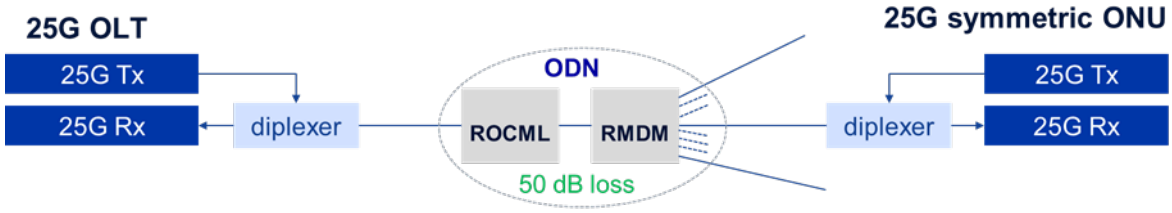


Figure 10 - 25G PON Over ROCML Network

7. Access Network Topology

The ROCML-RMDM is the optical access physical layer that connects the spine/leaf network located in the hubs/headends to the metro and backbone networks. It provides optical amplification to provide appropriate signal levels to the Remote PHY nodes for COX's HFC network. The current OCML is essentially a 10G DWDM optical physical layer which provides 10Gsignals to RPDs and remote OLTs.

The ROCML will transport 25G DWDM signals to Remote MACPHY nodes in a DOCSIS 4.0 ESD 1.8 GHz network and multi-wavelength 10G PON signals which can be transported over extended fiber links of 60+ km, obviating the need for remote OLTs in the outside plant.

7.1. DOCSIS 4.0 ESD 1.8 GHz Network

Figure 11 depicts a high level DOCSIS 4.0 ESD (1.8 GHz) network in the COX access networks in a node + x environment. Many devices will need to be upgraded and Figure 11 provides approximate numbers. To get further to an all-Internet Protocol (IP), DOCSIS4.0 10G network, the following need to be upgraded:

- Converged Cable Access Platform (CCAP) to Virtual Controller (or v mac manager)
- Converged Interconnect Network (CIN) to 25 Gbps network aggregation (via ROCML)
- Nodes need to be upgraded to Remote MAC Devices (RMDs).
- Nodes and amps need upgrades to ultra-high-split 1.8 GHz
- Taps need to be upgraded to 1.8 GHz (or better)
- D4.0 cable modems and gateways capable of ultra-high-split and 1.8 GHz

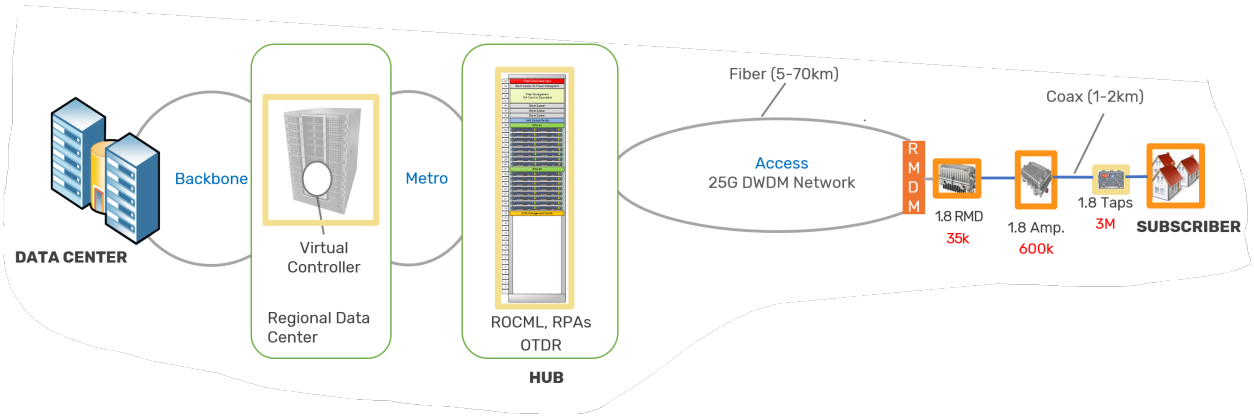


Figure 11 - DOCSIS 4.0 ESD 1.8 GHz Network

7.2. Fiber To The Home (FTTH) Network

Figure 12 shows how the ROCML - MDM can be utilized to create a true passive PON network and transport 10G signals to subscribers. The biggest Capex in a FTTH network will be trenching fiber to subscriber homes and some upgrades such as:

- CIN needs upgrade to ROCML
- Centralized OLTs
- ONUs at the subscriber

However, this is a one-time cost and future upgrades such as higher capacity 25G PON will only require changing equipment at the access fiber end points such as OLTs in the headends and ONUs at the subscriber homes. Such a true FTTH network requires less hardware, has no field actives, significantly improves reliability while reducing OPEx.

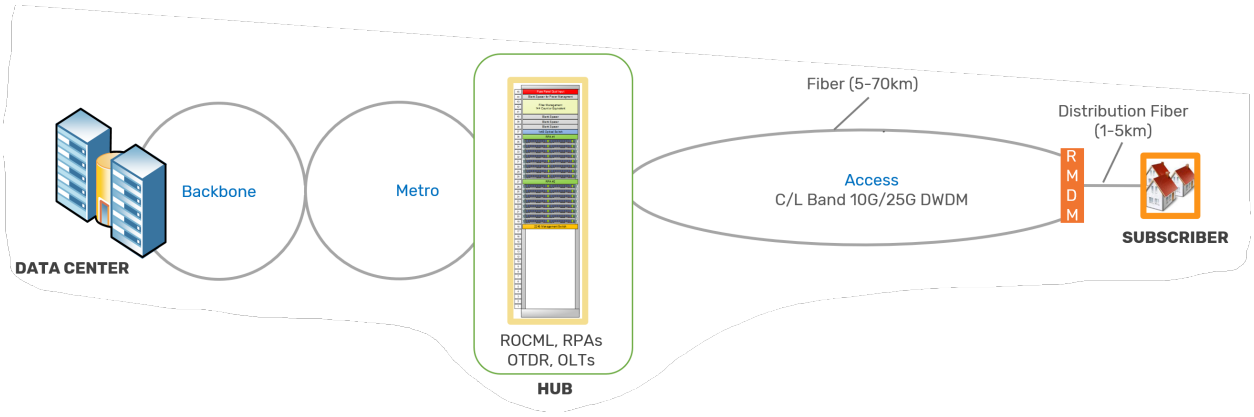


Figure 12 - FTTH ROCML PON Network

8. Conclusion

The ROCML - RMDM is a true universal access platform for the foreseeable future. It's a single integrated optical communication module which can be used for DOCSIS 4.0, long distance symmetrical 10G PON (scalable to 25G PON), 5G and enterprise services. Its innovative design allows for a 10G/25G DWDM direct detect transport over 60+ km fiber links to support RPDs or RMDs and its long-distance 60+ km PON capability creates a true passive optical network. The ROCML will allow COX to seamlessly transition from a DOCSIS 4.0 ESD 1.8 GHz network to an eventual FTTH network.

Abbreviations

5G	5th Generation Mobile Network
10G	10 Gbps
25G	25 Gbps
25GS-PON	25 Gigabit Symmetric Passive Optical Network
100G	100 Gbps
400G	400 Gbps
bps	bits per second
CCAP	Converged Cable Access Platform
CD	Chromatic Dispersion
CIN	Converged Interconnect Network
DAA	Distributed Access Architectures
DCM	Dispersion Compensation Module
DCF-DCM	Dispersion Compensating Fiber DCM
DML	Directly Modulated Laser diode
DS	Downstream
FBG-DCM	Fiber Bragg Grating DCM
FTTH	Fiber to The Home
DOCSIS 3.0	Data Over Cable Service Interface Specifications Version 3.0
DOCSIS 3.1	Data Over Cable Service Interface Specifications Version 3.1
DOCSIS 4.0	Data Over Cable Service Interface Specifications Version 4.0
D4.0	DOCSIS 4.0
DWDM	Dense Wavelength Division Multiplexing
EDFA	Erbium Doped Fiber Amplification
ESD	Extended Spectrum DOCSIS
FTTH	Fiber to the Home
Gbps	Gigabits per second
GHz	Gigahertz
HFC	Hybrid Fiber-Coax
Hz	Hertz
IM-DD	Intensity Modulated Direct Detect
IP	Internet Protocol
IPv6	Internet Protocol Version 6
L2TPv3	Layer 2 Tunnelling Protocol Version 3
MDM	Mux DeMux
MSO	Multiple System Operator
NGPON2	Next-Generation Passive Optical Network 2
NRZ	Non-Return-to-Zero
OCML	Optical Communications Module Link Extender
ODN	Optical Distribution Network
OFDM	Orthogonal Frequency-Division Multiplexing
OLT	Optical Line Terminal
OLM	Optical-Layer Management
ONU	Optical Network Unit
Op Ex	Operating Expenses
OSNR	Optical to Signal Noise Ratio

OTDR	Optical Time Domain Reflectometer
PON	Passive Optical Network
PtP	Point to Point
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency
RMD	Remote MACPHY Device
RMDM	Raman Mux DeMux
ROCML	Raman Optical Communications Module
RPD	Remote-PHY Device
SCTE	Society of Cable Telecommunications Engineers
UP	Upstream
XGS	10 Gigabit Symmetrical

Bibliography & References

1. Harj Ghuman, 2017. *DWDM Access For Remote-PHY Networks Integrated Optical Communications Module (OCML).* SCTE
2. Harj Ghuman, David Job, 2018. *Coherent Access Applications for MSOs.* SCTE
3. Harj Ghuman, 2020. *Next Gen 25G Access.* LightReading Cable Next-Gen Technologies & Strategies
4. Claudio DeSanti, et al, 2020. *Super-PON: An evolution for access networks.* Journal of Optical Communications & Networking
5. Zhengxuan Li, et.al, 2014. *Symmetric 40-Gb/s, 100-km Passive Reach TWDM-PON with 53-dB Loss Budget.* Journal of Lightwave Technology

Optimizing DOCSIS 3.0 Configuration in the Upstream through Applied Reinforcement Learning

A Technical Paper prepared for SCTE by

Kevin Dugan
Scientist 3, Enterprise Data Analytics & Data Intelligence
Comcast
1800 Arch Street, Philadelphia, PA 19103
719.493.2600
kevin_dugan2@cable.comcast.com

Maher Harb
Director, Data Science
Comcast
1800 Arch Street, Philadelphia, PA 19103
267.260.1846
maher_harb@comcast.com

Dan Rice
Vice President, Access Architecture and Technology
Comcast
1800 Arch Street, Philadelphia, PA 19103
720.512.3730
daniel_rice4@comcast.com

Robert Lund
Principal Engineer
Comcast
1800 Arch Street, Philadelphia, PA 19103
303.907.9690
robert_lund@comcast.com

1. Introduction

In 2020, Comcast deployed a Profile Management Application (PMA) system (1) for optimizing the DOCSIS 3.0 (D3.0) configuration across upstream channels in the network to achieve the proper balance between efficiency and robustness (fault tolerance). The established PMA system strengthens organizational objectives to increase network capacity, proactively respond to impairments, and support robust service optimizations for customers. At its core, the current system adopts a rules-based approach, in which a static policy (in the form of defined thresholds) for the different telemetry features (e.g., signal-to-noise ratio and codeword error rates) govern the choice of channel configuration.

Limitations within PMA exist when shaping telemetry thresholds to adapt to a wider range of environmental conditions. Currently, configurations are assigned to channels starting with conservative profiles and progressively moving toward efficient, yet less robust profiles. Further innovations within Comcast's PMA implementation focus on the delicate balance of applying intelligent, dynamic decision-making policies while preserving proper configurations for the diverse set of network devices.

A reinforcement learning (RL) approach for PMA allows, through experience, learning an optimal policy and therefore, enhancing the criteria used at various decision points. Simultaneously, RL simplifies policy management by consolidating permutations of telemetry boundaries into a single entity, called a 'state'. PMA efficacy improves with RL by reducing the latency of transitioning into optimal, efficient profiles, and doing so with increased confidence across varying network conditions (4). Inherent in this implementation is the risk reduction for operators to deploy more profile changes that maximize capacity without crossing the boundary that introduces disruption in service. This paper introduces a proof-of-concept RL-based PMA system along with performance study based on initial experimentation conducted in our laboratory.

2. Reinforcement Learning Design for PMA Systems

The purpose for implementing RL on upstream PMA is twofold: to manage PMA using a dynamic policy that continuously learns and to select optimal profile configurations with increased efficiency. An RL-based system provides the framework to facilitate these two objectives. A dynamic policy that is updated over each time step is more attuned to fluid network conditions than a static policy with fixed global thresholds. When a dynamic policy is updated, the decision criteria for choosing the optimal profile configuration under the current network conditions represents the best-known policy discovered by the system (2).

2.1. RL Concepts for US PMA

The RL sequence consists of an *agent*, selecting *actions* to take from a given *state*, and calculating the *value* of the action taken with respect to a *reward* system as the resulting state interacts with the *environment*. Figure 1 represents the sequence of interactions in a feedback loop. As the agent collects the rewards, it updates the policy using the *value function* (Equation 2) that satisfies the Bellman Equation (2) to continually refine the values for states-action pairs encountered. This process is called *value iteration* and is the basis of the dynamic policy.

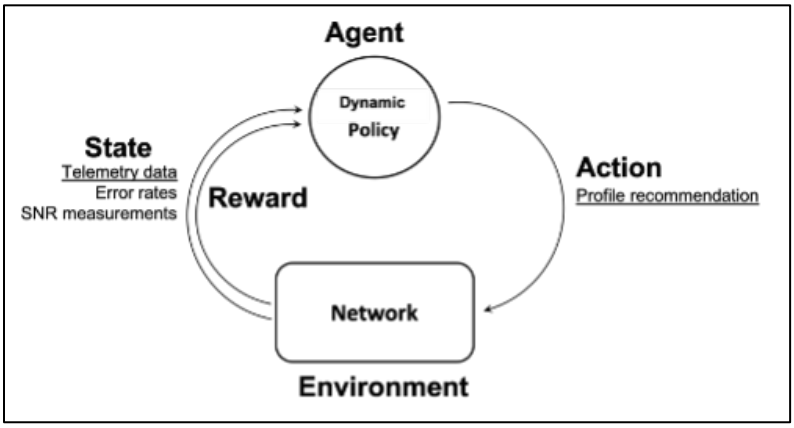


Figure 1 - RL Framework of the PMA System.

State-action values are numeric representations describing how valuable it is to take a given action from any state. They are calculated with respect to the next state-action pair, establishing the update process for sequential decision-making as Markov Decision Process (MDP). As the system builds experience, convergence occurs in the policy whereby updates to state-action pairs change the values over time in progressively smaller increments. The calculations of state-action values are based on the acronym SARSA. S, A, and R respectively denote State, Action, and Reward. With a subscript referring to the time step, SARSA can be represented with the following trajectory:

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, R_3, \ldots \quad (1)$$

In an MDP, the consequences of an action taken within the current state do not depend on maintaining the entire history of the trajectory; instead, just updating the state-action pair encountered maintains the legacy of experience already gained. The nature of the update considers how rewards are weighed, which influences how the policy is learned. For the proof of concept described in this paper, the Temporal Difference (TD) SARSA equation, $Q_\pi(s, a)$ in Equation 2 for online policy improvement is used to update the policy. The hyperparameters, α and γ, influence the learning rate and value of the future reward, respectively.

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha[R_{t+1} + \gamma Q\big(S_{t+1}, A_{(t+1)}\big) - Q(S_t, A_t)] \quad (2)$$

The α and γ terms are both floating point values between zero and one. α is a fixed learning rate where a value close to zero slows the learning and closer to one increases the learning rate. γ is a discount rate that determines how much to weigh immediate rewards and potential future rewards. As γ approaches one, the value of potential future returns influences the state-action pair estimates just as much as the immediate return; whereas a value closer to zero treats near-term rewards with more emphasis than future rewards. With respect to the US PMA effort, responding quickly to poor telemetry is critical in reducing customer impact. Finding an appropriate combination of these hyperparameters is an important objective for the problem at hand.

A dynamic policy is interesting to the PMA problem for its adaptive behavior that adjusts state-action pair values, and then uses those new values in real-time. With a dynamic policy in place, the RL system can adjust decision criteria under clean or adverse conditions. Consider a scenario where network impairments are causing high Uncorrectable Codeword error ratio (UCCW) rates and low signal-to-noise ratio (SNR). The RL policy, already having been exposed to this state, will recommend a profile change toward alleviating the side-effects of the larger problem (UCCW rate). If the issues persist, it is possible

for the RL policy to learn to take more aggressive action. The negative reward associated with the more conservative step will reduce its state-action value. Once an acceptable profile is reached, the state-action value is positively updated. At a point, if the cycle continues, the value of taking the aggressive action will become higher than the conservative action. Thus, the agent will then choose the aggressive action next time the poor telemetry is encountered.

In terms of α and γ, having parameters that reflect values associated with learning quickly is an intuitively logical approach. When an impairment occurs, the quicker the policy can adapt within the environment, the better decisions it will make. An example would be to set α to 0.8 and γ to 0.3. In this paper, multiple variations of the policy parameters are explored to inform of system short-term and long-term behavior implications associated with different policies.

2.2. RL Applied to US PMA System

Within the scope of an upstream PMA system, the states, actions, and a reward system were defined to represent characteristics of the current telemetry and configuration data to choose the optimal profile configuration for the present conditions (4). A state represents a distinct set of channel configurations and their associated network metrics. Each defined action is available for each state. The reward system is the mechanism used to influence how valuable it is to take an action from a state. As the system iterates over the timesteps, rewards are used to update the policy for the current action.

One key principle in RL systems is the trade-off between exploration and exploitation. Exploitation occurs when the highest-valued action previously encountered is taken from a state. However, that may not be the absolute best action to take for that state. To find the optimal action, the algorithm could select a random action to evaluate. When operating in production systems, random exploration is a risky endeavor because the policy could choose a transition that either introduces elevated UCCW rates, or transition to a slower profile when it is completely unnecessary to do so. In this POC, the evaluation of dynamic policies omits the exploration step and uses only exploitation to select the best actions available.

2.2.1. State

In the RL system for this paper, a state is considered a collection of telemetry metrics and configurations that are represented by discrete bins, with each state being unique. Continuous variables, like telemetry, were categorized by value ranges. Categorical variables fell naturally into their associated bins. Table 1 provides a breakdown of attributes to bins. In reality, only a small portion of the possible states will be encountered. As an example, not all CMTSs may be configured with six upstream channels and will never encounter any state where a fifth or sixth channel is represented.

A unique state is represented as a concatenated string of the bins as shown in Figure 2. In this example, the state of the channel on Remote PHY Device (RPD) HN1 is interpreted as a sub-optimal profile configuration that is experiencing a poor UCCW rate and needs to be downgraded to a more robust profile configuration.

Table 1 - Attributes of a State

Category	Attribute	# Bins
Telemetry	Uncorrectable Codewords (UCCW)	3
	Correctable Codewords (CCW)	3
	Signal to Noise Ratio (SNR)	6
Channel Configuration	Channel Width	3
	Modulation	5
	Profile Type	5
	Traffic Volume	2
	Channel Frequency	6
	CMTS	5
Total # Possible States		243,000

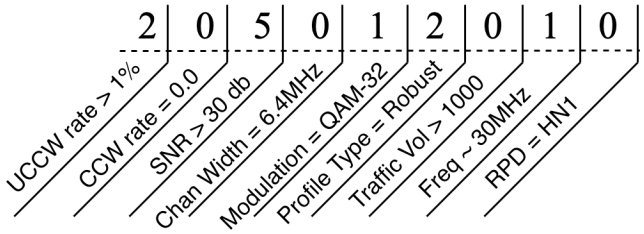


Figure 2 - Example of a Unique State ID

2.2.2. Actions

For each defined state, the complete set of actions is available to choose from when making profile recommendations. Like the states, many actions will never be encountered for certain states. For example, if the upstream channel was running on the optimal profile, there is no ability to move to a more efficient profile. Therefore, the state-action values for upgrades would all remain at zero.

In this RL system, three categories of actions were defined:

1. Upgrade or downgrade
2. Remain in the same configuration
3. Transition onto and off the most robust configurations, referred to as 'transient' profiles (these are designed to deal to dynamic impulse or burst noise)

Actions were limited to a maximum of four steps to restrict the dynamic policies from taking large action steps which would increase the likelihood of entering a poor state.

2.2.3. Reward System

The reward function serves to describe to the agent how it ought to behave. The key attribute to direct the agent in the simplest way is the UCCW rate. In this implementation, the reward system is boiled down to an evaluation of a Boolean condition that answers the question, "Is the UCCW rate greater than 1%?". If it is, a large negative reward (punishment) of -10 is given to the agent. Otherwise, the reward is $1 + profile\ speed\ gain$. A negative reward can still be observed if the UCCW rate is below 1%. For example, if a channel is in an optimal profile and the algorithm moves it to a transient profile, the loss of

speed is greater than the static bonus of +1 for keeping the UCCW rate low. This simple approach is designed to encourage upgrades in good conditions, and downgrades in poor conditions.

2.2.4. Policy Updates

The heart of the dynamic policy is the process of value iteration. Value iteration using TD SARSA from Equation 2 is the update algorithm that continues to adjust the state-action values until convergence (2). For each iteration, and each state and action encountered, the Bellman equation (2) is applied as an update rule in the form of Equation 2 and the profile recommendation logic uses the freshly updated policy in the same iteration.

3. System Architecture

3.1. Lab System

The lab system architecture was built off the existing Profile Management Application (PMA) Lab design. The primary functional groups consist of:

1. Impairment generation
2. RF switching matrix
3. Diplexing/combining components
4. Cable modem racks
5. Traffic generation
6. RPD nodes under test
7. vCORE and backoffice

The automation controller randomizes impairment profiles on the generator via an SCPI interface. It is also used to run GNURadio to create bespoke waveforms played back through a software-defined radio. The impairment sources are connected to an 8 x 16 port RF switch to steer or distribute the impairments to the appropriate devices under test. The impairments are combined into the appropriate points to feed either the RPD upstream burst receiver or the cable modems' downstream receivers. For the traffic generation loop, the network side interface is connected to the Distributed Access Architecture Switch (DAAS) and the CPE ports are VLAN'ed and connected to each cable modem on a high-density, mobile rack.

Five RPDs, with an average of 8 cable modems (CM), made up the population of devices for the trial. Each RPD contains a bonding group of either four or six D3.0 upstream channels, with a total of 24 upstream channels across the RPDs. Random impairments (none to severe) were introduced to evaluate the policies under both clean and adverse telemetry.

Table 2 - Lab Devices

RPD Name	Number of Channels	Number of CMs	Vendor
AS2	6	6	CMTS X
AN1	4	8	CMTS X
CN1	4	7	CMTS Y
HN1	4	12	CMTS Z
HS3	6	8	CMTS Z

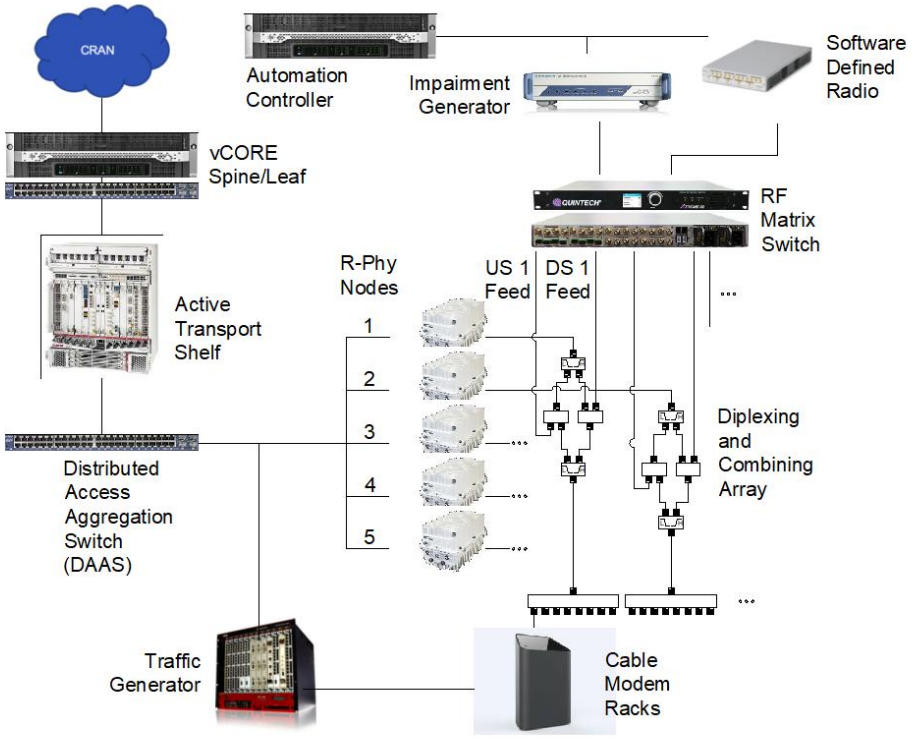


Figure 3 - Lab System Architecture

3.2. Data Pipeline

Each section of the pipeline is decoupled from the other sections to accommodate for experimentation and manipulation of policy parameters, states, actions, or other logic that would benefit from incremental changes. The sequence of the pipeline is shown in Figure 4:

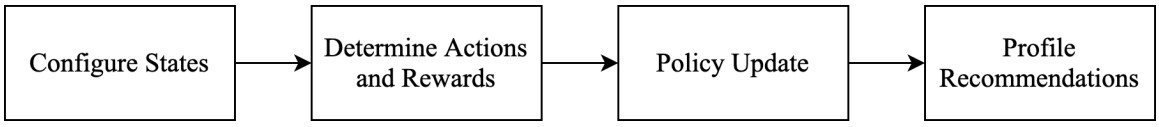


Figure 4 - Data Pipeline Sequence

The primary inputs are the attributes that make up a unique state (described in Section 2). Once collected, the current state can be assigned for each channel on the RPD. The data is associated with a timestamp that can be used, if necessary, to retrain a policy through the sequence of actions taken using historical data.

Arranging the data points by timestep for each upstream channel, the sequence of actions taken, and states encountered in order can be observed and rewards assigned. Being an independent step in the pipeline, this gives the opportunity to update the reward function or actions, and to be able to run the updates over the complete set of historical data or just the latest data points that need updated. In this pipeline, acquiring the action taken is not solely dependent upon the profile recommendation from the previous step, which makes determining the action taken delayed. A separate transaction manager must apply the recommendations first, and that function may or may not have succeeded.

In the policy update step, the latest timestep that needs to run through the SARSA equation with the full suite of the current state, action, reward, next state, and next action data is collected. The collection of all updated state-action pair values is preserved for use in recommending profiles.

Making profile recommendations is relatively trivial after the policy has just been updated; however, there are special cases to account for to either prevent abnormal behavior or change the degree of exploration desired. In most cases, the agent simply looks for the highest-valued action for the current state, also known as exploitation. Exploration is the selection of a random action a specified percentage of the time. In a production environment, either the exploration is very limited or not used at all. The profile recommendations in this paper for RL-based policies were all derived using exploitation – selecting the highest-valued action in all cases.

It is possible for the RL system to encounter states not seen before. When this happens, conventional logic is to select a random action. This could be detrimental to the channel's capacity on live CMTSs. In this described RL system, logic was added to directionally influence the profile recommendation based on the UCCW rate for the channel if a state is encountered for the first time. There is no need to downgrade a channel's profile if the telemetry is obviously adequate for consideration of upgrading.

3.3. Closed Loop

The closed loop system consists of the lab architecture and operation, the data pipeline, and the profile transaction manager. The lab generates telemetry and maintains configurations, which is processed through the pipeline, and the profile recommendations are applied onto the systems. Pipeline cycles are scheduled to complement the noise transition schedule in the lab by executing just prior to the noise transitions, ensuring that the next profile recommendations are based on the most current telemetry.

4. Building Dynamic Policies

Training a dynamic RL policy from an absolute baseline in this problem space would require more iterations over the pipeline than is realistically feasible to allow for adequate exploration of the state-action space in a live feedback loop. An approach to priming the state-action pair values in the policy is to leverage historical actions from the existing static policy. Starting from this point enables the policy tuning to occur in less timesteps, while also allowing for liberal exploration of the state-action space.

4.1. Initial Policy from Historical Data

The approach in using historical data for initially creating a RL policy follows concepts from imitation learning (IL) and inverse reinforcement learning (IRL). Both disciplines use a demonstration - a replication of behavior sequences in the problem space – to acquire experience and learn a policy. "The inverse reinforcement learning problem is to find a reward function that can explain observed behavior" (3). With a reward system in place, determining the right policy is a matter of exploring values of α and γ for the TD SARSA equation.

The static policy represents the human-generated data, taking directionally accurate (not necessarily optimal) actions under various network conditions. Doing so approximates state-action pair values toward their true values, which are refined over future value improvement iterations. The historical data is transformed into an ordered set of actions associated to states, as in Figure 5.

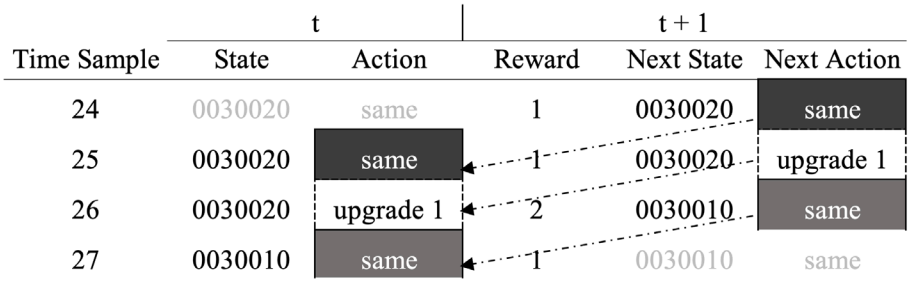


Figure 5 - Ordering Historical Data for Learning RL Policy

For each row, the $S_0, A_0, R_{t+1}, S_{t+1}, A_{t+1}$ observations needed to update the policy are arranged by windowing over the data and observing the sequence of profile transitions. This arrangement of historical data is the model used in future cycles when training RL policies using any decision-making policy. With the de-coupled implementation of the data pipeline, each step of data collections and transformations can be re-processed in their entirety. This becomes useful when adjusting a reward function or making any changes to the state or actions.

4.2. Tuning the Initial Policy

To build a true RL policy from the static policy's historical data, the RL policy itself is used to make profile recommendations and receive the resulting state condition through tens of thousands of state encounters. Over the iterations, the policy learns the highest-valued actions to take for many common situations using progressively restrictive variations in the trade-off between exploration and exploitation.

Initially, the policy was allowed to explore 100% of the time with the goal of starting to refine the state-action values. The exploration variable was decreased incrementally over time to the point where no exploration occurred to allow for the policy evaluation.

4.3. Multiple Policies

For each iteration, variations of policies can be trained on the actions taken by another policy. Values of α and γ are permuted to build multiple policies simultaneously. Since the parameters influence learning rate and reward weighting, the permutations of the parameters build policies that behave differently from one another, but still directionally appropriate.

5. Performance Study

The performance study focuses on comparing the static policy and the dynamic policies; specifically, evaluating profile speeds, profile sequences used, and policy responses to impairments. The objective is to identify a dynamic policy that can match or exceed the performance of the static policy.

5.1. Design

Every policy in the study will manage the profiles for five RPDs having either four or six upstream channels between 10.4MHz – 40.5 MHz along the spectrum. A minimum of 25 iterations for all five RPDs through the data pipeline will provide opportunities for the system to experience impairments from the randomized noise transitions from the lab. The dynamic policies are configured to not conduct any random exploration; they will behave in a similar manner as the static policy by always taking the best-

known action from a given state – even if the best-known action dynamically changes during policy updates.

Prior to each iteration, the profiles on all upstream channels are set to a baseline that mimics the configuration on a CMTS / RPD when it is onboarded into the PMA system. The channels lower on the spectrum begin in a transient profile configuration, while the upper 6.4MHz channels (above 27MHz) begin in sub-optimal profiles configured at modulation QAM-64. If the system has a sixth channel, it starts in a transient profile as well. For each dynamic policy pipeline run, updating the policy includes updating the other policies in consideration by applying the actions taken to different values for α and γ. In a sense, the decision-making policy is demonstrating actions to take for the other policies that get updated by evaluating decision-making policy's actions.

The four-channel configurations are all 6.4MHz wide, while the six-channel configuration adds a 3.2MHz channel below and a 1.6MHz channel above the four-channel configuration.

5.2. Policy Behavioral Expectations

Under good network conditions, the policies are expected to upgrade profiles for more efficiency and capacity until the most efficient profile is reached. The channel would ideally remain in the optimal profile if conditions are supportive. As UCCW rates climb over the 1% threshold, the policies are expected to downgrade the profiles to those more suited to handle noisy conditions. By doing so, the UCCW rate may recover at a point in the downgrade process, whereby the policy is expected to attempt periodic upgrades to check if the issue has been cleared. Otherwise, the policies are expected to learn to remain at a lower profile over time. This study was not designed to observe policy changes for long-running impairments, mostly due to time and resource constraints.

5.3. Performance Evaluation

Profile speed is a heuristic that can be used to represent the overall health of a bonding group (the collection of each upstream channels on each RPD). As the bonding group approaches the maximum profile speed, it is representative of the policy taking appropriate actions under good telemetry conditions. The maximum possible speeds are calculated by summing the optimal profile speeds for each upstream channel per system. Observing instances where maximum profile speed is not reached, determination of the cause falls into three categories: network impairments, poor decision by the policy, or an error external of the data pipeline (RPD channel error, transaction manager failure, etc).

UCCW rates are indicators where policies are expected to upgrade or downgrade profiles, if possible. Policies are expected to begin upgrading the profiles as conditions improve, downgrading under poor conditions, and remain in the best profile possible.

Achieving the best possible profile, as quickly as possible, is measurable by the number of timesteps it takes to achieve the optimal profile in a clean environment. To ensure a fair side-by-side comparison, the dynamic policies were limited to the largest transition step allowed in the static policy. The trajectory while transitioning off the baseline profiles is indicative of how assertive the policy is with respect to reaching the best-available profile.

5.3.1. Static Policy

The static policy has deterministic behavior, albeit through a combination of several thresholds. It was designed to reach optimal profiles by making transitions that are proportional to multiple telemetry metric

thresholds. Conservative decision criteria were purposefully built into the logic to greatly reduce introducing negative impact for customers.

5.3.1.1. Profile Speed Analysis

Observing the transaction history for each of the 25 timesteps, the static policy's ability to reach and maintain a steady state for the majority of RPDs indicates only few impairments were encountered. This allowed the policy to continue to use optimal profiles on many of the individual channels.

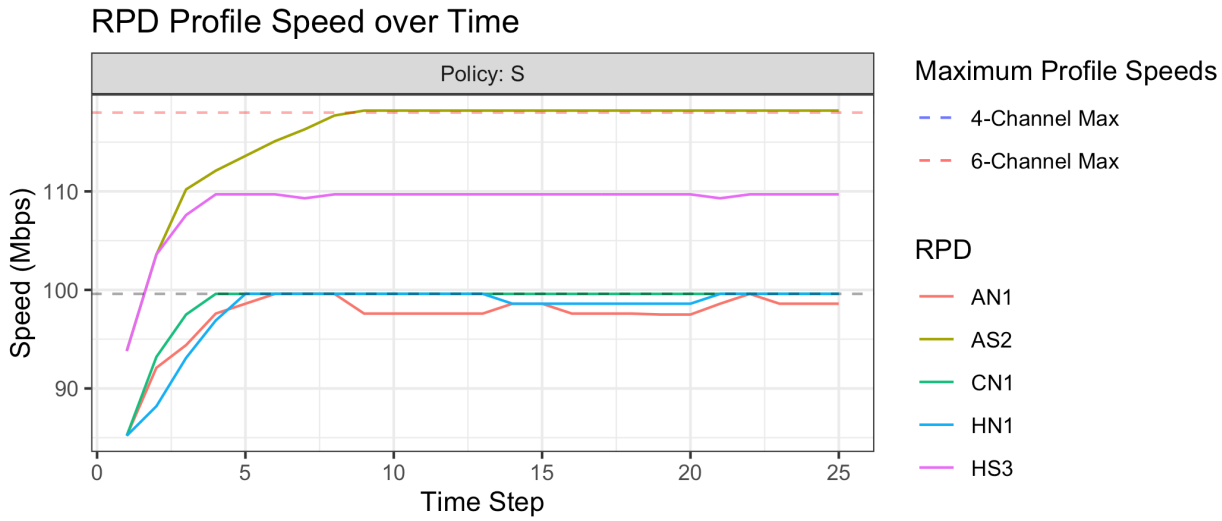


Figure 6 - Static Policy Profile Transitions in Terms of Speed

Under the static policy, the distribution of profile speed values is correlated to the UCCW rate values above and below 1%. Figure 7 affirms the behavior of the static policy achieving a steady state with mostly optimal profiles under the network conditions experienced in the iterations. AS2, CN1, and HN1 achieved the maximum possible speed for the bonding group and maintained during a significant portion of the study. A channel error occurred on HS3 that prevented the policy from achieving full speed for the bonding group. For that purpose, the policy achieved the maximum speed possible for the remaining five channels.

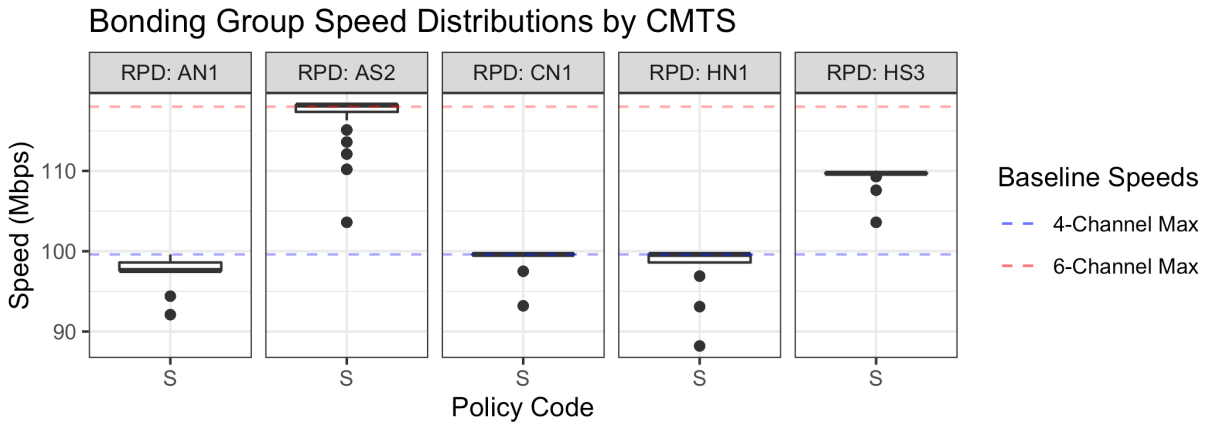


Figure 7 - Bonding Group Speed Distribution

Figure 8 illustrates how the RPDs were unaffected by the random noise generation from the lab. Despite experiencing majority clean UCCW rates, not all RPDs reached optimal profiles for all channels. A single channel on AN1 fluctuated between the top two profiles relating to reported SNR values and thresholds associated with SNR in the policy. Otherwise, the static policy behaved as expected, and under the conditions of the test environment, achieved the optimal profile available for the channel given the telemetry feedback.

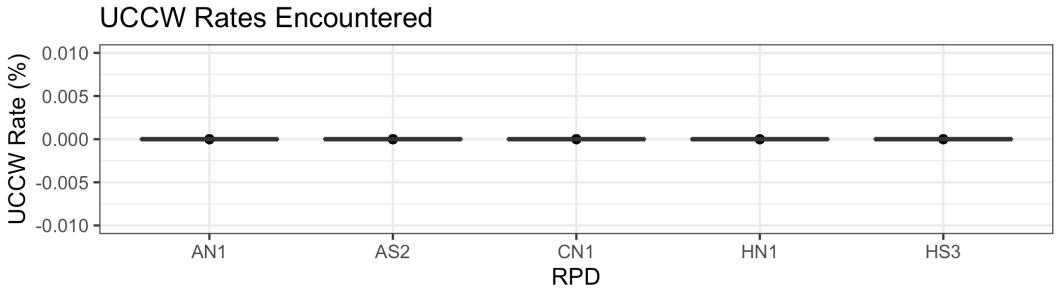


Figure 8 - UCCW Rates Encountered (Static Policy)

5.3.1.2. Latency to Optimal Profiles from Baseline

The number of timesteps the static policy took to reach a steady state operating on mostly optimal profiles for four-channel RPDs is represented in Figure 9. Omitting step zero, it took an average of four timesteps to get all three systems to a steady state on optimal profiles. Each RPD was upgraded at a slightly different rate, indicating a factor besides UCCW rate was affecting how large of a transition to make for each channel (SNR or CCW rates). The average SNR for RPDs CN1 and HN1 was below 30 dB for four of the six timesteps, while AN1 enjoyed an average SNR above 40 dB for the same timesteps. In spite of this, AN1 was the last to reach a steady state on optimal profiles.

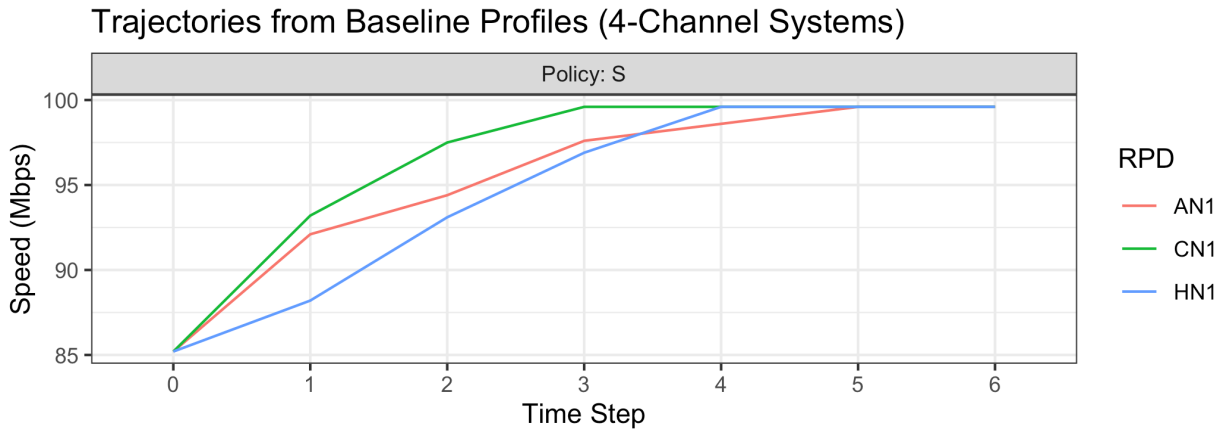


Figure 9 - 4-Channel Upgrade Trajectory from Baseline (Static Policy)

The six-channel RPDs demonstrated a wider disparity in terms of steps to reach a steady state. HS3 took two steps before leveling off, albeit notably by not upgrading the single channel that incurred reporting errors during the trial. AS2 reached a plateau after eight timesteps. This is attributed to the number of upgrades that had to occur for the upstream channel at 10.4MHz. Since it starts at the lowest possible profile, and limits upgrades to a maximum of four steps, this channel takes the longest to reach an optimal profile.

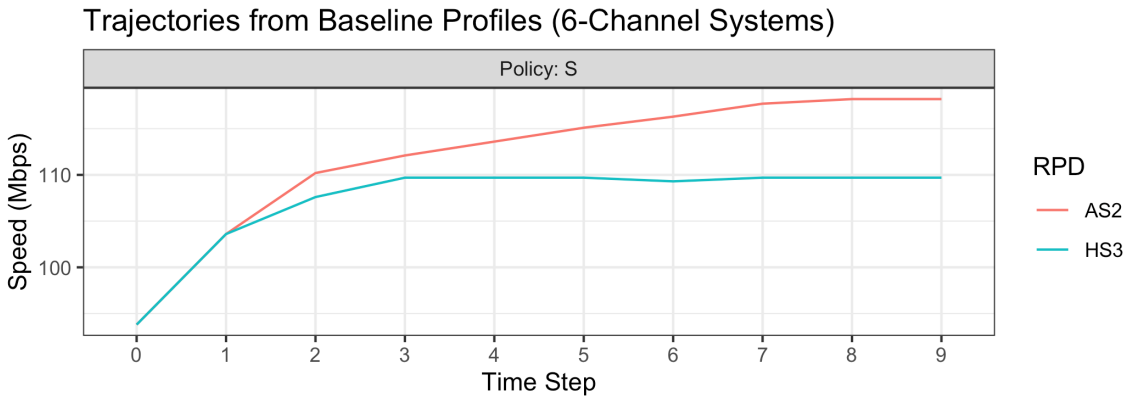


Figure 10 - 6-Channel Upgrade Trajectory from Baseline (Static Policy)

5.3.1.3. Summary Statistics

Over 25 iterations, the static policy achieved an average speed of 522 Mbps per iteration (all five RPDs), for an average bonding group (RPD) speed of 104.3 Mbps per iteration. Table 3 has the breakdown of profile usage:

Table 3 - Static Policy Profile Speed Metrics

Profile Type	% of Total Speed	% Profile Occurrences
Optimal	88.95%	85.15%
Sub-optimal	8.83%	8.44%
Transient	1.37%	2.24%
Below QAM-64	0.84%	4.17%

Also of interest is the relationship between which profiles were used in which locations along the spectrum during the trial. The transient profiles exist largely on the low end of the spectrum, as expected with clean telemetry throughout. Optimal profiles occurred often on the 6.4MHz channels from approximately 16MHz – 36MHz.

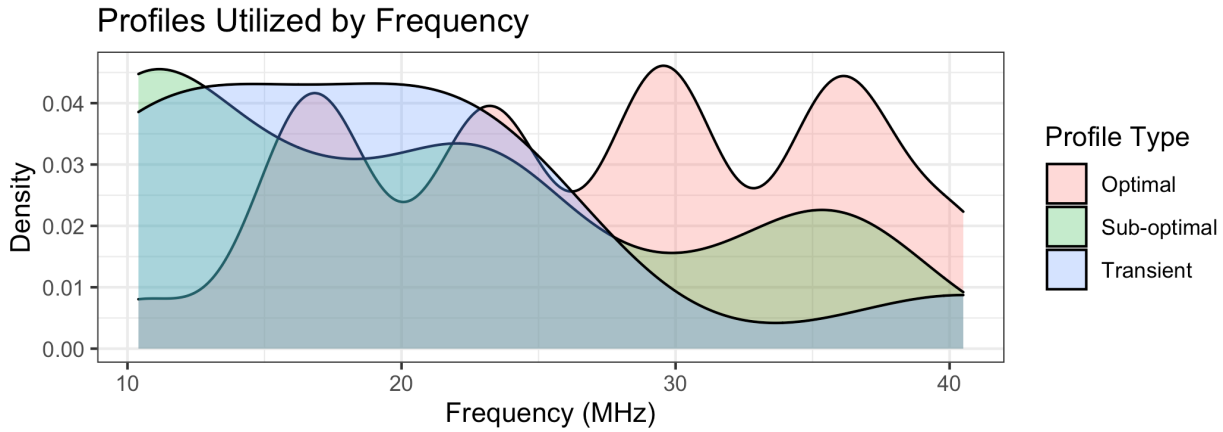


Figure 11 - Profile Type Utilization by Frequency (MHz) on Static Policy

5.3.2. Dynamic Policies

The four dynamic policies chosen for illustration in this paper were selected considering the values of α and γ that affect the learning rate and the weight placed on future rewards, respectively. Table 4 describes which values were used per policy, along with a high-level summary for how the policy generally behaved.

Table 4 - Dynamic Policy Information

Policy	α	γ	Policy Behavior
A	0.9	0.2	Achieved optimal profiles on 4 out of 5 RPDs
B	0.8	0.8	Indecisive, did not reach steady state, fluctuated profiles
C	0.3	0.8	Not as assertive as policy A, fluctuated profiles
D	0.8	0.2	Similar to policy A, optimal profiles on 3 out of 5 RPDs

5.3.2.1. Profile Speed Analysis

The dynamic policies experienced elevated UCCW rates on RPD HS3 during the trial. This prevented any of the policies from achieving optimal profiles in a steady form on that system. Policy A exhibited the most similar behavior as the static policy, maximizing the four-channel systems and reaching optimal speed on one of the six-channel bonding groups. Policies B and C have the most severe 'indecisiveness', fluctuating between profiles and largely not showing the ability to maintain a steady speed.

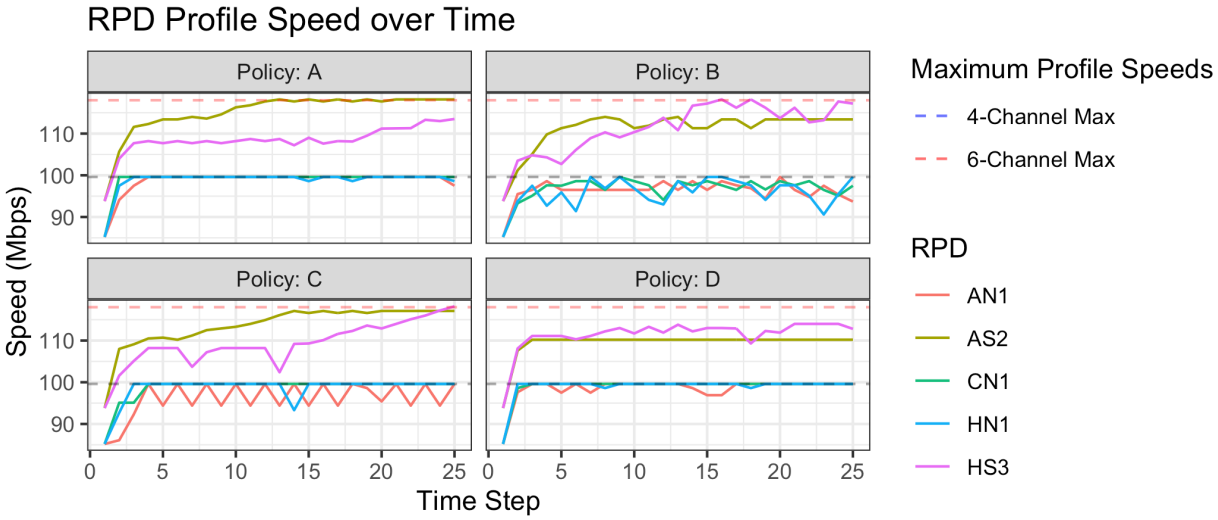


Figure 12 – RPD Profile Speed over Time (Dynamic Policies)

Figure 13 illustrates how the dynamic policies performed over the 25 iterations with respect to profile speeds distributions for each CMTS. The tight distributions observed on systems AN1, CN1, and HN1 indicate a lack of impairments over the iterations and affirm the observations in Figure 12. Certain policies quickly achieved optimal profiles and maintained throughout. The larger distributions are found in the six-channel systems, AS2 and HS3. Almost all the dynamic policies struggled to achieve consistent optimality on the two RPDs. Policy A had the most success and tended to be the most assertive of the policies.

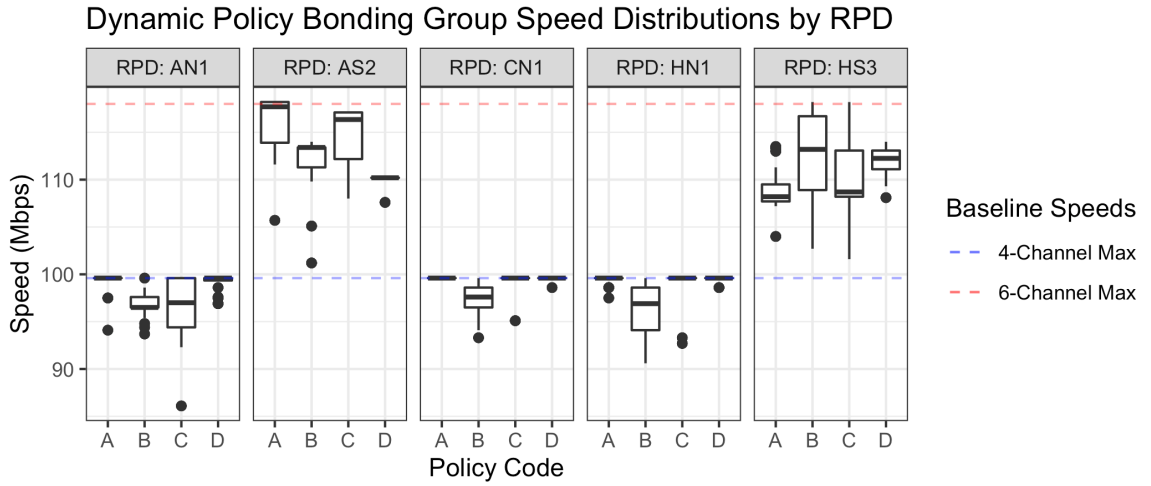


Figure 13 - Profile Speed Distributions by Dynamic Policy

UCCW rates did have an impact on system HS3 over the course of the trial for each of the dynamic policies. At one point, policy D experienced a channel with a severe UCCW rate of 100% (Figure 14). Since channels with severe impairments will not achieve the best profile, RL policies adapted to reaching profiles suitable for the impairment event.

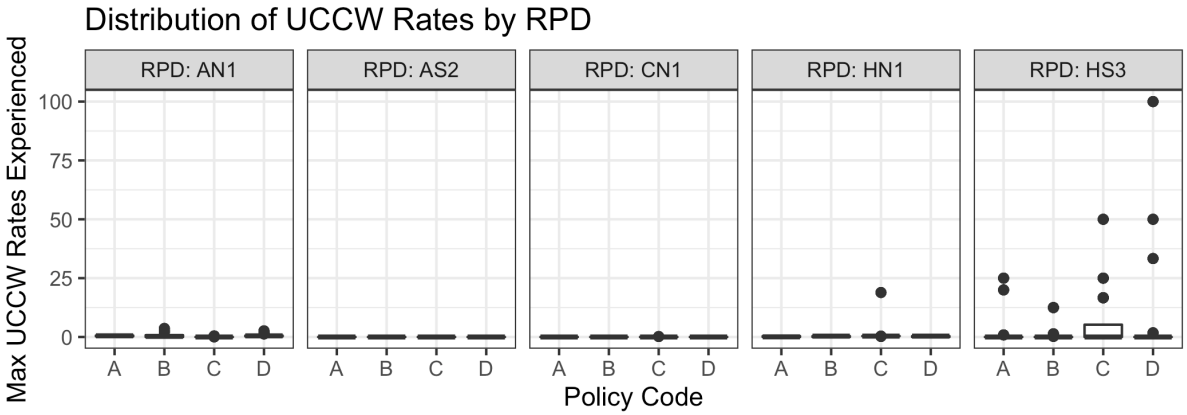


Figure 14 - UCCW Rates of RPDs by Dynamic Policy

Figure 15 highlights the UCCW rates observed on each iteration for each policy in relation to the overall profile speeds for each timestep on system HS3. In almost all cases, as a UCCW rate above 1% was detected, a decrease in profile speed for the bonding group indicates one or more channels downgraded to a slower, more robust profile. This is the expected behavior from both the static and dynamic policies. The change can be subtle for the channels less that 6.4MHz wide since those have less capacity to begin with, and the speed difference between profiles is a tighter distribution.

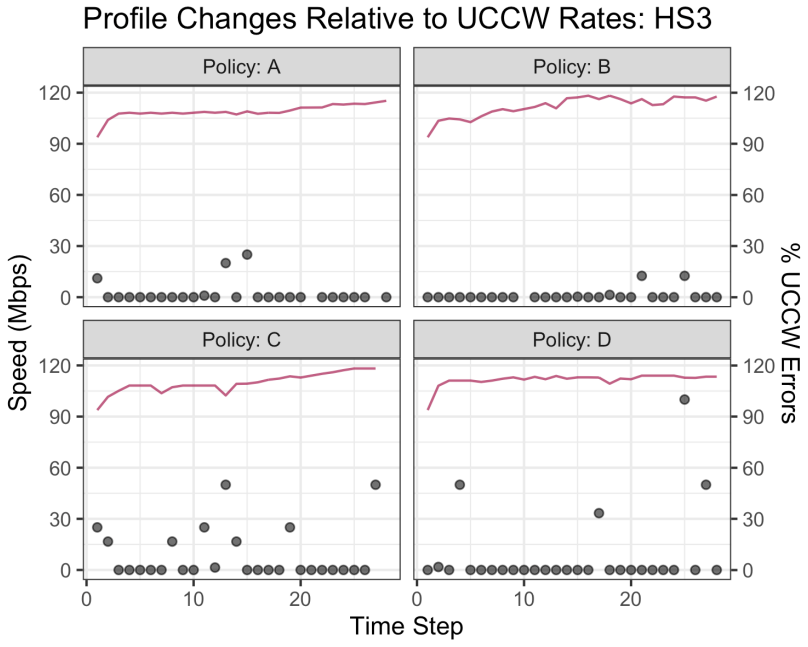


Figure 15 - Changes in Profiles with UCCW Rates on RPD HS3

As shown by policies A, B, and C, as UCCW rates improved, the policies continued to upgrade the profile configurations for more capacity. Observing the trends both with and without noise demonstrates the RL policies' ability to make directionally accurate decisions.

5.3.2.2. Latency to Optimal Profiles from Baseline

The dynamic policies took varying paths on the four-channel systems from the baseline profiles. Policy D took the most efficient route of the policies, reaching a steady state on optimal profiles in an average of two steps (actual = 1.7). Similarly, policy A averaged two steps, however, took a slightly less efficient route. Also notable, as policy D reached the optimal profile, a channel began fluctuating between the optimal profile and a two-step downgrade. This is a flaw within the policy. The same fluctuation activity is observed on policies B and C – both of which had more difficulty achieving and maintaining optimal profiles.

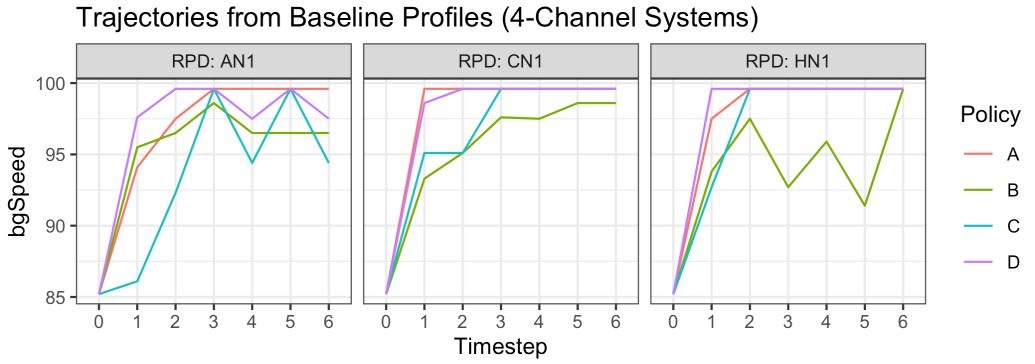


Figure 16 - 4-Channel Upgrade Trajectory from Baseline (Dynamic Policies)

The six-channel systems (AS2, HS3) had a more difficult path toward achieving optimal configurations. With the impairments on HS3, reaching the optimal profiles is not expected; however, reaching the best profile available is expected. The consistency between policies on HS3 indicates that impairments had a strong influence in keeping the profiles sub-optimal. The channels were influenced to use lower profiles to reduce the poor telemetry responses from the network in reaction to the chosen profiles. None of the policies on AS2 achieved optimal profiles, but the speeds were higher, indicating clearer telemetry. Policy D settled early after two timesteps, but at a sub-optimal overall speed. This indicates that the wrong action was valued highest for a particular state. The policy needs to explore different actions to overcome this.

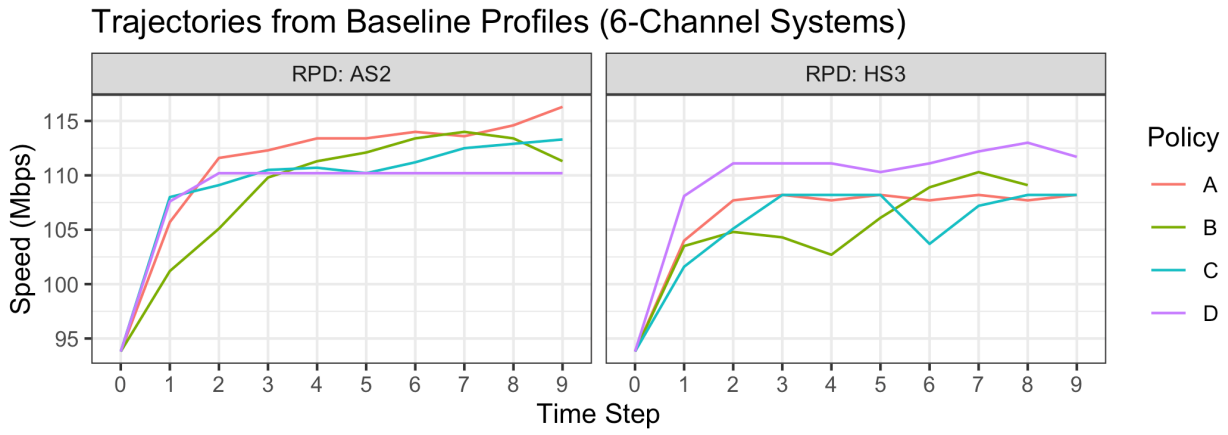


Figure 17 - 6-Channel Upgrade Trajectories from Baseline (Dynamic Policies)

5.3.2.3. Summary Statistics

Over 25 iterations, the dynamic policies achieved varying average bonding group speeds, as shown in Table 5 as an average speed per bonding group.

Table 5 - Dynamic Policy Raw Profile Speed Totals

Policy	Average per Bonding Group
A	104.0 Mbps
B	99.5 Mbps
C	103.2 Mbps
D	103.7 Mbps

Despite none of the dynamic policies achieving the 104.3 Mbps average, the measurement from the static policy, each of the dynamic policies experienced adverse UCCW rates that prevented the policies from upgrading into optimal profiles and remaining there. Yet, Policy A fell just short of attaining the static policy's benchmark.

Figure 18 has the breakdown of profile usage by profile type as it pertains to contributions to the overall speed values. Viewing the results by profile type is useful in understanding how the profiles were utilized across the iterations.

Notable in the profile density plot is the use of optimal profiles for policies A, C, and D in relation to the other profile types. Policies A and D exhibit the best consistency of optimal configurations.

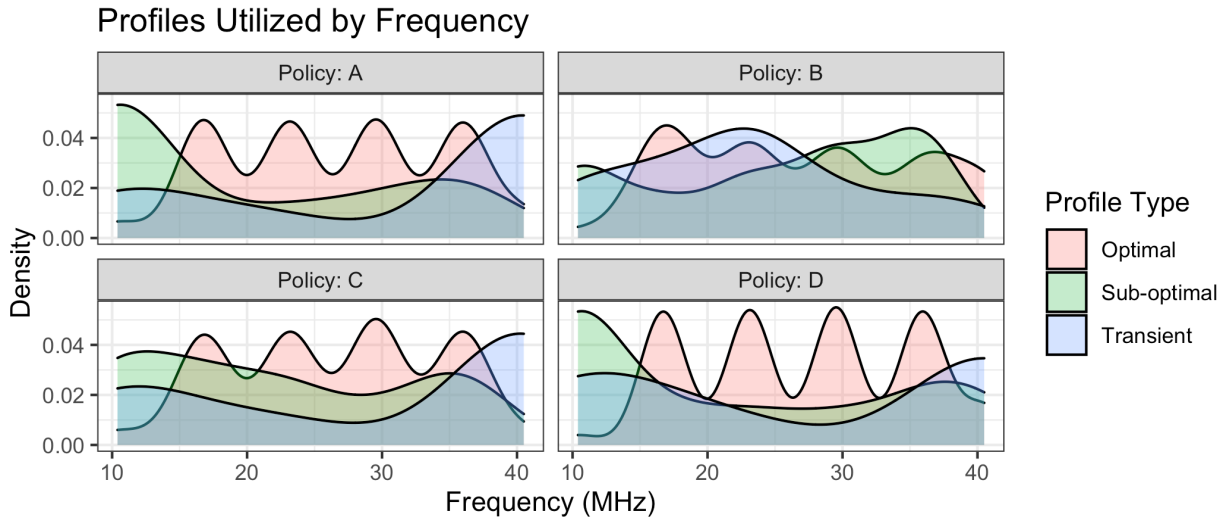


Figure 18 - Profile Type Utilization by Frequency (MHz) for Dynamic Policies

5.3.3. Summary Evaluation

Overall, policy A outperformed the other dynamic policies and will be the subject of a final extended study with respect to the static policy. Both policies ran for a total of 40 timesteps to provide additional insight into whether an RL based policy is definitively capable of being more performant than the static policy. Figure 19 shows the profile speed results comparing the two policies together over time.

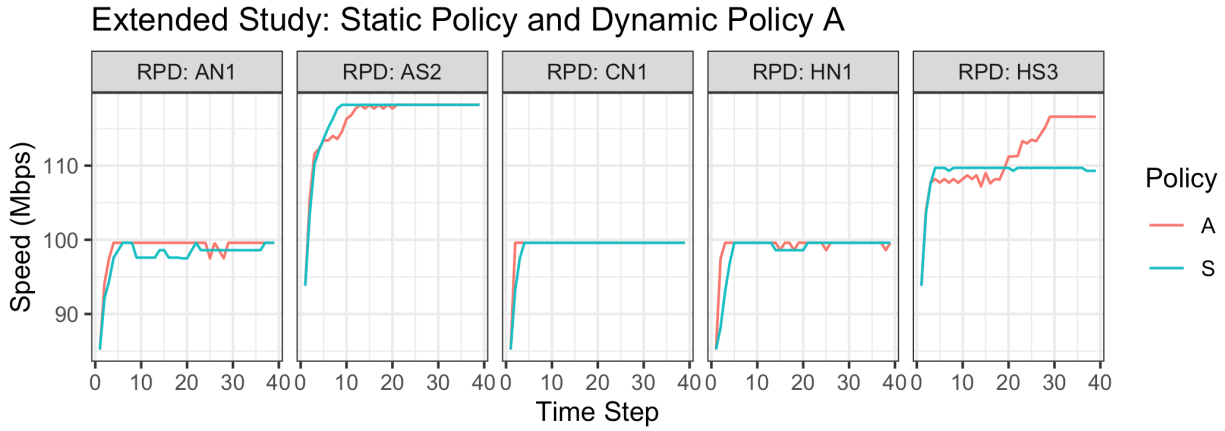


Figure 19 - Profile Speed over Time, Static vs Dynamic

A notable difference is on RPD HS3 when dynamic policy A runs were no longer encountering poor UCCW rates and saw a total increase of approximately 10Mbps from step 20 to step 30. The policy maintained that speed approaching the end of the study. The discrepancy of speed between the two polices is due to a channel reporting erroneous telemetry in the static policy iterations. It is unknown whether the static policy would have encountered poor telemetry or not as it upgraded, therefore it cannot be definitively proven by profile speed alone which is the better policy. Policy A increased to an average of 104.2 Mbps per bonding group of raw profile speed, while the static policy maintained an average of 104.6 Mbps. The high UCCW rates incurred on the systems during the trial runs for policy A explains much of the difference.

One other interesting point related to RL specifically occurred on RPD AS2. Figure 20 shows a zoomed in view of the change that took place. As policy A reached the optimal profile speed, a single channel had learned the highest valued action in that state was to downgrade by one step. Over four total cycles of fluctuating between the optimal profile and the sub-optimal one, the policy was getting updated by being penalized for moving down and rewarded positively by moving up. After those four cycles, the best action from the optimal profile for that channel changed to the action of remaining in the optimal profile under good telemetry conditions.

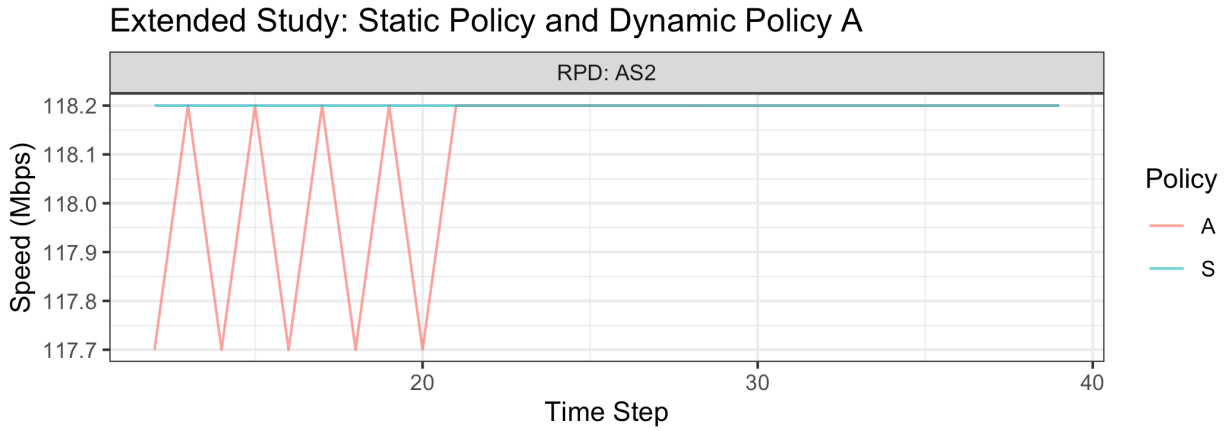


Figure 20 - Dynamic Policy Learning Better Action

Optimal profile configurations on policy A trials account for more of the overall aggregated speed than for the static policy. The static policy spent less time on transient profiles, however that is due to the differences in telemetry between the policy trials. Sub-optimal profiles – those that are between the optimal and transient profiles – accounted for almost twice as much of the total speed on the static policy than on dynamic policy A.

Table 6 - Profile Speed Metrics, Static vs Dynamic

	Policy A		Static Policy	
Profile Type	**% of Total Speed**	**% Profile Occurrences**	**% of Total Speed**	**% Profile Occurrences**
Optimal	91.67%	84.69%	88.95%	85.15%
Sub-optimal	5.11%	5.78%	8.83%	8.44%
Transient	2.42%	6.8%	1.37%	2.24%
Below QAM-64	0.79%	2.72%	0.84%	4.17%

Both policies have proven to behave very similarly in the interest of managing a US PMA system. Dynamic policy A only slightly edges out the static policy based on the following criteria:

- Policy A achieved optimal profiles on the four-channel systems from the baseline in an average of two steps, whereas the static policy took an average of four steps.
- Policy A utilized the optimal profile configurations more often than the static policy.
- Static policy achieved higher overall profile speed over the 40 iterations, with the caveat that the RPDs using policy A experienced more noise from the lab.

A key advantage the dynamic policy has over the static policy is the ability to learn on the fly. These policies are capable of correcting for changing conditions and adapting to what is normal for each RPD

system. As was demonstrated, policy A changed behaviors in the middle of the trial toward the more optimal action – this after experiencing several timesteps with poor telemetry.

While policy A, at a minimum, matches the behavior of the static policy, it also introduces the ability to build customized policies for individual RPD or CMTSs. Geography, weather, ingress, and other external factors impact network service to varying degrees – making a one-size fits all policy advantageous for many systems, but not applicable to all. Changes in one policy will not affect all systems, just the system for which it manages.

Balancing the many permutations of telemetry and configuration values using RL states removes complexities involved with tuning thresholds and applying the conditional logic in the algorithm for all thresholds.

5.4. Opportunities for Enhancement / Potential Future Steps

The proof of concept described in this paper is a promising step toward building more intelligent PMA systems. Below is a list of opportunities to improve the current policy building process and suggestions for future steps that would move this work forward.

- Training a dynamic policy from scratch – given enough time and resources, a policy trained from scratch would not be influenced by the initial historical data used in this POC.
- Synchronize noise settings from the lab such that each RPD and policy experiences the same sequence of impairments. Measuring responses of devices at each transition would provide a clearer picture of policy behavior differences.
- Improve the policy training process using n-step TD prediction methods, whereby more steps in the sequence and the rewards from those steps is used to estimate state-action values. This POC used the current state/action and next state/action to calculate values. Additional states and actions in the sequence can be bootstrapped for learning.
- Adaptation for scalability – architect solution capable of managing tens of thousands of RPDs. This may sound daunting at initial glance, however, the state-action space to maintain is significantly smaller than the collective set of possible states. In this study, .007% of the 243K possible states were encountered.

6. Conclusion

An upstream PMA system operating through a static policy is a proven effective strategy for configuring D3.0 channels across an entire network. By taking a cautious approach, the single policy caters more to the adversely impacted RPDs and CMTSs as the lowest common denominator when establishing thresholds that need to apply to a wide range of devices and conditions.

To improve profile configuration management, RPDs and CMTSs would benefit from a policy that best suits the individual operational environments. In fact, through the current implementation, a primary static policy manages the majority of Comcast's footprint, and a secondary policy manages a small set of devices with special requirements. Following a path of creating multiple static policies that manage different sets of devices would become difficult to manage.

One option for establishing self-managed dynamic policies is to apply an RL-based decision-making process that updates in real-time and is flexible enough to tune for either groups of devices or individual devices. The finer-grained management leads to confidently bolder profile transitions to reach steady state operations in less time than the static implementation. While both policies performed similarly in the trial,

the advantages a PMA system gets with an RL implementation may make the RL approach more appealing for large networks.

Abbreviations

CCW	correctable codewords
CM	cable modem
CMTS	cable modem termination system
CPE	customer premise equipment
DAAS	Distributed Access Architecture Switch
D3.0	DOCSIS 3.0
D3.1	DOCSIS 3.1
DOCSIS	Data Over Cable Service Interface Specification
dB	decibel
IL	imitation learning
IRL	inverse reinforcement learning
Mbps	megabits per second
MDP	Markov Decision Process
MHz	megahertz
α	alpha
γ	gamma
PHY	physical layer
PMA	Profile Management Application
POC	proof of concept
QAM	quadrature amplitude modulation
RL	reinforcement learning
RPD	Remote PHY Device
SARSA	state, action, reward, state, action
SCPI	Standard Commands for Programmable Instrumentation
SNR	signal to noise ratio
TD	temporal difference
UCCW	uncorrectable codewords
US	upstream
vCORE	core voltage
VLAN	virtual local area network

Bibliography & References

1. *Full Scale Deployment of PMA.* Harb, M, et al. s.l. : NCTA technical paper, 2020.

2. *Reinforcement Learning: As Introduction.* Sutton, R. S., Bach, F., & Barto, A. G.. s.l. : MIT Press Ltd, 2018.

3. *Algorithms for Inverse Reinforcement Learning*. Ng, A, Russell, S. Berkeley : s.n., 2000.

4. *A Reinforcement Learning Framework for Optimizing Throughput in DOCSIS Networks. Dugan, K., Harb, M., Rice, D.,* In Workshop on Flexible Networks Artificial Intelligence Supported Network Flexibility and Agility (FlexNets'21), August 27, 2021, Virtual Event, USA. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3472735.3473389

Additional References

ANSI C63.5-2006: *American National Standard Electromagnetic Compatibility–Radiated Emission Measurements in Electromagnetic Interference (EMI) Control–Calibration of Antennas (9 kHz to 40 GHz)*; Institute of Electrical and Electronics Engineers

The ARRL Antenna Book, 20th Ed.; American Radio Relay League

Code of Federal Regulations, Title 47, Part 76

Reflections: Transmission Lines and Antennas, M. Walter Maxwell; American Radio Relay League

Optimizing Value from Service Provider Wi-Fi in a Converged World

A Technical Paper prepared for SCTE by

Mike Darling
Principal Engineer
Shaw Communications
2728 Hopewell Place NE, Calgary AB T1Y 7J7
mike.darling@sjrb.ca

1. Introduction

Increasingly, subscribers want access to their broadband services at home, on the go and at their destinations. The value proposition of service provider Wi-Fi is to give subscribers access to their services when they are away from home at all types of destinations. In the absence of service provider Wi-Fi, subscribers must deal with authenticating to destination Wi-Fi on a case-by-case basis by asking for authentication codes, using their mobile data plans, or foregoing the use of their applications. Cisco estimates that the number of public Wi-Fi hotspots globally will quadruple from 2018-2023 [1]. Shaw operates a service provider Wi-Fi network in its cable footprint that is available free of charge to broadband wireline and wireless subscribers. This paper examines how subscribers use the service provider Wi-Fi network and seeks to answer some key questions around the value of the network, both to subscribers and to service providers.

2. Service Provider Wi-Fi Network

Shaw's service provider Wi-Fi network is made up of three types of access points (APs)—service provider APs expressly installed for service provider Wi-Fi connectivity, business Wi-Fi APs that are installed in business customer premises, and home hotspot APs installed in subscriber premises. Service provider Wi-Fi traffic is backhauled over fibre or coax to a wireless access gateway and into the core network.

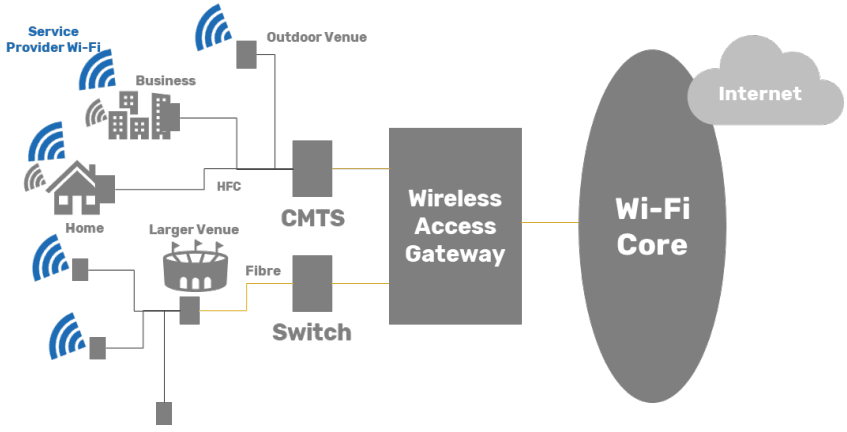


Figure 1 – Service Provider Wi-Fi Network

2.1. Types of Access Points

Service provider APs come in indoor and outdoor units with different backhaul options. The service provider network initially targeted the install of outdoor units that could be mounted on aerial infrastructure with power and backhaul provided via the hybrid-fibre coax (HFC) network. These APs are used to cover public areas such as parks or shopping districts. Indoor units were targeted at public spaces such as community centres and shopping malls and were backhauled via the HFC network or fibre optics, depending on the existing network infrastructure.

Figure 2 – Service Provider Wi-Fi Access Points

Business Wi-Fi APs are used for business Wi-Fi services, and secondarily used for service provider Wi-Fi. One or more SSIDs are set up for the business' use and service provider SSIDs are also broadcast for service provider subscriber use. These APs are used to allow clients of Shaw Business customers to use service provider Wi-Fi without needing to ask the business for authentication codes.

Home hotspot describes residential subscriber modems used for service provider Wi-Fi. In this case one or more SSIDs are set up for the subscriber and additional SSIDs are broadcast for service provider subscriber use. Similarly to business Wi-Fi, these APs allow service provider Wi-Fi users access to Wi-Fi in residential homes without needing to ask for authentication codes.

2.2. Network Access

Access to the service provider network is granted to all broadband and wireless subscribers. In terms of access, one difference is that the home hotspot network is only available to wireless subscribers. There are various methods of authenticating onto the network. Broadband subscribers can connect to the service provider network at an AP and enter their credentials into a splash page. Additionally, subscribers can add the MAC addresses of devices via a web portal, a feature that allows devices to be registered without being in range of the network. Wireless subscribers are automatically authenticated through their SIM cards. Once a device is authenticated it will automatically attach to the network when the device is in range. Broadband subscribers are allocated a 30Mbps downstream and 5Mbps upstream Wi-Fi speed tier, while wireless users have a 100Mbps downstream and 10Mbps upstream Wi-Fi speed tier.

2.3. Reporting Platforms

Shaw collects reporting data from several sources. Data is anonymized and aggregated to gauge performance of the Wi-Fi network.

Both APs and wireless access gateways send RADIUS accounting records with session data. The RADIUS accounting functions allow records to be sent at the start and end of sessions indicating the resources used during the session [2]. This analysis specifically looks at stop records, which are only sent at the end of a session.

To gather application-level statistics a deep-packet inspection (DPI) system was used. This system has a catalog of application signatures to which user traffic is mapped. The DPI system was also used to gauge bitrate statistics at the modem level for Wi-Fi APs backhauled over DOCSIS. This system takes 256ms data consumption samples and summarizes them into one-minute bitrate statistics. These statistics include average, mean, median, 95th percentile and peak throughput.

2.4. Subscriber Feedback

Shaw has undertaken consumer surveys of broadband and mobile subscribers to gauge the sentiment towards service provider Wi-Fi and to verify whether subscriber perceptions match the intended value.

81% of broadband subscribers and 86% of dual broadband and wireless subscribers who responded were aware that they had access to service provider Wi-Fi. Service provider Wi-Fi was ranked third in terms of factors impacting a subscriber's decision to keep their broadband services with Shaw, with 80% responding that it was impactful.

57% of respondents noted that they frequently or occasionally use service provider Wi-Fi, which ranked higher than other sources of Wi-Fi such as restaurants, coffee shops, and retail stores. Given the same options for connecting to Wi-Fi, 81% rated the quality of service provider Wi-Fi as good or very good, a higher percentage than all other options.

Overall, 86% of broadband subscribers and 95% of wireless subscribers who responded to the survey answered that they find value in service provider Wi-Fi.

3. Network Usage

Service provider Wi-Fi is meant to provide subscriber value. Anecdotally however, there had been reports that subscribers authenticating to the Wi-Fi network were unable to connect to the Internet. Additionally, there was a concern that individuals not subscribed to Shaw services were accessing the network. This section explores these issues, general network usage and performance and whether the network was being used as intended. The data presented in the following sections was gathered in 2019 and 2020, before the start of the COVID-19 pandemic.

3.1. Monthly Consumption

Average monthly Wi-Fi consumption was calculated per device for both wireline and wireless subscribers. As shown in Figure 3, average monthly Wi-Fi consumption was nearly identical for the two groups at approximately one gigabyte per month. One gigabyte is a small amount when compared to wireline monthly consumption but is comparable to wireless monthly consumption.

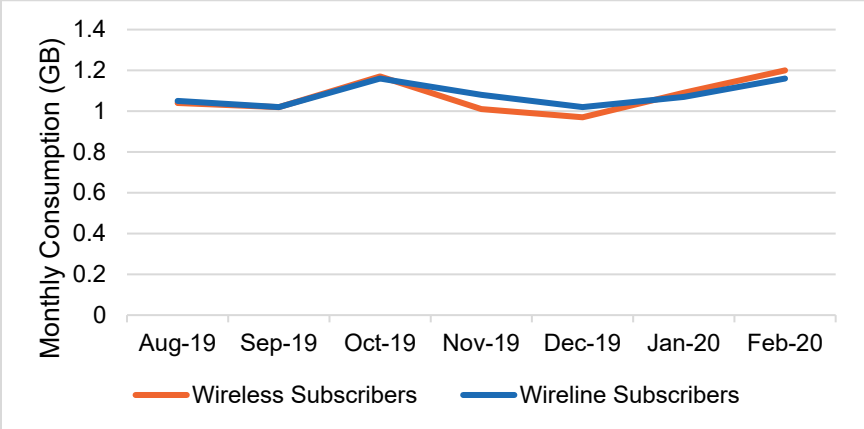


Figure 3 – Average Monthly Wi-Fi Consumption

Monthly Wi-Fi consumption averages hide the large disparities between subscribers. Although these disparities are observed in the wireline network, they are more pronounced in the Wi-Fi network, as shown in Figure 4.

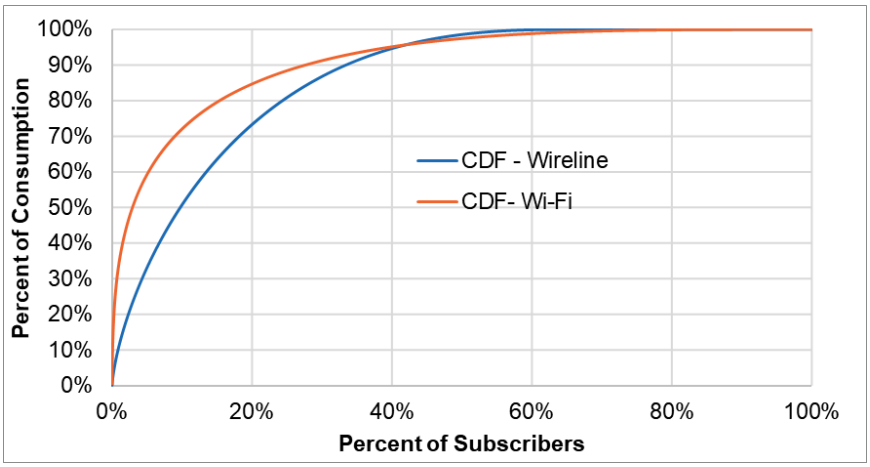


Figure 4 – Wi-Fi and Wireline Consumption Cumulative Distribution Function

This chart shows the cumulative distribution function (CDF) of Wi-Fi and wireline consumption. It can be observed that 20% of subscribers on the wireline network account for approximately 75% of consumption, while 20% of subscribers on the Wi-Fi network account for approximately 85% of consumption. There is no pronounced difference in the number of subscribers who account for little or no consumption, which would be expected if users were authenticating but not connecting to the Internet.

Figure 5 shows that there is an even spread in the percent of devices using Wi-Fi for a specific number of days per month. The data shows that 4.5% of devices access the network only a single day per month, while 4% of devices access the network every day of the month. This pattern fits with destination use because the number of days a subscriber visits destinations, such as restaurants, varies person to person. In contrast, subscribers access the wireline and wireless networks all or most days of the month.

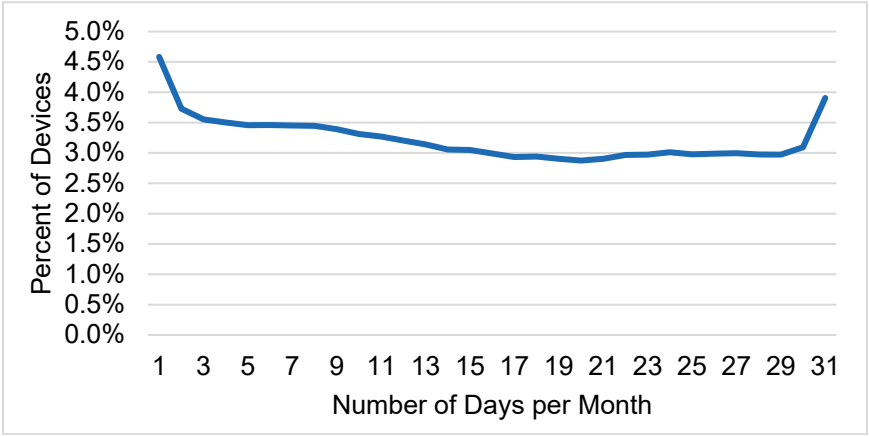


Figure 5 – Wi-Fi Days of Use per Month

Wi-Fi consumption was also compared to wireline and wireless consumption for individual subscribers. While wireline consumption is much higher, there is a correlation between high wireline consumption and high Wi-Fi consumption, as demonstrated by Figure 6. Wireline consumption is shown on the Y-axis in

gigabytes while Wi-Fi consumption is shown on the X-axis in megabytes. Note the logarithmic scales on both axes. The size of the bubbles indicates number of subscribers.

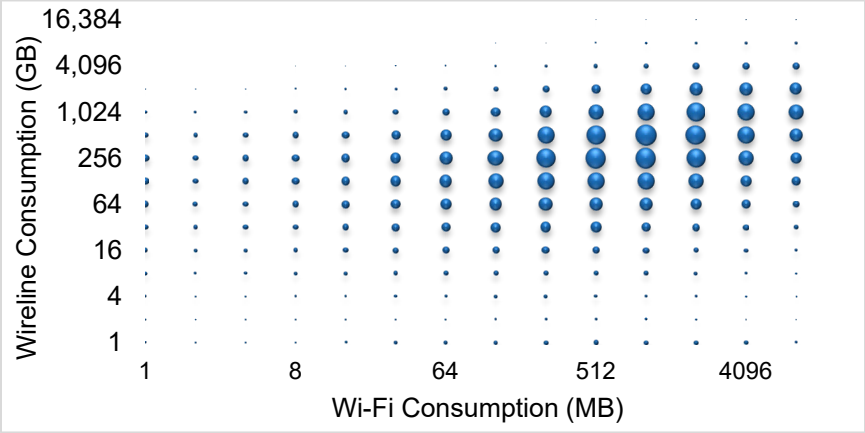


Figure 6 – Wi-Fi vs Wireline Consumption

This result points to Wi-Fi consumption being complementary to—rather than a substitute for—wireline consumption. There is not a significant number of subscribers with heavy Wi-Fi consumption and low wireline consumption.

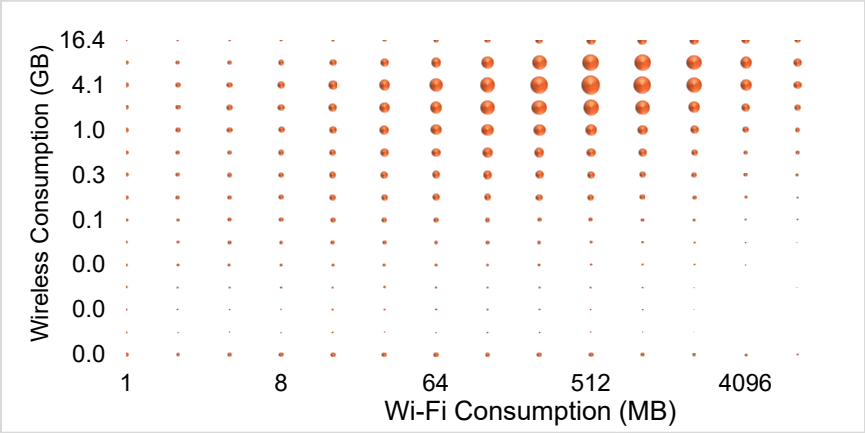


Figure 7 – Wi-Fi vs Wireless Consumption

Monthly consumption for wireless subscribers who used the service provider Wi-Fi network was also analyzed. The result showed a similar correlation to the wireline/Wi-Fi comparison in that subscribers who had high monthly wireless consumption tended to also have high monthly Wi-Fi consumption.

Wireline subscribers can register multiple devices on the Wi-Fi network, with limits set depending on which broadband tier they are subscribed to. Although there is potential for subscribers to allow people outside their household to access the network, this is not seen in the data. 70% of accounts have only one or two registered devices, which roughly matches the number of people per household.

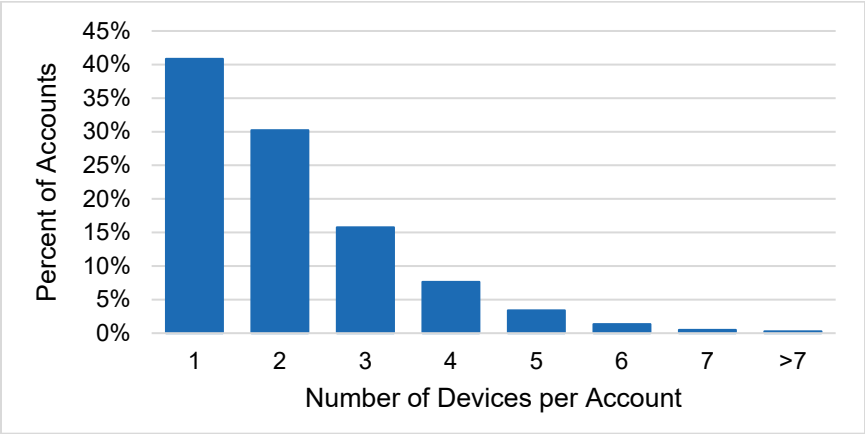


Figure 8 – Number of Devices Registered per Account

3.2. Bitrate Statistics

Average and 95th percentile bitrate is measured in the downstream and upstream every minute for a month for each DOCSIS backhaul modem. From those measurements the monthly 95th percentile bitrate is taken and graphed.

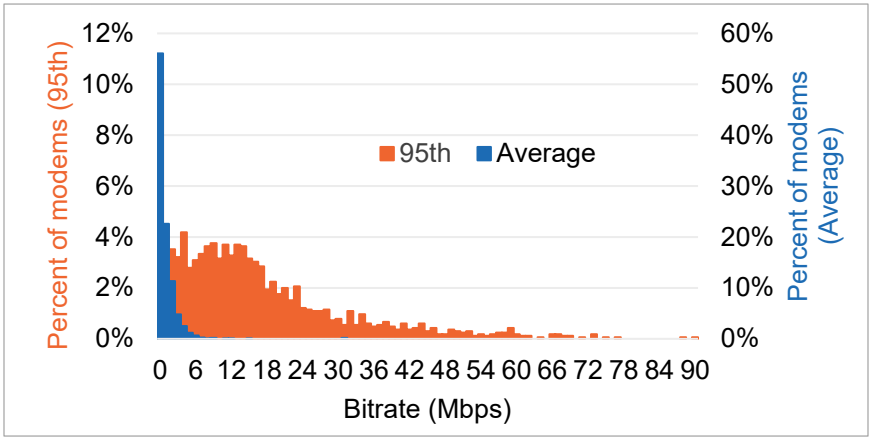


Figure 9 – Downstream Average and 95th Percentile Bitrate

On average, downstream backhaul modem bitrates are below 5Mbps, while the 95th percentile bitrate is more distributed.

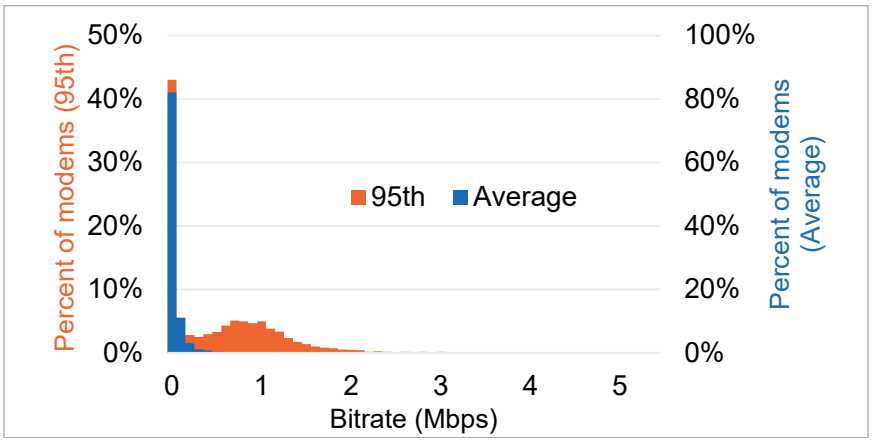


Figure 10 – Upstream Average and 95th Percentile Bitrate

The average upstream bitrate of the backhaul modem is below 1Mbps with the distribution of 95th percentile bitrates centered at approximately 1Mbps. These graphs show that the bitrate limitations placed on subscribers do not hinder their use.

3.3. Application View

The top applications ranked by consumption on the Wi-Fi network can be observed and contrasted with those on the wireline network using DPI systems.

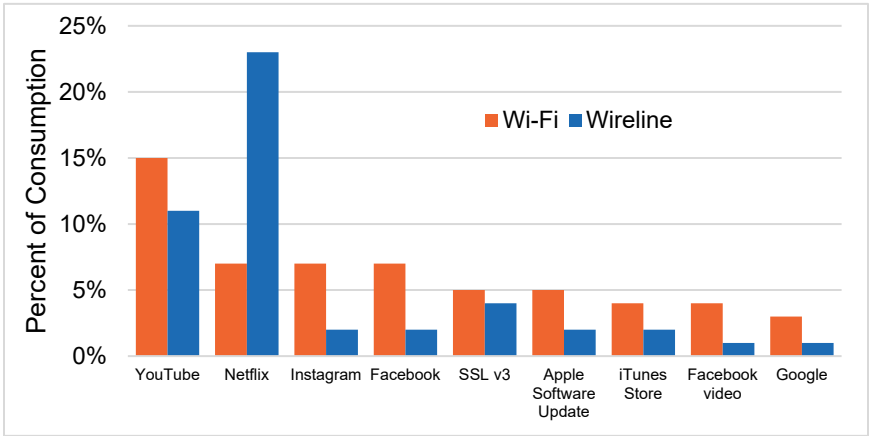


Figure 11 – Wi-Fi and Wireline Application Usage

Streaming media dominates data consumption for both wireline and Wi-Fi. In both cases, the top two applications are YouTube and Netflix. However, the order of these two applications is reversed for wireline and Wi-Fi, likely due to the nature of these applications. YouTube videos are generally shorter and can be more easily watched on a small screen at a destination connected to Wi-Fi, whereas Netflix is geared toward longer programming and is more commonly watched on a larger screen at home. Social media applications also account for a higher percentage of Wi-Fi consumption than wireline consumption as demonstrated by Instagram and Facebook.

As illustrated in Figure 12, while video dominates consumption, social media—specifically Facebook and Instagram—is used by a higher percentage of subscribers.

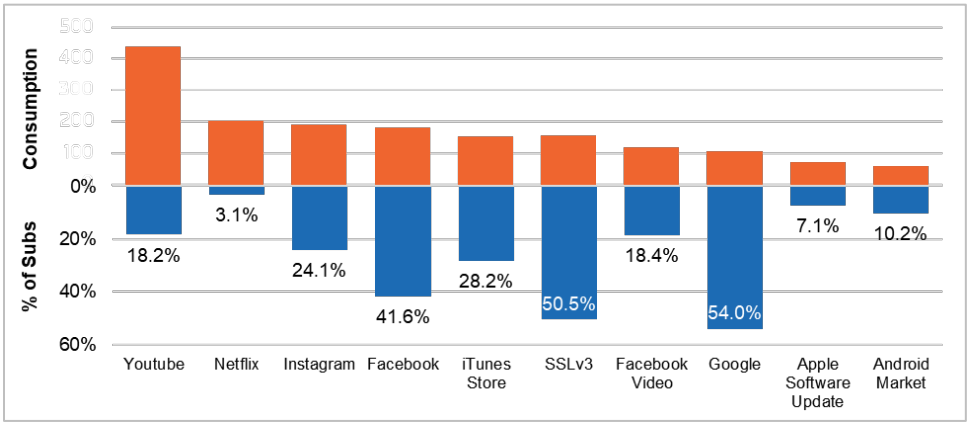


Figure 12 – Application Penetration and Consumption

When organized by application type and ranked by consumption, the top three categories are streaming media, web browsing, and messaging and collaboration (Figure 13).

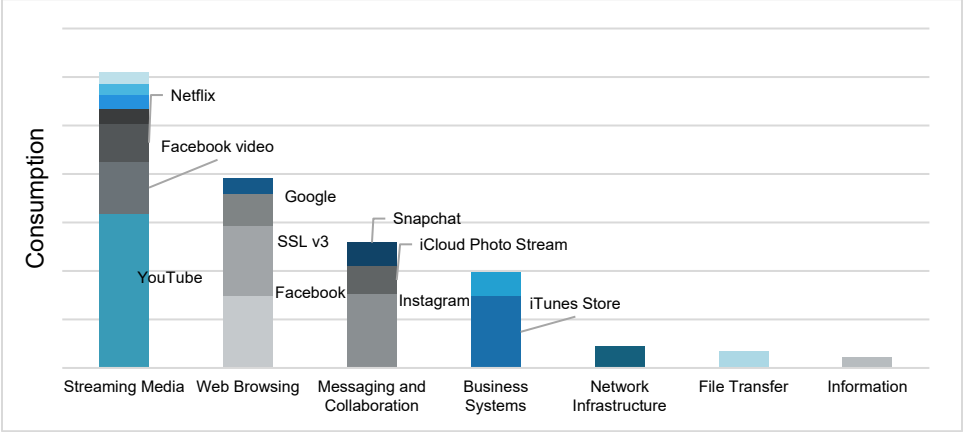


Figure 13 – Consumption by Application Type

3.4. Time of Day Trends

Wireline and wireless services have very different consumption patterns over the day. Wireline consumption peaks in the evening when subscribers are at home, while wireless consumption peaks in the late afternoon. Figure 14 shows that Wi-Fi time of day statistics are similar to wireless, but slightly delayed in time. This delay is likely to allow subscribers to get from home to their destinations where Wi-Fi is present.

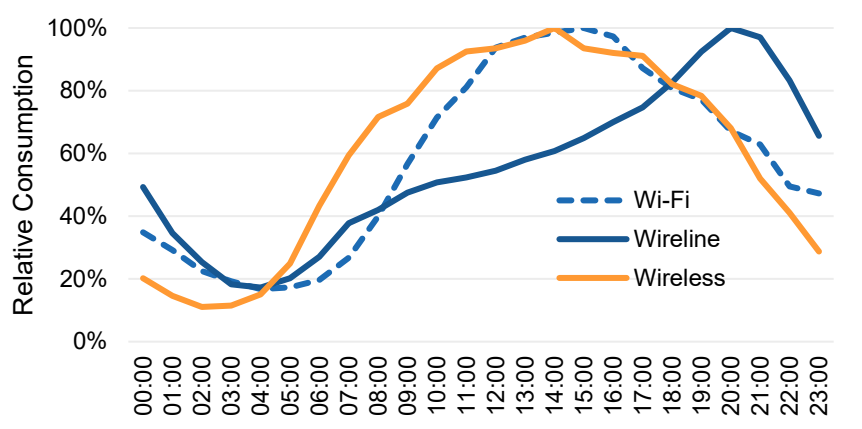


Figure 14 – Wi-Fi, Wireline and Wireless Network Time of Day Trends

Aggregate network consumption analysis points to subscribers deriving value from the Wi-Fi network and the network being used as intended. Per-device consumption, as well as time-of-day and monthly usage patterns, point to the Wi-Fi network being used to offload wireless data at destinations. The distribution of monthly Wi-Fi consumption per user was similar to the distribution of monthly wireline consumption per user. The correlations between Wi-Fi and wireless or wireline consumption demonstrates that subscribers have high consumption on both networks. Bitrate statistics and application usage suggest that devices connecting to the Wi-Fi network are mostly small-screen portable units. The number of devices per account roughly matches the number of people per household, indicating that there was no wide-spread sharing of access to non-subscribers.

After researching aggregate network usage, specific users and venues were analyzed for a more in-depth perspective.

3.5. Venue Examples

Wi-Fi APs are installed at many different types of venues, as shown in Figure 15. While most venues were similar in terms of consumption, transportation (transit platforms and similar venues), hotels and Shaw Wi-Fi Zones have higher consumption per AP.

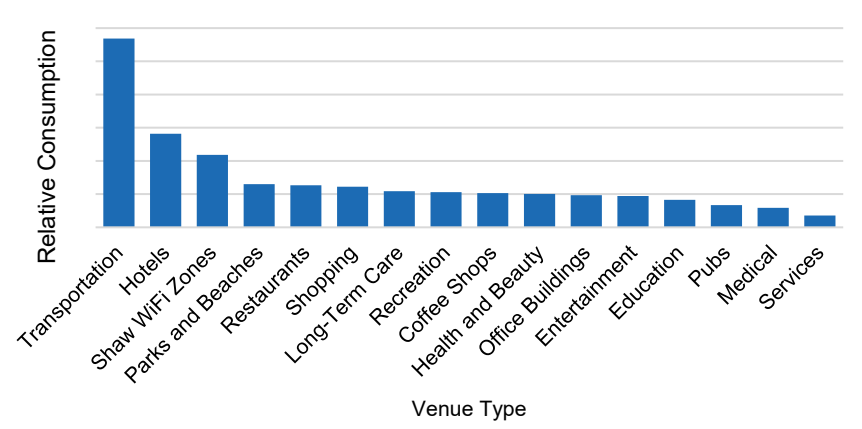


Figure 15 – Wi-Fi AP Consumption for Different Venue Types

One of the Shaw Wi-Fi Zones has previously been identified as potentially having poor performance, so an in-depth analysis was undertaken for that venue. Specifically, this venue is located close to a main thoroughfare and there is anecdotal evidence that subscribers in their cars would connect to Wi-Fi as they passed by or stopped at a stop light, disrupting any wireless session they may have had. Monthly consumption and unique device count were collected for each AP in the venue.

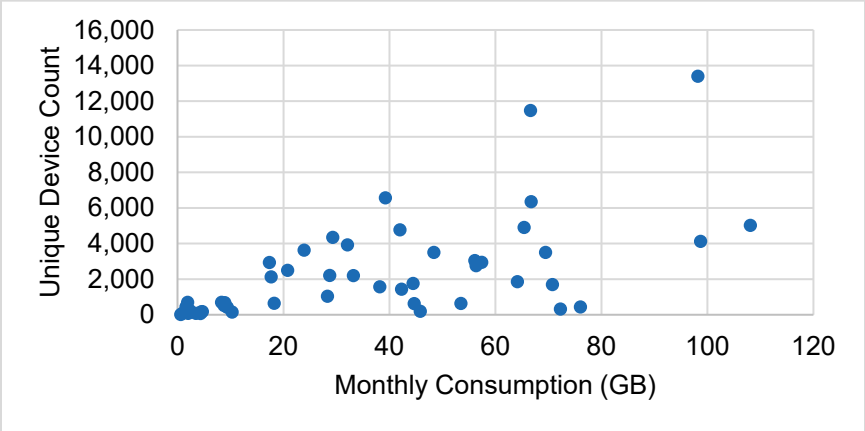


Figure 16 – Unique Device Counts vs Consumption for One Venue

As can be observed, there are large differences in consumption and unique device count between APs. APs with a high number of unique devices also have a high monthly consumption, as opposed to low monthly consumption, as might be expected with APs close to traffic lights. The correlation appears to be that APs close to popular locations such as restaurants and coffee shops have more unique devices and higher monthly consumption, while APs close to more lightly visited locations have fewer unique devices and lower monthly consumption.

One specific AP with a high number of unique devices and a large amount of consumption close to a popular restaurant and busy intersection was further analyzed. The device with the highest monthly consumption for that AP was 4GB, pointing to casual usage.

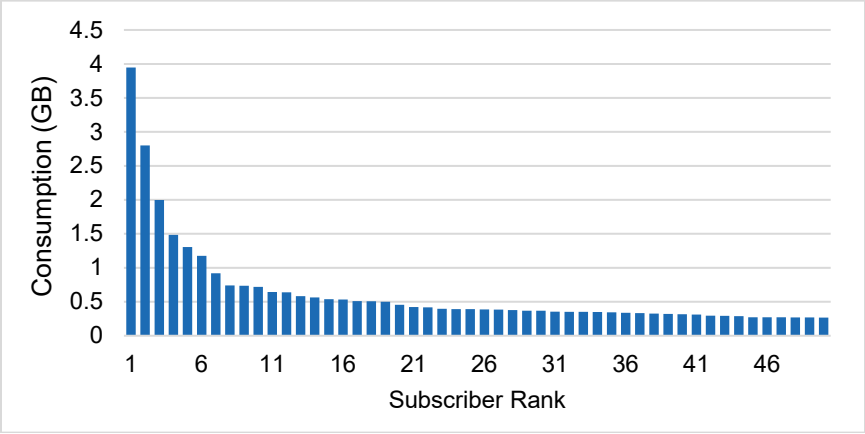


Figure 17 – Consumption of Top 50 Devices for One AP

As shown in Figure 18, the application distribution of this particular AP differed from the aggregate statistics. Video represented less consumption, with YouTube ranking tenth and Netflix ranking outside of the top ten applications. However, similarly to the aggregate statistics, social media ranked high on the

list. The top application by consumption for this AP was SSLv3, which is likely encrypted web browsing. The application mix points to short duration usage, as might be expected in a restaurant.

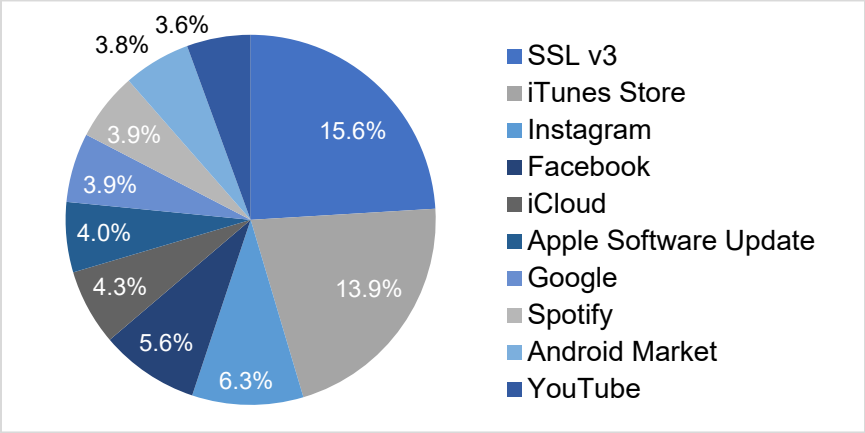


Figure 18 – Top 10 Applications for one Access Point

Five-minute throughput samples can be used to estimate bitrate over a longer duration. Max bitrates over five-minute durations tend to be low, indicating that few devices are being used to transfer large files or watch high quality video.

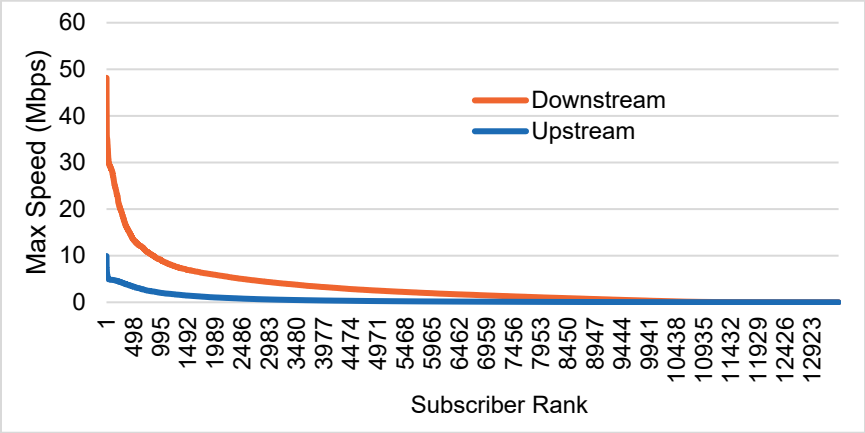


Figure 19 – Max Speed for All Users on one Access Point

While no evidence of a poor subscriber experience was found based on the data collected, it was determined that venue data can be used to prioritize the addition of new sites and potentially to reposition or remove APs that are being underutilized in their current locations.

3.6. Heavy Users

A dashboard was created to focus on a small number of heavy users whose monthly Wi-Fi consumption was greater than 100GB. In Figure 21, Wi-Fi consumption is plotted on the Y-axis and wireline consumption on the X-axis. The size of the circles denotes how many devices are linked to the account. Available data include device level statistics, as well as location and application-level information. Circles in pink have greater Wi-Fi consumption than wireline consumption and vice versa for circles in blue.

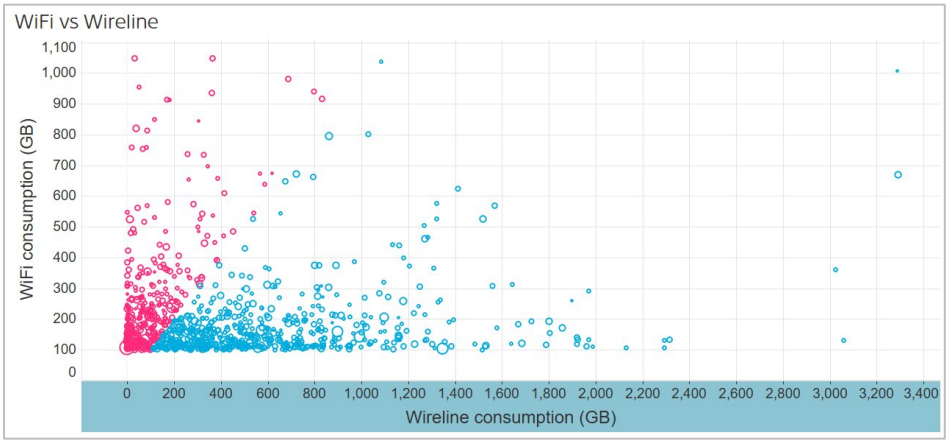


Figure 20 – Wi-Fi vs Wireline Consumption for Heavy Users

It was found that most consumption came from one or two connected devices, even for accounts with many connected devices. For most cases, consumption came from one or a small number of locations. This points to heavy but normal use of the Wi-Fi network, analogous to what is seen in the wireline network.

4. Network Optimization

Earlier in the year, a network optimization initiative was undertaken to reduce the occurrence of devices connecting to the 2.4GHz network at low signal strength. The intent was to reduce the 2.4GHz coverage to match the 5GHz coverage in order to decrease instances of devices authenticating but not connecting to the Internet and allow more devices to see the 5GHz network sooner as they approached an AP.

Figure 21 – 2.4GHz Coverage Optimization

When looking at the data for a single market, a large reduction in the number of devices connecting to the 2.4GHz network can be observed after the optimization was made. Importantly, there was no drop in consumption for the 2.4GHz network, meaning that the optimization was performing as intended and not preventing subscribers from accessing the network. At the same time there was an increase in consumption on the 5GHz network, pointing to more usage of the network.

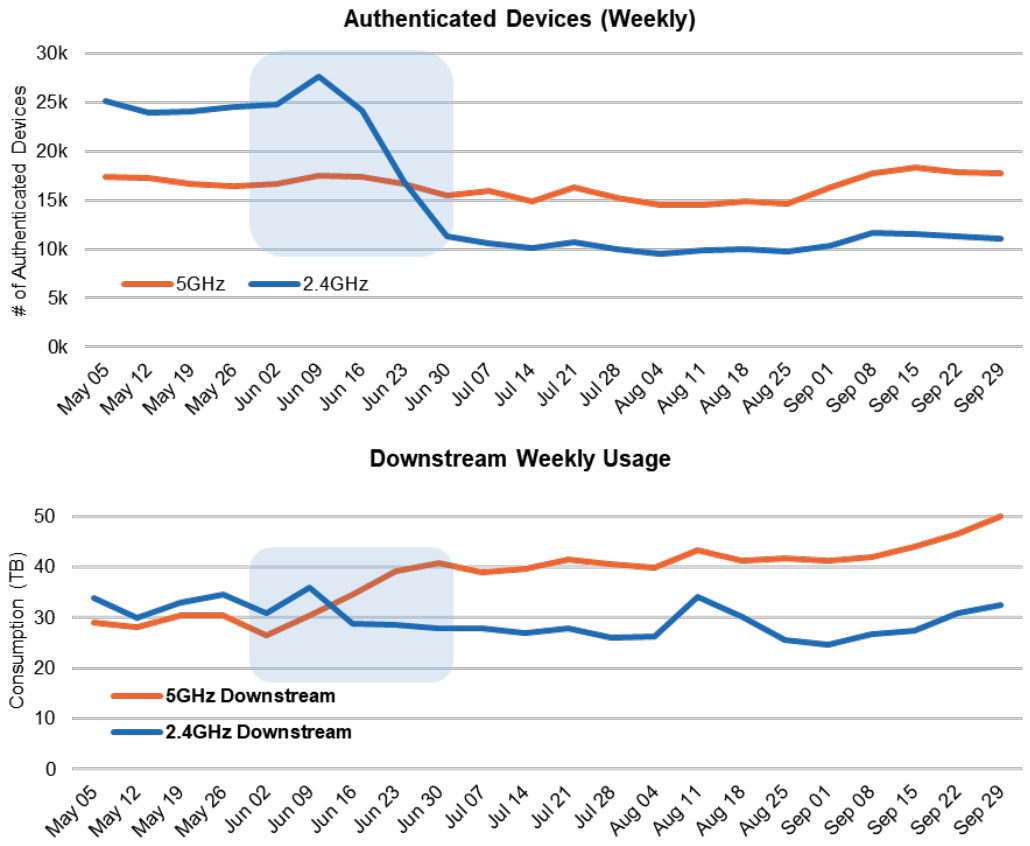


Figure 22 – Impact of 2.4GHz Optimization on Device Counts and Consumption

This optimization covered the area investigated and prevented users from authenticating but not connecting to the Internet. Given its effectiveness, it was eventually rolled out across the entire service provider Wi-Fi network.

4.1. Upgrade Strategy

Initial service provider Wi-Fi APs were Wi-Fi 5 and were reaching their end-of-life, which raised the question of if and how these APs should be upgraded. Wi-Fi 6 provides several new features (Figure 24) which combine to deliver greater capacity and a better performance with a high density of users [3].

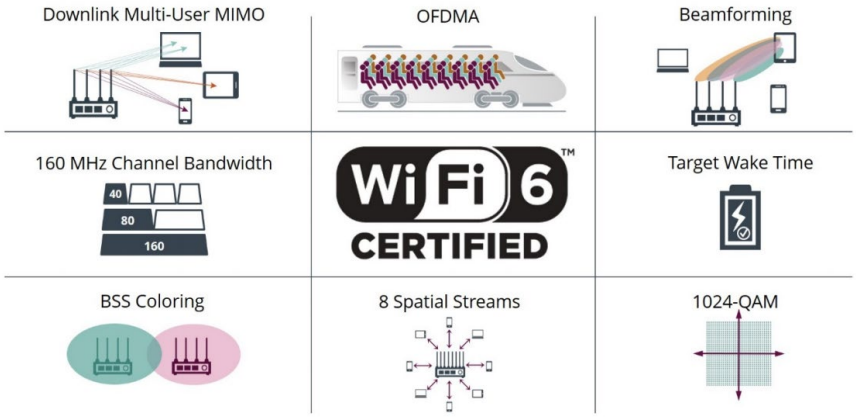


Figure 23 – Wi-Fi 6 Benefits (Source: Wi-Fi Alliance)

The experience analyzing venues revealed that there was significant differences between APs in terms of unique devices and monthly consumption. A methodology was created to separate APs into three different categories—Priority Upgrade, Maintain, and Minimize Support. The categorization was made based on the number of unique devices connected to the AP that consumed over 50MB of data in a month. 50MB was used because it provides tangible data offload and is likely to be above the data threshold that would cause a subscriber to seek out Wi-Fi.

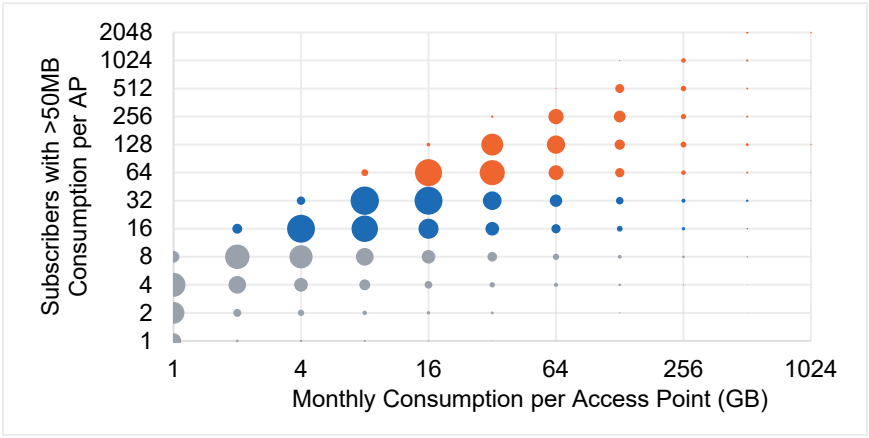


Figure 24 – Devices with >50MB of Consumption per AP

Service provider Wi-Fi APs with more than 64 subscribers with greater than 50MB of consumption were put on the list to upgrade to Wi-Fi 6 as they came to end of life. APs with between 16 and 64 subscribers were to be maintained, while APs below 16 were to have operational support minimized. This upgrade strategy allows investments to be focused where subscribers get the most value while reducing operational costs elsewhere. A simple strategy that uses the total consumption per device could prioritize APs where a single device consumes a large amount of data. Similarly, a strategy that uses total unique devices could prioritize APs where many devices connect but do not consume data.

4.2. Small Cell

Small cells can be used to increase wireless capacity and coverage in localized areas [4]. The coverage area is small compared to a macro cell, as shown in Figure 26.

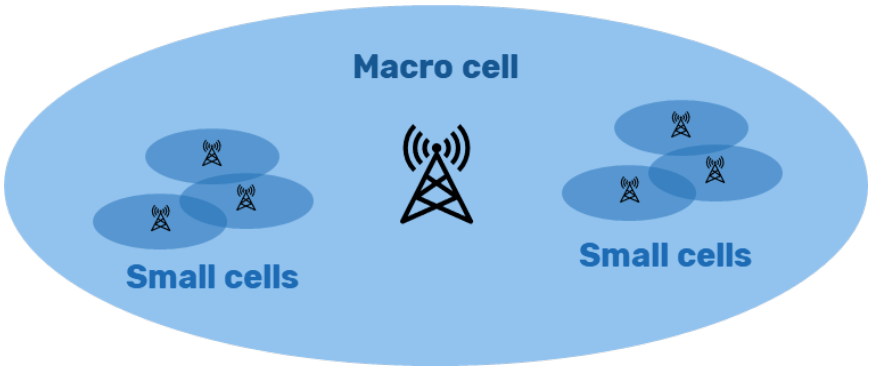


Figure 25 – Small Cell vs Macro Cell Coverage

The coverage area of a small cell will depend on the spectrum used. New mmWave spectrum used for 5G will have RF coverage areas similar to Wi-Fi. mmWave spectrum, or high-bands, represent a large capacity over small coverage areas [5], as shown in Figure 27.

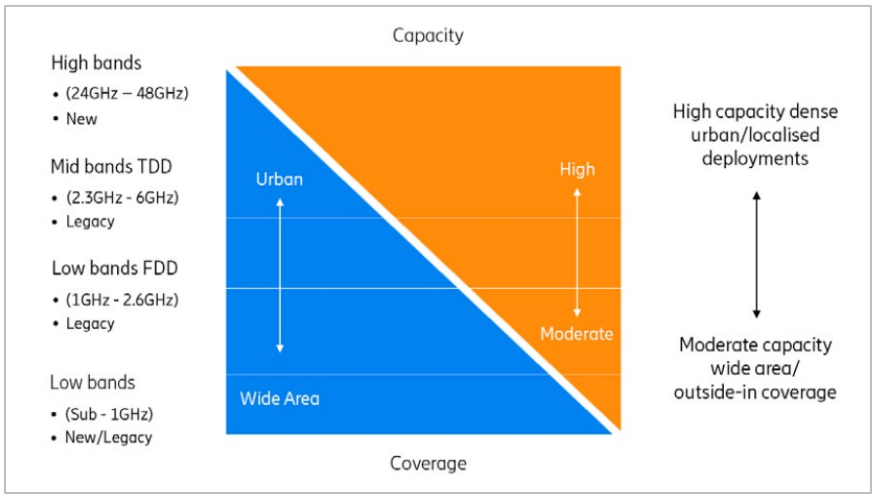


Figure 26 – Capacity vs Coverage for Different Spectrum Types (Source: Ericsson)

These new networks will require space, power, and backhaul. The service provider Wi-Fi network has all of these resources and in the future may be augmented or replaced with 5G small cells.

Shaw has deployed small cells in its wireless network, initially for trials and in advantageous locations. One such location had both service provider Wi-Fi and small cells deployed. The networks were installed separately from each other and individual coverage planning was done. The location of Wi-Fi APs and small cells were similar, however, allowing for a direct comparison. There were 26 Wi-Fi APs and 23 small cells in the facility. It was found that wireless subscribers had four times as much consumption on the Wi-Fi network compared to the small cell network. Figure 28 includes all Wi-Fi traffic and shows consumption per AP or small cell.

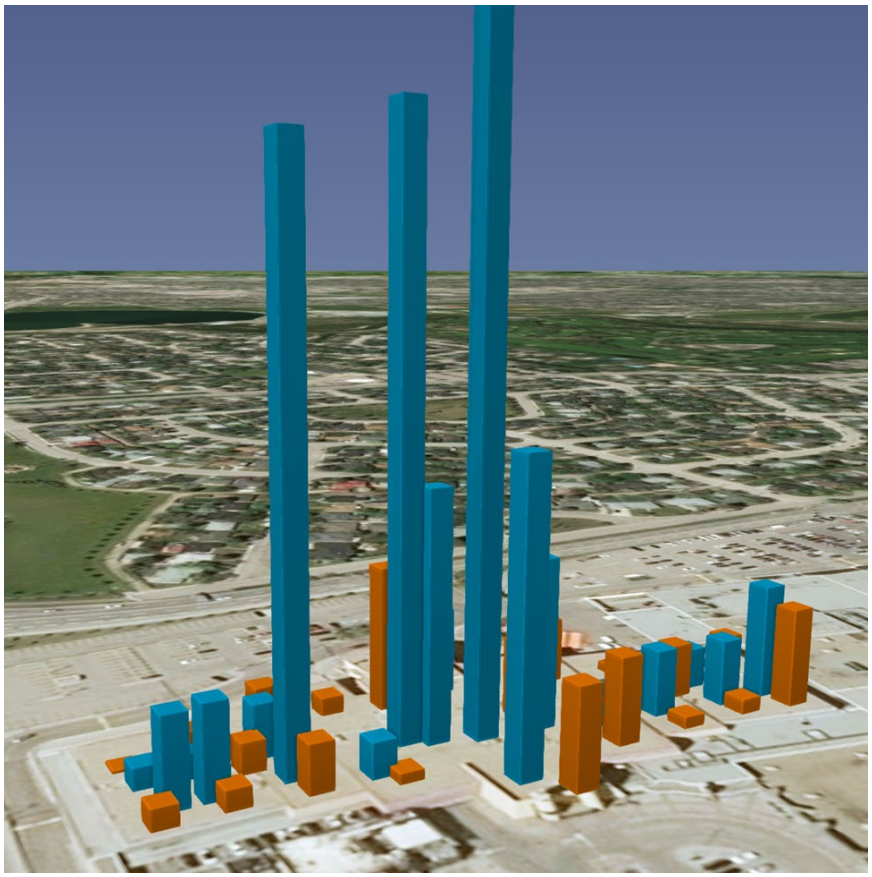


Figure 27 – Consumption for Wi-Fi APs and Small Cells

Wi-Fi consumption is shown in blue and small cell consumption in orange. The venue is a cinema with consumption highly centered in the lobby and the highest consumption at the AP near the entrance. The consumption for this AP was much higher than the others and the representative bar was cut off in the figure to allow the consumption of the other APs to be observed. The low small cell use raises the question of whether it makes sense to deploy both service provider Wi-Fi and small cells in the same locations for offload. However, wireless coverage is still a strong justification for placing small cells, especially for legacy devices that require 3G connectivity for voice.

Technology advancement, as well as business strategies, will dictate whether these networks will be replaced with 5G or whether Wi-Fi will continue to operate in parallel. Licensed-Assisted Access (LAA) is an LTE feature that allows mobile devices to use the unlicensed 5GHz band without having to switch to Wi-Fi. This relieves devices of the burden of deciding which network to prefer and under what conditions to switch from one network to the other. However, Wi-Fi may be the only option for devices without SIM cards such as laptops, or the preferred network in situations where LAA traffic is counted toward a subscriber's data usage. In scenarios where both networks exist, efforts such as CableLabs' Intelligent Wireless Network Steering (IWiNS) project aim to improve the user experience by steering traffic to the most appropriate network given the specific context.

5. COVID-19 Update

As with all networks, the service provider Wi-Fi network was heavily impacted by the COVID-19 pandemic. While wireline networks saw large increases in consumption, wireless networks saw

reductions, owing to stay-at-home orders or similar limitations on movement. Daily consumption on the Wi-Fi network fell approximately 50% in March of 2020 at the start of the pandemic, slowly climbing back to pre-COVID-19 levels over the summer (Figure 29).

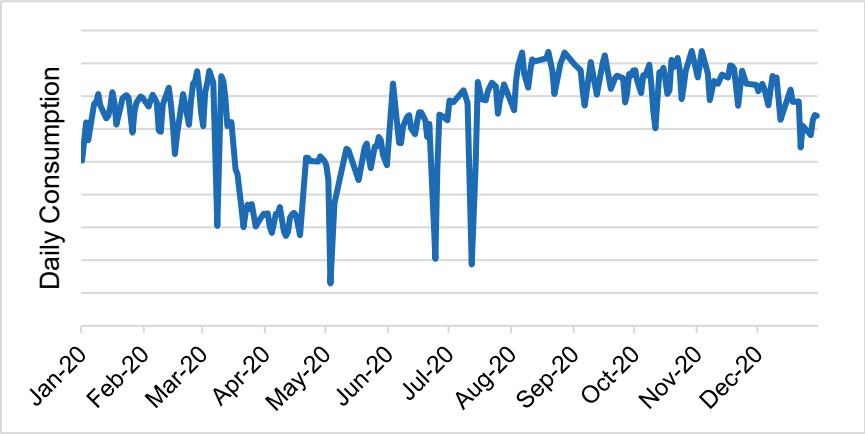


Figure 28 – Daily Consumption During COVID-19

Relative application use did not change significantly, with the notable exception of increased Facebook and decreased Instagram consumption (Figure 30). This makes sense as Instagram is used to document experiences, which are rarer during stay-at-home orders, while Facebook is used to search for news and check in with social networks.

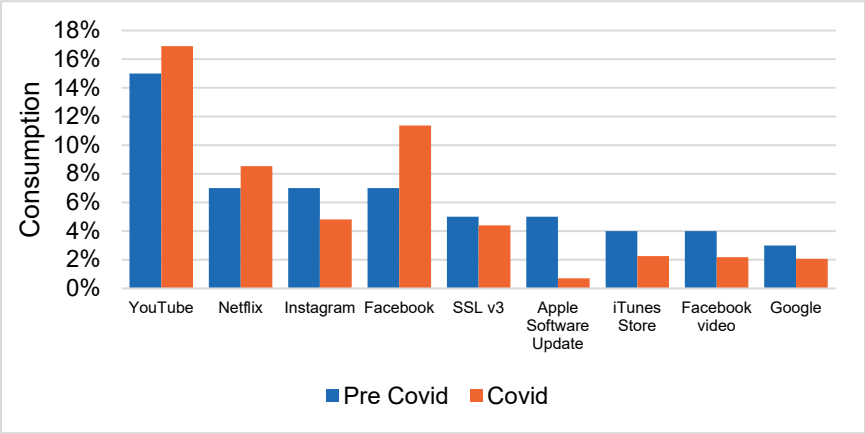


Figure 29 – Application Use During COVID-19

Home hotspot APs saw slightly increased consumption at the beginning of the COVID-19 pandemic in opposition to other Wi-Fi APs (Figure 31). This also makes sense as people stayed home or in their neighbourhoods where home hotspots are more common.

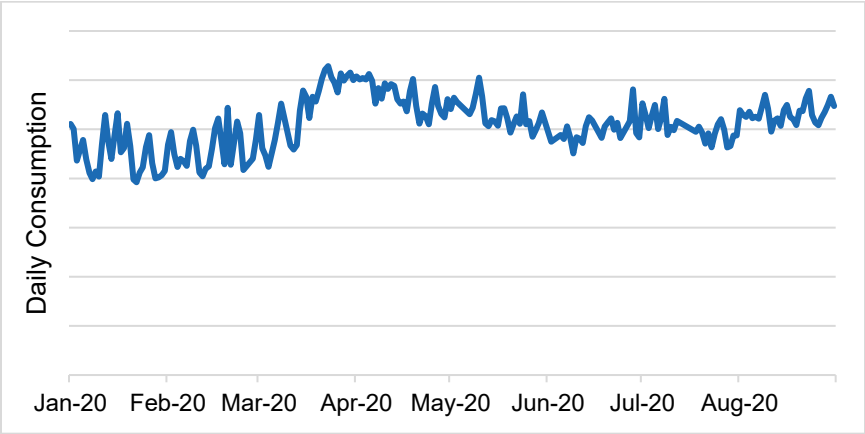


Figure 30 – Home Hotspot Consumption During COVID-19

6. Conclusion

As networks converge and become ubiquitous, consumers will expect to be able to access their broadband services at all times, wherever they are. Service provider Wi-Fi can deliver an important piece of this puzzle, extending subscribers a quality connection at their destinations. In general, service provider Wi-Fi is used as an extension of the in-home wireline network, geared towards shorter durations and smaller devices. The implementation of the service provider Wi-Fi network will have implications on the customer experience. In particular, it was found that devices could authenticate to the 2.4GHz network at low signal levels, but not be able to connect to the Internet. This can cause mobile sessions to end, resulting in frustrated subscribers turning off their Wi-Fi. Matching 2.4GHz and 5GHz coverage reduces this occurrence, leading to a better customer experience. The wealth of AP usage data can inform a network upgrade strategy, but care must be taken to choose metrics that optimize subscriber value. A strategy that uses the number of unique devices that consume 50MB or greater allows for AP upgrades to be targeted where subscribers find the most value. In the future, as small cells and 5G are deployed, the service provider Wi-Fi network will offer a wealth of data to optimize deployments and provide readily available space, power and backhaul.

Abbreviations

5G	Fifth Generation Wireless Standard
AP	Access Point
BSS	Basic Service Set
CDF	Cumulative Distribution Function
DOCSIS	Data over Cable System Interface Specification
DPI	Deep Packet Inspection
GB	Gigabyte
Gbps	Gigabit per second
GHz	Gigahertz
HFC	Hybrid Fibre Coax
IWiNS	Intelligent Wireless Network Steering
LAA	Licensed-Assisted Access
LTE	Long Term Evolution
MAC	Media Access Control
MB	Megabyte

Mbps	Megabit per second
MIMO	Multiple-input and Multiple-Output
mmWave	Millimeter Wave
Ms	Millisecond
OFDMA	Orthogonal Frequency-Division Multiple Access
QAM	Quadrature Amplitude Modulation
RADIUS	Remote Authentication Dial-in User Service
RF	Radio Frequency
SIM	Subscriber Identification Module
SSID	Service Set Identifier
SSL	Secure Sockets Layer
TB	Terabyte
Tbps	Terabit per second

Bibliography & References

[1] Cisco, "Cicso Annual Internet Report (2018-2023)," 2018

[2] Cisco, "How Does RADIUS Work?," 19 January 2006

[3] Wi-Fi Alliance, "Wi-Fi CERTIFIED 6: A new era for Wi-Fi," 2019

[4] GSMA, "Improving wireless connectivity through small cell deployment," May 2017

[5] Ericsson, "Leveraging the potential of 5G millimeter wave," 2021

Optimum Load Shaping

Charging Electric Vehicles and Batteries with Renewable Energy Sources

A Technical Paper prepared for SCTE by

Frank Sandoval
Principal
Pajarito Technologies LLC
1650 Gaylord St, Denver CO 80206
+1-720-338-1988
frank@pajaritotech.com

Dr. Robert F. Cruickshank III
CTO
GRIDIoT® by RCA
132 Cruickshank Rd #269, Big Indian, NY 12410
+1-703-568-8379
rfciii@cruickshank.org

Laurie Asperas Valayer
CSO
GRIDIoT® by RCA
16 Wincott Dr, Melville, NY 11747
+1-631-335-9197
asperasvalayer@gmail.com

1. Introduction

Traditionally, electric utilities forecast electricity consumption, aka load, and then generate enough power to supply demand at any point in time. But the world is changing, and this model is becoming tenuous at best, and perhaps will be wholly insufficient to meet our needs for low-cost reliable power. The grid is aging and becoming less reliable as evidenced by recent years' California wildfires and the February 2021 Texas Power Crisis. As the need to reduce carbon emissions is becoming more urgent, significant amounts of distributed power generation from solar and wind are coming online, and demand is increasing with the penetration of electric vehicles (EVs).

These changes present opportunities for innovation, and one area of activity attempts to shift the demand for energy (via load shaping) to the most efficient, cheapest, and cleanest sources of energy supply during the daily cycle. As significant consumers of electricity, cable companies have much to gain by adopting load shaping strategies for facilities, fleets, customers, and the outside plant.

Several working groups within the SCTE have been exploring technologies that could aid in actively shaping electrical loads. A specific use case is the battery-backed EV charging station. Cable fleets of tens of thousands vehicles will become electrified and there is ample motivation to drive down the anticipated costs of EV charging. Using utility-provided Optimum Load Shaping (OLS) signals—based on pricing, low carbon generation, and weather—each charger may be controlled to time its charging periods to ensure the least cost.

Other examples of load shaping include the orchestration of loads within a micro-grid that powers a cable facility. In this scenario, a micro-grid controller may obtain OLS signals and other data to drive algorithms that optimize the load within its domain, using solar, battery, fuel cell, and building management assets.

The outside plant consumes most of the electricity in a cable operation, and there are interesting concepts to leverage batteries, renewable generation, and consume energy in proportion to RF bandwidth powering.

2. Limits to Supply

Since electricity became available in 1882, the grid was simply assumed to grow to meet demand. The one-way model of energy generation, transmission, and distribution has generally been taken for granted, and as cities and towns grew—and more and more appliances and devices were powered up—electricity suppliers simply worked to satisfy all new demands. There are natural limits to this approach, and we've hit them.

2.1. Diverse Generation

The Public Utility Regulatory Policies Act of 1978 complicated the traditional one-way model of the grid by allowing non-utility generators to market their power to utilities. Energy supplies and pricing became more dynamic, and the sourcing and movement of electricity became more complex and less reliably predictable.

More recently, Distributed Energy Resources (DER), such as industrial, community, and household solar, wind, and emerging battery systems add vastly more energy generation sources, deeper into the distribution system, further complicating pricing and management systems. Repurposing the same infrastructure developed for stable one-way generation and transmission to a highly dynamic bi-directional system has worked remarkably well, it reminds one of how cable networks were fortuitously able to accommodate Data traffic overlaid upon infrastructure developed for broadcast TV. However, the grid was not designed for the uses its being asked to serve. Renewable generation fluctuates with the weather, complicating forecasting of fossil fuel generation. Regulations, financing models, distribution systems, management and maintenance systems, and infrastructure—all struggle to adapt without the benefit of a modern intentional design or purpose-built systems.

2.2. Carbon Pollution

Politics aside, carbon pollution appears to be a present and growing threat that communities and governments around the globe are reacting to. Utilities are adopting external or internal mandates to de-carbonize and, in many regions, vigorous efforts are currently underway to replace fossil fuel generation with less carbon-intensive alternatives.

While energy suppliers strive to replace and surpass current fossil fuel capacity, we consider ways to shape and blunt the growth of demand. If the rate of demand growth slows, then the rate of replacement of carbon intensive fuels by renewables accelerates, helping suppliers reach carbon targets more quickly.

2.3. Inefficiencies

As engineers, we might feel offended by how inefficient we've been with fossil generation. As shown in Figure 1, fully 2/3 of the energy released from burning fossil fuels is wasted as heat! Simply put, 2/3 of the carbon we've released into the atmosphere by electrical generation has served humanity not one bit of good. We can—and of course should—do better.

In Figure 1, line widths are proportional to flows. Given total usage of 100 quadrillion BTUs, all numbers on the chart denote approximate percentages, e.g., at left, solar power provided for 1% of U.S. Energy use, and at right, 33% of all energy provided for useful work and 67% was rejected as waste heat—via smokestacks, heat exchangers, and tailpipes.

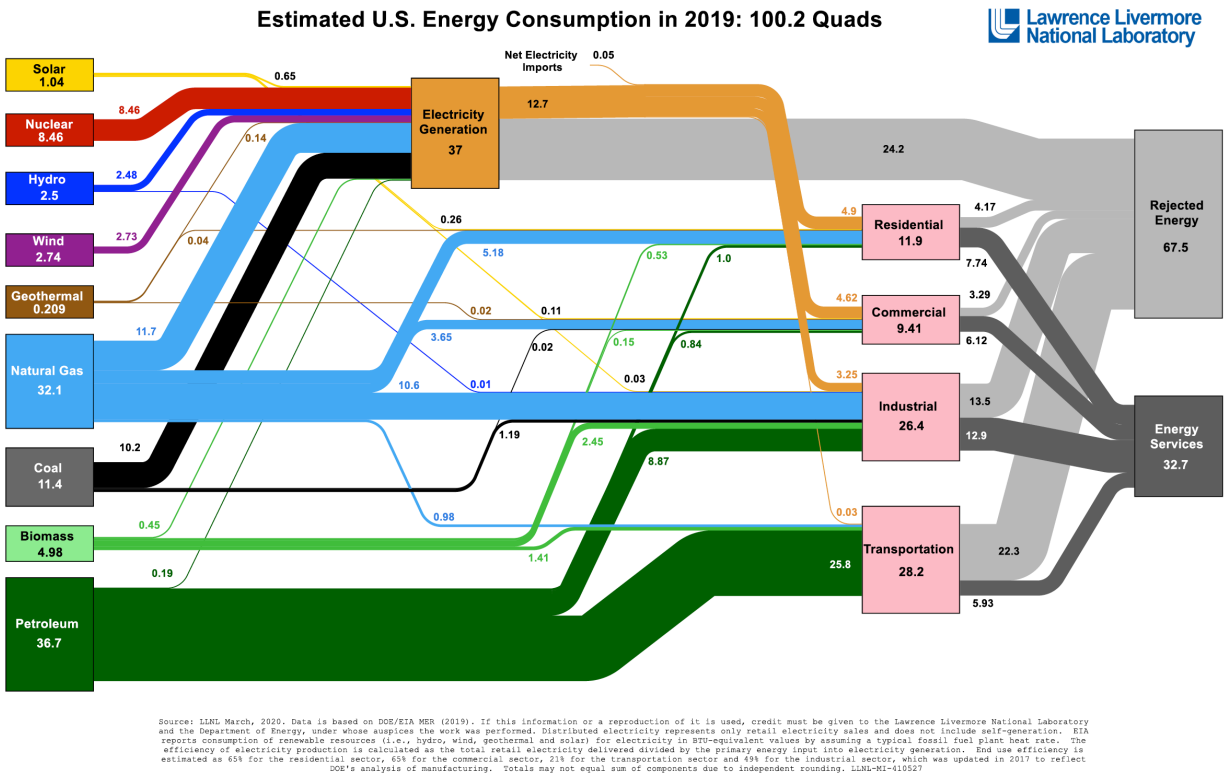


Figure 1 - U.S. Energy Consumption

2.4. Aging Infrastructure

A failed C hook, purchased circa WWII for half a dollar, caused the Camp Fire in California in 2018, the most destructive fire in the state's history to that point, causing $16.8B in damages and 85 lost lives. Images illustrative of Pacific Gas and Electric's failing infrastructure, responsible for starting the fire, are shown in Figure 2.

Figure 2 - PG&E Failing Infrastructure

The mean life expectancy of grid equipment is 65 years, yet, on average, equipment is in service in the field 68 years, with components as old as 108 years!

This leads to an inherently fragile and costly grid that will require time and money to rebuild and repair. Diminishing the maximum capacity that the grid is meant to support will lessen the overall costs and time required to retrofit.

2.5. Severe Weather

Recent history and unexpected trends indicate more frequent and severe weather events that can overwhelm the grid. The unprecedented cold in Texas earlier this year led to catastrophic grid failures. Increasing drought and heat, especially along the West coast is spiking demand for air conditioning.

A cruel irony of higher temperatures is that it makes fossil generation less efficient—the necessary cooling and condensation of water vapor is less effective leading to a higher ratio of fuel to kilowatt-hours (kWh) generated.

2.6. Electrification

Most car manufacturers have by now made it clear that Internal Combustion Engine (ICE) vehicles will be as quaint as horse drawn buggies in a few decades. Electric Vehicles (EV) alone will cause an estimated 25% surge in electricity demand, growth that today's infrastructure is simply incapable of handling.

Other factors will lead to increased electrical demand, for example rising temps will induce increased use of air conditioning. Continued efforts to reach carbon neutrality will necessitate transitions to electric heating, cooling, and cooking.

Figure 3 reminds us of the many of today's sources of fossil fuel consumption by transportation and buildings; these all will transition to electricity over time.

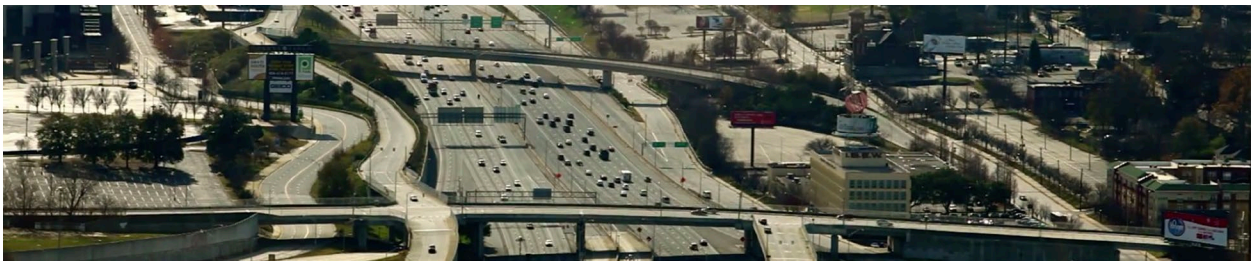


Figure 3 - Energy Consumers

2.7. Cyber threats

Increasingly, various elements of our national infrastructure, such as recently publicized ransomware attacks on fuel lines and meat processors, are under attack via software hacks. There are known instances of successful infiltrations of energy grids. While it appears intruders have to date not exploited such attacks to great effect, there is risk that that they could by activating current breaches or introducing others. Our ability to more fully monitor our energy infrastructure, effectively distribute generation, and dynamically respond to adverse conditions, whether natural or man-made, are signature elements of future grids. Mitigation strategies include shifting or curtailing loads by switching to local generation or postponing demand when the grid is attacked or otherwise failing.

3. Limit Demand

All of the stressors on supply described above require tremendous costs and years of effort to address, and fortuitously all of them can be mitigated to some extent by controlling the amount and timing of demand.

As noted, the general paradigm in electricity services has been that supply chases demand that grows unconstrained by any overt controls—and that electricity costs are so low that it presents very little impediment to growth in demand. The obvious upside of low costs and reliable access is accelerated economic growth and minimally constrained personal use. The obvious downsides are waste, pollution, and over-investment in fuel and infrastructure.

We consider here ways to diminish the rate of growth of electrical consumption, if not one day decreasing the absolute amount of electricity consumed.

There are two categories of demand side control:

- Lowering absolute consumption
- Time-shifting consumption

3.1. Frugality

There will always be people who are sensitive to wasteful consumption and don't particularly mind trimming their own use by turning down the thermostat in winter and hanging their laundry out to dry—but let's face it, most folks will opt first for convenience and comfort. Making it easy for people to lead more efficient lives will probably require seamless technologies and higher prices.

We mention this simply to acknowledge that we all might become more energy aware but concede that this will not likely be sufficient to eliminate all supply issues.

3.2. Pricing

Electricity has been relatively inexpensive since it was first made available 140 years ago. Most homeowners are not enthusiastic about seeking reduction in their consumption as it represents such a small percentage of monthly budget. Renewables are proving to be less expensive than coal and natural gas, and we may find ourselves paying even less for electricity over the long term. In the short term however, a carbon tax or other price signal might be introduced to spur transition to a net carbon neutral economy.

3.3. Building and Device Efficiency

Energy Star appliances, LED light bulbs, strict building codes, improved materials, better batteries, and many other improvements in all of the thousands of technologies we rely upon have had significant impacts on our overall energy footprint. Recall the cable voluntary agreement to deploy energy efficient set-top boxes. Many fossil fuel generation plants have not been build as the result of inexorable improvements in our material culture.

But these gains are significantly offset by other energy-hungry advancements. The growth of mobile and other personal devices, crypto-mining, expanding air-conditioning, and many other factors erode the gains produced by efficiencies.

3.4. Utility Incentives

For some time, various mechanisms have been developed to help utilities avoid brownouts and blackouts by encouraging customers—typically large facilities such as factories—to reduce or avoid consumption during times of expected peak load. Imagine a heat wave during which a utility forecasts a supply deficit in the late afternoon, as some businesses are still operating but many people are coming home and cranking up the AC. This utility may have pricing agreements with energy consumers to encourage strategies such as peak-shaving and load shedding to help avoid grid failures.

Time of use pricing seeks to encourage consumption in periods during the day when prices are low. To succeed, accurate and timely accounting of energy consumption must be collected and reported. This is a form of time-shifting demand, as opposed to lowering absolute demand.

Time-of-use pricing leads us to consider Transactive Energy models. Such systems promise to create highly efficient markets as producers and consumers use real-time and dynamic pricing to balance supplies and demand. Such systems are technically and operationally complex and have yet to be widely deployed.

While certainly valuable, given the scale of supply challenges, these strategies have not proven to be entirely sufficient to mitigate supply shortages or disruptions.

3.5. Direct Controls

A number of strategies have been developed to allow utilities to directly control system loads. You might have seen or even enrolled in a program that allows your electricity provider to manage your home's thermostat. Such programs are bespoke to address specific targets for a specific provider.

The Automated Demand Response (ADR) protocol has been standardized to spur adoption of such systems by enabling interoperability between utilities that issue control messages and the devices and management systems that respond to them. While this lowers technological barriers it doesn't necessarily lead to mass scalability. Business and product development of ADR programs remain time consuming and costly in order to meet specific goals of specific utilities.

4. Optimum Load Shaping

A close cousin to ADR, Optimum Load Shaping (OLS) provides a mechanism that may encourage time-shifting of energy consumption on a massive scale. The control signal is not a specific on/off command or a retail price signal, but simply a table of values that represent the optimal percentage of a day's electrical load that should be consumed within a given period, typically an hour. The set of values describe a curvilinear shape, and are intended to accordingly shape the load on the system.

The Optimum Load Shape protocol has been standardized by SCTE and published as ANSI SCTE 267 2021. The specification was developed within the SCTE micro-grid working group and specifically addresses the cable industry's emerging EV fleet and battery charging systems, however, the protocol can be applied to any flexible load.

4.1. Theoretical background

The economic value of flexible load shaping has been studied, using various data sets and modeling strategies. A study of related literature and a description of a simulation of the Texas ERCOT system, using historical data, is provided by [Cruickshank].

A key element of the simulation methodology is the Unit Commitment Model, used within the Generic Algebraic Modeling System (GAMS), and expressed in Figure 4.

The mathematical formulation of the Unit Commitment model constructed in GAMS consisted of a cost-optimization objective function based on production costs:

$$\min \left\{ \sum_{t=1}^{n} \left\{ \sum_{i=1}^{N} v_i c_i(p_i) \right\} \right\}, v_i \in \{0,1\} \qquad (11)$$

which was constrained by the need to match supply and demand in each time period:

$$\sum_{i=1}^{N} p_i = d \qquad (12)$$

where:

n was the number of time intervals in each optimization step

N was the total number of generators

v_i was the binary variable indicating whether a generator is committed (1) or not (0)

c_i was the operating cost of generator i ($US/MW)

p_i was the power generation of generator i (MW)

d was the system demand (MW)

Figure 4 - Unit Commit Model [Cruickshank]

Simulations were run to 1) calculate the cost and carbon emissions of meeting all load through traditional fossil-fueled generation, 2) calculate the cost and carbon of an optimally flat fossil-fueled generation scenario, and 3) to overlay renewable generation onto the flat fossil-fueled

scenario. Flat constant-output fossil-fueled generation is more efficient than variable output, as it eliminates the inefficiencies of starting and stopping generators, ramping generators up and down, and running at less than ideal capacity, i.e. less than maximum efficiency. Flat fossil-fueled generation does not, however, meet demand that exceeds that static supply, hence renewables and storage are overlaid onto that scenario to create an Optimum Load Shape.

The result of the third simulation indicates that costs and carbon emissions can be significantly reduced if fossil-fueled generation is held flat and load follows the shape produced by overlaying renewables.

4.2. Architecture

The OLS system follows the common client-server pattern, using the term Producer instead of server, and Consumer instead of client; an OLS producer makes available OLS signals, and any number of consumers may acquire the signals as shown in Figure 5.

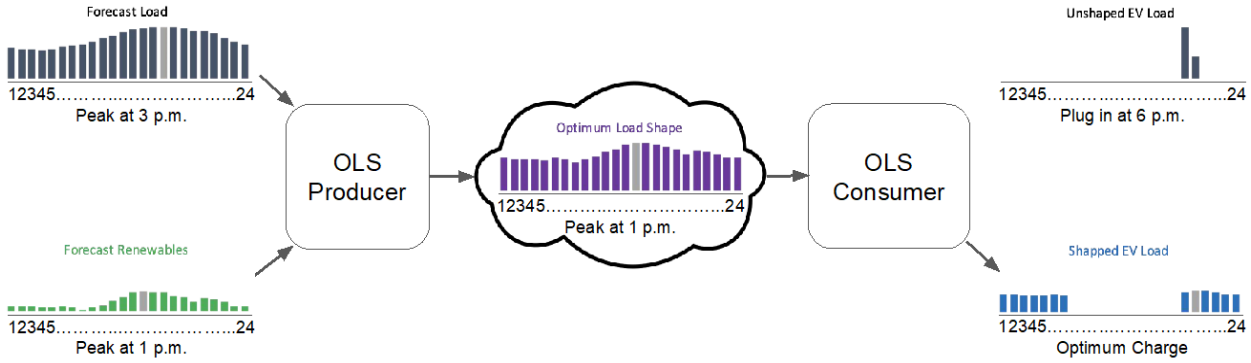


Figure 5 - OLS Architecture

An OLS signal is a datagram that simply contains a table of timestamped values. Timestamps represent the start of a period sometime in the future, typically the top of an hour of the day following the generation of the signal, and as mentioned above, the values represent the percentage of daily use that should ideally be consumed for that period.

The producer of the signal may use any algorithm to create the signal, incorporating input such as fossil fuel generation costs, availability of renewables with weather forecast, machine learning inputs from historical usage data and anything else those crafty nerds can make use of. In its simplest form, the signal can be made available to everyone within a service area—since the shape applies to the supply-side and does not target specific users.

The left side of the diagram in Figure 5 illustrates a strategy to optimize the use of renewables by flattening fossil fuel generation and overlaying the forecast renewable generation. The closer the aggregate load matches this shape, the more efficiently the mix of renewable and fossil fuel generation will be.

Consumers of OLS are not concerned about why the signal takes the shape it takes; they simply respond as best they can to the signal. Some electrical consumers, say the lighting of a retail store, cannot time-shift, but many others, say a battery or thermostat, often have flexibility to determine the times at which to pull power from the grid.

Each consumer autonomously decides how best to respond to OLS signals, if at all. For example, an EV charging system may define flexibility constraints, such as blocking certain hours during the day in which a car will be in use and not available for charging—and a specific hour, say 7 am, by which the car must be fully charged. Given these constraints, the OLS consumer can strive to match the OLS shape as best it can. This still can produce dramatical results, for example, avoiding load spikes at the close of the workday by smoothing the load across a population of EVs.

The right side of the diagram in Figure 5 illustrates examples of unshaped load (at top) and shaped load (at bottom). Figure 6 provides a more detailed analysis of the concepts in Figure 5 and depicts ~20% savings at far right.

	Supply-side: Two forecasts [GWh] feed Optimum Load Shape (OLS) Producer							Demand-side: OLS Consumers modulate load			Forecast Load	Electricity Price and Cost		
	Forecast Load	–	Forecast Renewables	=	Net Generation	→ Flat Net Generation	→ Optimum Load Shape	EV Optimum Load Shape	→ Shaped EV Load	Unshaped EV Load		Retail ¢/kWh	Shaped ¢/h	Unshaped ¢/h
Midnight	36.1		6.0		30.1	35.2	3.7%	7.3%	0.73	-		15.1	11.1	0.0
1	34.6		5.8		28.8	35.2	3.6%	7.3%	0.73	-	Forecast Renewables	14.4	10.5	0.0
2	33.7		5.4		28.3	35.2	3.6%	7.2%	0.72	-		14.2	10.2	0.0
3	33.4		4.6		28.8	35.2	3.5%	7.1%	0.71	-		14.4	10.2	0.0
4	34.7		4.3		30.4	35.2	3.5%	7.0%	0.70	-		15.2	10.7	0.0
5	37.5		6.7		30.8	35.2	3.7%	7.5%	0.75	-	Net Generation	15.4	11.5	0.0
6	38.2		4.6		33.6	35.2	3.5%	7.1%	0.71	-		16.8	11.9	0.0
7	40.0		0.4		39.6	35.2	3.7%					19.8	0.0	0.0
8	43.2		4.5		38.7	35.2	3.5%					19.4	0.0	0.0
9	47.0		7.8		39.2	35.2	3.8%				Flat Net Generation	19.6	0.0	0.0
10	50.7		14.2		36.5	35.2	4.4%					18.2	0.0	0.0
11	53.8		18.0		35.8	35.2	4.7%					17.9	0.0	0.0
Noon	56.5		23.6		32.9	35.2	5.2%					16.5	0.0	0.0
13	58.4		25.0		33.4	35.2	5.3%				Optimum Load Shape	16.7	0.0	0.0
14	59.7		24.7		35.0	35.2	5.3%					17.5	0.0	0.0
15	60.2		23.5		36.7	35.2	5.2%					18.4	0.0	0.0
16	59.7		20.0		39.7	35.2	4.9%					19.9	0.0	0.0
17	58.0		17.8		40.3	35.2	4.7%				↓ Deliver OLS over any network ↓	20.1	0.0	0.0
18	55.9		11.6		44.4	35.2	4.1%	8.3%	0.83	7.00	Shaped EV Load	22.2	18.5	155.3
19	55.0		16.7		38.3	35.2	4.6%	9.2%	0.92	3.00		19.1	17.7	57.4
20	52.3		14.8		37.4	35.2	4.4%	8.9%	0.89	-		18.7	16.7	0.0
21	47.5		11.5		36.1	35.2	4.1%	8.3%	0.83	-		18.0	15.0	0.0
22	43.0		6.2		36.7	35.2	3.7%	7.4%	0.74	-	Unshaped EV Load	18.4	13.5	0.0
23	39.5		6.0		33.6	35.2	3.6%	7.3%	0.73	-		16.8	12.3	0.0
Totals →	1,128.7		283.6		845.1	845.1	100%	100%	10.00	10.00		Cost→	$ 1.70	$ 2.13
	GWh		GWh		GWh	GWh	Unitless	Unitless	kWh	kWh		Savings→	$ 0.43	$ 157.03
Assumptions & notes:													per/day	per/year

The vehicle is in use and unavailable for charging from 0700 - 1800 hours.

The charger takes unavailability into consideration and autonomously adjusts the Optimum Load Shape that it received from the supply-side.

1. Data are illustrative of load for the serving area of the Electric Reliability Council of Texas, and represent ~10% of U.S. nationwide electricity usage on a hot peak summer day.
2. Model can be similarly applied to any vertically integrated serving area, e.g., for micro and nanogrids. Modeling wholesale markets and on-site combined heat & power (CHP) is TBD.
3. In power systems, generation = load. In optimization, loads from electric vehicles (EVs), facility batteries, and smart devices are modulated , i.e., shifted forward & backward in time.
4. The OLS is normalized between zero and one by dividing the generation at each time step by the sum of generation over the simulation horizon, in this case 1-day.
5. The unshaped EV Level 2 charger draws 7 kW for 1.43 hours = ~10 kWh of energy for ~100 miles of travel. No opt-in or two-way communications is required; OLS can be broadcast.

Figure 6 - EV Charging Example Detail

4.3. Protocol

The ANSI SCTE 267 2021 standard defines the Data Model for an OLS signal in YANG format. YANG is a human and machine-readable schema language designed to model hierarchical data structures. YANG can be converted via software tools, or by hand, into various transport and encoding formats, such as REST and Protobuf.

The standard provides an HTTP REST OpenAPI specification of an API based on the Data Model. REST is arguably the most common protocol in use today to support client-server architectures.

Human and machine-readable assets are available on Github at:
https://github.com/cablelabs/scte-ols

There could be many non-exclusive paradigms for communicating OLS signals, including pull, push, and notifications.

The lightweight specification of the payload and non-normative guidance on transport/encoding formats is intentional and encourages innovation. Both OLS producers and consumers are expected to iteratively develop and improve algorithms to generate and respond to signals, based on learnings in the field.

4.4. Monitoring

The OLS specification currently provides no definitions for how OLS consumers expose their behavior in response to OLS signals. This greatly simplifies the entire platform, compared for instance to Transactive Energy models. While one might expect that OLS consumers will more aggressively adopt OLS if given new incentives, which be performance based, these business issues are not explicitly addressed in the current standard.

Additional definitions could be developed and standardized to support monitoring of OLS consumers. For now, we expect the market to innovate and discover where then real needs lie. In the meantime, when OLS is used with increasingly popular dynamic electricity rates and existing back office, interval metering, and billing infrastructures— lower cost of operations electricity bills are the result.

4.5. Use Case: Cable EV Fleet Charging

The initial use cases that led to the development of OLS within the SCTE micro-grid working group are concerned with lowering costs, increasing network reliability, and enhancing operations during the transition of cable fleets from internal combustion engines to EVs. In the absence of load controls, local load spikes may occur when numbers of EVs stop service for the day and attempt to recharge at the same time. Such spikes can easily overwhelm a local grid endpoint, such as a transformer, leading to expensive and time-consuming upgrades to the distribution grid. Such software-based mitigations are called ‘non-wired solutions’ in the utility industry and are highly prized as they avoid capital outlays.

With OLS, the EV charging load can be shaped to follow renewables and smoothed out over several hours, as described above. This provides cable operators more flexibility in when to recharge fleets and could lower overall costs where Time-of-Use dynamic pricing or other incentives are provided by utilities.

Figure 5 illustrates an example where pricing follows expected load. In this particular example, load shifting can reduce OLS consumer costs by approximately 20%.

The effort to create the OLS standard benefited from having specific use cases identified, but the resulting standard has applicability to any scenario in which a load can be shifted in time.

The authors have built a web-based API that provides SCTE-compliant OLS signals for nearly 22,000 grid transmission interconnection regions across the US. The OLS signals reduce the cost of purchasing electricity. Broadband providers may consume an OLS and further target and optimize to specific sub-regions or consumers. For example, a micro-grid controller might consume a regional OLS and subsequently pass further-optimized OLS signals to sub-components like battery walls, EV charging systems, building management systems, thermostats, and other flexible loads.

5. Conclusion

ANSI/SCTE 267 2021, Optimum Load Shaping, is intended to benefit the cable industry in lowering energy costs and improving operations. It's meant to directly address the transition of the fleet of service vehicles to EV and can also provide value by being applied in many other use cases, including many beyond the cable industry.

Besides cost reductions and operations improvements for cable operators the benefits of OLS include:

- Reduce fossil fuel utilization
  - Improve efficiency by reducing peaks, starts/stops, up/down ramping
  - Reduce fuel costs, water consumption, pollution
- Accelerate transition to low-carbon economy
- Reduce investments in electrical infrastructure
  - Provide non-wires solutions to local congestion
- Lower costs to electrical consumers where time of use or other incentives are offered
- Improve Service Continuity for cable operators as grid failures are reduced

Abbreviations

ADR	automated demand response
EV	electric vehicle
HTTP	Hypertext Transfer Protocol
ICE	internal combustion engine
kWh	kilowatt hour
OLS	Optimum Load Shape
Protobuf	protocol buffer
REST	representational state transfer
SCTE	Society of Cable Telecommunications Engineers
YANG	Yet Another Next Generation [Model]

Bibliography & References

ANSI/SCTE 267 2021: *Optimum Load Shaping for Electric Vehicle and Battery Charging*

R. F. Cruickshank*, Estimating the Value of Jointly Optimized Electric Power Generation and Residential Electrical Use, Ph.D. thesis, University of Colorado, Boulder, Colo (September 2019).*

OLS Blog, https://optimumloadshape.com/

ORAN, The Future Of Wireless Architectures

A Technical Paper prepared for SCTE by

Bill Beesley, Principal Solutions Architect, Fujitsu
2801 Telecom Parkway
Richardson, TX 75082
bill.beesley@fujitsu.com
972.479.2098

1. Abstract

As MSOs continue to transform their networks to be more wireless centric, many are considering the architectural map provided by the Telecom Infrastructure Project's Open Radio Access Network or ORAN project. ORAN has defined a set of open interface specifications that allow vendor neutral disaggregation of the radio access network software and hardware that delivers increased flexibility, via best of breed vendor options and eliminates the issues of vendor lock in that currently burden the wireless infrastructure ecosystem. This paper will outline what ORAN is and how it can be implemented in an operationally sustainable fashion sharing examples from existing operator implementations. The paper will provide how MSOs can and should use an ORAN architecture to increase flexibility in vendor selection, take advantage of virtualization and containerization of software, provide innovation via adoption of new technologies and maximize supply chain diversity

2. Introduction

When Alexander Graham Bell (or Elisha Gray depending on which side of the historical fence you stand) created the first telephone in 1876, it took another 75 years for the telephone to reach 50 million users. Television, which started broadcasting in 1929 took 33 years to reach 50 million users. The World Wide Web took only 4 years, and the popular app Pokemon Go only took 19 days to reach 50 million users. What we've seen happen over the last 100 years is that the pace at which people adopt technologies has accelerated astronomically. Much of this increased pace is arguably because of the Internet, which hit 50 million users in 7 years, itself a technology evolved on the telephone lines created by Bell. It is hard to imagine what Bell would think of what his invention has evolved into and even harder to imagine he would have believed that most of that evolution has happened in the last ten years.

Transport networks have also had to evolve to keep up with the expectations of users. Network function virtualization has increased the agility of networks, made them more programable and adaptable, and ultimately made them less expensive to own and operate. Open optics, open line systems, open ROADM, DOCSIS, and other open standards work has further allowed innovations in networks. Globally, ORAN is beginning to bring these same evolutions to the wireless mobility ecosystem. ORAN is allowing operators to evolve their networks to be more flexible, capable and agile. But here in the US, there is still much skepticism on the viability of ORAN, even among operators who are currently creating greenfield networks and are not constrained with having to own and operate an existing closed ecosystem.

3. What is ORAN

Generally speaking, ORAN represents efforts to standardize the radio access network by disaggregating the hardware from the control software and defining standard interfaces to allow a multi-vendor network ecosystem. For the purposes of this paper, the author is intentionally conflating the work being led by two groups into a single term, "ORAN. The first is the O-RAN Alliance and the Telecom Infrastructure Project's OpenRAN working group. O-RAN with the hyphen is an alliance founded in 2018 by AT&T, China Mobile, Deutsche Telekom, NTT DOCOMO, and Orange with a mission to reshape the RAN industry towards more open and competitive ecosystem. Next is TIP's OpenRAN program comprised of both suppliers and operators who collaborate to create open, multi-vendor interoperable products. Traditionally, radio access networks or RANs were closed ecosystems, often with all the necessary equipment and management and control software coming from a single vendor.

The most recent version O-RAN architecture specifications divide the RAN into the radio side, and the management side. On the management side, there is the service management and orchestration components as well as the Non-Real time RAN Intelligent Controller (RIC) which is primarily responsible for high level intent-based control and optimization of RAN resources. The non-real time RIC is analogous to a traditional element management system providing element visibility, high level management and reporting. On the radio side, the near-real time RIC is responsible for supporting more granular control and data, generally via microapps delivering functions such as anomaly detection and mitigation The radio side also defines the radio unit (RU) providing RF interfaces and protocols, the centralized unit (CU) supporting higher layers of the stack (eg SDAP, PDCP, and RRC) and the distributed unit (DU) supporting the lower layers (eg RLC, MAC, and PHY). These abstracted elements can be delivered by either a single supplier or the operator can select to obtain these from multiple suppliers.

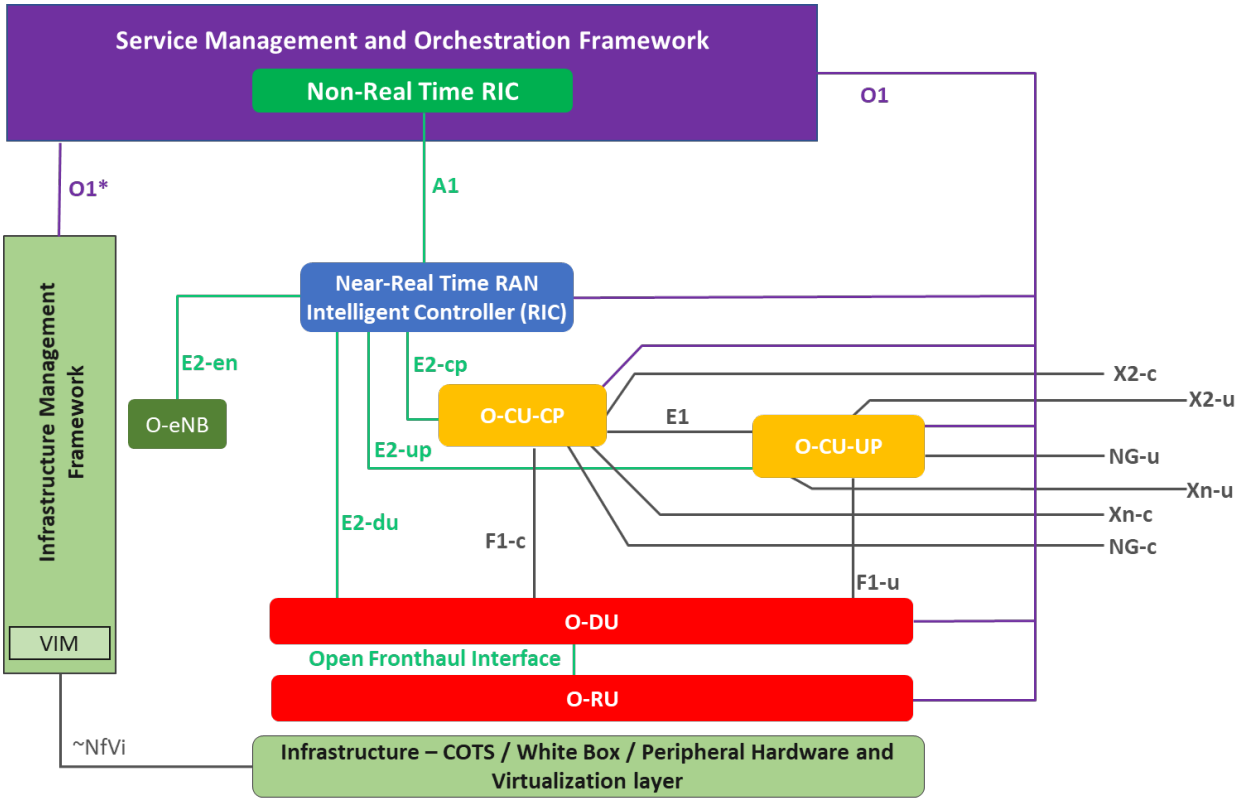


Figure 1 - O-RAN Alliance Logical Architecture

4. Arguments against ORAN

Historically, open systems such as the original IBM PC, DOCSIS, Open ROADM, and others have driven down costs to own and operate their respective technologies, but more importantly they spurred innovations in those industries that allowed the original use cases to grow. So, why would anyone be resistant to ORAN? Frankly, there are very valid concerns about how ORAN changes how wireless networks are built and operated. The current 4G and 5G technology that allows us to have near ubiquitous access to communications and information are very complex systems with a lot of moving parts that must operate in a consistent and coordinated manner. Operators who are skeptical of ORAN

generally have two reasons as to why they are resistant to the new paradigm. The first is supply chain simplicity. We often hear this phrased as “one throat to choke”. By having a single vendor ecosystem, there are fewer components to manage in the network. There is one company to call on for support, technicians have fewer systems to be trained for, and warehousing of equipment for deployments or spares is simpler. The second argument builds upon the first, closed ecosystems are more fully integrated whereas ORAN aside from TIPs interop lab work has little integration. There is a valid concern among operators that they will be left with total responsibility to integrate systems from multiple vendors and if problems arise, will be left in between suppliers who are pointing their fingers at each other. There is also the obvious concern that integration in a network does not happen just one time. Over the life of the network, as elements are updated and patched, this integration may have to happen multiple times. Thus, they take what is perceived as the less risky approach of selecting a vendor to supply most if not all the components necessary to build their radio network on the premise that vendor will be the “only throat to choke” to ensure systems are fully integrated, secure and operate as envisioned. Most of these operators are taking a “wait and see” attitude towards ORAN and opting to continue deploying traditional closed ecosystem networks. Recognizing the very valid concerns of operators who must manage and maintain networks at scale, the remainder of this paper will show how these concerns can be managed as well as a number of reasons why the risk in moving to the new paradigm is worth the benefits gained.

5. Supply chain advantages

The counter argument to the supply chain simplicity argument is one of supply chain integrity. Diversity of suppliers lowers the risk to your operations and gives you more flexibility. The current pandemic has brought us far too many stories of multiple supply chain disruptions. Historically events like natural disasters, disease, and even industrial accidents have caused supply disruptions of one kind or another. The telecommunications industry is currently experiencing supply shortages that are creating component lead times that may have traditionally been days or weeks to currently running into years. To mitigate this and ensure their networks can continue to grow and maintain the pace of demand mentioned earlier in this paper, operators are having to risk buy equipment that they won’t receive until months or a year or more later or risk disrupting business growth. ORAN allows for vendor diversity that de-risks supply chain disruptions by allowing the operator to select alternate vendors should their primary supplier not be able to deliver.

Selecting a single vendor for your network also constrains your ability to deliver new features or capabilities as they become available in the market. With a single supplier you are limited by their roadmap and their view of what technologies are important to bring to market and when. With an open ecosystem of suppliers, you are not locked into a single vendor’s vision of the network. When innovations become available, if another supplier brings them to market first you can take advantage of the innovation because you are not locked in. Open systems allow innovations to come to market and begin delivering value to your network and your customers faster.

Lastly, it is very difficult to imagine that any single vendor is going to be the best at everything that makes up a network as complex as the modern RAN. Having an open ecosystem allows you to select best of breed suppliers and make fewer compromises in your network.

6. Systems integration

While it is generally true that ORAN presents new integration challenges, it is not entirely true to suggest that if an operator elects to go with an open network, they are forced to take on all the integration

challenges themselves. First, there is nothing stopping an operator from deploying an open ecosystem in partnership with a single supplier who takes on all the integration responsibilities. This is a bit like the old adage, eating your cake and having it too. By having a single supplier responsible for the final solution, the operator takes on less effort and risk themselves while still allowing them to add in alternative technologies and products in the future should they decide to shift. Having the flexibility to pivot at some point in the future in this way is extremely difficult with vendor specific systems.

Operators can also take advantage of the work that TIP has been doing to create the OpenRAN Exchange which is designed to accelerate the creation of an ORAN ecosystem. The marketplace has a wide list of suppliers of RU, CU/DU, transport, software, consulting services and other components that have been demonstrated to provide best of breed technology and interoperability that is the promise of ORAN. More information can be found on their website listed in the references at the end of this paper.

Much of the concern of solutions integration for ORAN are essentially based on the fear of the unknown and operators are apprehensive about being able to not just deploy a service once, but ensuring the ability to consistently deliver a service in a repeatable manner thousands of times. Fortunately, wide industry support for deploying ORAN is available and growing daily.

7. ORAN and Security

There is a lot of back and forth over whether closed or open ecosystems are more or less secure. Generally speaking, the most adopted and popular systems, closed or open will tend to get the most attention of bad actors attempting to exploit flaws and gain access. It is also important to remember that if you are depending on a closed system being more secure because the internals are less widely known, there is an old adage that obscurity is not security. The advantage that open systems have over closed is the sharing of information about faults and exploits is much faster and much more transparent. Issues tend to be identified and brought to community awareness more quickly. There is also much more flexibility in developing workarounds more quickly and the operator often does not need to wait for one to be delivered as they have the option of implementing their own. Lastly, bugs and other security issues tend to get resolved faster in open systems because there are more developers contributing. A community of five thousand developers can outpace a company who may have as few as five developers working on a project.

8. Network Agility

As mentioned earlier in the paper, modern telecommunication networks must consistently evolve to meet the ever-increasing demands users bring. What Bell started as a communications network has evolved to become an information and content delivery system that even just a few short years ago was unimaginable. Emerging use cases for the RAN such as Industry 4.0, IoT, gaming, autonomous vehicles, AI/ML, augmented reality and others require not only increased bandwidth and lower latency, they also require networks that are adaptable and agile. The modern radio network will need to be envisioned as software that can be automated and programmable.

In addition to this, to achieve the latency and bandwidth promises of 5G, orders of magnitude more radios will need to be deployed so that the serving groups can have smaller numbers of users. Much like fiber deep technologies like Remote PHY are delivering making more bandwidth available per subscriber, highly distributed antenna architectures will soon become the norm. One major MSO has publicly

shared at trade conferences a vision of a radio network that is not the traditional large macro cell towers but thousands upon thousands of small strand-mounted radios deployed deep into the neighborhoods.

While closed ecosystem networks can support all of the above 5G features, having the most flexible, agile network will become a competitive necessity to support them. Open systems are much more adaptable, agile and flexible.

9. Opportunities for new revenue

Admittedly, the concerns raised over migrating to an ORAN architecture are valid even if they use appeal to possibility or fear of the unknown as their basis. We can't blame operations executives and their teams for being hesitant to dive into the unknown, especially when they are the ones held accountable should the network or systems not operate in a consistent and reliable manner. As a NOC expert friend of mine used to quip, "there are no optimists in Operations".

But the industry has successfully implemented open architectures in other areas of the network – DOCSIS is one example that comes to mind – and these implementations not only saved the industry both expense in purchasing the network and operating it, they drover further innovations that created new revenue streams.

Ultimately this is where ORAN will provide the biggest benefit to our industry. It supports an open and agile network that allows for innovations to bring new revenue opportunities to the network more quickly.

One of the most intriguing new service offerings that operators are expressing interest in is E5G or enterprise 5G. This is the delivery of private 5G networks, primarily to large enterprise and heavy industry to support use cases such as deterministic band for IoT and sensors in manufacturing, deterministic low latency/high bandwidth networks to support autonomous vehicles for industries such as mining, and the ability to slice out private networks so that mobility devices can access local network resources in a secure and reliable manner. Again, as stated 5G deployments can be accomplished with closed ecosystem networks but to provide competitive offerings that take advantage of 5G technologies such as network slicing, function virtualization and programmability, ORAN provides a clear differentiator. To be competitive in the emerging E5G space, operators will need to offer highly flexible and tailored offerings as no two enterprise networks have the same internal requirements. By selecting an open architecture, operators will be able to create offerings that appeal to the small to medium enterprise sweet spot that makes up 92% of the businesses in the United States and where the MSOs have been wildly successful in the past. ORAN will allow the operator to create highly adaptable offerings while minimizing capital and operating costs, especially for the edge radio and EPC functions as there will be more down-scalable products that minimize the cost to deploy.

Finally, it seems appropriate to give a mention to another upcoming open specification known as SCTE GAP. GAP is a working group that has created a standard set of mechanical specifications for outdoor equipment that allows for component reuse and the creation of a supplier ecosystem that will drive down costs and increase innovation much like the IBM PC standard did. As of the time of the writing of this paper, the final draft of the GAP spec is undergoing the final voting and comment process at the SCTE and should be finalized by the time this paper is published. What is innovative about GAP is that the service delivery modules can be mixed and matched to support multiple use cases at the edge. This means a GAP deployment could contain a cable modem module for transport, a CBRS small cell module and a compute module to push virtualized compute resources right to the subscriber edge. This opens up a number of use cases such as the aforementioned E5G. IoT for smart cities, edge gaming, or even

offering real-time access to the radio network through slicing and virtualization. This would allow the MSOs to offer access to highly distributed radio and compute resources to provide something similar to an MVNO offering but with a much higher degree of granularity and flexibility. It is not hard to imagine that given enough time to deploy thousands and thousands of distributed radios in the MSO network, which already has space, power and transport to the edge, that at some point in the future a cable operator could become a dominant wireless provider in the areas they serve. But using legacy closed systems won't create the next generation network that can support this vision.

10. Conclusions

As previously stated, ORAN doesn't mean that the operator has to become the wireless system integrator managing multiple supplier relationships and taking on all the risk for issues that arise. MSOs can select single suppliers to either provide a fully integrated ORAN solution or they can work with ecosystem suppliers that are emerging via working bodies such as TIP to build, deploy and maintain their infrastructure. What is important to remember is that ORAN means flexibility. Flexibility in defining what capabilities you want to have in your network. Flexibility to select best of breed suppliers, whether now or at some point in the future. And ultimately, ORAN provides the type of agile and programmable network that will be a competitive differentiator in our 5G IoT world

11. Abbreviations and Definitions

11.1. Abbreviations

5G	Fifth generation
AI/ML	Artificial intelligence/machine learning
CU	Central unit
DOCSIS	Data over cable service interface specifications
DU	Distributed unit
E5G	Enterprise 5G
GAP	Generic Access Platform
IoT	Internet of things
MSO	Multi service operator
MVNO	Mobile virtual network operator
NOC	Network operations center
ORAN	Open Radio Access Network
PC	Personal computer
PHY	Physical layer
RAN	Radio access network
ROADM	Reconfigurable optical add drop multiplexer
RU	Radio unit
SCTE	Society of Cable Telecommunications Engineers
TIP	Telecom Infrastructure Project
US	United States

11.2. Definitions

Network slicing	A network architecture that enables the multiplexing of virtualized and independent logical networks on the same physical network infrastructure
5G	Fifth generation network technology standard for broadband cellular networks
Industry 4.0	Fourth industrial revolution
Open line system	Optical line system allowing vendor interoperability
Open ROADM	Interoperability specifications for ROADMs
O-RAN Alliance	Alliance between wireless operators to shape open RAN standards

12. Bibliography and References

How Long Does It Take to Hit 50 Million Users? Jeff Desjardins, Visual Capitalist, June 8th, 2016 https://www.visualcapitalist.com/how-long-does-it-take-to-hit-50-million-users/

O-RAN Alliance https://www.o-ran.org/

ORAN Alliance Logical Architecture – obtained from https://docs.o-ran-sc.org/en/latest/_images/o-ran-architecture.png

TIP Exchange https://exchange.telecominfraproject.com/

The GAP, Ed Dylag, July 21th, 2021 SCTE Broadband Library https://broadbandlibrary.com/the-gap/

Overlaying Mid-Band Spectrum Backhaul/Fronthaul onto HFC

A Symbiotic Convergence of Cable & Wireless

A Technical Paper prepared for SCTE by

John Ulm
Engineering Fellow, Broadband Systems
CommScope – CTO Network Solutions team
Moultonborough, NH 03254
+1 (978) 609-6028
john.ulm@commscope.com

Dr. Martin Zimmerman, PhD
Engineering Fellow
CommScope – Outdoor Wireless Networks
2400 Ogden Ave, Suite 180, Lisle IL 60532
+1 (815) 341-4364
martin.zimmerman@commscope.com

Stuart Eastman
Principal Systems Eng
CommScope – Access Technology team
101 Tournament Dr. Horsham PA 19044
+1 (215) 323-1140
stuart.eastman@commscope.com

Zoran Maricevic, PhD
Engineering Fellow
CommScope – CTO Network Solutions team
15 Sterling Drive Wallingford, CT 06117
+1 (203) 303-6547
zoran.maricevic@commscope.com

1. Introduction

Cable 10G and Wireless 5G may seem to be at odds. However, when combined, they offer an evolutionary strategy with much synergy.

5G uses a collection of different frequency bands, each with unique characteristics. Recent developments in C-band, CBRS (Citizens Broadband Radio Service) and Wi-Fi provides some new mid-band spectrum (i.e. 3 - 6 GHz) that is offering a middle ground that may be the future wireless workhorse. Its reach covers a significant number of mobile users with substantial data rates. But its deployments may need many more densely packed cell sites than current 4G LTE macro-cells. This presents an opportunity for MSOs to leverage their existing HFC infrastructure for providing both backhaul and power to those new cell sites.

The paper presents a basic tutorial on mid-band wireless technologies in the 3-6 GHz range that includes C-band, CBRS and Wi-Fi 6E. It covers MIMO antenna systems from 2T2R to 64T64R and when and where each is appropriate. ORAN (Open Radio Access Network) standards help to virtualize the 5G infrastructure, identifying backhaul, midhaul and fronthaul interface options.

The many choices for the mid-band wireless system can vary bandwidth requirements from 100's Mbps to many 10's Gbps. The paper shows which configurations can easily be supported on DOCSIS 3.1 while others might require DOCSIS 4.0 and some may need direct fiber connect.

Some case studies are provided where potential mid-band small cells are mapped to actual HFC networks. Results from a CBRS design show its potential reach. This data is used to map cells to several existing HFC nodes. The nodes under study vary from dense urban nodes (i.e. >250 HP/mile) down to sparse rural nodes (i.e. <40 HP/mile). Various trade-offs are considered in cell site placement on the HFC.

HFC appears to be ideally suited to support this Mid-band xHaul infrastructure. A strategy is laid out for cable plants of varying densities. D3.1 midhaul can be leveraged extensively in the early days to get wide coverage quickly. Very dense urban areas will eventually require complex antenna/MIMO systems with fiber fronthaul. This integrates nicely with an N+2 fiber deep strategy. But even then, cells with DOCSIS xHaul will be needed to fill the holes and hotspots. D4.0 then enables even higher capacities at these cable cell sites.

In the end, Cable and Mid-band wireless (C-Band, CBRS, Wi-Fi 6E) are much stronger together and are at the core of a next gen converged network evolution.

2. 5G Midband and Wireless for Cable Dummies

For many cable operators, wireless in general and 5G in particular is a brand new, if not foreign, technology. This section provides a tutorial to help educate cable technologists in this area.

Mobile Wireless Services have been deployed since the mid-1980s through a succession of generations:

- **1G**: 1980s – Analog signals, typically 1 Tx port and 1-2 Rx ports on the radio (diversity improved signal reception). Frequency bands were 850 MHz in US and 900 MHz RoW (rest of world). Voice only
- **2G**: early 1990s – Digital signals, typically 2 ports on radio one for Rx, one for Tx/Rx. Initially Voice only but later technologies such as GPRS and EDGE allowed data to be encoded as if it were voice for early data transmission. First introduction of mid-band spectrum 1900 MHz in the US, 1800 MHz RoW. Fairly quickly all existing 1G services were converted to 2G.
- **3G**: 2000 – Digital signals, designed for data transmission, again 2 ports on the radio, only one doing Tx. New frequency bands were added, specifically at 2100 MHz (slightly different bands in US and RoW)
- **4G**: 2009 – (a.k.a. LTE, Long Term Evolution) Digital signals, designed for data transmission. MIMO introduced (Multiple Input Multiple Output) which improved data capacity by using multiple transmitters and receivers. In general, radios had 2 or 4 ports with all ports capable of Rx and at least 2 Tx ports. New frequency bands including 700 MHz and 2600 MHz in the US, 800 MHz, 2300 MHz and 2600 MHz RoW. VoLTE (Voice over LTE) encoded voice as data (similar to VoIP) to allow 4G systems to handle voice traffic. This allowed the decommissioning of 2G and 3G networks to commence.
- **5G**: ~2020 – Digital. Improved efficiency compared to 4G. In addition, 5G is meant to be more flexible, so that it could in theory replace not only existing mobile wireless communications standards (e.g. 4G) but also those for fixed wireless, vehicle anti-collision radar, Wi-Fi, Bluetooth, etc. To date 5G has really only been applied to achieve faster mobile wireless, but other Use Cases remain to be explored.

There are several aspects in which mobile wireless networks differ from fixed wireless networks such as Wi-Fi. Mobile wireless networks are designed such that a service area is broken up into a collection of cells and the network transitions the user from one cell to the next automatically and seamlessly as the user moves geographically. Therefore, mobile wireless is often referred to as cellular service. With Wi-Fi this transition is handled manually.

Another key difference is that Wi-Fi uses shared spectrum. This results in limits in terms of both antenna gain and maximum transmitted power, as well as their sum EIRP (Effective Isotropic Radiated Power). EIRP is connected to coverage – higher EIRP means that the signal strength will be adequate for data transmission over a wider area.

Most mobile wireless spectrum is licensed within geographic regions and within the region the license holder has exclusive use of the spectrum. This means that the license holder does not need to worry about interfering with other users. This enables them to achieve much higher EIRP levels. For example the UNII bands (which includes the 5 GHz spectrum that Wi-Fi uses) has a maximum allowed EIRP of 36dBm while most mobile wireless networks deploy radio and antenna systems with maximum output in the range of 65-75dBm (so a factor of 1,000 to 10,000 higher EIRP). The new CBRS band is a unique case. The spectrum is shared, but the power levels are higher, with a maximum of 47dBm EIRP per 10 MHz channel.

2.1. 5G Midband – What is all the hype?

5G is a collection of different frequency bands, each with unique characteristics. Recent developments in C-band, CBRS & Wi-Fi provides some new mid-band spectrum (i.e. 3 - 6 GHz) that is offering a middle ground that may be the future wireless workhorse. Its reach covers a significant number of mobile users with substantial data rates. But its deployments may need many more cell sites then current LTE macro-cells.

Citizens Broadband Radio Service (CBRS) is a first-of-its-kind effort to get maximum utilization out of spectrum. It refers to 150 MHz of spectrum in the 3550 MHz to 3700 MHz range that the FCC has designated for sharing among different tiers of users. The 3.5 GHz band has been identified as a critical band for wireless mobility. This frequency is low enough to have good propagation characteristics, particularly in comparison to extremely high frequency millimeter waves (mmWave). But it is also high enough so that advanced antennas using M-MIMO (massive multiple input multiple output) technology are small enough to meet zoning restrictions and be deployed.

However, once 3.5 GHz was proposed for usage for mobile wireless, the US found itself with a problem. The spectrum right around 3.5 GHz was already being used, by the US military in coastal regions and by incumbent Wireless Internet Service Providers (WISPs) inland. Most countries would have given up on this spectrum, but the FCC came up with a clever plan to maximize usage involving a 3-tier hierarchy.

The highest priority tier goes to US military applications and (for now) other legacy incumbents. SAS (Spectrum Access System) serves as a traffic cop, telling other users to shut down when the US military is using the spectrum. However, since the US military application is primarily ship-borne radar, the usage is mostly confined to coastal regions, particularly a few spots where US naval ships are based and even in those regions the usage is sporadic.

The second priority tier goes to PAL (Priority Access License) license holders who have paid to have exclusive use of 10 MHz channels within a specific geographic region (for CBRS these regions are counties). In any county, a single entity can own up to 4 PAL licenses which guarantees 40 MHz out of the 70 MHz available for PAL license holders.

The lowest level priority tier is General Authorized Access, GAA. GAA users have free access to the spectrum on a first-come, first-served basis. Since no more than 70 of the 150 MHz can be licensed under PAL, GAA users are assured that at least some spectrum will always be available for GAA use.

Since GAA usage is free, this allows end users to build and run mobile wireless networks for a fraction of the cost that would be required if licensed spectrum were being used. This puts CBRS into the same category as other free spectrum services such as Wi-Fi or Bluetooth. However, since CBRS has a EIRP cap of 47dBm instead of 36dBm, individual sites can cover a much larger area then Wi-Fi. This can be very advantageous in a campus or office park setting. And since CBRS typically is deployed with mobile wireless technology (usually 4G or 5G, though some radios use proprietary systems), true mobile wireless service is available, and handoffs can be managed automatically. This gives CBRS a further advantage over Wi-Fi.

In 2020, several mobile network operators (MNOs), primarily Verizon, AT&T and T-Mobile, spent over $80 billion purchasing C-Band spectrum (3.7 – 3.98 GHz) at the U.S. Federal Communications Commission (FCC) auction. Based on an accelerated clearing schedule, 100 megahertz of the auctioned spectrum will be cleared in 46 of the top U.S. markets by December 2021. Verizon and AT&T won 60 MHz and 40 MHz, respectively, of the earliest available C-band “A” blocks. Verizon’s deployment plans initially call for turning up spectrum at existing macro sites focused on 46 markets. Rural fill-ins, small

cells and in-building are part of the picture down the line. By December 2023, the remaining 180 megahertz in these same 46 markets, as well as the full 280 megahertz in the other markets, will be cleared for use by 5G services.

Comparing C-Band with CBRS, the C-Band EIRP limits are much higher – the combined antenna and radio may generate 76dBm. But none of the spectrum is free and in fact the licenses for C-band were much more expensive on average than CBRS PAL licenses.

Meanwhile in the United States and other countries, 1.2 GHz of spectrum from 5.9 to 7.1 GHz has been set aside as unlicensed spectrum that is being used in Wi-Fi 6E. However, the higher 6 GHz frequency band and lower transmit powers will limit the outdoor range for Wi-Fi 6E.

2.2. Antenna 101 – Success starts with the Antenna

Outdoor wireless network success starts with the antenna. It is the equivalent of the speaker/headphones for an audio system. Specifically, the antenna transforms the guided RF (Radio Frequency) energy generated by the radio and carried by transmission lines into free space electromagnetic waves that propagate through the atmosphere.

It should be noted that base station antennas (BSAs) are not intended for point-to-point communications. The goal of a base station antenna is to provide relatively uniform coverage in an area where coverage is desired (inside the cell area) with a minimum of excess energy going outside the cell area.

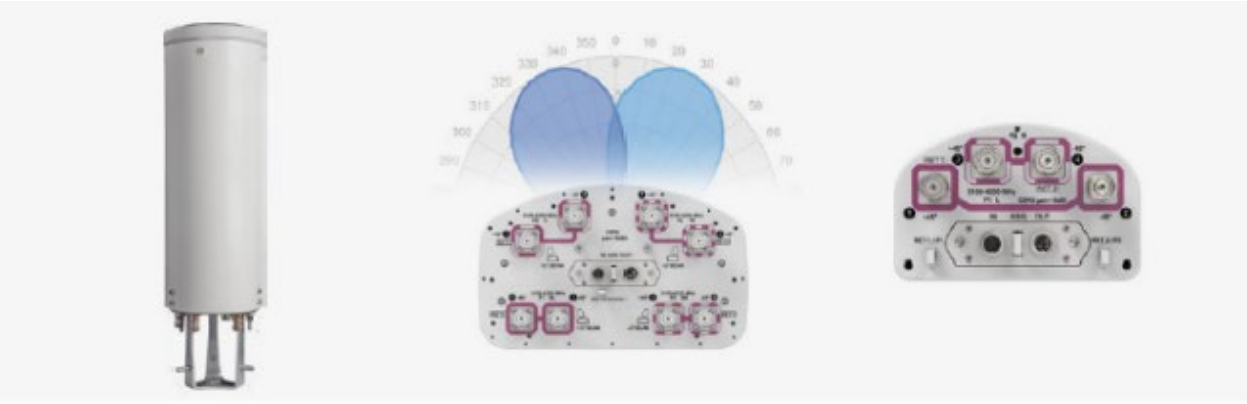
2.2.1. Omni vs. Sectored Antenna

There are different types of cell arrangements from the perspective of the base station equipment. In the 1G systems and even today in rural areas or for small cell applications, the site where the base station equipment is located will provide wireless coverage in all directions around the site using an omni-directional antenna. This is considered a single-sector site.

Note that in the wireless field, “omni-directional” refers only to the pattern in the azimuth plane. This is in contrast to academia where “omni-directional” is often thought of as identical to “isotropic” which indicates that energy is uniformly radiated throughout the 4π steradians of space. So in mobile wireless, one common antenna type is a “high gain omni” which marries an omni-directional azimuth pattern with a narrow, highly-directive elevation pattern.

The most common cell configuration today is a 3-sectored site (Figure 2). In this case the area surrounding the base station tower is divided into 3 120-degree sectors, each with its own set of radios and antennas. Typically, the antennas used will have a HPBW (half power beamwidth) of roughly half the sector size, so in this case 65° azimuth HPBW antennas are usually used. Since modern cellular systems have 100% frequency re-use, a 3-sectored site offers 3 times the data capacity of a 1-sectored site for the same geographic area.

Finally, in cases where very high capacity is required, a commonly chosen option is a 6-sectored site (Figure 2), which can offer up to twice the capacity of a 3-sectored site. For these sites, typically antennas are used with a 35° azimuth HPBW. In many cases operators will use special “twin-beam” antennas, where a single antenna is designed to provide coverage for two of the 60-degree wide sectors. This minimizes the amount of clutter at the top of the tower. Images of antennas for 1-sector, 6-sector and 3-sector sites are shown in Figure 1.



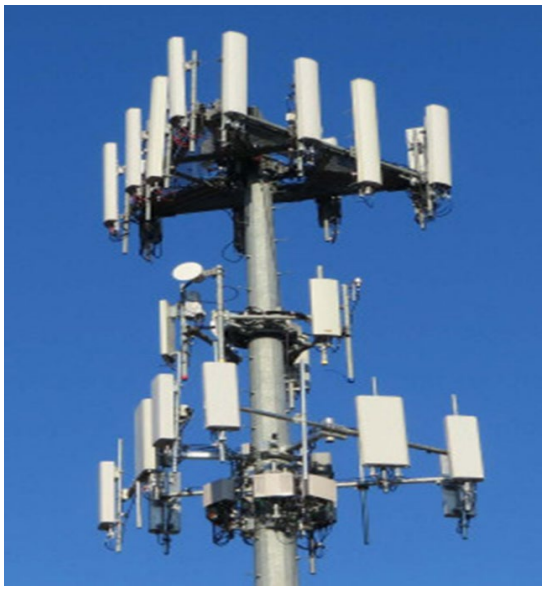


Figure 1 – Antenna technologies

Figure 2 – Sector antennas – Three to Six Sectors

Note that the full theoretical capacity may not be realized due to the overlap of radiated energy between sectors causing interference. In a cellular network there is always a balance that must be maintained between coverage and capacity. If there is too little overlap in coverage of the individual sectors, then holes in coverage may appear. But if there is too much overlap then interference builds to a level where capacity is reduced. Capacity generally is a function of SINR (Signal to Interference and Noise Ratio) which measures this level of interference relative to the primary signal strength. So, a goal of antenna designers is to make antennas that maximize the amount of energy that goes into the sector relative to energy going outside the sector. And the goal of an RF planner (network designer) is to deploy the antennas in such a way to further improve focusing of energy into the sectors.

2.2.2. EIRP considerations

A unique feature of CBRS is the medium level EIRP cap of 47 dBm. In Wi-Fi systems with a 36 dBm cap, it is typically assumed that the antenna gain will be quite low as well as the transmitted power. In other licensed bands, such as C-band, the EIRP cap is so high that it is effectively never reached, and both passive and active antenna systems are designed without any concern for the EIRP cap. The only limit on gain is the fact that the antenna must cover a certain region and so a very narrow fixed beam might not be appropriate.

For macro cell applications, C-band radios can offer up to 320W RF power and 21-25 dBi antenna gain, implying EIRP in the 78 dBm range. In a small cell or strand mount application, the radio power might be 40W (i.e. 4 channels @ 10W) with a 10-12 dBi antenna, so more like 58 dBm EIRP. This makes CBRS competitive for small cell applications, especially since the CBRS cap is per 10 MHz channel. Therefore, if 40 MHz of spectrum is available, the actual cap is 53dBm, not 47dBm.

EIRP equals the sum of transmitted RF power and antenna gain. It describes the energy density level of transmitted signals and is distance agnostic. However, given EIRP, one can calculate energy density at a specific distance from the site.

For wireless communications, one often talks about the link budget, which is the calculation of energy that makes it from the transmitter to the receiver. Per Figure 3, the link budget is positively influenced by transmitter power P_{TX}, transmit antenna gain G_{TX}, and receive antenna gain G_{RX}. It is negatively affected by path loss L_{FS} and other system losses L_{OTHER}. The first two terms P_{TX} and G_{TX} make up the definition of EIRP. Given that EIRP is limited, one can only improve the link budget further by increasing the receive antenna gain G_{RX} or reducing path loss (e.g. install antennas at a taller height to help eliminate blockages).

Let us digress for a moment to discuss antenna gain. Gain is a measure of an antenna's ability to focus the radiated energy in a particular direction. Typically, this is measured in dBi, dB relative to an isotropic radiator, which uniformly radiates energy in every direction. Since the antenna may radiate different amounts of energy in each direction, this is expressed as a bi-variate function Directive Gain = $D(\theta,\varphi)$ where θ and φ are variables describing one's angular position relative to the antenna. Since this relative to an isotropic radiator if the antenna has high Directive Gain at some values of (θ,φ) then it must have low values (< 0dBi) in many other directions. The maximum value of Directive Gain is called Directivity. Finally, antenna gain = Directivity – Antenna Losses. Examples of antenna patterns with low gain and high gain are shown in Figure 4.

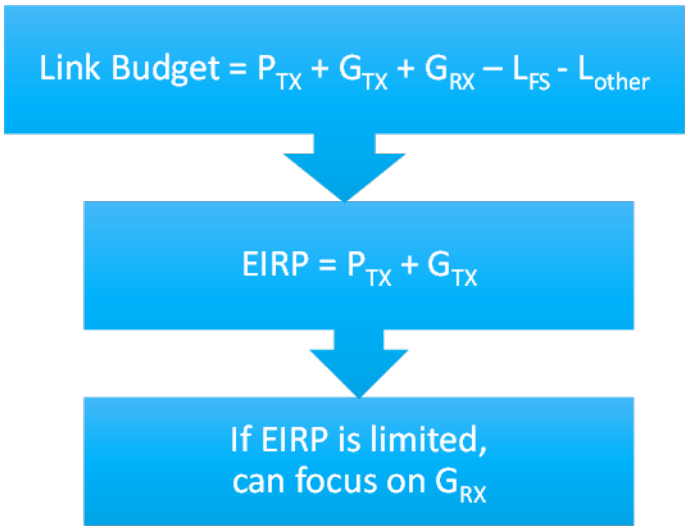


Figure 3 – Link budget vs EIRP

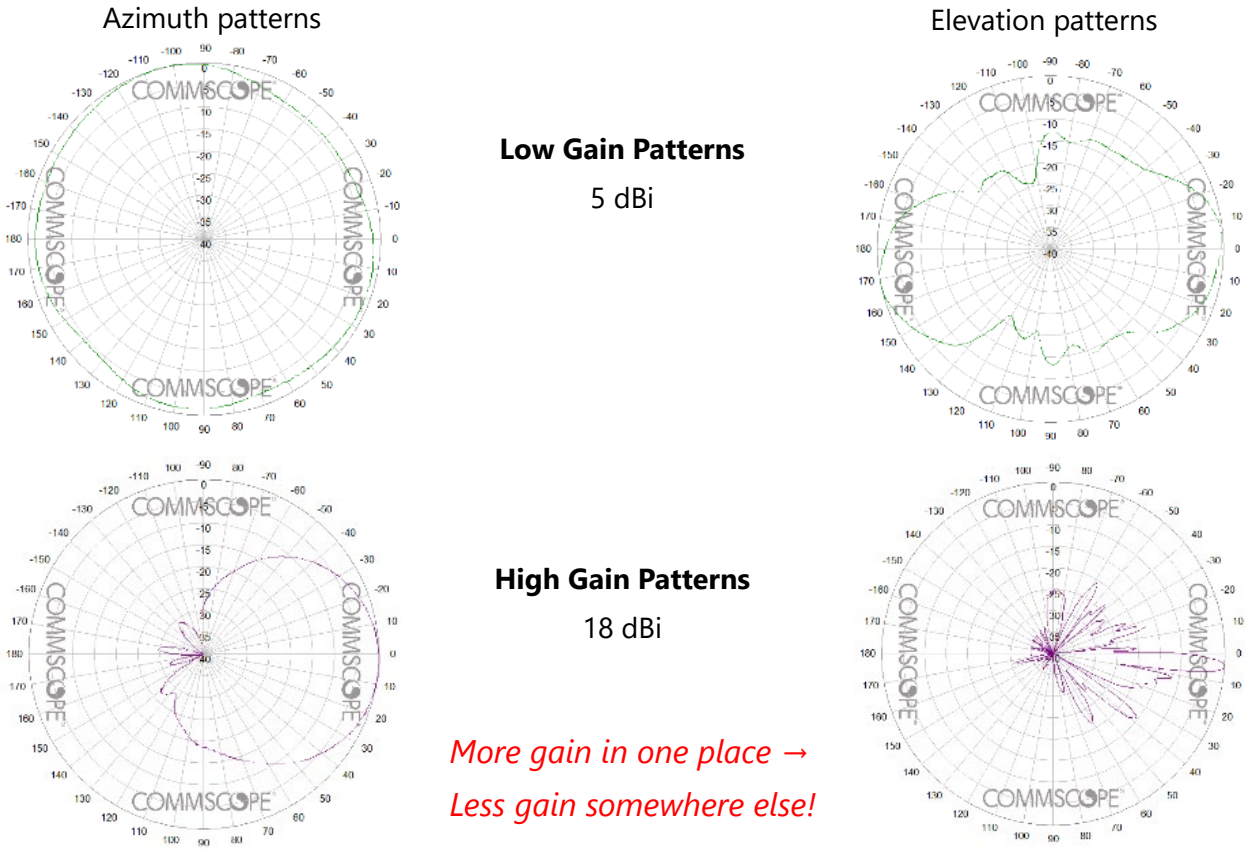


Figure 4 – Antenna Gain

The link budget equation is normally applied to point-to-point antennas which are aimed so that each has its maximum directive gain pointed at the other antenna. But as mentioned earlier, base station antennas cover an area and at the various points in the area the link budget between the base station antenna and the User Equipment (UE) antenna is better described by

$$P_{TX} + D(\theta,\varphi)_{TX} + D(\theta,\varphi)_{RX} - L_{FS} - L_{other}$$

Since antennas with high gain by necessity must have low values of $D(\theta,\varphi)$ at many angles, one can see that it is not necessarily a good thing for the antenna gain to be high.

From the perspective of the downlink, it is better to have higher input power and lower antenna gain since this implies that the energy is more evenly spread across the sector of coverage. However, from the perspective of the Uplink, both the UE transmit power and the UE gain are quite limited. So, the only way to improve the link budget is by increasing G_{RX}, the gain of the receive antenna, which for the uplink is the base station antenna. Thus, the optimal antenna design depends on whether CBRS is used only for the downlink, or for both downlink and uplink. Note – if the CBRS band is only used for the downlink, then the UE is most likely using a low band for the uplink. The low band has less path loss L_{FS} which helps the UE link budget with lower P_{TX} and G_{TX}.

2.2.3. Antenna arrays

The word antenna can refer to a single radiating element or a collection of radiating elements, called an array that are fed from a common input. Antenna arrays can be one dimensional (e.g. a single column of radiating elements) or two dimensional (e.g. a rectangular array with M rows and N columns). Most MIMO base station antennas use a single column as the array.

The size of these arrays can vary substantially, for example from 2T2R to 64T64R. The larger arrays tend to be used at macro towers and in conjunction with massive MIMO (M-MIMO). The technology enables features such as beam forming in 3-dimensions (e.g. vertical for tall office buildings). Small cells, such as strand-mount and streetlight locations, will tend to have much smaller antenna arrays due to their location restrictions.

2.3. Pattern impact on capacity

To verify the above, CommScope ran some RF simulations. Three different antennas with varying levels of gain were examined for a 3-sector grid.

1. 17.3 dBi gain, 5.4° EL HPBW, 4° tilt
2. 16.7 dBi gain, 9.0° EL HPBW, 6° tilt
3. 13.8 dBi gain, 17.0° EL HPBW, 8° tilt

The simulation was repeated for inter site distances (ISD) of 1 mile and 0.5 mile and for three different rad center heights (the height of the antenna above ground) of 45, 100 and 150 feet. The antennas were sited at the far corners of the 3 sectors and pointed towards the center of the area so that the impact of interference between cells could be taken into account. The layout for the 1-mile cell radius case is shown in Figure 5.

For this layout, each sector covers 5 square kilometers. The environment of the three sectors can be described as follows:

- Sector 3 (top): moderate/high density, residential/light industrial

- Sector 4 (right): open, flat, rangeland/airport
- Sector 5 (left): sparse/moderate density, industrial

Figure 5 – One-mile cell radius suburban scenario

The small triangles represent the locations of the sites. In particular, consider the case with 1-mile ISD and 45-foot rad center height. The simulation looked at three parameters:

- Downlink coverage (RSRP)
- Downlink capacity (RS-CINR)
- Uplink capacity (UL Allocated bandwidth throughput)

Since the results vary depending on the placement of the User Equipment (UE) in the sector, they are typically portrayed statistically via a Cumulative Distribution Function (CDF). The 3 CDF curves are shown in Figure 6. Note that for all 3 graphs, better performance is indicated by data points that are higher (same level of performance over a larger area in the sector) and further to the right (higher level of performance over the same area in the sector).

From a downlink coverage perspective, the best results come from the antenna with the lowest gain (the green curve). This is because the EIRP is capped. The lowest gain antenna G_{TX} can use higher transmit power P_{TX} and overall provides more even coverage across the sectors. The middle gain antenna provides the next best coverage (red curve) and the highest gain antenna provides the worst overall coverage (the blue curve).

For downlink capacity, the order is the same, with the lowest gain antenna performing the best and the highest gain antenna performing the worst. However, the results change for uplink capacity. Now the mid-gain antenna performs the best, the highest gain antenna performs nearly as well, and there is a large gap separating the performance of the low gain antenna from the other two antennas. The key difference is that for the downlink the transmit power could be increased to compensate for the 3.5dB difference in gain. But for the uplink capacity case the UE transmit power will be unchanged, so the higher gain becomes more important. Note that for mobile wireless systems, the Uplink path is always the weaker one because the base station radio can transmit at much higher power levels than the UE.

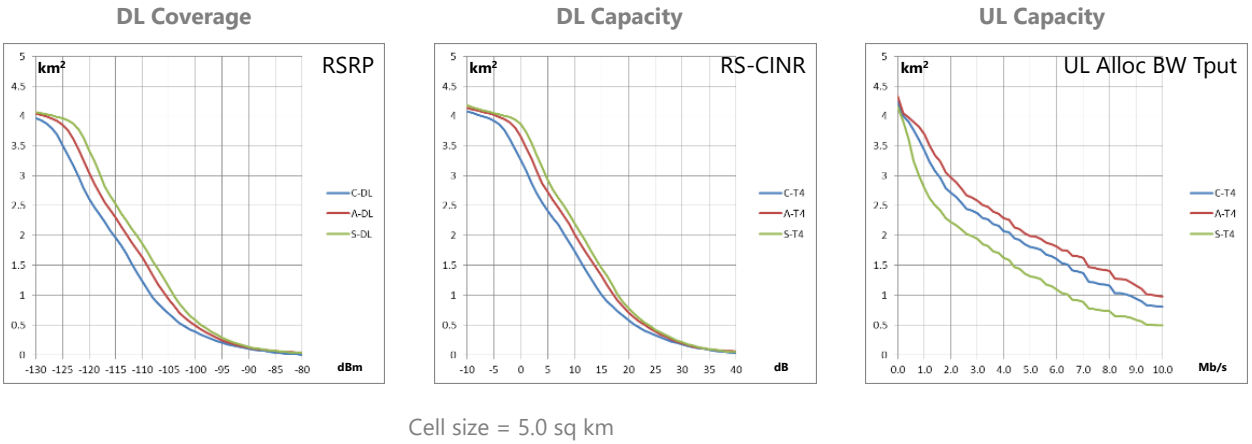


Figure 6 – One-mile cell radius suburban scenario - 45' Height Antenna

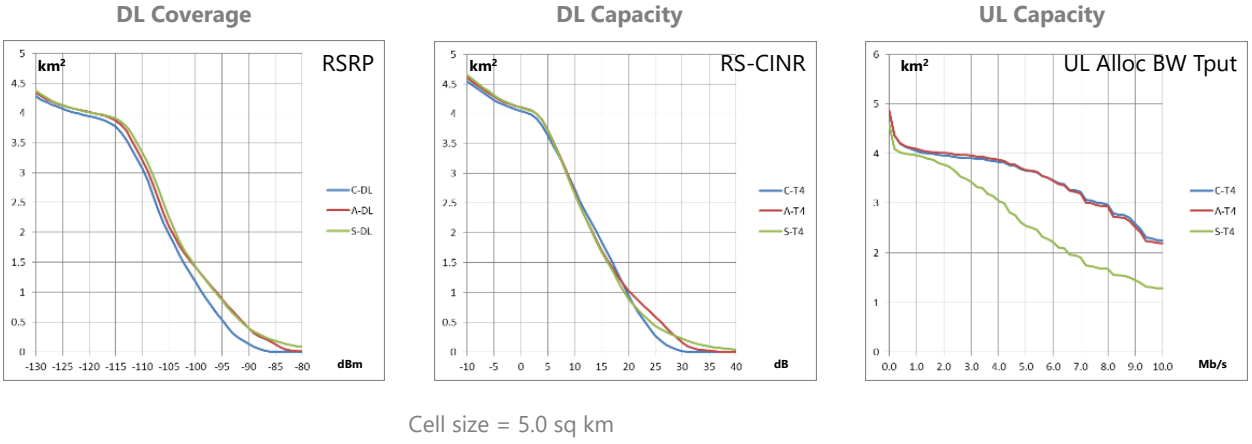


Figure 7 – One-mile cell radius suburban scenario - 150' Height Antenna

In Figure 7, the same 3 CDF plots are shown for the 1-mile ISD with a 150 foot rad center height. The results are generally the same except that the performance gap for DL coverage and DL capacity. The mid-gain and high-gain antennas perform identically for UL capacity, but there is a huge gap in performance between those two antennas and the low gain antenna.

Summarizing these results lead to the following conclusions:

If using CBRS for supplemental downlink:

- Meet the EIRP limit by using more radio power with lower gain antenna
- More even coverage due to fatter elevation pattern → higher RSRP and SINR
- **Depending on site specifics, optimal EL HPBW ~ 15–20°, optimal gain 12–14 dBi**

If using CBRS for uplink and downlink:

- Uplink thruput increased with higher gain BSA
- **Elevation HPBW 5–10° brings improved performance, optimal gain 15–18 dBi**
- Assumes 3 sectors, 65° Az HPBW
- Can get increased gain/capacity via sectorization

2.4. Small Cell Coverage

For cable operators, the most enticing use case that takes advantage of their HFC infrastructure is the small cell deployment. Perhaps the two most obvious applications will be strand mounted small cells on aerial coaxial plant and small cells mounted on streetlights. Streetlights may be the best/only option for underground plants. Streetlights also have an advantage over strand mount as they are at a higher elevation (e.g. 45' vs. 30') which enhances the reach of the antenna.

Table 1 provides some small cell range estimates for C-Band, CBRS and Wi-Fi 6E. As shown previously, the coverage is also impacted by the transmit power which varies quite a bit between technologies. Wi-Fi 6E is also at a disadvantage as it is using the 6GHz which has higher path losses than the 3.5 to 4.0 GHz bands used by CBRS and C-Band.

Table 1 – Small Cell & Strand-mount Coverage Range Estimates

Mid-band Small Cell Ranges	EIRP	Mounting Location	Reasonable Range (more Urban)	Stretch Range (more Rural)
C-Band	52-58	Streetlight	600m (~2000')	900m (~3000')
		Strand	425m (~1400')	640m (~2100')
CBRS	47-53	Streetlight	340m (~1150')	500m (~1650')
		Strand	240m (~800')	360m (~1200')
6GHz Wi-Fi 6E	36	Streetlight	70m (~240')	100m (~325')
		Strand	50m (~175')	70m (~240')

Note that the above table contains several assumptions:

1) The C-band radios are 4 x 20W and the antennas are about 12dBi gain. Some variation in EIRP is allowed given that the antenna gain may vary.
2) Note that the actual antenna gain depends on whether an omni or directional antenna is used and the height of the antenna
3) Directional antennas normally have a higher gain than omni antennas, but you need 3 of them to cover a site for 360 degrees. So nominally the directional antennas take a 5dB hit in terms of coverage.
4) CBRS EIRP is shown as a range since the cap is per 10 MHz channel, so a user with more channels can increase their EIRP.
5) Small cell array length can be as high as 24" (600mm) for a streetlight antenna, but strand-mount arrays are normally under 8" (200mm) in height. Thus, assumed that in general the strand mount system would have 3dB less EIRP than a streetlight system. The difference in height might have a slight difference in propagation, but we assumed that this was negligible.
6) Propagation loss goes as the square of distance. So, assume that a 6dB increase in EIRP corresponds to a doubling of the range.
7) The range is for outdoor UE. There can be significant additional path loss when trying to penetrate inside buildings depending on building materials.

2.5. Backhaul, Midhaul, Fronthaul RAN Interfaces

Given the many diverse requirements that 5G networks must support such as high data rates, low latency and high reliability, the implementation of the Radio Access Network (RAN) has been under constant debate. Early proposals focused on the idea of Cloud-RAN with a dense network of cells. However, the 3GPP's 5G-R RAN2 specification included eight different functional split options. A discussion on these different interface options are detailed in [LARSEN_2018] and [ORAN_2020].

The functional splits for the different interface options from [ORAN_2020] are shown in the lower half of Figure 8. The industry has evolved to supporting a distributed RAN architecture that includes a Central Unit (CU), a Distributed Unit (DU) and a Radio Unit (RU) that might also be called a Remote Radio Unit (RRU). This architecture is shown in the top half of Figure 8. The term fronthaul refers to the interface between the DU and the RU/RRU. Midhaul refers to the interface between the CU and the DU.

Today, most 5G industry focus is on one of two options:

- Option 2: a high-level centralized unit (CU) and a distributed unit (DU) split which is essentially a separated control and user plane. In this implementation the DU and remote radio unit (RRU) are often combined into a single entity as a self-contained access point.

- Option 7.2 Cat A: a low-level split that allows for high reliability and low latency communications and near-edge deployment. This split takes place between the Hi-PHY (Physical Layer) and Low-PHY. In this split, only the Low-PHY and RF functions are in the access point.

With more complex antenna arrays (e.g. 64T64R) and massive MIMO (M-MIMO), the processing requirements for the DU increase significantly. Using Option 7.2x allows the DU to be moved to a more optimum location and it reduces the size and power requirements at the antenna site.

With simpler antenna arrays and reduced MIMO levels, the DU becomes simpler and can be more easily integrated with the RU/RRU. Small cells are a perfect example of this.

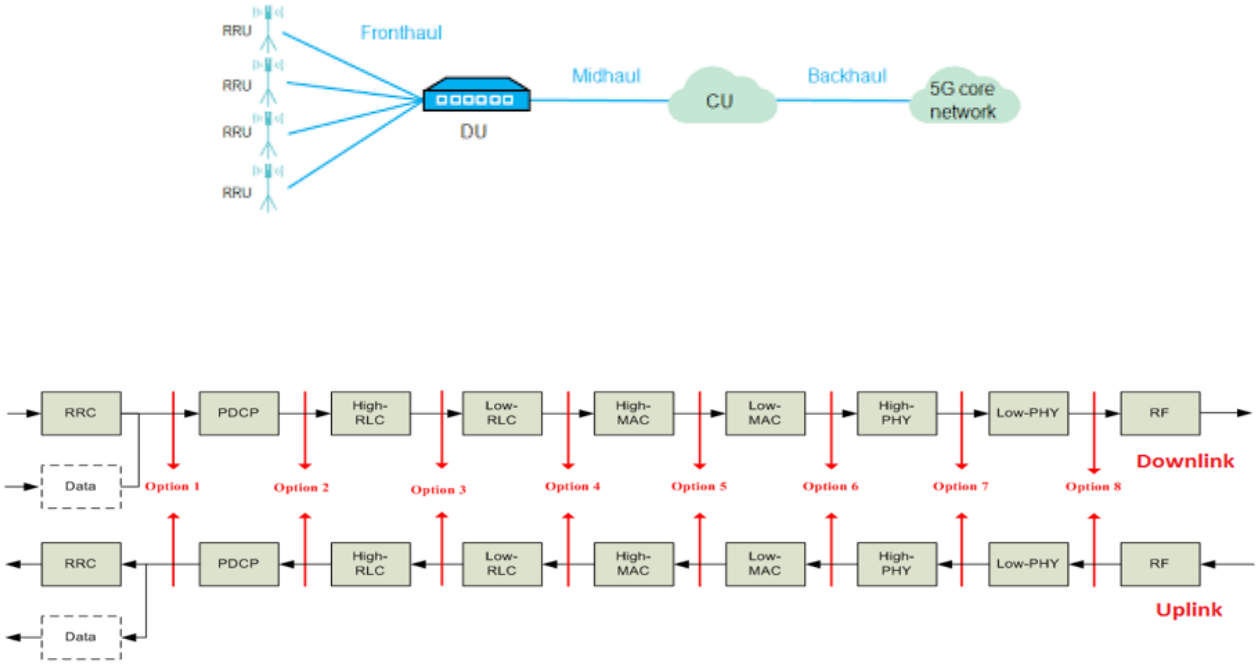


Figure 8 – Backhaul, Midhaul, Fronthaul RAN Interfaces

3. Capacity Planning for Midhaul or Fronthaul Cells

3.1. Midband Backhaul & Fronthaul Interface Capacity Requirements

The amount of capacity required for wireless xHaul varies significantly based on several factors. It becomes a function of the number of antennas, MIMO level, channel bandwidth, number of sectors per cell and the RAN interface used (e.g. Midhaul or Fronthaul). Table 2 shows some capacity examples for various configurations that could be seen from HFC-based small cells to Macro tower base stations.

Table 2 – xHaul Capacity Estimates for various Antenna configurations

Antenna	MIMO	Location	Channel Bandwidth	Sectors per cell	Midhaul DL	Midhaul UL	Fronthaul DL	Fronthaul UL
2T2R	2x2	Strand or Streetlight	40MHz, DL only	1	525 Mbps	-	1.9 Gbps	-
			40MHz	1	420 Mbps	62 Mbps	1.9 Gbps	2.0 Gbps
4T4R	4x4	Strand or Streetlight	40MHz, DL only	1	1050 Mbps	-	3.8 Gbps	-
			40MHz	1	840 Mbps	125 Mbps	3.8 Gbps	4.1 Gbps
			100MHz	1	2.2 Gbps	320 Mbps	9.7 Gbps	10.6 Gbps
8T8R	BF 2x2		40MHz	3	0.6 – 1.1 Gbps	90 – 165 Mbps	2.8 – 5.0 Gbps	3.1 – 5.5 Gbps
	4x4	Mini-Macro	40MHz	3	1.3 – 2.2 Gbps	180 – 333 Mbps	5.7 – 10 Gbps	6.2 – 11 Gbps
	4x4		100MHz	3	3.3 – 5.8 Gbps	500 – 850 Mbps	15 – 26 Gbps	16 – 28 Gbps
64T64R	8x4	Macro	100MHz	6	10 – 23 Gbps	0.7 – 1.7 Gbps	44 –104 Gbps	24 – 56 Gbps

The above table assumes that the downlink (DL) is operating at its best modulation of 256-QAM. This may be generous given real-world conditions, but wanted to show a worst case capacity estimate. The uplink (UL) is assumed to be operating at 64-QAM modulation. For NR-TDD, it is possible to configure the mix between DL and UL. Most of the rows use a DL:UL ratio of 80:20. Two of the small cell rows are configured for 100% DL operation only. In these scenarios, it is assumed that the weaker UL signal is using more robust Low-band frequencies (e.g. <1 GHz).

As can be seen by this table, the Option 2 Midhaul interface has significantly lower bandwidth capacity requirements then the Option 7.2 Cat A Fronthaul interface. Capacity requirements also increases with the channel bandwidth and the number of sectors per cell. As added data point, using an IPv6 backhaul with IPSec would consume about 10% more capacity than the midhaul; while actual UE data consumption would be about 89% of the midhaul capacity.

3.2. CBRS RF Simulation Case Study

A recent case study performed a CBRS RF simulation analysis for a North American metro area. This analysis covered a region with approximately 40K addresses, and roughly 800 radios.

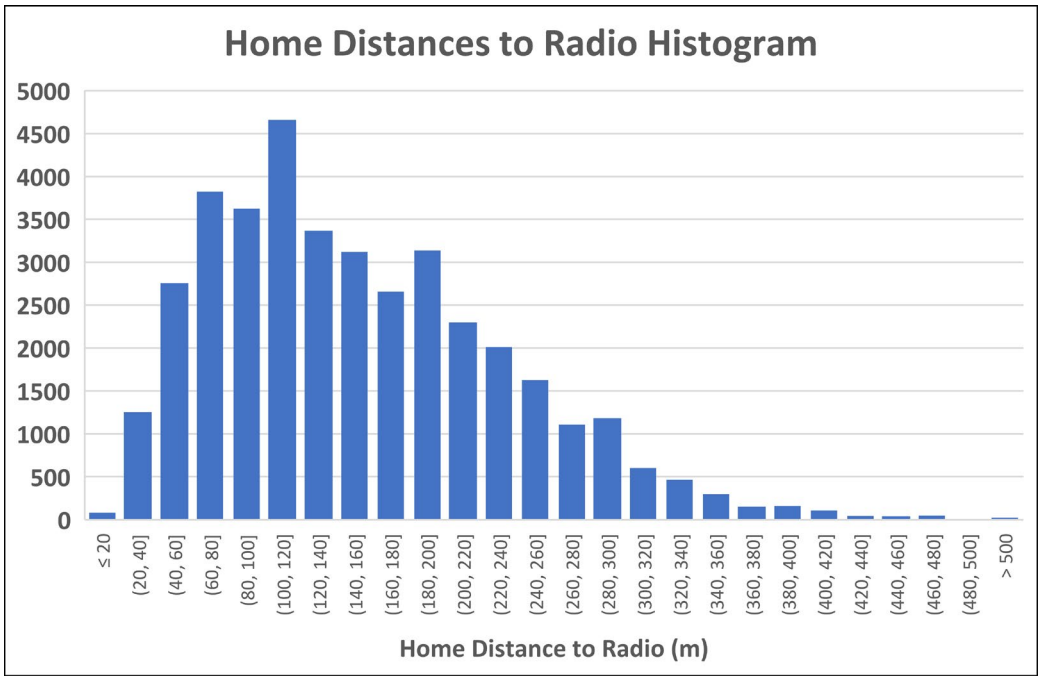


Figure 9 – CBRS Metro area Study – Home Distances to Radio histogram

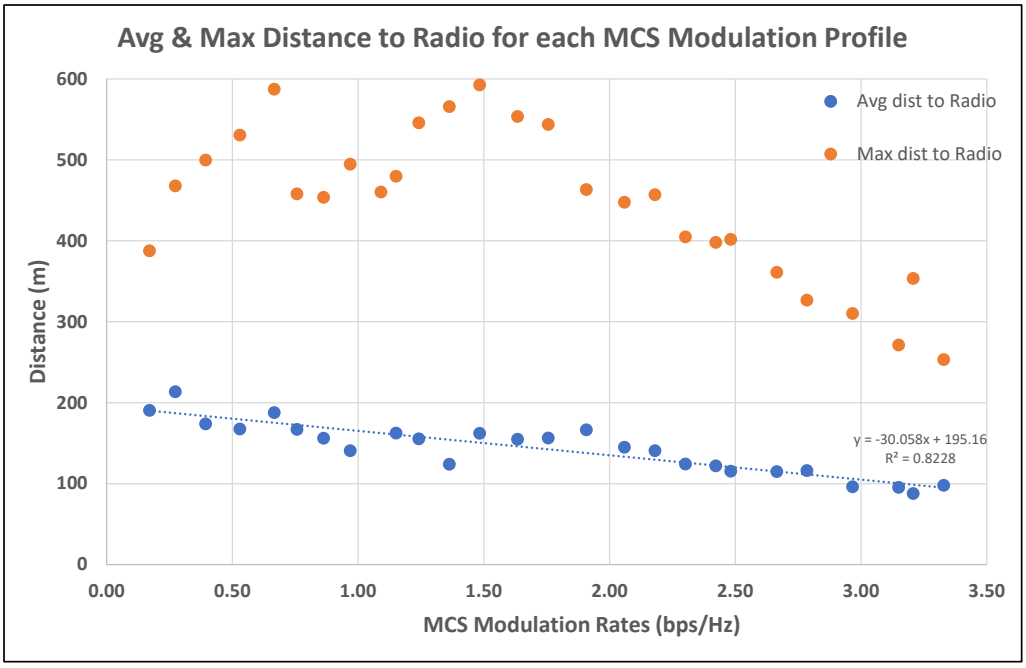


Figure 10 – Average & Max Distance to Radio for each MCS Modulation profile

Some of the key statistics from the study around the distribution of home addresses to the radio include:

Address Distances to Radio:

- Average = 153m
- Maximum = 592m
- Median = 140m
- 95% ≤ 300m
- 99% ≤ 385m

Figure 9 shows that most home addresses are within 200m of a radio and 95% are within 300m. However, there are still several addresses that could be up to 600m away.

Figure 10 shows the average and maximum distance to the radio as a function of the Modulation Coding Scheme (MCS) modulation rate. As the MCS modulation rate increases, there is a linear drop in the average distance to radio and an even more substantial in the maximum distance. For the best MCS rates, the max distance tends to be less than 300m.

Figure 11 is the inverse of Figure 10 where the average and maximum MCS rates are mapped as a function of the distance to the radio. The Max Distance for a given profile is stable up until ~1.5 bps/Hz but drops quickly above that. For distances up to 300m, average modulation (bps/Hz) drops with distance while max stays to MCS 27. Above 300m, average MCS rates stay flat while maximum MCS rates decline quickly with distance.

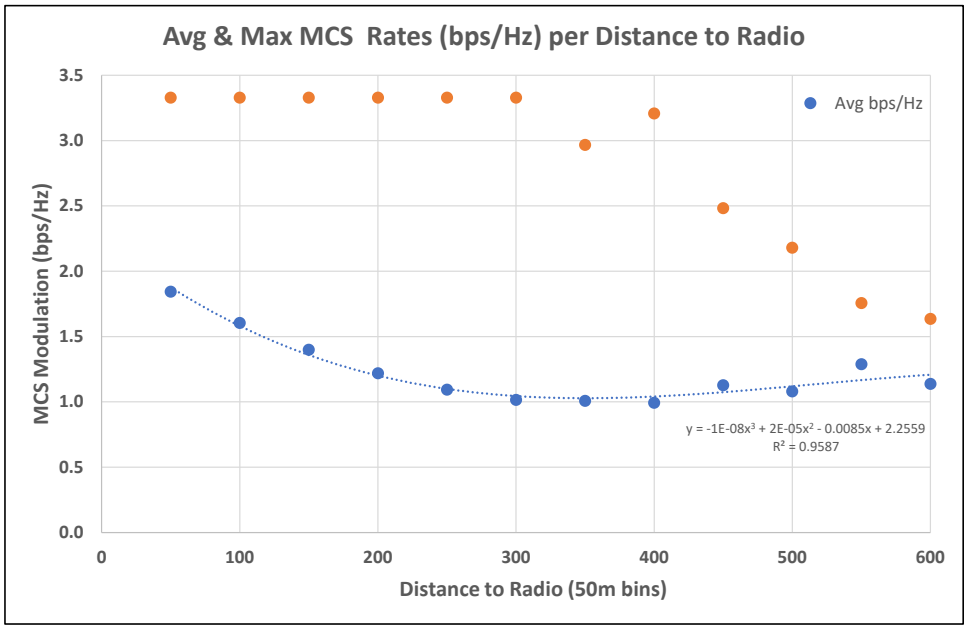


Figure 11 – Average & Max MCS Rates per Distance to Radio

Another observation from the CBRS case study was that in dense areas with >200 home addresses per radio, maximum distance was mostly <300m. Meanwhile, serving radios with fewer addresses (e.g. <60) have a wide range of max distances, even to 600m. Most Radios have fewer than two dozen addresses >300m, while a handful of radios (i.e. <1%) have 40-80 addresses >300m.

In general, the results from this case study reenforce the CBRS range estimates as shown in Table 1.

3.3. HFC Network Capacity Planning

3.3.1. DOCSIS 3.1 Capacities

DOCSIS 3.1 introduced OFDM/OFDMA technologies and increased frequency spectrum of up to 1218 MHz downstream and 204 MHz upstream. This means that the maximum capacity of an DOCSIS 3.1 HFC network could reach 9 Gbps downstream and 1.5 Gbps upstream. This assumes that the operator has retired most or all of its legacy video spectrum (i.e. converted to IPTV or SDV) and replaced most 2.0/3.0 modems with the newer D3.1 modem technology.

From an HFC perspective, the streetlight and strand-mount small cells with an Option 2 midhaul interface shown in Table 1 can be supported by a DOCSIS 3.1 system with an 85MHz mid-split. A 40 MHz DL could fit within 96 MHz OFDM channel or even multiplexed with the residential downstream data. And by accounting for some statistical multiplexing gains, it might even be possible to put a couple midhaul based small cells on a single DOCSIS 3.1 service group (SG). A 100 MHz DL:UL small cell with 2.2 Gbps DL and 325 Mbps UL capacity might fit better on a 1218/204 MHz HFC plant.

Implementing a fronthaul interface on the small cell is much more challenging. A 40 MHz DL-only 2x2 small cell could consume an entire 192 MHz OFDM channel. A 40 MHz DL-only 4x4 small cell then requires two 192 MHz OFDM channel. Turning on the UL will require multiple Gbps upstream capacity. This would require DOCSIS 4.0 as discussed in the next section.

3.3.2. 10G™ Capacities

10G was first announced in 2019 as a vision, or lighthouse beacon, to guide our industry roadmaps towards 10 Gbps services. Much progress has been made in the last 2-3 years. 10G includes multiple technologies including enhanced fiber optics as well as DOCSIS 4.0 technologies.

Figure 12 from [CableLabs_10G] shows the 10G vision for DAA services operating over cable, FTTP and wireless technologies while sharing a common converged fiber optical network through an aggregation node. [CableLabs_10G] states:

> *"The 10G optical network is the backbone of the distributed access architecture and will provide the industry with opportunities for true service convergence that leverages the flexibility and tremendous capacity provided by fiber optics. This year, CableLabs released an update to the 100 Gbps point-to-point coherent optics specification and released a new 200 Gbps specification – both intended to support the aggregation requirements of the distributed access architecture. While operators currently deploy 10G passive optical network technology (PON) where fiber to the premise is preferred, the IEEE standard for next-generation 25G-PON and 50G-PON technology remains on track for mid-2020 completion."*

As can be seen in the figure, 10G has the vision of supporting Mobile Backhaul/Fronthaul and Fixed Wireless Access (FWA) over the service provider's optical infrastructure. The P2P Coherent optics has the reach (up to 80km) and the capacity (to 200 Gbps) to support these applications. It even supports the fronthaul capacity requirements for the Macro cell shown in Table 2.

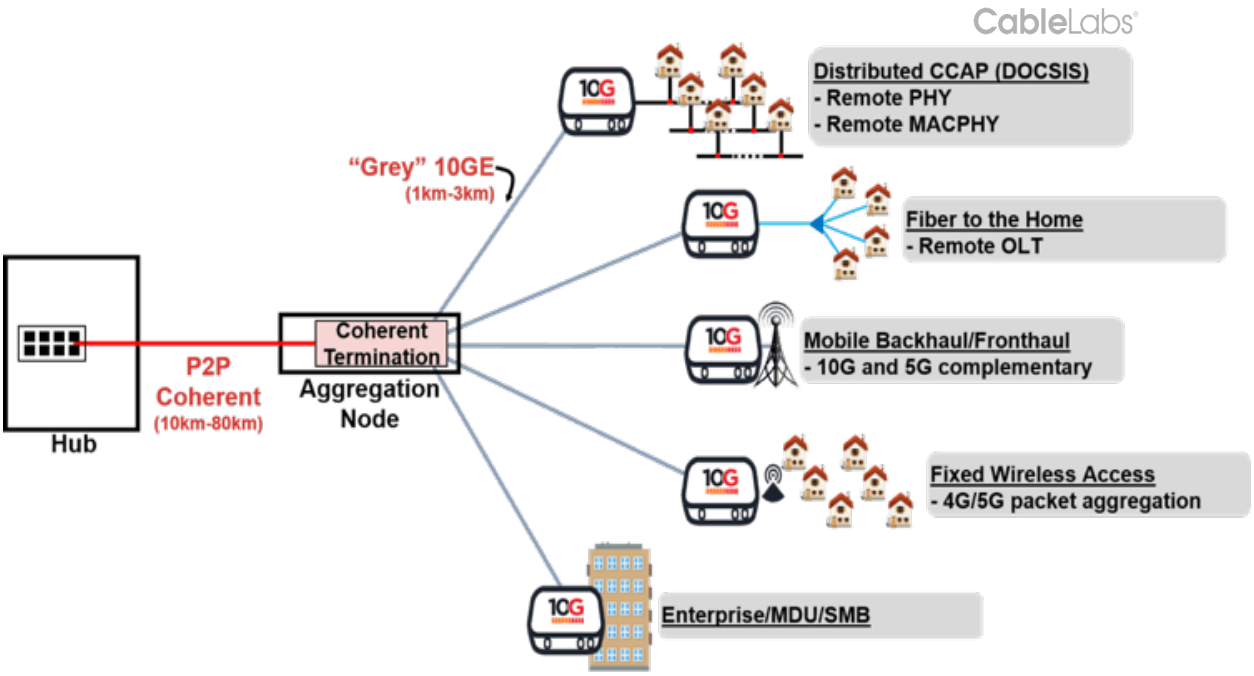


Figure 12 – 10G™ Converged Optical Network – Distributed Access Architecture vision

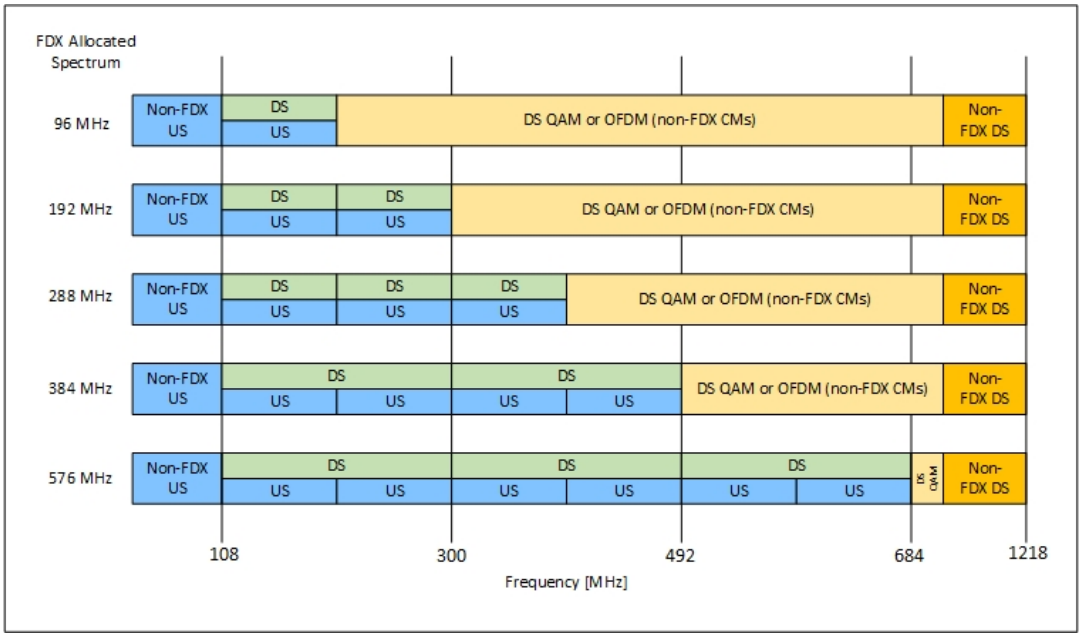


Figure 13 – Full Duplex DOCSIS (FDX) Spectrum Band Options

However, it may not always be possible to co-locate the small cell or mini-macro cell adjacent to the HFC fiber node or along the fiber path. This might result in the cells being connected to the coax portion of the HFC. For small cell RRU with fronthaul interfaces, this might necessitate the use of DOCSIS 4.0. [ULM_2019-1] discusses the capacities that DOCSIS 4.0 can enable and the migration path to 10G.

One area of focus at CableLabs is a technology called Full Duplex DOCSIS (FDX). FDX leverages echo canceller technology to allow simultaneous upstream and downstream operation in the FDX band. FDX is targeted at a fiber deep Node+0 DAA environment. FDX is now part of the new DOCSIS 4.0 specification [FDX_PHY].

The FDX capability offers a fundamental benefit that permits upstream spectrum expansions to occur without causing reductions in downstream spectrum. FDX proposes to have downstream and upstream transmissions occurring in the same frequency band at the same time. In the FDX specification, the overlapping frequency bands are shown in Figure 13.

On a fiber deep Node+0 plant, the upstream OFDMA channel might net capacity of as much as 8-10 Mbps per MHz. This means that a full spectrum 108-684 MHz FDX system might support ~5 Gbps US.

With the Node+0 architecture, the fiber node is now within 300m to 500m of every home. This means the need for using FDX over coax may be minimized.

The other facet of DOCSIS 4.0 is Extended Spectrum DOCSIS (ESD) which supports 1.8 GHz plant with different potential upstream splits in a Node+X plant. In this scenario, the fiber node may be a couple kilometers from the furthest home so it may become necessary to put small cells on the coax.

The extended 1.8 GHz downstream adds much needed DL capacity that could be used for fronthaul RRU devices. The upstream split can then be adjusted based on the desired UL capacity.

So, fronthaul RRU small cells might make a good early use case for DOCSIS 4.0, especially ESD.

4. Using Outside HFC Plant for 5G – considerations and logistics

[ULM_2019-2] took an in-depth look into supporting High band mmWave cells over HFC. At these high frequencies, the wireless cell is limited to very short distances on the order of <100m to 200m which creates coverage issues. The following sections highlight some of the HFC findings from that paper.

Across all the various wireless options, there is a driving need for much smaller cell sizes. To make this happen requires an infrastructure that supports both *the power and the backhaul* to the small cells. The cable industry Hybrid Fiber Coax (HFC) networks are ideally positioned to support this. The HFC networks might support the addition of attached in-line small cells at various demarcation points on the HFC plant. These cells can be added to the DOCSIS network to support 5G, Wi-Fi and/or CBRS/LTE over the HFC.

Figure 14 shows the additional level of density variability that must be considered in cell placement. This example is from a suburban city in New England. In Figure 14, lot sizes vary from 1 acre on the left, ½ acre lots in the middle to <¼ of acre on the right. Cell placement must also account for open spaces and office campus space too.

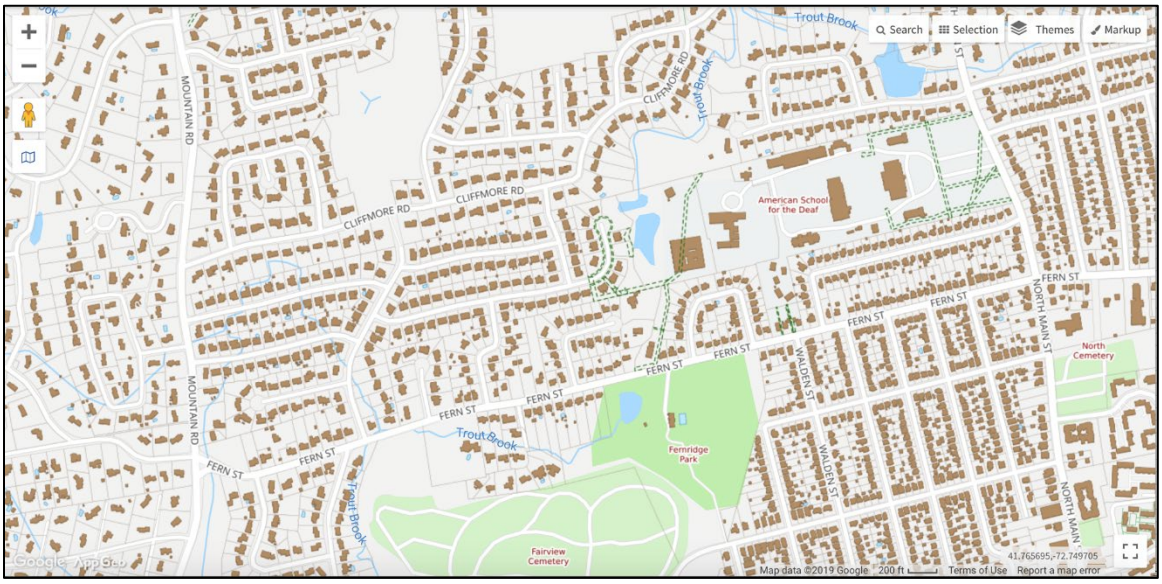


Figure 14 - New England suburb, illustration for variability of lot sizes within

5. HFC Case Study for N+3 nodes of varying Homes Passed densities

The HFC case study in [ULM_2019-2] considered 5 node examples that varied from low homes passed density in a rural area to a high urban node with many homes. Table 3 provides the key statistics for each of the five nodes. In general, these are N+3 nodes, except the highest density node being N+2. The homes passed per coaxial mile (HP/mile) ranges from 37 to 274. This paper now investigates the implications of mid-band small cells overlaid onto these same HFC nodes.

5.1. Mapping Mid-band Small Cells to N+3 HFC Plant

This section looks at mapping mid-band small cells to HFC N+3 nodes of various densities. The first step is to co-locate the small cell with the fiber node to ensure that the small cell has fiber backhaul. This gives the operator the flexibility to choose whether to implement Option 2 midhaul interface or the Option 7.2x fronthaul interface. If the operator is also implementing a DAA strategy with either a Remote PHY Device (RPD) or Remote MACPHY Device (RMD) in the node, then the small cell can potentially share the 10G long haul Ethernet link as well. After that, additional scenarios may be shown by adding other small cells located on the coax segment adjacent to one of the HFC plant active components.

For High and Med-High density nodes, a cell radius in 250m to 350m range is used. For Low and Med-Low density nodes, a cell radius can stretch to 500m range for handful of homes. This aligns with the CBRS small cell ranges from Table 1. Note that C-band small cells would have an even bigger coverage area due to their increased power budget.

Table 3 – Statistics of 5 HFC N+3 nodes of various densities

N+3 NODE Case Study:	Low	Med-Low	Medium	Med-High	High
Coax Plant Mileage	4.17	6.16	3.54	2.51	1.90
Aerial	0.82	4.72	3.44	2.18	1.72
Underground	3.35	1.44	0.10	0.33	0.18
Total Actives	21	30	21	19	14
Actives/Mile	5.0	4.9	5.9	7.6	7.4
Cascade	N+3	N+3	N+3	N+3	N+2
Total Passings	153	352	398	469	520
Aerial Passings	27	269	383	200	500
UG Passings	120	83	0	0	0
Comm/MDU passings	6	0	15	269	20
HP/Mile	37	57	112	187	274

5.1.1. High Density Node Example

The high-density node example shown in Figure 15 is N+2 with 520 total homes passed (HP) with a density of 274 HP/mile. The fiber node is located on the top side of the area. But even with that, most of the homes are easily within the 250m inner radius.

In this example with the node at one edge of the area, it may make sense to use a 180-degree directional antenna rather than an omni-directional antenna. This can provide some added gain for reaching the outer fringes of the node area.

In Figure 16, the CBRS small cell is placed on the coax adjacent to the first level amplifier. This provides a very central location for an omni-directional amplifier. Almost all of the homes are within 200m of the small cell. The only drawback might be that this limits the small cell to use an Option 2 midhaul interface.

5.1.2. Med-High Density Node Example

The medium-high density node in Figure 17 is N+3 with 469 total HP and 187 HP/mile. While most of the homes are within the 250m radius of the small cell, there are still a good number of homes that are in the 250m to 350m range. This means that the CBRS small cell might better be streetlight mounted for the extra elevation and coverage rather than strand-mount.

So again, it appears that this density node can be serviced by a single CBRS small cell that is co-located with the fiber node.

5.1.3. Medium Density Node

The medium density suburban node in Figure 18 is N+3 with 398 total HP and 112 HP/mile. This is an oddly shaped node, unlike the nicely packed previous two examples. Because it is so stretched, it cannot be covered by a single CBRS small cell. We do use this example to show how an operator might place four small cells scattered within this node to give coverage to both this node and neighboring nodes. One small cell is co-located with the fiber node and the other three are on the coax.

Note – if this node also implements a 2x2 RMD/RPD, then it is possible to arrange the CMTS service groups such that there are no more than two small cells on any given DOCSIS network.

This node shows why an operator needs to look holistically across a multi-node region when deciding on small cell placements.

5.1.4. Med-Low Density Node Example

The medium-low density node in Figure 19 is in a residential development and has N+3 with 352 total HP and 57 HP/mile. Because this is a more rural setting, the figure also now includes a 500m radius for the enhanced range for a streetlight mounted CBRS small cell.

The 500m radius appears to cover a majority of the homes in the node. This may be acceptable if the goal is to off-load as much traffic as possible without the requirement for 100% coverage. If the operator has partnered with a C-band MNO to deploy C-band small cells, a single cell should cover this entire node area.

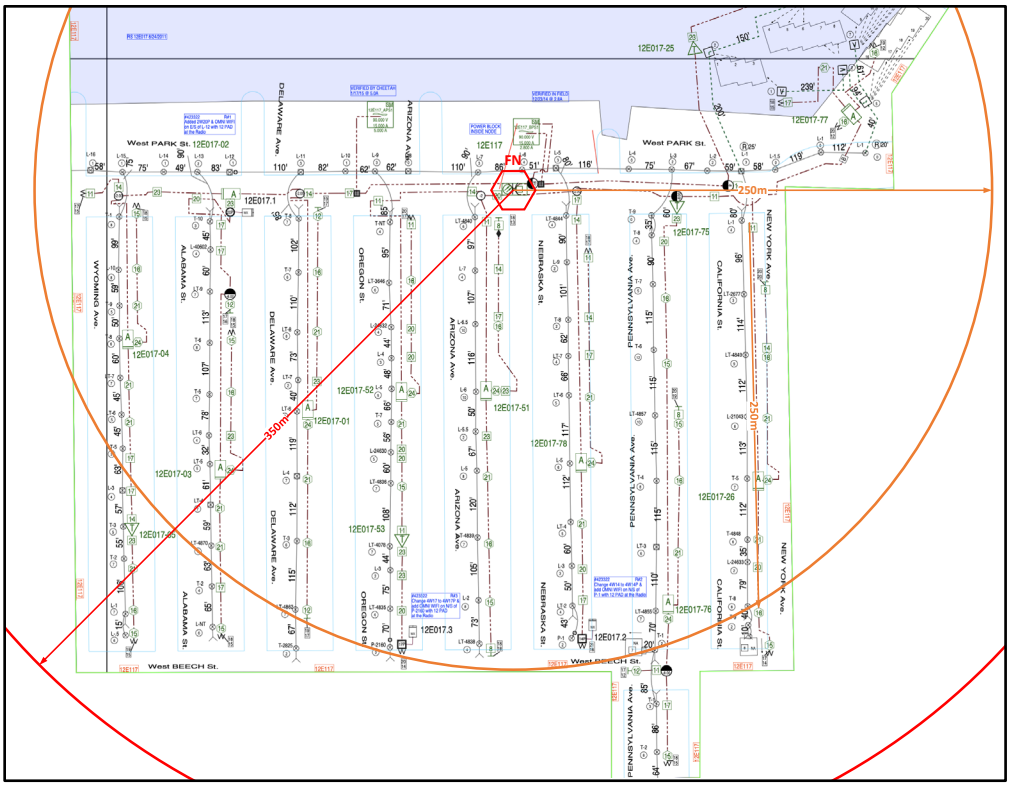


Figure 15 – High Density Node (274 HP/mile) with Small cell at Fiber Node

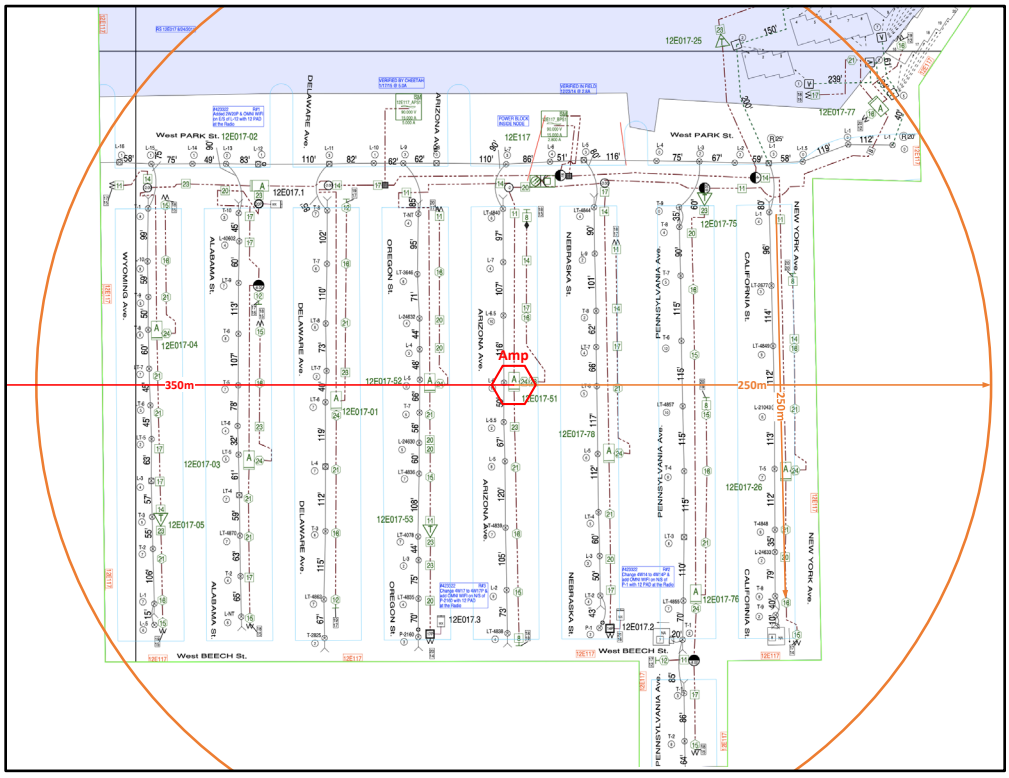


Figure 16 – High Density Node (274 HP/mile) with Small cell at HFC Amp location

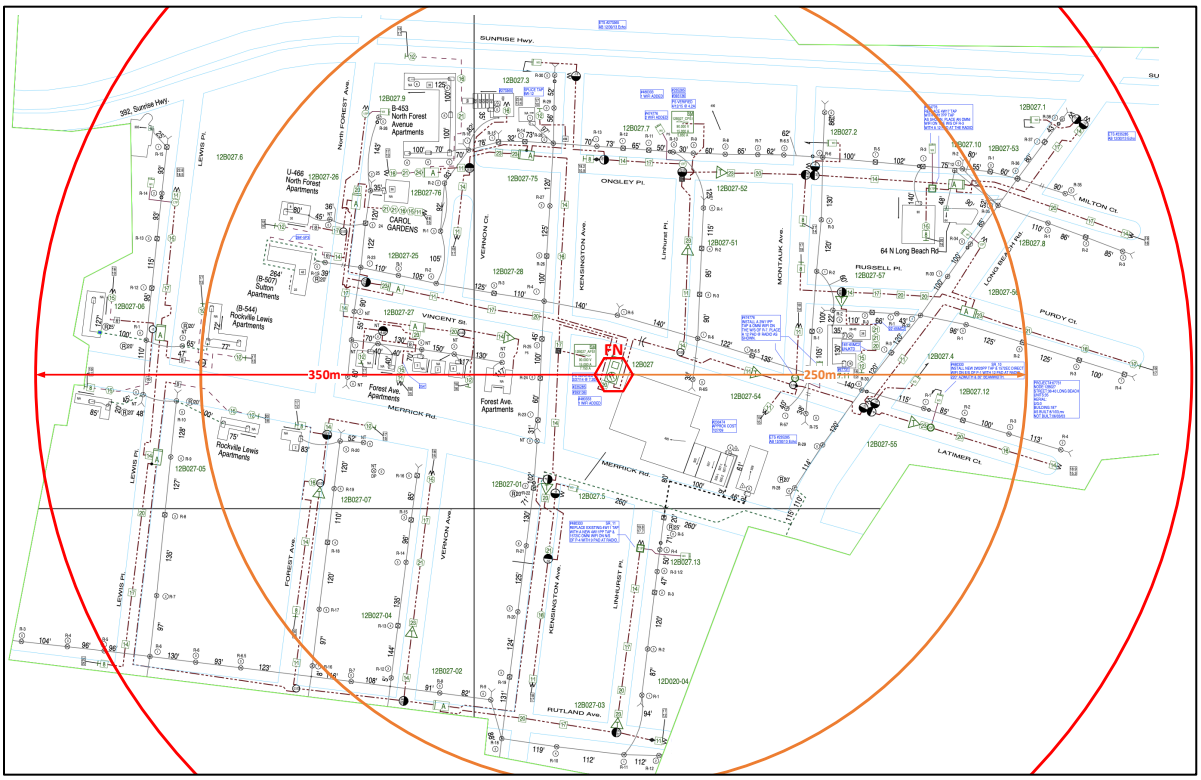


Figure 17 – Med-High Density Node (187 HP/mile) with Small cell at Fiber Node

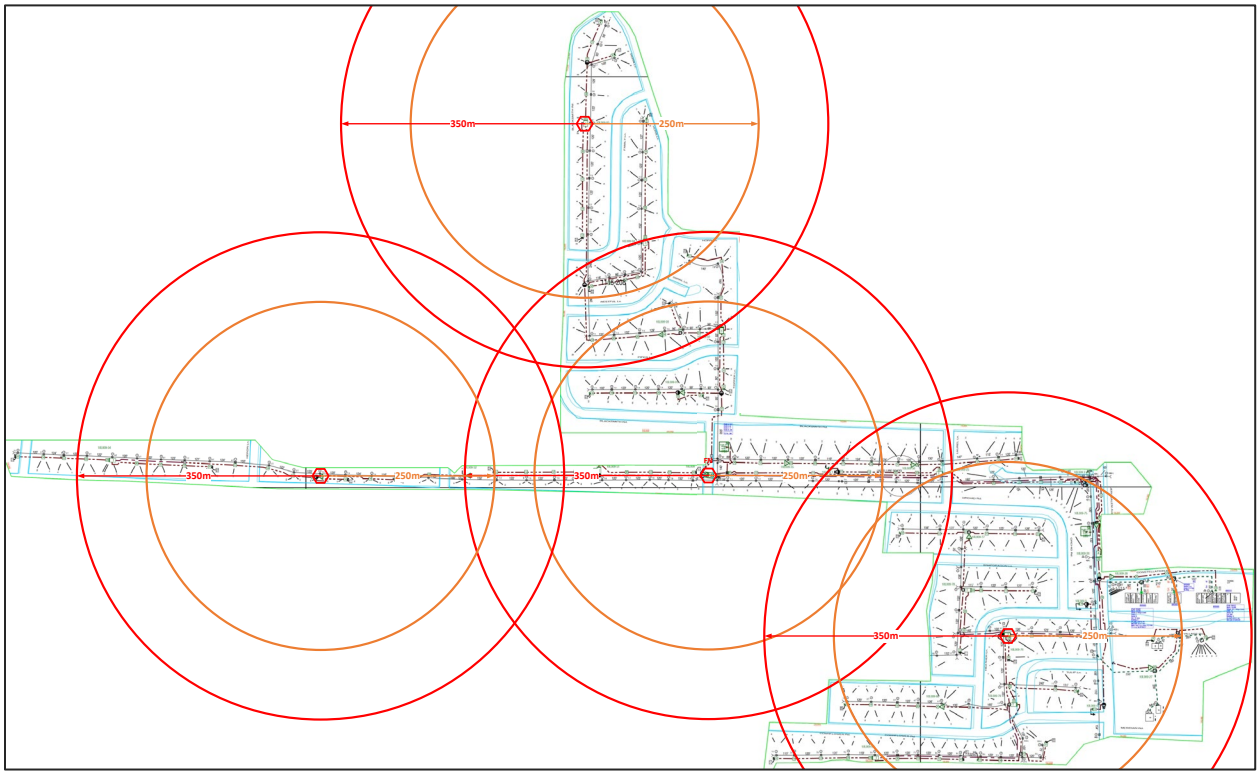


Figure 18 – Medium Density Node (112 HP/mile) with Small cell at Node + HFC Amps

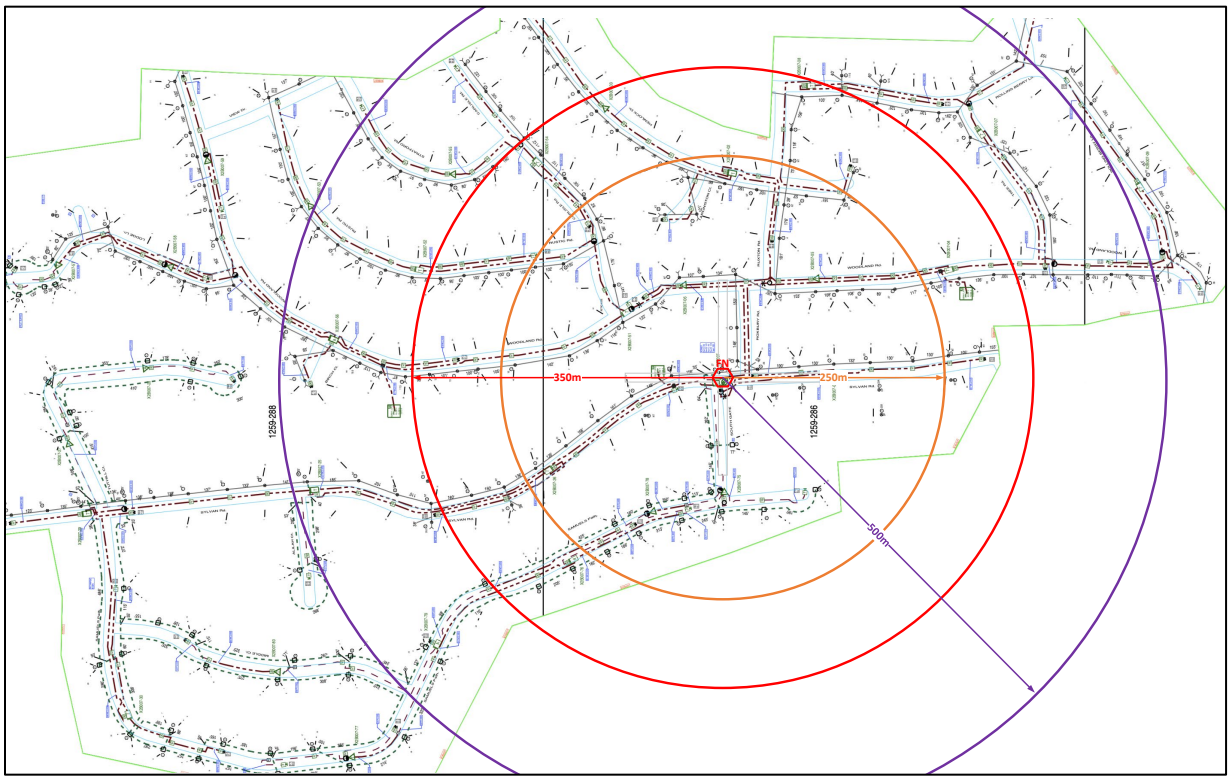


Figure 19 - Med-Low Density Node (57 HP/mile) with Small cell at Fiber Node

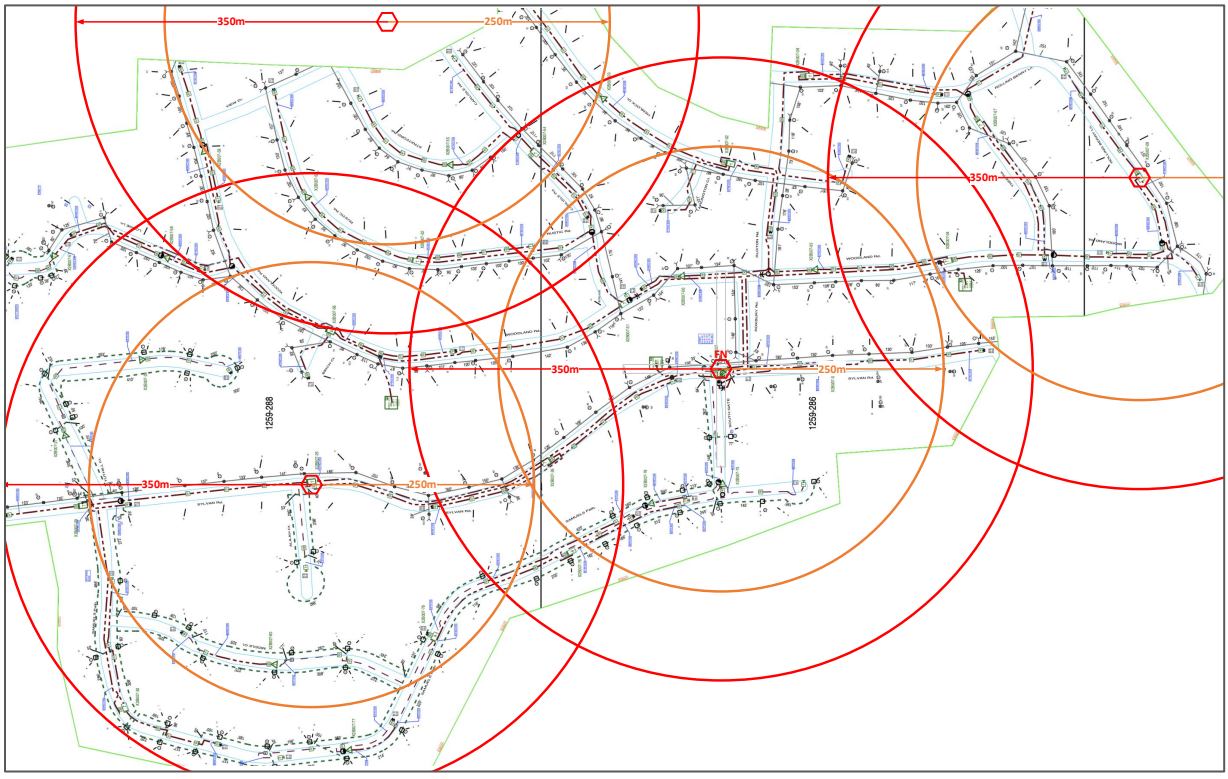


Figure 20 - Med-Low Density Node (57 HP/mile) with Small cells at Node + HFC Amps

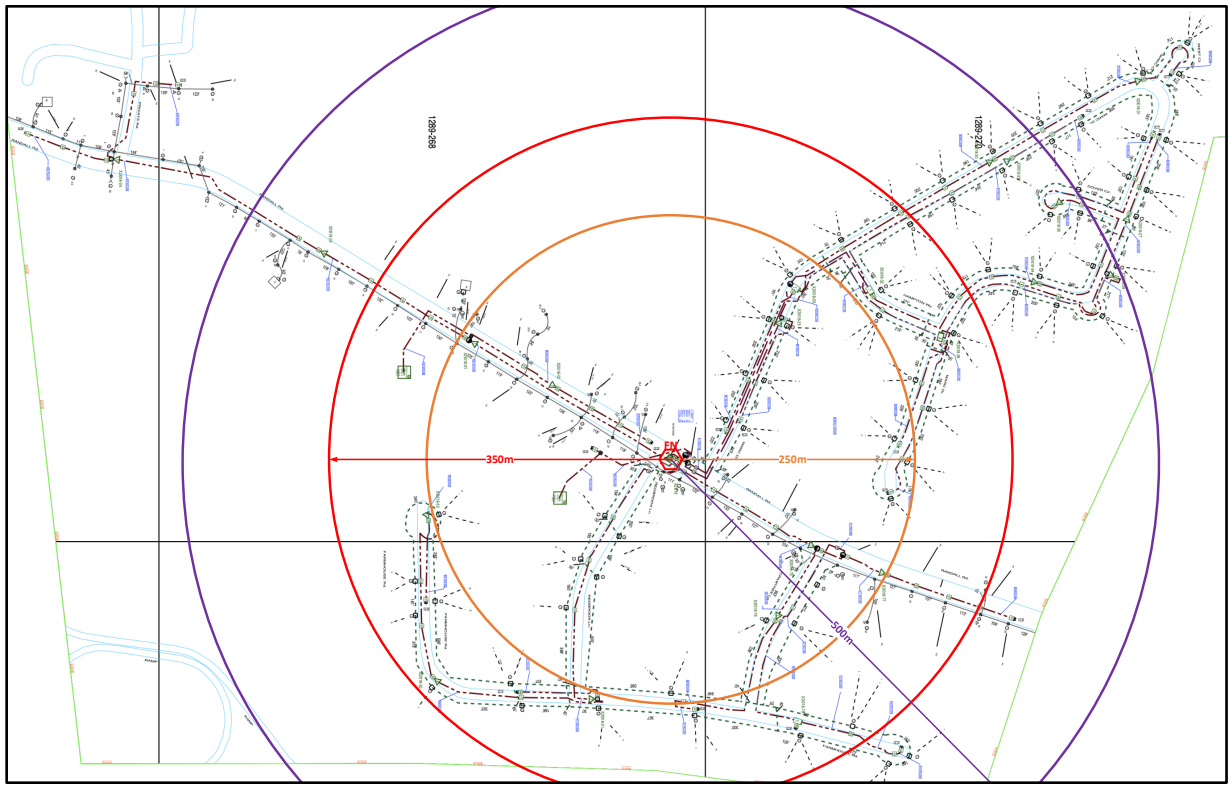


Figure 21 – Low Density Node (37 HP/mile) with stretched Small cell at Fiber Node

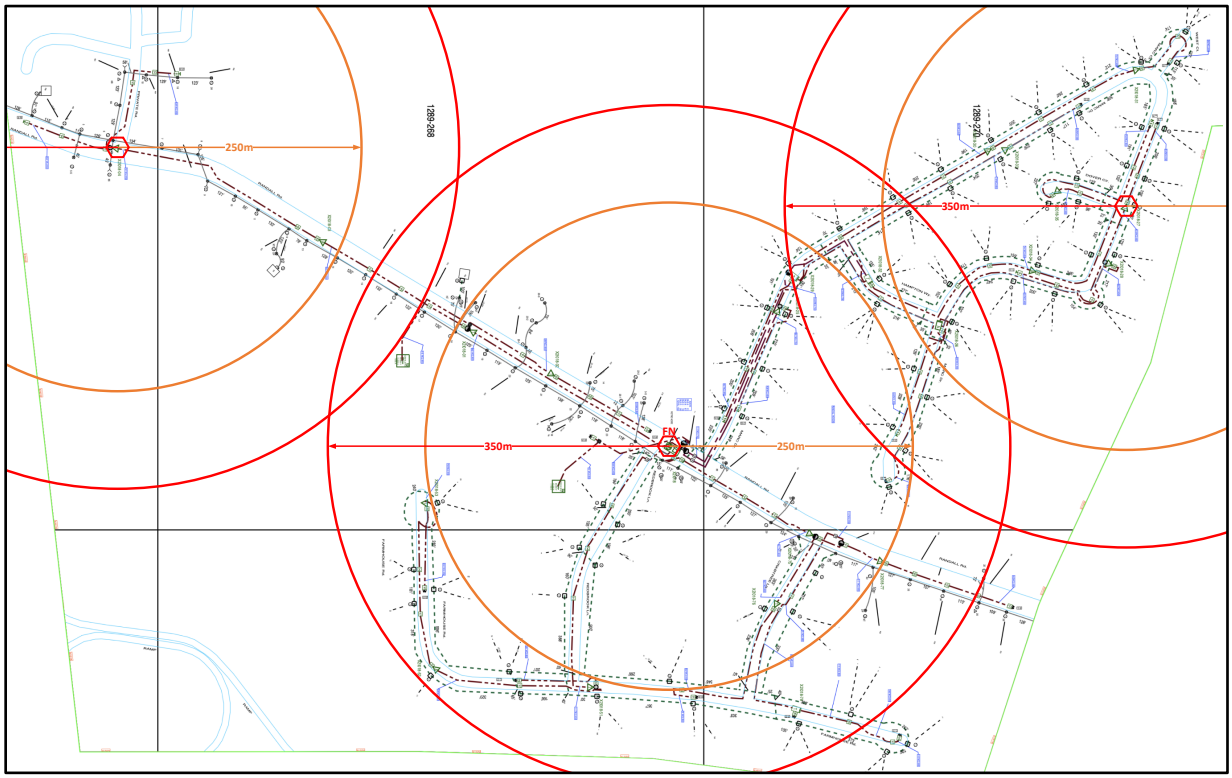


Figure 22 – Low Density Node (37 HP/mile) with Small cells at Fiber Node + HFC Amps

Figure 20 shows a couple of additional CBRS small cells added to the node. If these are all strand-mounted, then maybe their reach will be limited to the 250m to 350m range. In this example, one small cell has fiber access while the other two have coax backhaul. With a 2x2 RMD/RPD, they could be on separate DOCSIS networks.

5.1.5. Low Density Node Example

The low-density rural node in Figure 21 is N+3 with 153 total HP and 37 HP/mile. A streetlight mounted CBRS small cell with ~500m radius does a good job of covering most of the node area, but it still needs some help at the fringes. A C-band small cell should cover this node area without any problem.

Figure 22 shows some strand mounted CBRS small cells with 250m to 350m range. One is co-located with the fiber node while the other two or on the coax in opposite directions (i.e. probably separate RF legs and potentially separate DOCSIS networks.

5.2. Summary – Mapping Mid-band Small Cells to N+3 HFC Plant

After looking across an extremely wide range of homes passed densities (i.e. from 37 to 274 HP/mile), a single CBRS small cell that is co-located with the fiber node is sufficient to cover most of the homes in N+3 plant. For those nodes that need some additional small cells to achieve full coverage, this could be accomplished with only one or two small cells per DOCSIS networks.

Operators with larger HFC cascades (e.g. N+5, N+6) will obviously need additional small cells to achieve their coverage. But this case study shows that N+2/N+3 might be an optimal HFC design target for operators thinking of 5G mid-band convergence. As time progresses and bandwidth needs continue to rise, an operator might want to migrate from a DOCSIS based backhaul to a fiber backhaul. So, the operator might consider how they will eventually pull fiber to these small cells on the HFC coax plant as part of their overall fiber deeper strategy.

6. Mapping Mid-band Small Cells across multiple N+6 Fiber Nodes

6.1. Multi-node N+6 case study

The previous case study had certain limitations. First, many plants have longer amplifier cascades such as N+5/N+6 with fiber not as deep as the N+3 case study. Second, the expanded range with the mid-band small cells now potentially covers parts of multiple nodes at a time. The previous N+3 study gave us a wide range of densities, but only viewed a single node at a time. The next case study expands this to look at a much larger area from a North American metro suburban area to measure the impact across a multitude of nodes with varying densities.

Table 4 shows statistics for a ~3.5 square mile area consisting of 9 adjacent nodes. In addition to the statistics for the entire area in column one, the next four columns show the statistics for some select nodes: i.e. the highest and lowest density ones, as measured by number of homes passed per mile.

Table 4 – Statistics of Metro-suburban HFC N+6 nodes of various densities

N+6 Case Study:	Overall Area (9 nodes)	Node #1 Low Density	Node #2 Low Density	Node #3 High Density	Node #4 High Density
Coax Plant Mileage	**59.6**	9.56	6.58	4.3	2.24
Aerial	**36.1**	3.87	5.39	3.34	1.59
Underground	**23.5**	5.69	1.19	0.96	0.65
Total Actives	**381**	61	45	32	13
Actives/Mile	**6.4**	6.4	6.8	7.4	5.8
Cascade	**N+3 – N+6**	N+6	N+5	N+4	N+3
Total Passings	**5,740**	724	502	628	370
HP/Mile	**96**	76	76	146	165

Figure 23 displays the entire area. The node boundaries are shown on the map with magenta lines. The green lines show the HFC fiber routes. In addition to connecting the nine nodes to the hub site, the fiber backhaul also connects to two Macro tower base stations in the upper quadrant.

6.1.1. CBRS Small Cells at Fiber Nodes Only

Figure 23 also displays nine CBRS small cells located next to or near a fiber node. This location provides access to power and fiber backhaul in case the operator wants to use the Option 7.2x fronthaul interface. The small cell range is roughly shown as concentric circles of 250m and 350m coverage radius. This corresponds roughly to the coverage area of strand-mount and streetlight mount respectively.

Note that two of the fiber nodes were within 250m of each other. Putting a small cell at each node site would have resulted in too much overlap and interference. Rather than eliminating a small cell, we chose to move them a short distance away from the node (i.e. 500’ and 860’). The small cells are still on the fiber backbone and hardline coax to get access to power plus fiber backhaul. Again, the fiber backhaul allows for a fronthaul option 7.2 Cat A interface to be used if desired.

Figure 23 – N+6 Suburban area with Small cells at Fiber Node

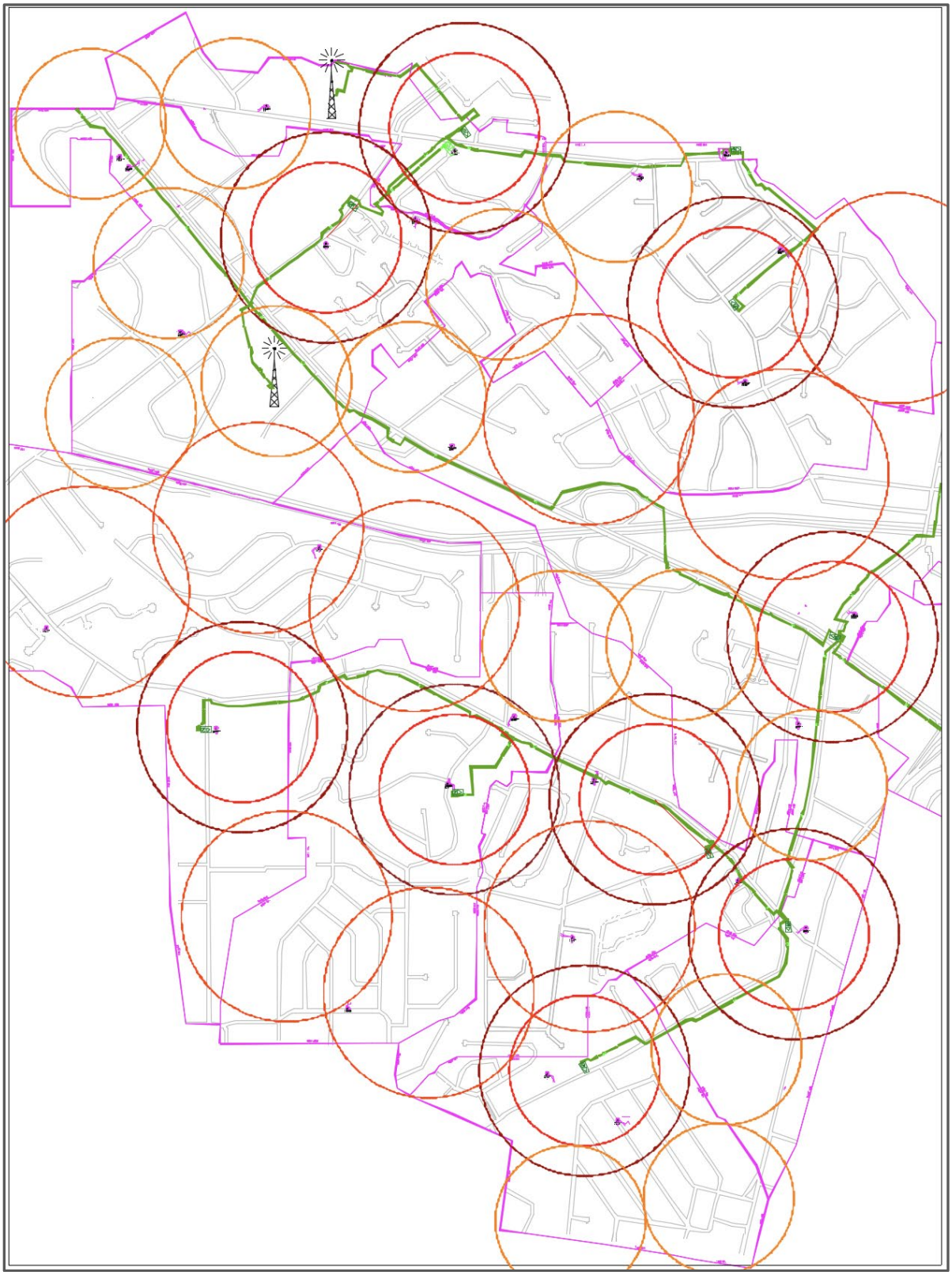


Figure 24 – N+6 Suburban area with Small cells at Fiber Node + HFC coax

The rectangle captured in Figure 23 and Figure 24 is about 1.9 x 2.6 miles (~ 5 square miles), while the actual area covered by our 9 nodes is ~3.5 square miles. As can be seen in Figure 23, less than half of the area has coverage, even with the extended 350m streetlight range. If the operator's goal is just to off-load some mobile data onto their network, this might be good enough. Also note how this coverage on N+6 plant is significantly less than the coverage seen on N+3 plant in the previous section.

6.1.2. CBRS Small Cells at Fiber Nodes and Coax locations

The next step is to fill in the area coverage with strand-mount CBRS small cells with ~250m to ~350m range. This is shown in Figure 24. These small cells would use a DOCSIS backhaul. For this analysis, any lower density nodes with <85 HP/mile were assumed to support ~350m cell range, while higher densities >85 HP/mile would only get ~250m range.

This requires 23 additional small cells to get reasonably complete coverage with a small number of residences just outside the cell radius. Care is taken that these small cells are centered either on the hardline coax strand or on a pole. Note that using streetlight mounting increases the coverage area to 350m to 500m and might eliminate a third of these coax-connected cells.

Overall, there are roughly three to four small cells for every fiber node in this N+6 HFC example. But this can vary quite a bit from node to node. Some nodes only have a single coax-based small cells while others need four more coax-based cells. It turns out that this is a function of the node's homes passed density.

Figure 25 – A zoom-in into one of the high-density nodes

Figure 25 zooms in into the two highest density nodes (i.e. Node #3 and #4 from Table 4). These are located in the lower-right corner of Figure 24 and have densities of 146 – 164 HP/mile. The high-density nodes tend to have shorter cascade lengths and only need one coax-based small cell in addition to the one at the fiber node to cover their area. It took a total of 4-5 cells total to cover this two-node area.

Figure 26 zooms into the lowest density nodes (i.e. Node #1 and #2 from Table 4) in the middle-left of Figure 24. These nodes have a density of 76 HP/mile. These low-density nodes need five coax-based small cells (with expanded 350m coverage) in addition to the two at the fiber node cells to complete the coverage of the 2 nodes' area. Even though there are many more small cells on the DOCSIS network, it should not stress the capacity of the system. The five DOCSIS small cells on low density nodes covers roughly the same number of homes passed as the two DOCSIS small cells on the high-density nodes. So, it is expected that the total DOCSIS load would be similar between both scenarios.

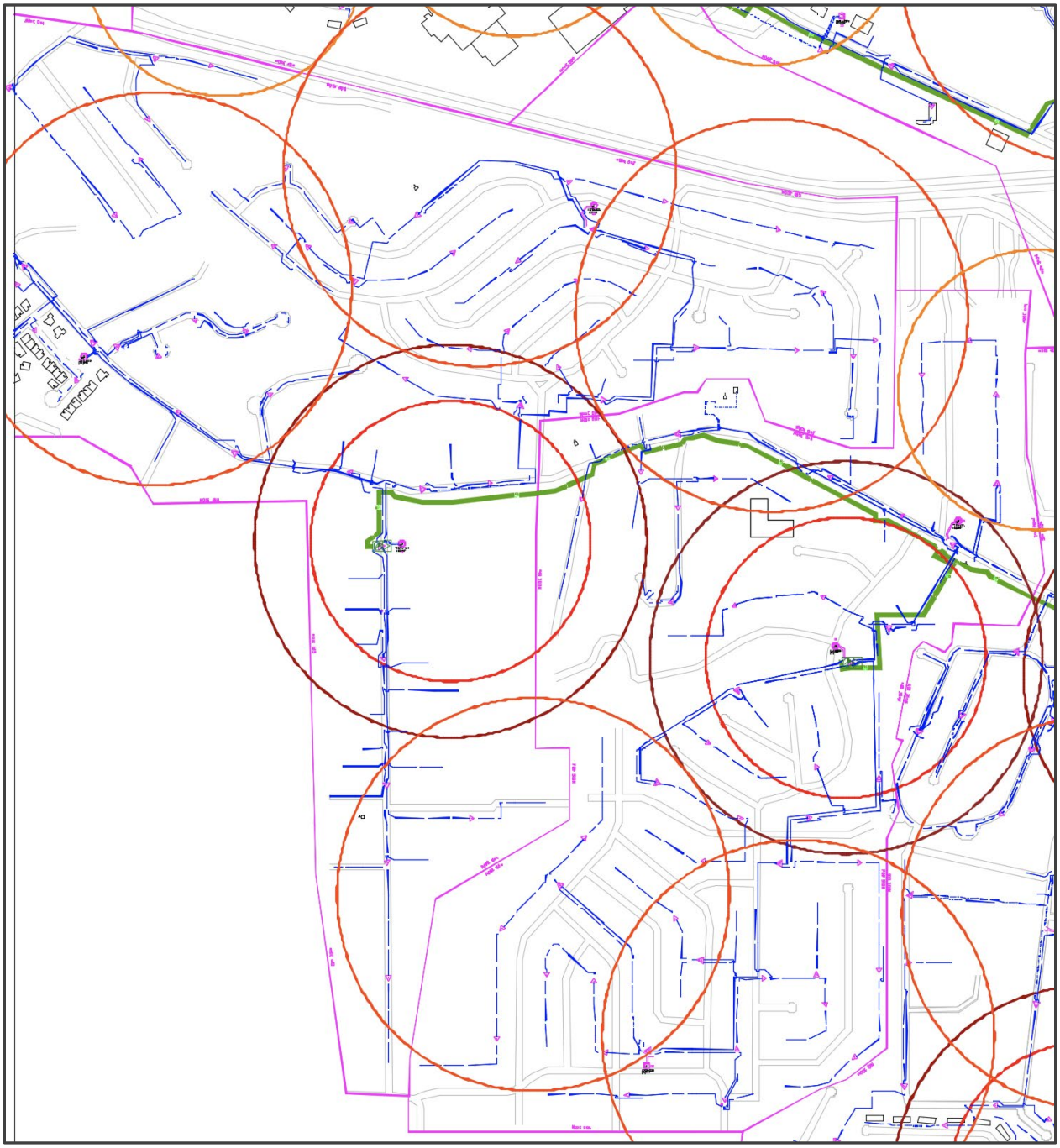


Figure 26 – A zoom-in into one of the lower-density nodes

One strategy that an operator might consider is to go to a lower 2x2 MIMO (instead of 4x4) to extend the cell reach and reduce the total number of small cells needed. The operator is effectively trading off coverage versus user capacity.

Note – In Figure 25 and Figure 26, hardline coax is shown in blue; and actives are visible as well.

6.2. N+0 Upgrade case study

The CommScope HFC design team did a N+0 upgrade design for this case study area. Table 5 shows the statistics for the N+0 upgrade compared to the original N+6 HFC plant. This upgrade pushes fiber much deeper into each node area. Total fiber mileage for this area would increase from 8.55 miles to 24.8 miles.

Figure 27 shows an update to the small cell placement from Figure 24. Note the additional fiber runs shown in green. As it turns out, fiber now passes next to 17 of the original 23 coax-based small cells. The remaining 6 cells are now within 100m – 150m of fiber. With a little additional effort, every CBRS small cell could eventually have a direct fiber connection. Note in the N+0 design that there are now 110 fiber nodes compared to a total of 32 small cells for this area.

In general, N+0 upgrades tend to be relatively expensive because of the amount of fiber being pulled. As shown, N+0 designs also push the fiber much deeper than is needed for CBRS small cell coverage. Our estimates are that an N+2 upgrade might have fiber node placements that align nicely with CBRS small cell placements. An interesting future study would be to look at a more economical N+2 upgrade to see how it aligns with the CBRS small cells.

Table 5 – Statistics of Metro-suburban HFC N+6 nodes vs. N+0 Upgrade

N+6/N+0 Case Study:	N+6 (original area)	N+0 (upgraded area)
Total Fiber Nodes	9	110
Total Fiber Mileage	8.55	24.8
Coax Plant Mileage	59.6	
Aerial	36.1	
Underground	23.5	
Total Actives	381	110
Actives/Mile	6.4	1.85
Cascade	N+3 – N+6	N+0
Total Passings	5,740	5,740
HP/Mile	96	96

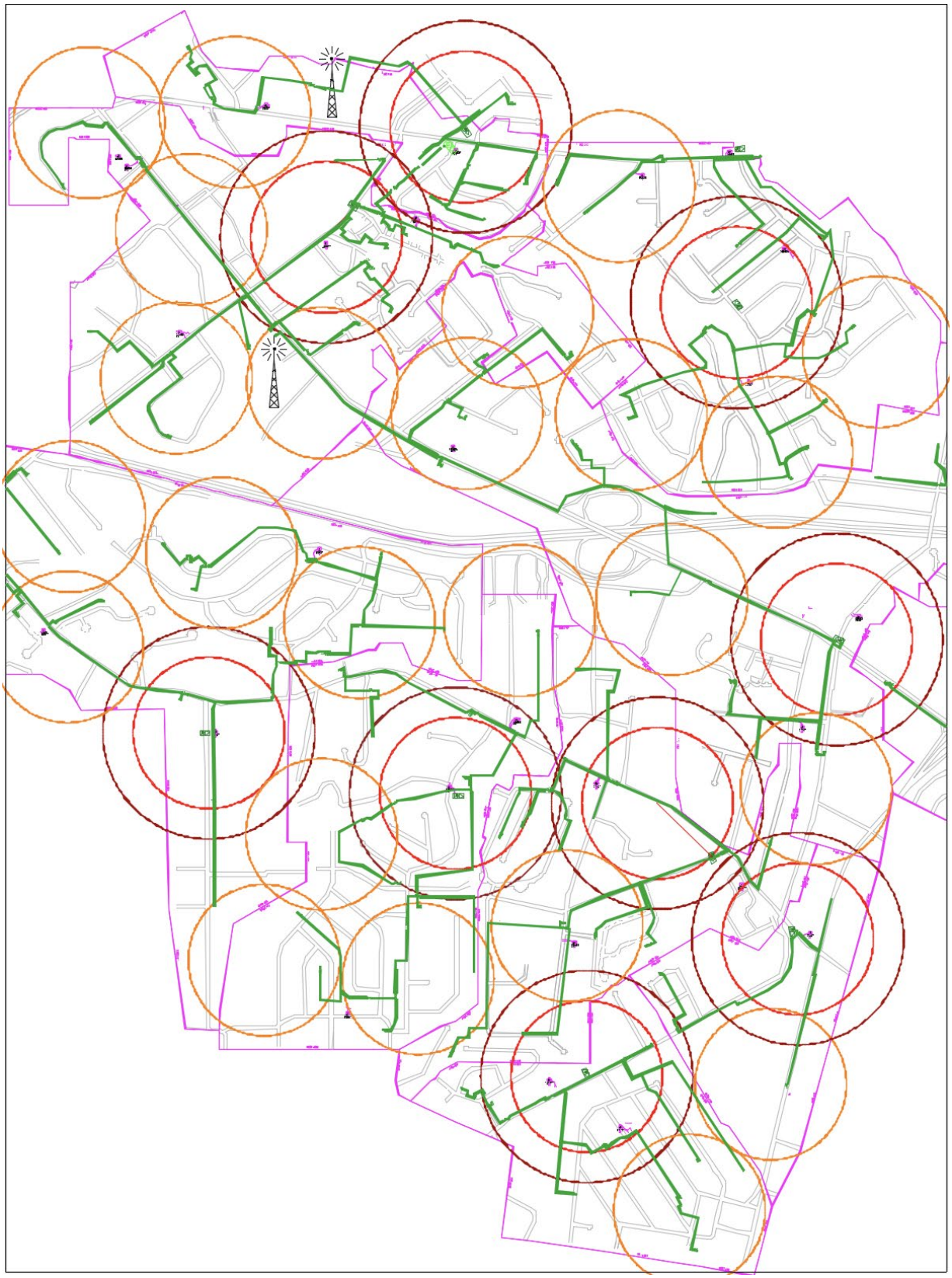


Figure 27 – Suburban area with N+0 fiber upgrade, Small cells at Fiber Nodes

7. Summary

Combining Cable 10G and Wireless 5G can offer an evolutionary strategy with much synergy. Recent 5G developments in C-band and CBRS provides some new mid-band spectrum (i.e. 3.5 – 4 GHz) that is offering a middle ground that may be the future wireless workhorse.

The paper presented a basic tutorial on mid-band wireless technologies; looks at the 5G mid-band capacity requirements; and then showed several case studies on how CBRS small cells might overlay various HFC nodes of varying homes passed densities.

The many choices for the mid-band wireless system can vary bandwidth requirements from 100's Mbps to many 10's Gbps. The paper shows which configurations can easily be supported on DOCSIS 3.1 while others might require DOCSIS 4.0 and some need direct fiber connect.

Two case studies are provided where potential mid-band cells are mapped to actual HFC networks. The nodes under study vary from dense urban nodes (i.e. >250 HP/mile) down to sparse rural nodes (i.e. <40 HP/mile). Various trade-offs are considered in cell site placement on the HFC.

7.1. Lessons Learned

Here is a collection of key takeaways from this paper:

7.1.1. Mid-Band Small Cell Coverage Range

- Small cells will most likely have a 2T2R or 4T4R omni-directional antennas supporting either 2x2 or 4x4 MIMO.
- CBRS strand-mount small cells might have approximately 240m reach in an urban setting with a greater reach (e.g. 360m) in a more rural setting.
- CBRS small cells on top of a streetlight have increased reach, perhaps up to 340m reach in an urban setting with a greater reach (e.g. 500m) in a more rural setting.
- C-Band small cells have even further reach then CBRS small cells thanks to its higher EIRP.
- Wi-Fi 6E range will be hampered to less than 100m reach due to its higher 6 GHz frequency and lower transmit powers. However, it is still expected to rule inside the home as the Mid-Band frequencies may struggle getting inside buildings.
- Some 5G small cells might be downlink (DL) only, using more robust Low-Band frequencies for the weaker uplink (UL) signals

7.1.2. Small cell Midhaul/Fronthaul capacity requirements

- Option 2 Midhaul interface substantially reduces bandwidth capacity requirements compared to Option 7.2 Cat A Fronthaul interface.
    - Option 2 requires more electronics at the radio site (i.e. DU + RU combined)
    - Option 7.2x allows for more sophisticated algorithms (e.g. beam forming) to be done in the edge cloud.
- In general, Option 7.2 Cat A interface would need a direct fiber connection
    - TBD whether DOCSIS 4.0 could meet all of the capacity and strict timing requirements for the Fronthaul interface
- DOCSIS 3.1 capacity appears to comfortably handle Mid-Band small cells with Option 2 Midhaul interface
    - 100 MHz of Mid-Band spectrum might need 1218/204 MHz HFC

7.1.3. Mapping Mid-band Cells to HFC – Key Takeaways

- Locate first small cell at or near the fiber node to leverage both power and fiber backhaul.
    - Maximum flexibility, including the choice of using Option 7.2x Fronthaul interface
- Add additional small cells with Option 2 Midhaul interface as needed along the HFC coax to access power plus DOCSIS network.
    - Over time, can pull fiber to any small cells whose capacity outgrows DOCSIS
- N+2 HFC appears to align nicely with CBRS small cells at fiber node location
- Higher density areas (in HP/mile) tend to require fewer coax-based small cells
- Lower density areas (in HP/mile) tend to require several more coax-based small cells
    - But capacity requirements are also lower due to smaller HP/mile
    - Optionally could support a lower 2x2 MIMO (instead of 4x4) to extend cell reach and reduce number of small cells needed.

7.2. DAA Synergies

- Small cells near fiber node can share 10G Ethernet connection with RMD/RPD.
- A distributed DU in the field that aggregates 6-12 small cells with Option 7.2x interfaces fits nicely into the CableLabs 10G DAA architecture
    - DU in the field greatly reduces the long range backhaul bandwidth capacity requirements (e.g. from 100's of Gbps down to 10's of Gbps)
    - Aggregation node with CableLabs coherent optics provides plenty of bandwidth capacity for the Mid-Band wireless network distribution.
- RMD works best for distributed DU in the field
    - RPD would require DOCSIS MAC core to be located near DU, not in the cloud

7.3. Potential Mid-Band Business Opportunities for Cable Operators

Considering these lessons, what makes sense for HFC service providers? CBRS/C-Band small cell reach covers a significant number of mobile users with substantial data rates. But its deployments need many more cell sites then current LTE macro-cells. This presents an opportunity for MSOs to leverage their existing HFC infrastructure for both backhaul and power.

First considering rural locations, MNOs have typically used Low-Band frequencies to maximize the distances between their macro tower base stations. This spacing may not be suitable for C-Band delivery from the tower to reach the entire community. As 5G bandwidth needs increase, MNOs may decide that deploying C-Band small cells is more economical than building more macro towers. The cable operator can provide the location (i.e. strand-mount) along with power and backhaul as a service to the MNO that owns the C-Band spectrum.

Alternatively, the cable operators might want to deploy their own 5G mobile network leveraging CBRS over their own infrastructure. In rural locations, there is a good probability that the cable operator can get access to a decent chunk of the 150 MHz CBRS spectrum via GAA.

If the cable operator is also acting as a virtual MNO (vMNO) by partnering with one of the leading MNOs, it might choose to place a handful of CBRS small cells in the busiest locations as an off-load strategy. Or the cable operator might build out the CBRS small cells across its HFC for more complete

coverage of the area. Note that 100 MHz of spectrum can enable UE data rates of 2 Gbps down by 300 Mbps up. Very impressive.

In urban settings, 4G capacity needs have already forced the MNOs to locate macro towers much closer together. So, they will be in a much better position to offer C-Band from macro towers leveraging sophisticated antenna arrays (e.g. 64T64R) and beam forming algorithms. The need for C-Band small cells will be much smaller; but it may still be needed in hot spots to help reduce congestion.

Therefore, the CBRS small cell will be key for cable operators in urban settings. But there will potentially be many others competing to get CBRS spectrum in these settings. Comcast, Charter and Verizon have all been active in acquiring CBRS PAL licensing so they can be guaranteed their share of the 150 MHz spectrum. Remember that PAL license holders can only consume up to 70 MHz total (up to 40 MHz per single PAL license holder), so at least 80 MHz will still be available for GAA users.

7.4. Conclusion

In conclusion, HFC is ideally suited to support this Mid-band xHaul infrastructure. A strategy is laid out for cable plants of varying densities. D3.1 midhaul can be leveraged extensively in the early days to get wide coverage quickly. Very dense urban areas may eventually require complex antenna/MIMO systems with fiber fronthaul, which integrates nicely with an N+2 fiber deep strategy. But even then, cells with DOCSIS xHaul will be needed to fill the holes and hotspots. D4.0 then enables even higher capacities at these cable cell sites.

In the end, Cable and Mid-band wireless (C-Band, CBRS, Wi-Fi 6E) are much stronger together and are at the core of a next gen converged network evolution.

Bibliography & References

[CableLabs_10G] “The Path to 10G: 2020 Update”, Mariam Sorond, CableLabs blog – https://www.cablelabs.com/path-10g-2020-update, CableLabs 2020

[FDX_PHY] “DOCSIS 4.0 Physical Layer Specification”, CM-SP-PHYv4.0-D01-190628, CableLabs 2019

[FDX_XSD_IBC] “Full duplex DOCSIS & Extended Spectrum DOCSIS Hold Hands to Form the 10G Cable Network of the Future”, by F. O’Keeffe et. al., IBC 2019

[LARSEN_2018] P. Larsen, et. al., “A Survey of the Functional Splits Proposed for 5G Mobile Crosshaul Networks”, IEEE Communications Surveys and Tutorials, 2018

[ORAN_2020] ORAN-WG9.Transport.0-v00.12 Technical Specification, “O-RAN Open Xhaul Transport Working Group 9 Xhaul Transport Requirements”, O-RAN Alliance, 2020

[ULM_2019-1] J. Ulm, T. J. Cloonan, “The Broadband Network Evolution continues – How do we get to Cable 10G?”, SCTE Cable-Tec Expo 2019, SCTE

[ULM_2019-2] J. Ulm, Zoran Maricevic, “Cable 10G vs. Wireless 5G – Foe or Friend? A Survey of Next Gen Network Directions”, SCTE Cable-Tec Expo 2019, SCTE

[ULM_2016] “Adding the Right Amount of Fiber to Your HFC Diet: A Case Study on HFC to FTTx Migration Strategies”, John Ulm, Zoran Maricevic; 2016 SCTE Cable-Tec Expo

Abbreviations

3GPP	3rd Generation Partnership Project
4G, 5G	4th, 5th generation (wireless)
10G	10 gigabit platform (cable)
AP	access point
BSA	base station antennas
bps	bits per second
BW	bandwidth
CAPEX	Capital Expense
CBRS	Citizens Broadband Radio Service
CDF	Cumulative Distribution Function
CMTS	Cable modem termination system
CPE	Consumer Premise Equipment
CU	Central unit
DAA	Distributed Access Architecture
DOCSIS	Data Over Cable Service Interface Specification
DL	Down link
DS	Downstream
DU	Distributed Unit
EIRP	Effective Isotropic Radiated Power
EM	Electro-magnetic
ESD	Extended spectrum DOCSIS
FCC	U.S. Federal Communications Commission
FDX	Full Duplex (i.e. DOCSIS)
FTTP	Fiber to the Premise
FWA	Fixed Wireless Access
GAA	General authorized access
Gbps	Gigabits Per Second
GHz	Gigahertz
HPBW	half power beamwidth
HFC	hybrid fiber-coax
HP	Homes Passed
Hz	Hertz
IPTV	Internet Protocol Television
ISBE	International Society of Broadband Experts
IEEE	Institute of Electrical and Electronics Engineers
LOS	Line of sight
LTE	Long term evolution
MAC	Media Access Control interface
Mbps	Megabit per second
MCS	Modulation Coding Scheme
MDU	Multiple Dwelling Unit
MHz	Megahertz
MIMO	multiple-input and multiple-output
M-MIMO	Massive MIMO
MNO	Mobile Network Operator
MSO	Multiple System Operator

MVNO	Mobile Virtual Network Operator
N+0	Node + 0 actives
N+X	Node + X actives where X = 1 or greater
NCTA	National Cable and Telecommunications Association
nLOS	Near line of sight
Nsub	Number of subscribers
OFDMA	Orthogonal Frequency Division Multiplexing Access (Upstream)
OFDM	Orthogonal Frequency Division Multiplexing
OPEX	Operating Expense
ORAN	Open Radio Access Network
P2P	Point to point
PAL	Priority access license
PHY	Physical interface
PON	Passive Optical Network
QAM	Quadrature Amplitude Modulation
RAN	Radio Access Network
RF	Radio frequency
RMD	Remote MAC-PHY device
RoW	Rest of world
RPD	Remote PHY device
R-PHY	Remote PHY
RRU	Remote Radio Unit
RU	Radio Unit
Rx	Receive
SAS	Spectrum Access System
SCTE	Society of Cable Telecommunications Engineers
SDV	Switched Digital Video
SFU	Single family unit
SG	Service Group
SINR	Signal to Interference and Noise Ratio
TDD	Time division duplexing
Tx	Transmit
UE	User Equipment
UL	Up Link
US	Upstream
VoLTE	Voice over LTE
WISP	Wireless Internet Service Providers

Practical Implementation of Profile Management Application (PMA) to Improve Data Throughput in the Presence of Impairments

A Technical Paper prepared for SCTE by

Brady Volpe
Founder and CEO
NimbleThis and The VolpeFirm
3000 Old Alabama Rd. Suite 119-434, Alpharetta, GA 30022
+1-404-954-1233
brady.volpe@volpefirm.com, brady.volpe@nimble-this.com

1. Introduction

Speed. We never seem to get enough of it. This is certainly true when it comes to high-speed data. Over the past 20 years the data over cable system interface specification (DOCSIS®) publications have undergone several iterations to accommodate more and more speed. DOCSIS 3.1 introduced new technologies specifically to support even more speed improvements. Specifically, DOCSIS 3.1 introduced orthogonal frequency division multiplexing (OFDM) in the downstream and orthogonal frequency division multiple access (OFDMA) in the upstream. OFDM and OFDMA are unique in their implementation of frequency allocation because they use very narrow slices of bandwidth in the RF spectrum to transmit data, called sub-carriers. A single OFDM channel is made of many sub-carriers. For example, a single 192 MHz OFDM block can contain 7600 subcarriers. These subcarriers are only 25 kHz or 50 kHz in bandwidth. Contrast this to our traditional single-carrier quadrature amplitude modulated (SC-QAM) channel which is typical 6 MHz (or 8 MHz for some non-North American systems).

Each subcarrier can have its own modulation from 16-QAM up to 4096-QAM (even high order modulations may be supported. Again, we can contrast modulation to legacy SC-QAM, which are limited to 64-QAM or 256-QAM. In the comparison of SC-QAM to OFDM we see a drastic difference between bandwidth (6 MHz vs 25 kHz) and modulation (256-QAM vs 4096-QAM).

As previously mentioned, in OFDM each subcarrier can have its own modulation. This means one subcarrier can be running at 4096-QAM while its adjacent subcarrier could be running at 16-QAM. This is a very power feature when impairments are present because each subcarrier can be optimized to ensure modems can receive data from each subcarrier no matter how bad the impairment. But how could one possibly know how to configure 8000 subcarriers? It's not possible. This falls into software called the profile management application or PMA coupled with proactive network maintenance (PNM).

This paper will explain how PNM and PMA work together to optimize the OFDM downstream and later the OFDMA upstream to maximize the data throughput in the presence of downstream and upstream impairments.

2. What is PNM?

Proactive network maintenance (PNM) has been steadily growing in acceptance and popularity over the past several years with adoption amongst cable operators from Tier 1 such as Comcast all the way to the smallest Tier 4 operators — both in the U.S. and internationally — as a tool for optimizing their work force and improving subscriber quality of experience (QoE). No longer is it seen as a shiny gimmick or a novelty to detect the poltergeist in the network before it goes bump in the night. It is a go-to network maintenance tool with the added benefits of workforce optimization.

PNM DOCSIS Overview

PNM in DOCSIS collects several key metrics including equalization data from cable modems, downstream spectrum captures from modems and upstream spectrum capture from the cable modem termination system (CMTS). Along with traditional SNMP metrics, this data is used to identify physical layer impairments often not identifiable with legacy monitoring systems and meters. Two key takeaways with a properly orchestrated PNM system are that it provides visibility into impairments most operators have lacked visibility into, and it provides clear, visible, and actionable insights, such as whether the impairment is in the subscriber's home or in the outside plant. This alone drives key decisions as to which resource to send (and where to send it!) to resolve a problem; intelligent problem-solving leads to resource optimization.

Types of Problems Proactively Found

Proactive tools are rapidly seen as must-have tools as the need for nearly every industry to resolve issues before they occur is increasingly under demand. DOCSIS PNM specifically focuses on the physical RF plant. This means coax cable, connectors, and passive & active devices. Everything from the CMTS to the cable modem (CM).

Interestingly, the CMTS and CM are rarely the sources of problems. In fact, systems that use PNM see a rapid decrease in unnecessary modem replacements — the modem was never bad in the first place. The most frequently found problem is typically bad or improperly installed F connectors, such as the one in Figure 1.

Figure 1 - F-connector not flush with the cable's dielectric

When connectors are not properly installed, this can create several problems:

- Micro-reflections
- Intermittent connections
- Signal ingress
- Signal egress

The problems are difficult to detect with a traditional signal level meter (SLM) but can cause intermittent issues with cable modems and set top boxes resulting in repeat calls to the subscriber's home. Further, improperly installed connectors allow signal egress, causing possible regulatory issues with FCC Part 76

leakage regulations, as well as upstream and downstream ingress, potentially impacting many (or all!) subscribers on the same fiber node.

Figure 2 - Trunk line outer shielding cut by technician when coring cable for connector, illustrates another common issue identified with PNM — outside plant impairments.

In **Figure 2** we see that the outer shield on a hardline coax has been cut the entire way around when a technician was overzealous during coring. When the shielding is compromised, the coaxial cable's impedance is no longer 75 ohms. This results in a major impedance mismatch, or micro-reflection. Further, ingress and egress will result from this exposed section of coax.

In the examples of **Figures 1 and 2**, these impairments may or may not be immediately obvious to the technician depending on the severity of the damage. However, one fact we know is that cable does not get better over time. Water ingress will create more corrosion and these impairments will result in eventual outages of either single modems or clusters of modems. The impact will be dissatisfied subscribers, and technicians in firefighting mode to find and fix the root cause.

What if there were a better method? A way that we could find these impairments before the customer was impacted. There is. It's **PNM**.

Impact to Home Techs

PNM starts with home installation technicians. They are the front line, installing new modems every day or visiting subscribers who are calling about issues with their modems. Having worked with PNM since 2012 I can tell you two facts: 1) the majority of return path ingress comes from individual subscribers'

homes, and 2) most hard-to-find impairments causing repeat truck rolls to come from subscribers' homes. And this is where the home techs spend most of their time.

Given the new visibility into micro-reflections and group delay, home techs can now see impairments previously hidden to them. Micro-reflections are typically the dominant issue in subscribers' homes due to bad wiring, such as the example shown in **Figure 3.**

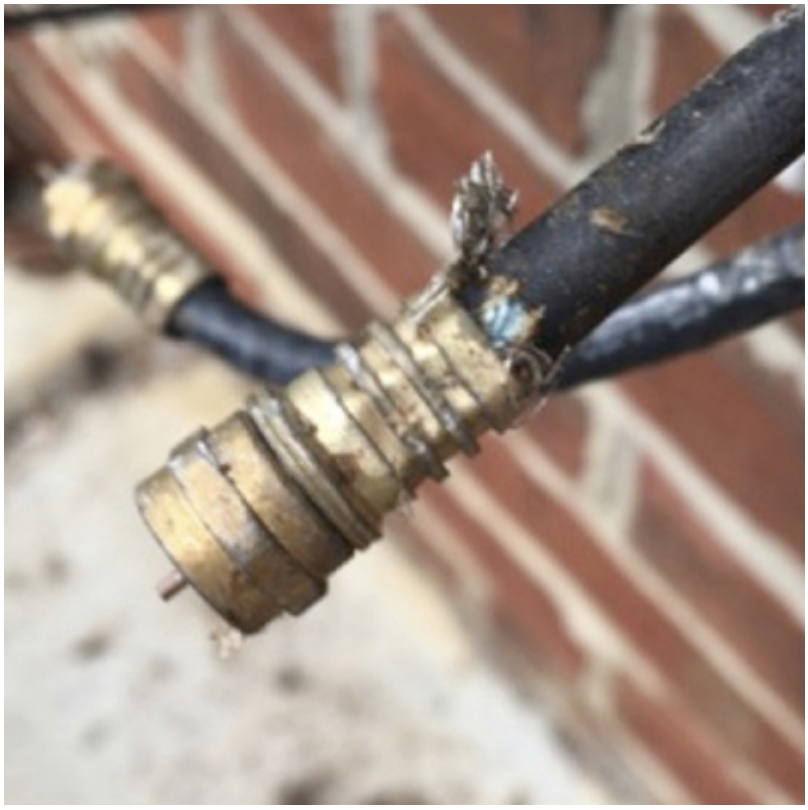


Figure 3 - Crimp-on connector showing poor grounding and corrosion on shielding.

The example shown in **Figure 3** is often difficult to find, especially when it exists under crawl spaces or trailers. However, the impacts can be maddening for technicians trying to resolve modem issues for frustrated or unhappy subscribers. Modems will report T3 timeouts, slow speeds, uncorrectable codeword errors and intermittently drop offline. Often the modem will be replaced by the tech in futility because traditional SLMs will not show micro-reflections. PNM will quickly show the micro-reflection and even estimate the distance to the micro-reflection, quickly showing the tech that there is a problem and where to look for the problem.

Every time a field tech installs a new modem or has an opportunity to gain access to a customer premise, this is an opportunity to be proactive. PNM gives the home tech the ability to test the subscribers' homes for several typical performance parameters, such as power levels and modulation error ratio (MER). And even more importantly they can test PNM metrics such as group delay, micro-reflections, and full band capture. These additional metrics go far to ensure every modem installed will perform well for the subscriber and will not allow ingress into the plant provided the tests pass. If the tests fail, the home technician should resolve the issues or escalate them to a more senior technician.

Impact to Maintenance Technicians

Maintenance technicians face the same issues as field technicians: impairments that are invisible with traditional tools. PNM provides visibility. Maintenance techs can now also see micro-reflections, damage to the outside plant, and group delay.

By identifying groups of modems affected by the same outside plant damage, we can form a Correlation Group. A Correlation Group is a cluster of modems seeing the same impairment. This cluster of modems can help triangulate and localize the outside plant impairment, thus identifying its location. Now maintenance technicians know there is an outside plant impairment, how severe it is, where it is located, and how many subscribers are impacted.

Frequently, outside plant impairments are made visible before subscribers notice the plant impairment. This is key to being proactive. The plant damage can be fixed on a schedule convenient to the operator or can be ignored until it degrades to the point that subscribers start to call CSRs. This becomes a company decision as to how they handle proactive activities.

PNM, DOCSIS 3.1 and RxMER

DOCSIS added many more PNM tests with the release of DOCSIS 3.1. One very power test added for OFDM and OFDMA is called RxMER. RxMER per subcarrier provides a modulation error ratio (MER) measurement for each subcarrier in the OFDM or OFMA channel. This means if an OFDM channel has 7600 subcarriers, the tester will receive 7600 MER results, one corresponding to each subcarrier. When plotted, it looks like a spectrum analyzer chart, but has different meanings. See Figure 4 for an example plot of OFDM RxMER per subcarrier data.

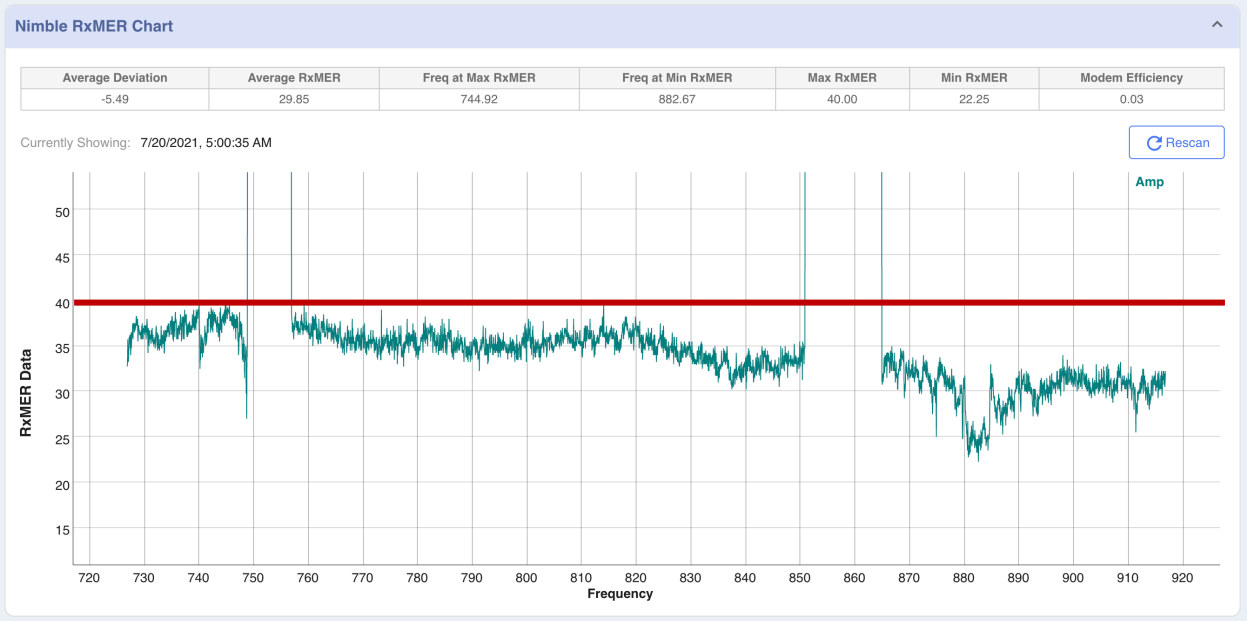


Figure 4 - RxMER per subcarrier plot of OFDM channel from live plant.

In **Figure 4** there are some unique observations one can make. First, the average RxMER is 29.85, as shown in the top of the chart. A red line is added at 39 dB MER, which is the threshold for which 4096-QAM is required. Ideally, all RxMER datapoints should be above the red line to provide the highest level of data speed to the subscriber. If the CMTS was configured to provide 4096-QAM, this modem would not be able to receive any data. See Table 1 for the mapping of RxMER to OFDMA modulation. Finally, there are two

sections of the RxMER plot where the chart goes above 50 dB. One between 749 and 757 MHz, and another between 851 and 865 MHz. These are called exclusion bands. Exclusion bands are configured in the CMTS to disable subcarriers where the cable operator knows high levels of RF ingress exists, such as LTE interference. Exclusion bands are helpful to disable subcarriers which we know will be severely impacted no matter how hard we work to clean up the HFC plant.

RxMER can be directly mapped to the modulation level which can be supported by a receiving cable modem. These mappings are defined in [PHYv3.1] (*see Table 46 - CM Minimum CNR Performance in AWGN Channel*) and the [CM- OSSIv3.1] (*see Table 72 - CmDsOfdmRequiredQamMer Object*). Table 1 shows these mappings up to 16384-QAM, but field testing has indicated that these mappings are conservative. As indicated in Figure 4, 39 dB is often used as the limit for 4096-QAM rather than the CableLabs recommendation of 41 dB for 4096-QAM.

Table 1 - Mapping of Downstream RxMER to supported QAM Level [1].

Constellation/ Bit Loading	CNR/MER (dB)
16 QAM	15.0
64 QAM	21.0
128 QAM	24.0
256 QAM	27.0
512 QAM	30.5
1024 QAM	34.0
2048 QAM	37.0
4096 QAM	41.0
8192 QAM	*46.0*
16384 QAM	*52.0*

3. What is PMA?

The profile management application (PMA) ingests RxMER data from a given node, analyses the RxMER per subcarrier data, and outputs the following:

- A modulation profile which can be applied to the CMTS which is optimized for all modems in the presence of RF impairments. This single profile is used across the entire OFDM / OFDMA channel, but does not account for frequency specific impairments, such as band roll-off or LTE interference, as an example,
- A modulation profile per subcarrier, which can then be used to create segmented profiles on the CMTS. This is a more refined method that allows the cable operator to optimize their profiles based on frequency specific impairments and is particularly useful where impairments are not consistent

across the spectrum. There are many use cases for this, but some obvious ones are high frequency roll-off, LTE ingress, suckouts, and low frequency noise,

- A per-modem profile which can be applied to each modem by MAC address. This is an aggressive method of profile management, which is not supported by all CMTS vendors at the time of writing of this document. However, future support is expected. In this method, each cable modem can have a specific profile applied specific to the modems' RF impairments, thus optimizing the end users experience and ensuring every subscriber has maximum data throughput in the presence of changing RF impairments.

The objective of the above process is to not only increase network capacity, but to also prevent outages. If a subscriber's modems is offered a profile higher than what the modem can sustain in the presence of RF impairments, little or no data will flow over the OFDM channel to the modem. In most CMTS implementations, the OFDM channel is set to the "primary" or "preferred" channel. When the primary channel is down, the subscriber experiences no data or from their perspective, an outage. This results in customer service representative (CSR) calls and likely truck rolls. This is costly to the cable operator. The PMA can function in a similar way to pre-equalization in the upstream, by changing the CMTS profiles to prevent unnecessary CSR and truck rolls, thus saving the cable operator money. Further, the PMA saves the subscriber preventable outages and makes for happier subscribers.

Today's CMTSs support only a few profiles, which does vary from vendor to vendor. Let's look at an example vendor that supports four profiles, which can be called profiles 0, 1, 2, and 3.

A CMTS engineer may configure profile 0 to operate with every subcarrier at 256-QAM. Then profile 1 to operator at 1024-QAM, profile 2 to operate at 2048-QAM and profile 3 to operate at 4096-QAM. On the CMTS, these profiles would look something as follows:

- ofdm ds-profile 0 default-modulation 256qam
- ofdm ds-profile 1 default-modulation 1024qam
- ofdm ds-profile 2 default-modulation 2048qam
- ofdm ds-profile 3 default-modulation 4096qam
- ofdm frequency low-edge 834000000 high-edge 1026000000 plc-block 842000000

Every subcarrier in the above profiles is assigned the default-modulation. For example, for ds-profile 0, every subcarrier is configured for 256-QAM; and for ds-profile 3, every subcarrier is assigned 4096-QAM. A subscriber's cable modem can dynamically choose which profile to use based on the impairments between the CMTS and the cable modem. High modulations will allow higher data speeds and lower modulation will allow lower data speeds.

Finally, notice the very last line. This indicates the OFDM start frequency of 834 MHz and stop frequency of 1026 MHz. It also shows a physical layer link channel (PLC) frequency of 842 MHz. The PLC is a special narrow channel of 400 kHz wide that carries signaling and boot-strapping information (e.g., OFDM channel parameters and MAC management messages). This PLC can be easily recognized: it lies in the middle of a specially defined 6 MHz wide range containing 8 pilot subcarriers [2]. See Figure 5 for an example of the PLC in an OFDM channel.

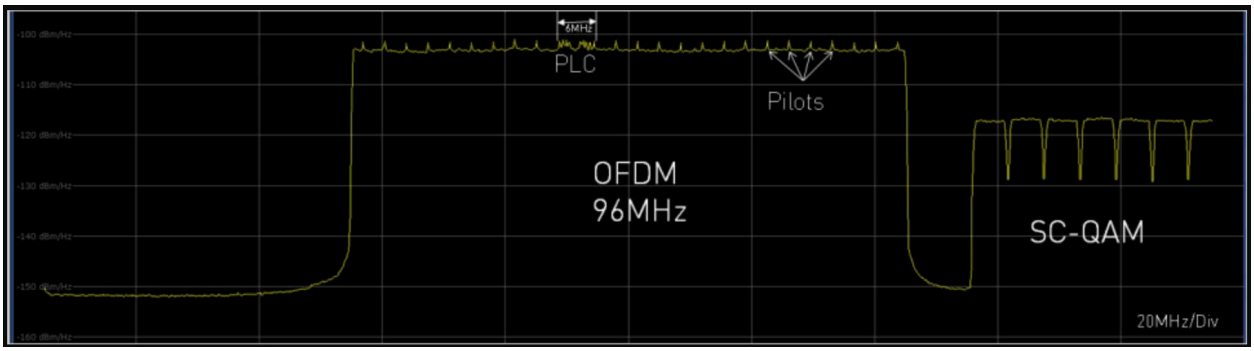


Figure 5 - Example OFDM channel highlighting the PLC [2].

The PLC is of critical importance because it delivers signaling and MAC management messages to all cable modems. If the PLC is impaired, it will cause all cable modems to go offline. We will discuss more on this in the next section.

Some of today's CMTSs permit only four modulation profiles as shown in the example above, while others support more. But what happens if RF impairments exist which are non-optimal for the configured modulations? As an example, what happens if there are significant impairments which permit modems from using 1024-QAM, but 512-QAM would be sufficient? Because 512-QAM is not an option in our selected set of profiles, modems must drop down to 256-QAM. Thus, we are leaving bandwidth on the table. Further, impairments vary over time. So even if we re-programmed the CMTS to support 512-QAM to regain our bandwidth, we may find a day later it needs a new update for different modulations. What's more is that most operators program their CMTSs with a flat profile, such as 4096-QAM across the entire OFDM channel. Considering that OFDM channels can be 192 MHz wide, one can imagine that impairments such as roll-off at high frequencies would make 4096-QAM unusable at high frequencies. It is possible to program mixed modulations in segments. An example of this would be where lower frequencies that have high RxMER are configured for 4096-QAM. Similarly, higher frequencies in the roll-off band with low RxMER are configured with lower modulations such as 512-QAM or lower. But again, how do we even know any of this information such as what frequency bands should get which modulations?

On top of this, how do we know the ideal location for the PLC on every node on every CMTS in the network? This can become a substantial amount of work when dealing with hundreds or thousands of nodes and a network that is constantly changing. We need a solution, and the solution exists – it's called PMA.

PMA is processing engine which analysis the RxMER data from every cable modem in each node. After analyzing the RxMER data, the PMA engine will provide the optimal modulation for each subcarrier and the optimal frequency location for the PLC. This takes the guess work out of determining which profiles and PLC location each node requires. Further, PMA can be fully automated so that every node can be optimized multiple times per day to compensate for the ever-changing RF impairments.

Figure 6 shows a high-level concept of the PMA architecture as it is implemented in multiple MSO systems. It generally has four components as described next.

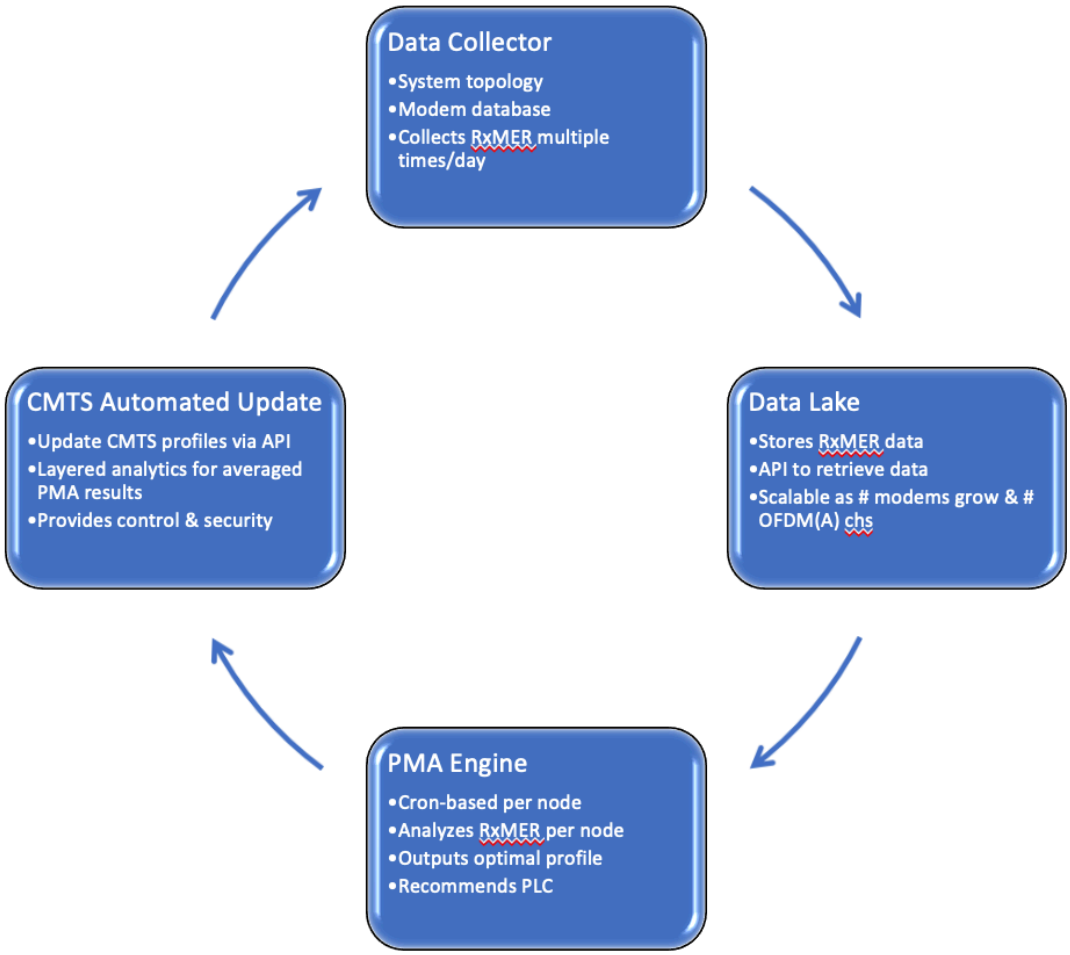


Figure 6 - PMA System Architecture.

Data Collector: The PMA process begins with the collection of RxMER data. To do this, the collector must first communicate with the CMTS to obtain a list of all cable modems and obtain topology awareness. Topology awareness simply means a mapping of each cable modem to its respective fiber node. The data collector then communicates with each cable modem to obtain the OFDM RxMER file. It also obtains additional telemetry information from the modem and CMTS to perform downstream and upstream PMA.

Data Lake: PMA can be performed on the fly; however, it is much more beneficial to make analysis on averaged data. Therefore, the data collector stores its data in a large database. Application programming interfaces (APIs) are implemented to easily retrieve data from the database.

PMA Engine: The PMA engine runs on a routine basis to pull the latest data from the Data Lake and process the data. It does so using an API to retrieve RxMER per subcarrier and other data for each fiber node in the network. Once done processing, the PMA engine outputs a file that contains the optimal profile set and PLC placement for the given node, along with the CM to profile assignments.

CMTS Automated Update: The CMTS automated update is responsible for pulling the output files created by the PMA engine and applying the profiles to the CMTS and CMs. This procedure ensures the cable

operator has full control of when and how profiles are applied. Further, profiles can by modified by the cable operator via analytics so that they are further optimized according to the limitations of the cable operator's CMTS equipment.

In the next section we will see the benefits than can be achieved through the PMA.

4. What are the benefits of the PMA?

The PMA has several benefits, which are immediately realizable as follows.

- The PMA will provide the optimal profile for a fiber node:
    - The bandwidth gains in running a well-designed set of profiles can be anywhere from 15% to 40% capacity increase on a channel, compared with running the whole channel at 256-QAM. This can translate to a solid 200 to 400 Mbps extra capacity on each OFDM channel! [3]
- The PMA will reduce individual subscriber and likely entire node outages:
    - When running the PMA on each node on a regular basis, profiles are continuously being optimized for new impairments. Profiles are downgraded and upgraded, similar to how upstream dynamic modulation profiles are used for SC-QAM channels to accommodate dynamic changes in the upstream. But the PMA will do the same for the downstream and the upstream. This helps ensure subscriber traffic continues to flow over the OFDM and OFDMA channels during dynamic impairment changes.
- PMA will provide PLC placement recommendations to ensure the PLC is not operating in an area with impairments:
    - This is often overlooked as a critical feature of PMA. If the PLC is placed in a region where impairments (think about a suckout or LTE ingress) exists, this will result in an outage. PMA will aid in preventing such instances from occurring. Again, the PMA can help prevent critical fiber node outages resulting in significant cost savings from truck rolls and subscriber grief.

The PMA is not only beneficial in increasing extra capacity in the DOCSIS network, but also in reducing CAPEX expenditures through increased network reliability.

These immediate benefits may seem obvious and exciting, but they have a bottom-line impact to total system capacity, which results to OPEX savings and preventative CAPEX investment.

In 2019, Comcast developed a PMA system for generating and transacting D3.1 downstream profiles tailored to the conditions of each OFDM channel in its network. Some point-in-time metrics from Comcast's deployment of PMA indicate its realized value [4]:

- 34.3% capacity gain in OFDM profiles (Division A),
- Raw gain of 6020 Gbps for Division A,
- 91.0% CM success rate (percent of CMTSs that were successfully configured with updated profiles).

Comcast considered the PMA a huge success and subsequently released public press releases [5].

5. PMA in Action

To see a PMA engine in action, it is first important to understand the relationship between modulation efficiency, which is the bits per MHz, and required MER to support the given modulation. This relationship can be developed by adding the modulation efficiency to Table 1, which results in Table 2.

Table 2 - Modulation Efficiency vs Bit Loading vs MER

Modulation	RxMER per Subcarrier	Modulation Efficiency
No Data	<12	0
QPSK	12.0	2
16 QAM	15.0	4
64 QAM	21.0	6
128 QAM	24.0	7
256 QAM	27.0	8
512 QAM	30.5	9
1024 QAM	34.0	10
2048 QAM	37.0	11
4096 QAM	41.0	12
8192 QAM	46.0	13
16384 QAM	52.0	14

The PMA outputs both modulation and modulation efficiency (also called bit loading) when RxMER per subcarrier data are provided to it. Modulation efficiency is easier to plot and manipulate than modulation. An easy lookup table can be generated to reference modulation efficiency to RxMER per subcarrier. This allows one to compare the RxMER per subcarrier data to the recommended MER profile for a given modulation.

Example RxMER per subcarrier data from all modems in a fiber node is shown in Figure 7. As is typical with RxMER data of many fiber nodes, a portion of modems are showing very high level RxMER (>40 dB). These modems will be capable of supporting 4096-QAM across the entire spectrum. However, as can be seen in Figure 7, more than 75% of the modems fall below 41 dB MER with many modems falling below 30 dB MER. In fact, 25% of the modems on this fiber node fall below the 30 dB MER threshold across the entire spectrum. Modems falling below the 30 dB spectrum will only support 256-QAM or less! Keep in mind that legacy SC-QAM had a maximum modulation profile of 256-QAM.

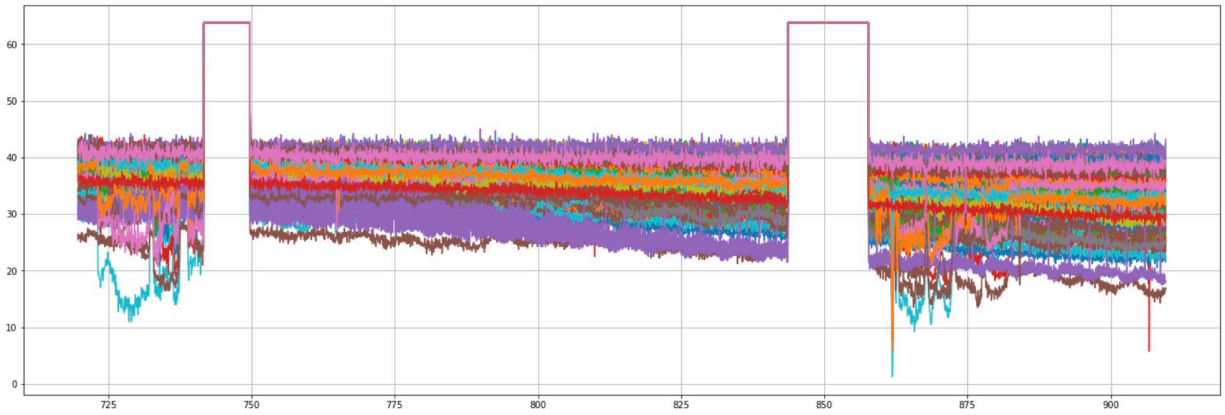


Figure 7 - Sample set from a typical fiber node with varied impairments.

This presents a challenge for cable operators to determine what modulation profiles should be assigned to this node. Without a PMA it is a guessing game. With a PMA, very inciteful information can be gathered which will allow the cable operator to configure the CMTS with the most optimal profiles to ensure all cable modems are well supported with a finite number of profiles that are supported in the CMTS.

Figure 8 shows a typical recommendation from the PMA output given the RxMER input of Figure 7. The blue dotted line in Figure 8 indicates the absolute MER for each subcarrier, which translates to the respective modulation in Table 2. The red line indicates the maximum and minimum recommended modulation profiles from a PMA based on the RxMER per subcarrier input. The red lines are broken into segments which can be applied to a CMTS that support mixed modulations in a segmented method. Additionally, the suggested profiles in this section are examples only and should not be considered as recommendations for any production system without prior testing.

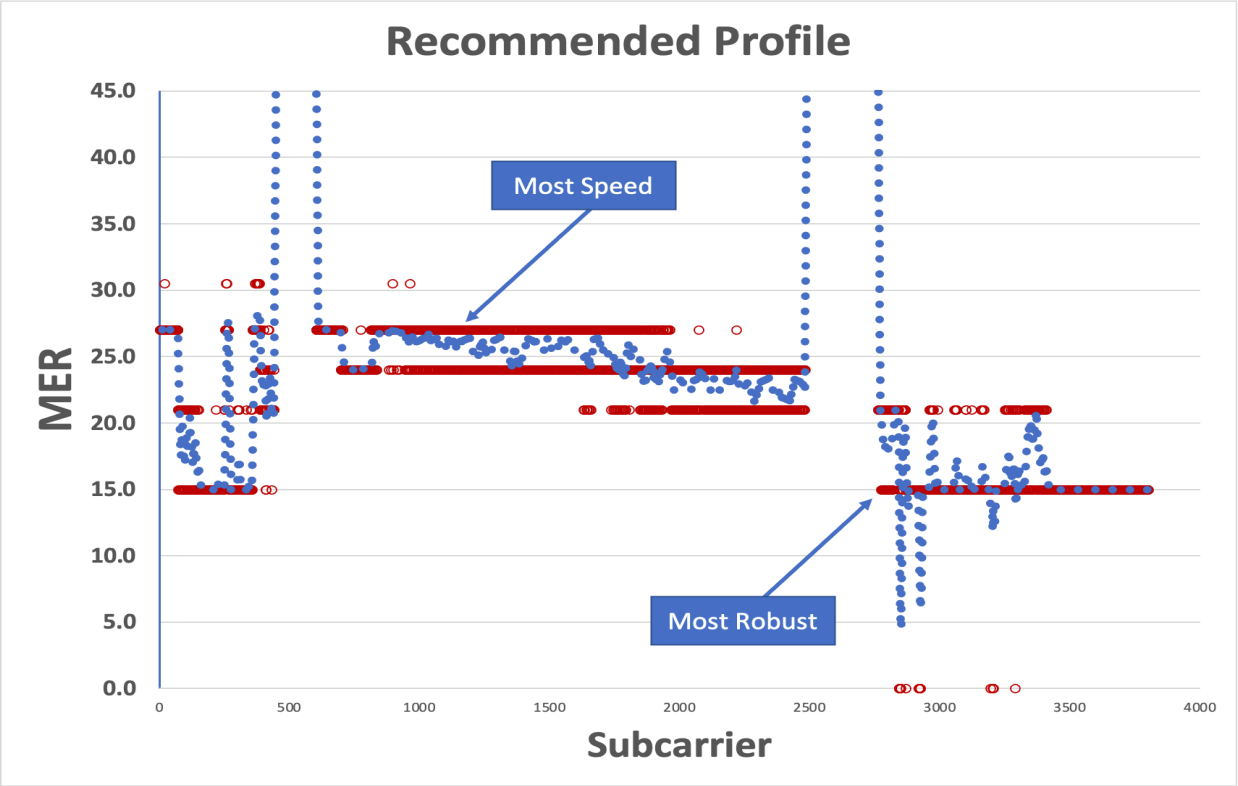


Figure 8 - PMA profile recommendations for fiber node in Figure 7.

In Figure 8, the upper limit red line is indicated as the "most speed" line. This means if the cable operator were to choose the upper end of the recommended profile, cable modems would likely achieve high data throughput, but would suffer more uncorrectable codeword errors. On the other hand, the "most robust" line indicates that choosing the lower modulation will ensure cable modems will have lower data speed but will likely have very few if any uncorrectable errors because the modulation is set low and therefore resilient to errors in the presence of impairments.

Ultimately, the data that is being presented are minimum and maximum profiles which can be updated in the CMTS to support all impaired modems in the fiber node and still maintain maximum data throughput.

Let's consider the original configuration on the CMTS as follows.

- ofdm ds-profile 0 default-modulation 256qam
- ofdm ds-profile 1 default-modulation 1024qam
- ofdm ds-profile 2 default-modulation 2048qam
- ofdm ds-profile 3 default-modulation 4096qam

First, we know that keeping ds-profile 3 is a good idea, because there are modems that support >40 dB RxMER across the band. These modems will be able to use ds-profile 3 and obtain the maximum throughput of the OFDM channel.

Next, there are some very impaired modems. To these modems always can transmit data profile 0 will be downgraded to 64-QAM, which is a very robust profile. Profile 1 will also be downgraded to 256-QAM for slightly less impaired modems. Profile 2 will be modified to a segmented profile to align the PMA profile output shown in Figure 8. There are three segments aligning with the three primary tiers in Figure 8 of 64-QAM, followed by 512-QAM and then back to 64-QAM. The segmented modulation in profile 2 represent the optimal profile from the PMA output that will give the maximum data throughput for modems. Finally, profile 3 of 4096-QAM is enabled for modems not experiencing impairments. Modems not experiencing impairments will be able to use profile 3 and will see the maximum data throughput the OFDM channel can support.

- ofdm frequency low-edge 807000000 high-edge 903000000 plc-block 850000000
- ofdm ds-profile 0 default-modulation 64qam
- ofdm ds-profile 1 default-modulation 256qam
- ofdm ds-profile 2 low-freq-edge 807000000 high-freq-edge 817000000 64qam
- ofdm ds-profile 2 low-freq-edge 817000000 high-freq-edge 877000000 512qam
- ofdm ds-profile 2 low-freq-edge 877000000 high-freq-edge 903000000 64qam
- ofdm ds-profile 3 default-modulation 4096qam

This updated profile provides an example of how one can update a CMTS profile based on PMA output to ensure that modems at the high RxMER will be able to support very fast data speeds. At the same time, the modems experiencing RF impairments will be able to toggle between the segmented profile 2 or drop to even lower modulations supported in profile 1 or 0. This will result a good quality of experience for every subscriber on the fiber node. In essence, the CMTS has been modified for this fiber node to provide maximum data speed when possible and maximum robustness when required. Every subscriber should be able to continuously receive downstream data regardless of outside plant impairments or bad in-home wiring thanks to the PMA output results.

The profiles above can be updated in an automated method each time the PMA is run on a given node to update profiles based on changing plant conditions.

6. Future Considerations

Obtaining RxMER data from modems is time consuming. It requires the use of simple network management protocol (SNMP) to configure the modems to send RxMER data back to a centralized collector. Once configured, the modems send a file using the trivial file transfer protocol (TFTP) to a TFTP server. Because plant impairments vary, it is important to gather RxMER data multiple times per day from each modem and update the CMTS accordingly. Taking multiple RxMER samples per cable modem and averaging these samples together for a single modem can make the PMA engine even more accurate. This is because a single RxMER sample from a modem could be impacted from an impairment, but multiple samples will provide a more concise data sample. However, doing so takes tremendous effort to gather RxMER data rapidly. A future implementation on the cable modem side would enable the modems to stream RxMER data back to the collector or to perform the averaging on the modem itself. As a side note, full band capture averaging is performed by most cable modem vendors today, so RxMER averaging should be doable capability. This would enable the collection of RxMER data more frequently without the addition of costly and energy hungry servers. Today, cable operators rely on servers in their data centers to collect RxMER data, which as previously mentioned, this is a CPU intensive process and

not green for the environment. As the industry focuses on optimizing data throughput, it must also focus on green initiatives. These initiatives start with pushing computations at edge devices such as cable modems rather than server farms collecting data from modems.

As shown in this paper, today it is possible to create segments in profiles. This is very effective for addressing frequency specific impairments to optimize throughput and help ensure subscribers do not experience outages. There is a limitation with this, however. Segmented profiles will often not address specific in-home wiring issues where an individual subscriber's modem is severely impacted. For this, CMTS vendors must support modem-based profiles. The PMA will output a profile for each modem. At least one CMTS vendor currently supports this functionality. As an industry it is important that achieving per modem-profile support in the CMTS is adopted. This will further improve capacity gains and the reduction of CSR calls and truck rolls for subscribers with bad in-home wiring. Like upstream pre-equalization, the PMA will output a profile to temporarily address subscribers with in-home wiring issues and ensure that their modems remain online until such a time that their wiring issues can be addressed. From this standpoint, the PMA can be considered a proactive application as it can flag individual subscriber issues having low, but still working modulation profiles.

7. Conclusion

In conclusion, a PMA is a useful add-on to PNM and DOCSIS 3.1 with OFDM. No cable plant is without RF impairments. Whether the impairments are in the outside plant or in subscriber homes, these impairments change over time. To provide reliable high-speed service over an OFDM channel it is important to configure the CMTS with the optimal profiles to maximize data throughput. A PMA provides operators the necessary visibility to ensure they are deploying appropriate profiles for each node. Further, it is evident that the PMA can be quite effective at reducing CSR calls and truck rolls via its inherent ability at improving the robustness of OFDM and OFDMA channels. This results in a cost savings for cable operators as well as keeping subscribers happy and online.

A properly developed PMA engine will create profiles for each node every time RxMER per subcarrier data are polled from DOCSIS 3.1 modems in each node. Keeping in mind the architecture diagram of Figure 6, one can understand the closed loop process of a PMA and how this plays out. Data collection, data storage, PMA analysis, and finally profile updates on the CMTS. Rinse and repeat. This process should be repeated at least four times per day for every node in the network, and more frequently if possible. Why? Because RF impairments change during the day and maximizing data throughput and customer quality of experience (QoE) are critical in today's world where people work, learn, and play at home with a high dependency on data provided over DOCSIS networks.

While this paper primarily focused on the PMA in the downstream, a PMA can work equally well, if not better, in the upstream. The upstream is the Achilles heel of the HFC network, having even more RF impairments than the downstream. As cable operators roll-out more OFDMA channels in the upstream, it is anticipated that a PMA will become a critical component to maximizing upstream data throughput and subscriber QoE.

Abbreviations

AI	Artificial Intelligence
API	Application programming interface
bps	Bits per second
CM	Cable Modem
CMTS	Cable Modem Termination System
CNR	Carrier to noise ratio
DOCSIS	Data Over Cable Service Interface Specification
D3.0	DOCSIS 3.0
D3.1	DOCSIS 3.1
FBC	Full-Band Capture
FEC	forward error correction
HD	high definition
HFC	Hybrid fiber coax
Hz	hertz
ISBE	International Society of Broadband Experts
JSON	JavaScript Object Notation
MER	Modulation error ratio
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PLC	Physical link channel
PMA	Profile management application
PNM	Proactive Network Maintenance
QAM	quadrature amplitude modulation
RF	Radio Frequency
SCTE	Society of Cable Telecommunications Engineers
SNR	Signal to noise ratio
bps	bits per second
FEC	forward error correction
HD	high definition
Hz	hertz
ISBE	International Society of Broadband Experts
QoE	Quality of experience
SCTE	Society of Cable Telecommunications Engineers

Bibliography & References

[1] Data-Over-Cable Service Interface Specifications DOCSIS 3.1 Physical Layer Specification (CM-SPPHYv3.1-I16-190121); Cable Television Laboratories

[2] https://www.excentis.com/blog/docsis-31-take

[3] https://www.cablelabs.com/tprofile-management-application-optimizing-docsis-3-1-networks

[4] Full Scale Deployment of PMA, Lessons Learned from Deploying the Profile Management **Application System at Scale and Considerations for Expanding the System Beyond OFDM,** A Technical Paper prepared for SCTE•ISBE by Maher Harb, Bryan Santangelo, Dan Rice, Jude Ferreira, Comcast

[5] https://www.nexttv.com/news/comcast-tech-chief-werner-says-network-is-handling-load-just-fine

Additional References:

"A Machine Learning Pipeline for D3.1 Profile Management", M. Harb, J. Ferreira, D. Rice, B. Santangelo, and R. Spanbauer, NCTA technical paper, 2019.

Convolutional Neural Networks for Proactive Network Management: Developing Machine Learning Models to Detect and Classify Impairments in D3.1 OFDM Channels, SCTE Cable-Tec Expo 2020, Ferreira, Harb, Subramanya

Code of Federal Regulations, Title 47, Part 76

https://www.cablelabs.com/tprofile-management-application-optimizing-docsis-3-1-networks

DOCSIS 3.1 Profile Management Application Technical Report, CM-TR-PMA-V01-180530, CableLabs, 2018

https://www.cablelabs.com/increase-upstream-reliability-and-capacity-with-optimized-profiles

\

Preparing For DOCSIS® 4.0 Upstream

A Technical Paper prepared for SCTE by

Nader Foroughi
Senior Network Architect
Shaw Communications
2728 Hopewell Place NE, T1Y 7J7
403-648-5937
nader.foroughi@sjrb.ca

1. Introduction

DOCSIS® 4.0 technology was created as a part of the 10G roadmap to increase capacity in both upstream (US) and downstream (DS). As we prepare to deploy this technology, the attention is rapidly shifting towards upstream, which was accelerated due to COVID-19. Last year, a noticeable jump in upstream utilization was realized by almost all multiple service operators (MSOs) worldwide, emphasizing the need for higher throughputs in the upstream. The shift to working remotely, learning from home, video conferencing and increased gaming activity demonstrated the need to increase bandwidth in the upstream to meet evolving consumer appetites.

Currently, many multiple system operators (MSOs) operate in 42, 65 or 85 MHz plant. The smaller bandwidth benefits the modems and amplifiers operating in the return spectrum. The same operators are planning on expanding the upstream spectrum bandwidth to 396 or 492 MHz in the near future, with 204 MHz as an intermediate step. Many operators are also planning on deploying DOCSIS 4.0 capable equipment in existing plant without re-spacing, which means there are a few key challenges to be considered.

In this paper, we will evaluate the DOCSIS 4.0 plant models by addressing these key challenges, including modem transmit capabilities, upstream amplifier performances and potential upstream performance expectations for various node and serving group architectures. The goal is to highlight the main areas that should be prioritized to ensure optimal DOCSIS 4.0 upstream performance in the access network.

Furthermore, this paper primarily aims to shed light on areas of the US that we need to focus on, along with providing new insights into performance of nodes and serving groups based on their characteristics and properties. The DOCSIS specifications define what can be expected from the cable modem (CM) and the cable modem termination system (CMTS). However, it does not specify what performance MSOs can expect in various plant architectures. This paper outlines the most important areas of focus for the optimal approach to deploying DOCSIS 4.0 technology in the US and sets expectations for performance in the access network.

2. Baselines and Assumptions

2.1. Distributed Access Architecture (DAA)

One of the main assumptions made in this paper is that DOCSIS 4.0 technology will be deployed in a DAA environment. There are many benefits to upgrading nodes to DAA that will not be discussed here, such as power savings in head-ends and hub-sites. In this paper the baseline assumption for required power levels and signal quality relies on DAA nodes being deployed as a part of DOCSIS 4.0 upgrades.

2.2. CM Transmit Capability

In order to find any potential shortfalls for the access network from the CM transmit channel set (TCS) back to the amplifier and DAA node port(s), the following table has been considered. In the DOCSIS 4.0 specification for modem transmit (Tx) power there are three different tilt options provided (8 dB, 10 dB and 12 dB). For the purpose of this paper only the 10 dB tilt option will be addressed.

Table 1 – DOCSIS 4.0 CM Tx /1.6 MHz

Upstream Centre Frequency	108 MHz	684 MHz	Spectral tilt (dB)
Upstream Reference Power Spectral Density (PSD) (dBmV/1.6MHz)	33	43	10

Converting the power levels above from 1.6 MHz reference power spectral density (PSD) to 6.4 MHz equivalent levels, the modem Tx power is shown below.

Table 2 – DOCSIS 4.0 CM Tx /6.4 MHz

Upstream Centre Frequency	108 MHz	684 MHz	Spectral tilt (dB)
Upstream Reference PSD (dBmV/6.4MHz)	39	49	10

Table 2 has been illustrated in the graph below:

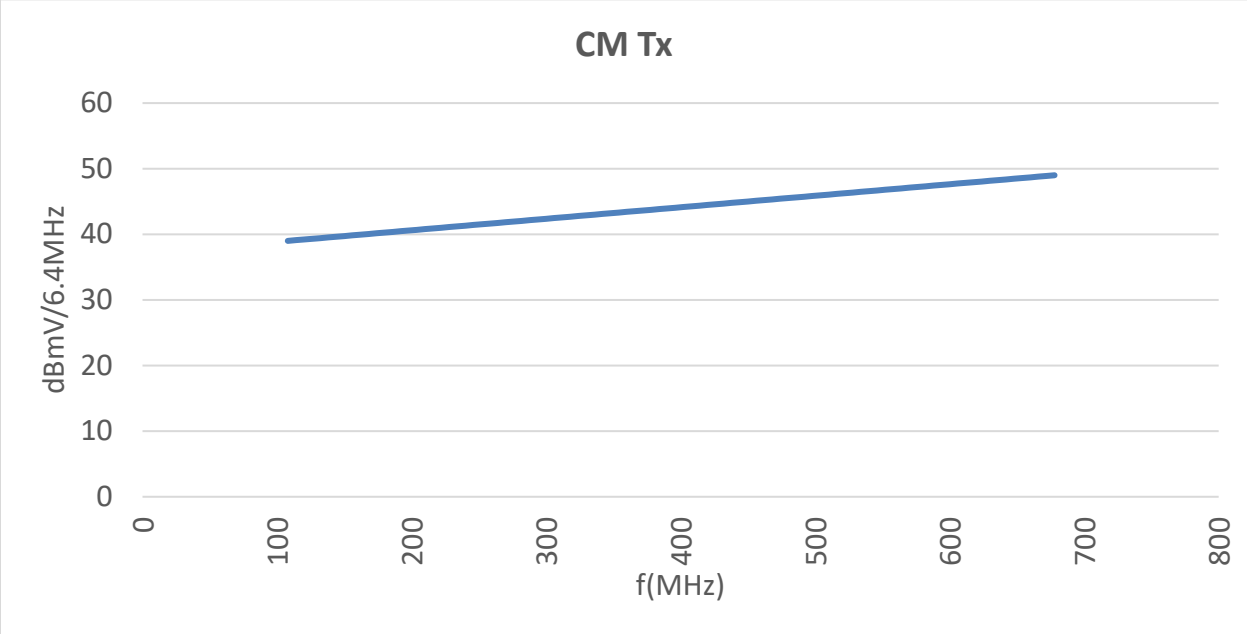


Figure 1 – DOCSIS 4.0 CM Tx Levels

2.3. CM Tx Modulation Error Rate (MER)

According to DOCSIS 4.0 CM PHY specifications, the following CM MER can be assumed:

Table 3 – DOCSIS 4.0 CM Tx MER

Grant	Tx MER
100% Grant (all OFDMA Mini Slots Used)	Each mini-slot MER ≥ 42 dB
Under-Grant Hold Bandwidth (UGHB)	Each mini-slot MER ≥ 47 dB

2.4. Modulation Order vs. Power and Carrier to noise

In order to have a baseline for achievable modulation orders throughout the plant in the US, Table 39 from the DOCSIS 4.0 PHY specification has been utilized. DOCSIS 4.0 PHY Table 39 outlines the performance that can be expected based on carrier to noise ratio (CNR) and power level received at the DAA node. Throughout this paper, power levels and CNR have been addressed to determine if one (or both) are limiting factors in performance.

The DOCSIS 4.0 PHY Table 39 does not account for the internal loss of the DAA node. For this reason, the required power levels for each modulation order have been increased by 3 dB to account for the additional insertion loss from the port of the amplifier to the remote-PHY-device (RPD) and/or remote-MAC/PHY-device (RMD) module, demonstrated in Table 4.

Table 4 – Constellation vs. Power vs. CNR

Constellation	CNR (dB)	Set Point (dBmV/6.4 MHz)
QPSK	11.0	-1
8-QAM	14.0	-1
16-QAM	17.0	-1
32-QAM	20.0	-1
64-QAM	23.0	-1
128-QAM	26.0	3
256-QAM	29.0	3
512-QAM	32.5	3
1024-QAM	35.5	3
2048-QAM	39.0	10
4096-QAM	43.0	13

CNR in the table above is otherwise referred to as signal to noise ratio (SNR) or MER.

SNR and MER have been used interchangeably throughout this paper. In Section 2.10 carrier to composite noise (CCN) has also been used as a method to estimate SNR/MER.

Section 5 explores expected performance in various plant architectures.

2.5. Amplifier Noise Figures

In order to determine the capability and performance in a cascade line, the noise figure of the return path amplifiers should be specified. The table below demonstrates this.

Table 5 – Return Path Amplifier Noise Figure

Return Path (US) Amplifier NF	6 dB

2.6. DOCSIS 4.0 Plant Models

The following plant models were created as a part of the DOCSIS 4.0 project:

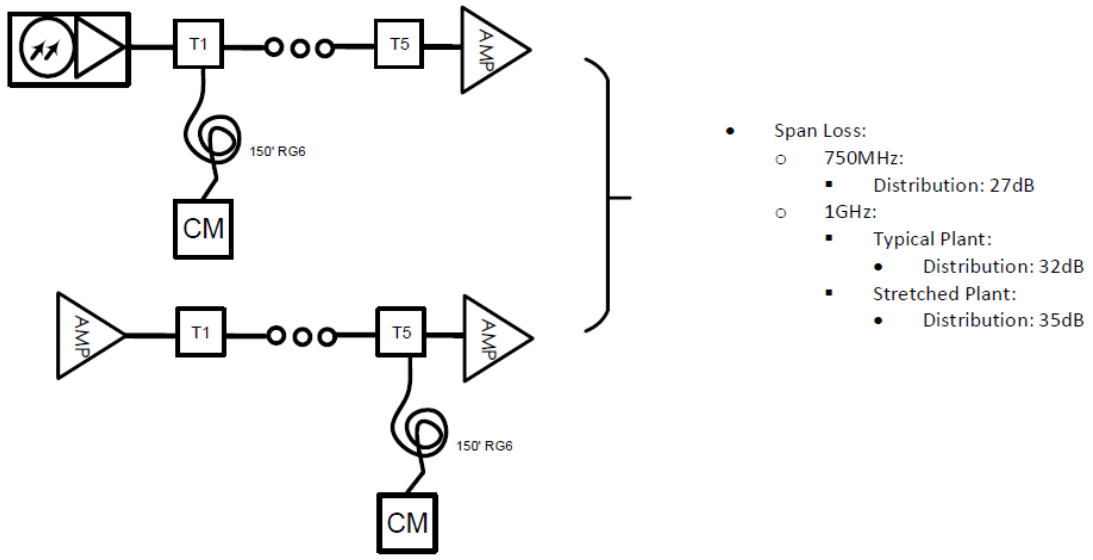


Figure 2 – DOCSIS Plant Models

In order to create a 'reasonable worst-case scenario' plant to analyze which essentially covers a higher percentile of the architectures MSOs might encounter in the outside plant (OSP), a distribution plant with 35 dB of span loss at 1 GHz has been considered. Refer to Section 3 for further details.

2.7. Noise

Designing a cascaded system for optimal CNR is always a top priority for an operator. One of the biggest contributors in network design is the receive (Rx) power at the amplifier, given that it is one of the primary drivers for achieving higher CNR throughout the distribution plant.

The minimum thermal noise power can be calculated using the following formula:

$$n_p = kTB$$

Where:

- n_p = noise power in watts
- k = Boltzmann's constant (1.34×10^{-23} joules/K)
- T = absolute temperature in K
- B = bandwidth of the measurement in Hz

The thermal noise in ~16.7 °c expressed in dBmV/6.4MHz is:

$$N_p = 57.1\ dBmV$$

The conversion of bandwidth from 6 MHz to 6.4 MHz has been rounded up from 0.28 dB to 0.3 dB.

From the equations above, carrier to noise can be calculated using the following formula:

$$C/N\,(dB) = \ C_i(dBmV) + 57.1 - NF(dB)$$

Where:

- C_i = input signal
- NF = noise figure of the amplifier

The equation above shows the significance of the Rx power versus noise figure of the amplifier, in network design.

The overall cascade C/N for amplifiers operating at different output levels can be derived from the following equation:

$$C/N_{total}\,(dB) = -10log\left\{10^{\frac{-C/N1}{10}} + 10^{\frac{-C/N2}{10}} + \cdots + 10^{\frac{-C/Nn}{10}}\right\}$$

Where, C/N_x is the carrier to noise of each amplifier calculated independently.

When cascading identical amplifiers operating at the same output level, the following approximation is typically used:

$$C/N_{total}\,(dB) = C/N_x - 10log_n$$

Where:

- C/N_x = the carrier to noise of a single amplifier
- $n =$ the number of identical amplifiers in cascade

2.8. Distortion

Distortion products from an amplifier or cascade of amplifiers have historically been characterized by measuring second order (CSO) and composite triple beat (CTB) on analog carriers. These distortion products are harmonics of the primary signal. Today, however, MSOs primarily use digital carriers. Digital carriers' distortion products do not appear similar to those of analog carriers. Instead, they appear very similar to a raised noise floor. For this reason, composite intermodulation noise (CIN) is the best way to characterize the distortion performance of amplifiers today.

The rate of accumulation of CIN products is dependent on many factors but two primary factors in how fast CIN accumulates are the amount of total composite power (TCP) utilized in the gain chip and the output level. In this paper, it is assumed that CIN products will accumulate at a 10*log rate.

2.9. Carrier to Composite Noise (CCN)

In this paper, CCN has been used as the primary method of determining signal quality. Although SNR and MER can be measured using meters and measurement equipment, there are inconsistencies in these types of measurements, especially when using different measuring equipment [4]. Some of these are:

- Equalized vs. unequalized MER measurements
- Measuring device noise floor (NF)
- Input level to the measuring equipment
- Temperature

In order to remove these inconsistencies, this paper aims to calculate CCN as a more consistent representation of signal quality based on input power levels, carrier to thermal noise (CTN), and CIN from each equipment. The formula used for calculating CCN is outlined below:

$$CCN(dB) = -10log\left\{10^{\frac{-Starting\ CCN/SNR/MER}{10}} + 10^{\frac{-CTN_{Total}}{10}} + 10^{\frac{-CIN_{Total}}{10}}\right\}$$

3. Plant Model Created

Based on the assumptions made for DOCSIS 4.0 plant models in Section 2.7, the plant model below has been created as a 'reasonable worst-case scenario' for analysis. This plant model results in roughly 35 dB of span loss at 1 GHz, which can be considered 'stretched' for today's standards of deployment, as per Section 2.6. Consideration that the following plant model encompasses a higher percentage of the scenarios that one might encounter in the outside plant has been covered.

Figure 3 – Analyzed Plant Mode – 35 dB Span Loss

As demonstrated above, a 26 dB tap value has been considered for an ultra-high-split (396 and 492 MHz) scenario and a 23 dB tap has been considered for a high-split scenario, both of which are analyzed further in this paper.

4. Analysis

4.1. Upstream Spectrum Bandwidth

Further to Section 2.11, in order to accurately set a baseline for the input power to the return path amplifier, the noise-power ratio (NPR) of the amplifiers must be studied. Due to lack of availability of NPR data at the time of writing, the following assumption has been made:

$$P_i(new) = P_i(legacy) - 10 * \log\left({}^{new\ BW}/_{legacy\ BW}\right)$$

Where:

- $P_i(new)$ = the new input power /6.4 MHz into the return path amplifier
- $P_i(legacy)$ = the current input power /6.4 MHz into the return path amplifier

An assumption has been made that the current input level to the return path amplifiers in an 85 MHz plant is 16 dBmV/6.4MHz. In this case the following table can be calculated using the formula above.

Table 6 – Return Path Expected Levels for US Splits

Return Path BW	Input Power to Return Path Amp. (dBmV/6.4MHz)
85 MHz	16
204 MHz	12
396 MHz	9
492 MHz	8

Historically MSOs have tried keep CM Tx levels as high as possible to keep carrier-to-ingress-noise as high as possible, but at the same time may have exhausted the transmit capability of the CM TCS. With that said, reducing Rx power levels into the return path amplifiers should not be a significant concern, assuming that regular plant maintenance and plant hardening practices are carried out.

Moreover, high transmit levels out of the modem can potentially increase the risk for interference between adjacent homes, as demonstrated in Section 4.3.

4.2. CM TCS's

DOCSIS 4.0 CMs will have two TCS's, one of which is capable of transmitting up to 204 MHz and the other from 108 MHz up to 684 MHz. It should be noted that there is an overlap region between the two TCS frequency bands.

The overlap region between the two TCS's can be used to the operator's advantage for near-term high-split deployment and long-term ultra-high-split deployment. MSOs have always tried to keep CM Tx levels as high as possible, both for achieving high SNRs and high carrier-to-ingress-noise. As discussed in the previous section, although the input to the return path amplifier has to be reduced by roughly 4 dB when expanding the return spectrum from 85 MHz to 204 MHz, that does not necessarily mean that modem transmit levels have to be reduced.

Amplifiers have the ability to pad the signal prior to the return (and forward) path amplifiers. Additionally, many MSOs condition their taps in the distribution line to 'force' CMs to transmit with high levels, increasing their source SNR/MER. In other words, CM transmit power, source (CM) MER and input to the return path amplifier chips have to be balanced by the MSO. Therefore, MSOs should try to optimize the amount of TCP available in the CM TCS without leaving 'unused power', which can help increase the CM Tx MER.

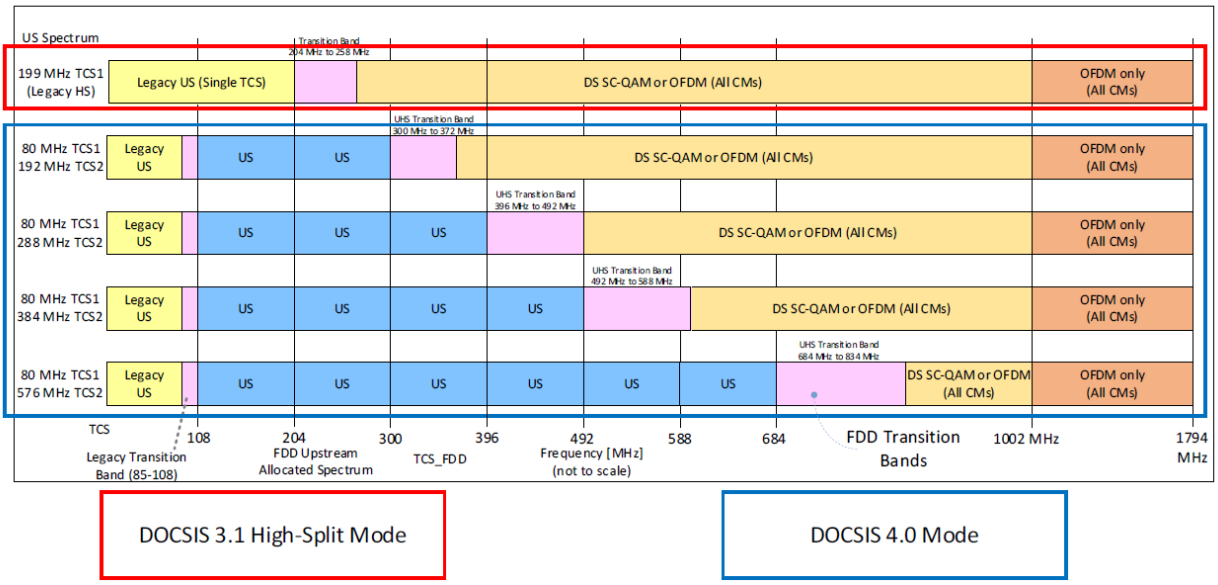


Figure 4 – Configurable FDD Upstream Allocated Spectrum Bandwidths

4.3. 204MHz Analysis

Based on the plant model discussed in Section 3, the following loss values can be calculated from the CM at the end of a 150' RG6 drop, installed as a point of entry device (PoE) to the amplifier or node port:

Table 7 – Plant Model Loss Values

Loss from each CM to Amp Port:					
Freq. (MHz)	**Tap 1 (23)**	**Tap 2 (23)**	**Tap 3 (17)**	**Tap 4 (14)**	**Tap 5 (11)**
5	23.87	22.39	20.91	19.83	19.55
30	24.77	23.78	22.79	22.20	22.40
50	25.29	24.54	23.80	23.45	23.90
83	25.93	25.48	25.03	24.98	25.71
108	26.34	26.07	25.82	25.95	26.87
150	26.91	26.96	27.03	27.47	28.70
204	27.54	27.96	28.38	29.19	30.76

Referring to Table 6, although the input to the 204 MHz return path amplifier needs to be reduced by 4 dB, the receive levels at the port of the amplifier is kept at 16 dBmV/6.4 MHz to determine if the CM has enough power to transmit to the amplifier port in this plant model. This is shown in Table 8.

Table 8 – CM Tx Levels to Amplifier or Node Port

CM Tx/6.4MHz to Port – 16 dBmV/6.4MHz Rx Level					
Freq. (MHz)	**Tap 1 (23)**	**Tap 2 (23)**	**Tap 3 (17)**	**Tap 4 (14)**	**Tap 5 (11)**
5	39.87	38.39	36.91	35.83	35.55
30	40.77	39.78	38.79	38.20	38.40
50	41.29	40.54	39.80	39.45	39.90
83	41.93	41.48	41.03	40.98	41.71
108	42.34	42.07	41.82	41.95	42.87
150	42.91	42.96	43.03	43.47	44.70
204	43.54	43.96	44.38	45.19	46.76

Knowing that the TCS is capable of 65 dBmV of TCP means that the modem can transmit 50 dBmV/6.4 MHz from 5 MHz to 204 MHz. As can be seen in the table above, the transmit levels are well below the TCP limit of the CM TCS. This can help the operator keep CM transmit levels 'consistent' with mid-split levels.

This is only true if the MSO is operating in 204 MHz return plant. If the return path spectrum is expanded to any frequency higher than 204 MHz, the first TCS will be limited to 108 MHz.

Another item to note is that the modem can in fact transmit with an 'up-tilt' (Table 9). This will be used in Section 4.4 for evaluating adjacent channel interference (ACI).

It can also be seen that we are now operating closer to the CM dynamic range window (DRW). Although many operators will not deploy any carriers below 15 MHz due to noise concerns, the modem can still transmit with up-tilt, as high as ~9 dB of tilt. This can cause concerns with the DRW of the CM, given that the current CMs have a maximum DRW of 12 dB.

4.4. Near Term Risk Analysis

In [3], a method for estimating the level of risk for ACI is outlined. Figure 5 demonstrates this method. It can be seen that with higher modem transmit levels—up to 204 MHz and beyond—the potential for energy leakage between adjacent tap ports increases. This additional energy in the adjacent home, assumed to be a 'legacy' mid-split home, can interfere with set top boxes (STB) and CMs. When the delta between the interferer and the downstream received signal for the legacy device goes above a certain threshold, depending on the front-end design of each device, it can cause degraded service.

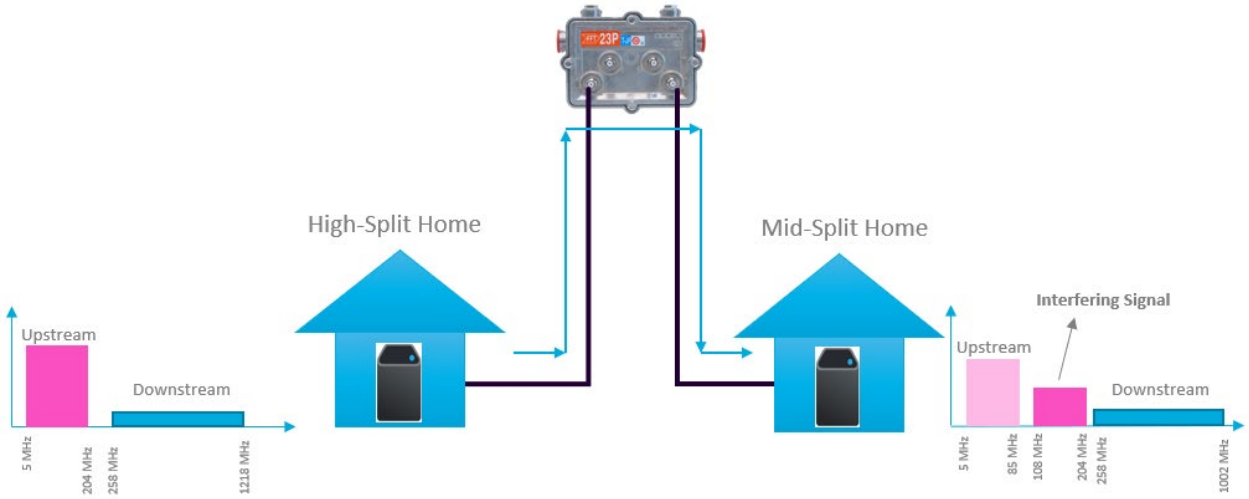


Figure 5 – Adjacent Home Interference – 204 MHz

A few modifications have been made to the methodology used in Dr. Prodan's paper to better outline the level of risk when an operator moves to high-split and ultra-high-split.

Prior to the analysis, the following should be noted:

- Based on Table 9 (below), the CM can transmit with 3 dB higher power per measured BW.
  - This has been determined by averaging the transmit levels from Table 9 from 85 MHz to 204 MHz.
- Tx levels close to 85 MHz have been gathered from the mid-split capable modems in the field.
- Rx levels close to 258 MHz have been gathered from the mid-split capable modems in the field.
- The high-split and mid-split devices are installed behind two-way splitters.

With that in mind, the following formulas can be used to evaluate the level of risk:

$$P_{loss} = (tap\ port\ to\ port\ isolation) + 2 * (drop\ cable\ attenuation + in\ home\ splitter\ loss)$$

$$Adjacent\ Channel\ Interferene\ (ACI) = CM\ Tx\ Level + 3(dB) - P_{loss}$$

$$Carrier\ to\ Adjacent\ Channel\ Interference\ Ratio\ (CACIR) = CM\ Rx\ Level - ACI$$

Along with the assumptions above, let us also assume that the tap port-to-port isolations can be any of the following:

- 20 dB
- 25 dB
- 30 dB

Table 9 outlines the assumed loss values for this analysis.

Table 9 – Averaged Loss Values

Loss Value Table	Drop (100' RG6)	Splitter	Port-to-Port Iso
	4 dB (Averaged from 108 – 204MHz)	3.5 dB (2-way splitter insertion loss)	20, 25, 30 dB

Tap port-to-port isolations are dependent on the frequency in which they are measured, along with the internal splitting formation of the tap ports. The values in the table above are assumed to be averaged across the spectrum.

Based on these numbers, the following figure was produced by sorting the calculated carrier to adjacent channel interference ratio (CACIR) value for each modem in an ascending order:

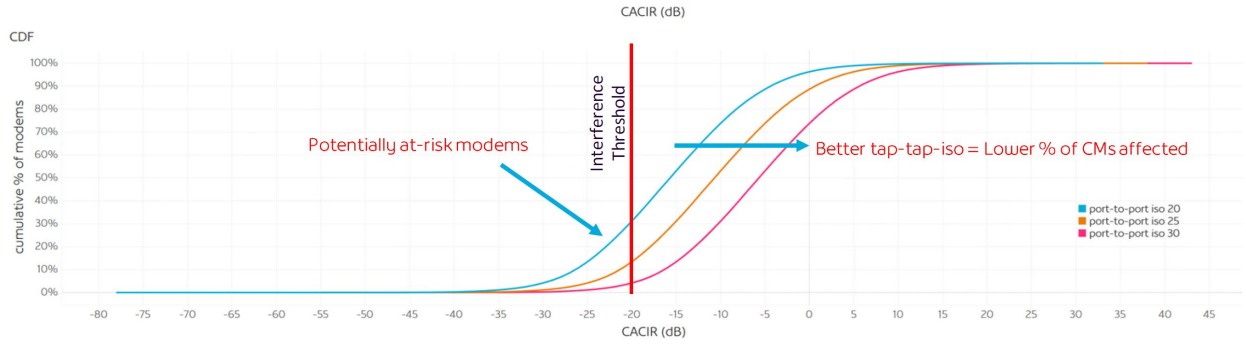


Figure 6 – CACIR Cumulitive Density Function

As seen in the figure above, the tap-to-tap isolation values of 20-, 25- or 30-dB result in vastly different levels of risks for interference in the field.

The CACIR threshold of -20 dB is set as a baseline for this study. Various devices in the field, including CMs with different front ends, can change this threshold and by extension the level of risk.

4.5. Ultra High-split Analysis

To determine shortfalls in the upstream with regards to the modem transmit power, it is assumed that all of the modems in the analyzed plant model are transmitting with their maximum capability (see Figure 1). Note that the focus of this analysis is on ultra-high-split TCS in the DOCSIS 4.0 CM. Refer to Section 4.2 for further information.

Figure 7 on Receive Power /6.4 MHz at the node and amplifier ports was created based on the required levels in Table 6, along with CM Tx capabilities and the assumed plant model, in Sections 2.2 and 3 respectively.

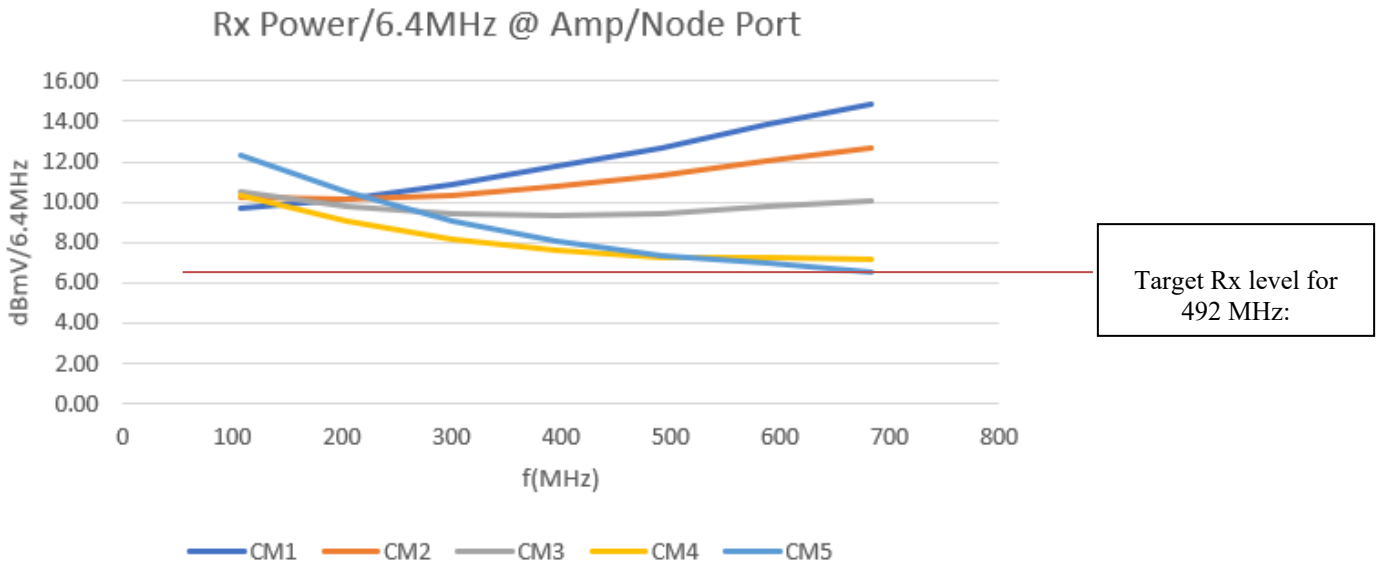


Figure 7 – Receive Power /6.4 MHz at Node and Amplifier Port

It can be observed that the modems installed at the end of the drop from Taps 1-3 have plenty of headroom in comparison to the target levels assumed at the port. It can also be seen that 'lower' receive levels at the amplifier port are only a concern for higher frequencies in modems that are in lower value taps (farther away from the node and amplifiers).

Figure 7 shows that frequency allocations (frequency stacking) appear to be an appealing approach for CMs. Essentially, although the DOCSIS 4.0 CMs will bond to all the Orthogonal Frequency Division Multiple Access (OFDMA) channels available in the spectrum, the CMs that are closer to the node/amplifier can use higher frequencies more efficiently given the 'headroom' available to them. Consequently, CMs that are farther away from the node and amplifiers can use lower portions of the spectrum.

As data usage patterns today outline, CMs will rarely use all the channels that they have bonded to. They will only do so in instances when large files are being uploaded or a speed test is being performed. For a majority of the usage cases, frequency stacking seems to be an appealing option for MSOs to optimize their return path, assuming the CMs do not do this automatically.

4.5.1. Splitter Added to the Plant Model

In this section the effect of additional flat losses such as splitters in the mainline will be discussed. When designing the plant, all additional insertion losses must be taken into account. This means that if a two-way splitter was to be added to the plant model, the 5 dB additional insertion loss must be deducted from the overall span loss, equaling to 30 dB of overall loss. As a result, the total span length of the original plant model could be reduced by roughly 200 feet. Figure 8 demonstrates this concept.

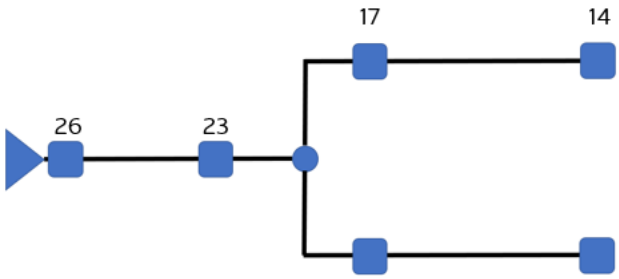


Figure 8 – Plant Model with Added Two-Way Splitter

Based on the new plant model, the receive levels at the port can be revised, as illustrated in Figure 9.

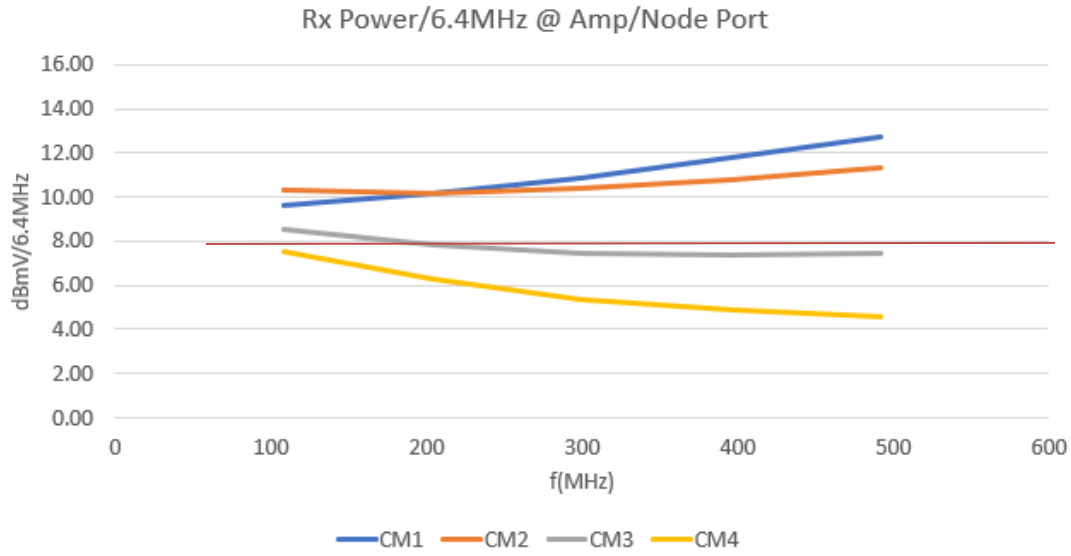


Figure 9 – Receive Power /6.4 MHz at Node and Amplifier Port

It can be observed that the upstream signals from CMs behind the mainline two-way splitter are received with less power than the target level of 8 dBmV/6.4 MHz for 492 MHz of return bandwidth (BW), based on Table 7. This may seem concerning at first but knowing that many MSOs will not be utilizing the entire 684 MHz of return BW, the CM can theoretically allocate the unused TCP from the higher portion of the spectrum to lower parts. For the purpose of this example, an assumption has been made that a maximum return BW of 492 MHz has been planned as a part of DOCSIS 4.0 upgrades. With most of the TCP concentrated at the higher frequencies, the CM can theoretically 'raise' the original transmit power by roughly 4 dB. However, in order to avoid concerns with spurious emissions, this paper assumes that the modem has 3 dB of additional power per channel, shown in the figure below.

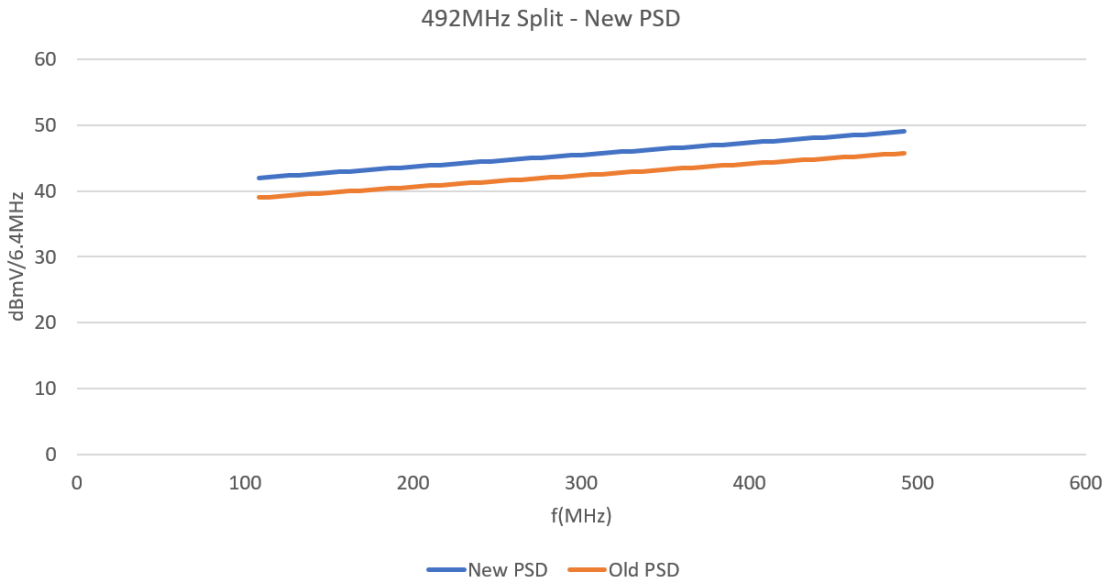
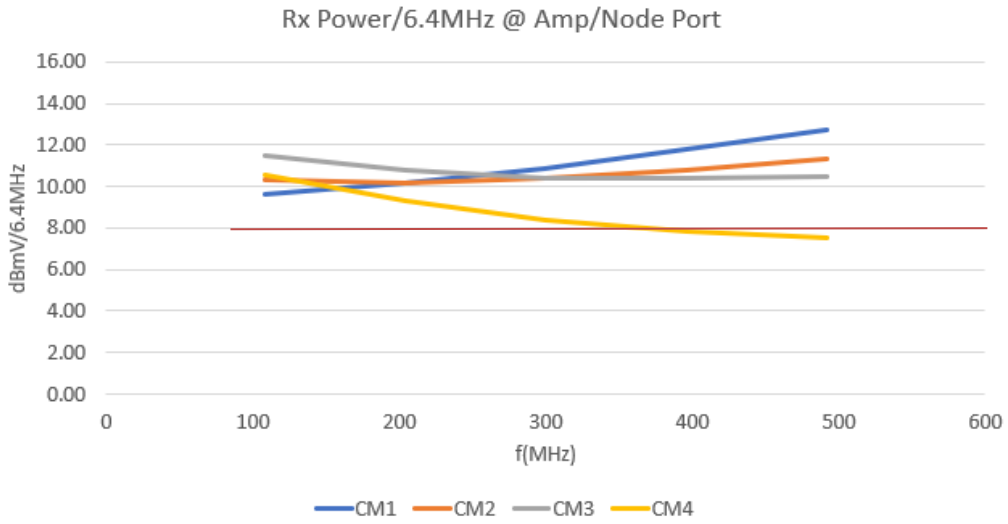


Figure 10 – CM Tx Level with 3 dB Boost for 492 MHz Split

Applying the new CM PSD to the same plant model, the following receive levels can be expected at the port of the node and amplifiers:



Target Rx level for 492 MHz:

Figure 11 – Receive Power /6.4 MHz at Node and Amplifier Port

The raised CM PSD and increased power levels from the CM can increase the risk for ACI. Close examination of this option in various plant models and scenarios prior to deployment is encouraged.

It should be noted that receiving below the rated level for each upstream split (outlined in Table 7) will not always result in a noticeable MER degradation. CCN and MER are highly dependent on the source MER and the NF of each amplifier in the return path, along with the cumulative noise and distortion products from each amplifier. This will be explored more in Section 5 of this paper.

4.5.2. Flat Losses – The Limiting Factor

One of the most significant results from the studies done in Section 4.5, and particularly in Section 4.5.1 when a two-way splitter was added to the plant model, is that high flat losses appear to be the most limiting factor in the upstream.

The plant model analyzed was 35 dB of span loss at 1 GHz, which can be considered quite 'stretched' for today's OSP span losses. It was observed that CMs in the plant model at the end of 150 feet of RG6 cable should be able to make it back to the amplifier and node ports, approximately within the range of 396 or 492 MHz, which are being considered by many MSOs. In other words, coaxial loss is something that has been taken into consideration in the CM design with the ultra-high-split TCS.

Flat losses can be somewhat challenging to overcome, especially in the upstream. Flat losses, as the name suggests, affect the entire spectrum in the same way, meaning that it cannot be overcome by tilt. It should be noted that this analysis was undertaken with only a two-way splitter. There are other types of OSP equipment currently deployed by MSOs that have much higher amounts of flat loss across the spectrum, namely, couplers and multi-dwelling-unit (MDU) style indoor splitters. The additional flat loss can cause concerns due to the modem having to transmit at close to maximum across the entire spectrum for the carriers to be received at the target level. This can potentially cause the CM to go into partial service mode due to insufficient Tx power. The overall upstream performance will be discussed further in Section 5.

5. Signal Quality and Noise Funneling

In order to quantify performance in the network both signal quality and power must be taken into consideration. Thus far, this paper has analyzed signal power in the plant models created for DOCSIS 4.0 networks. We discussed how the CM is able to overcome coaxial loss due to its increased TCP, output power level and tilt in the upstream. However, that does not indicate the signal quality that can be expected in the upstream. In this section, we analyze the possible upstream MER and outline limiting factors.

SNR and MER modeling in the upstream can be quite challenging due to the funneling effect. To define it broadly, noise funneling is the summation of all the unwanted noise and distortions in the return path. There can be many different sources of noise funneling, such as impulse noise, ingress noise and common path distortion (CPD). Generally, it is accepted that if the plant is free of physical impairments such as cuts in cable or unterminated taps, the funneling effects from the sources mentioned above can be minimized, if not resolved. For this paper, we focus instead on the amplifiers in the field and how they contribute to thermal noise (CTN) and distortion (CIN) accumulations in the upstream.

DOCSIS 4.0 technology is designed to be deployed in a cascaded environment. Although the number of amplifiers in cascade (series) is one of the most important factors in the downstream, for upstream, as it currently stands today, it is the total number of amplifiers. As an example, an N+6 plant with no splitters can have up to 24 amplifiers, assuming four outputs from the node. This number can increase dramatically with the addition of splitters in the topology.

In reality MSOs will deploy different amplifier types for their gain capabilities and the number of output ports, such as multi-port amplifiers and line extenders. However, this will not have a significant impact on overall performance, as will be demonstrated further in this section. In this paper, all amplifiers are assumed to be identical with the same NF mentioned in Section 2.5.

In order to quantify performance and determine the most important factor, the following has been assumed:

- Amplifier NF: 6 dB
- Amplifier CIN: 56 dB
- Input level to each amplifier in the return path: 6 dB flat across the spectrum
- Number of ports utilized in the node: 4
- CTN: All amplifiers on either the entire node or each leg that would contribute to signal degradation
- CIN: Only the amplifiers in series on each leg of the node that would contribute to signal degradation

Following the assumptions above, we need to consider the number of amplifiers that exist both in series and in total. On the one hand, the total number of amplifiers is what is considered for the cumulative effect of thermal noise funneling from each individual amplifier in the node, assuming that each leg of the node cannot be isolated and they all funnel into a single port in the return. The amplifiers in series, on the other hand, would be contributing to total CIN in each leg of the node along with the total CTN, assuming that port of the node has been isolated. Utilizing the formulas outline in Sections 2.8, 2.9 and 2.10, Tables 10 and 11 can be calculated.

Table 10 – Upstream Peformance – All Ports Funneled

Total Number of Amplifiers	CTN	CIN	Source MER (38 dB)	Source MER (40 dB)	Source MER (47 dB)
16	45.36	49.98	37.04	38.56	42.28
32	42.35	48.22	36.35	37.61	40.3
44	40.97	47.55	35.92	37.04	39.3
56	39.92	46.97	35.52	36.54	38.48
68	39.07	46.00	35.12	36.04	37.73

Table 11 – Upstream Performance – Single Amplifier Leg

Total Amps in Series	CTN	CIN	Source MER (38 dB)	Source MER (40 dB)	Source MER (47 dB)
4	51.38	49.98	37.55	39.31	44.29
6	48.37	48.22	37.26	38.87	43.05
7	46.99	47.55	37.08	38.61	42.4
8	45.94	46.97	36.9	38.37	41.84
10	45.1	46.00	36.68	38.07	41.19

To better compare the results, the MER values for each table have been plotted in Figures 12 and 13.

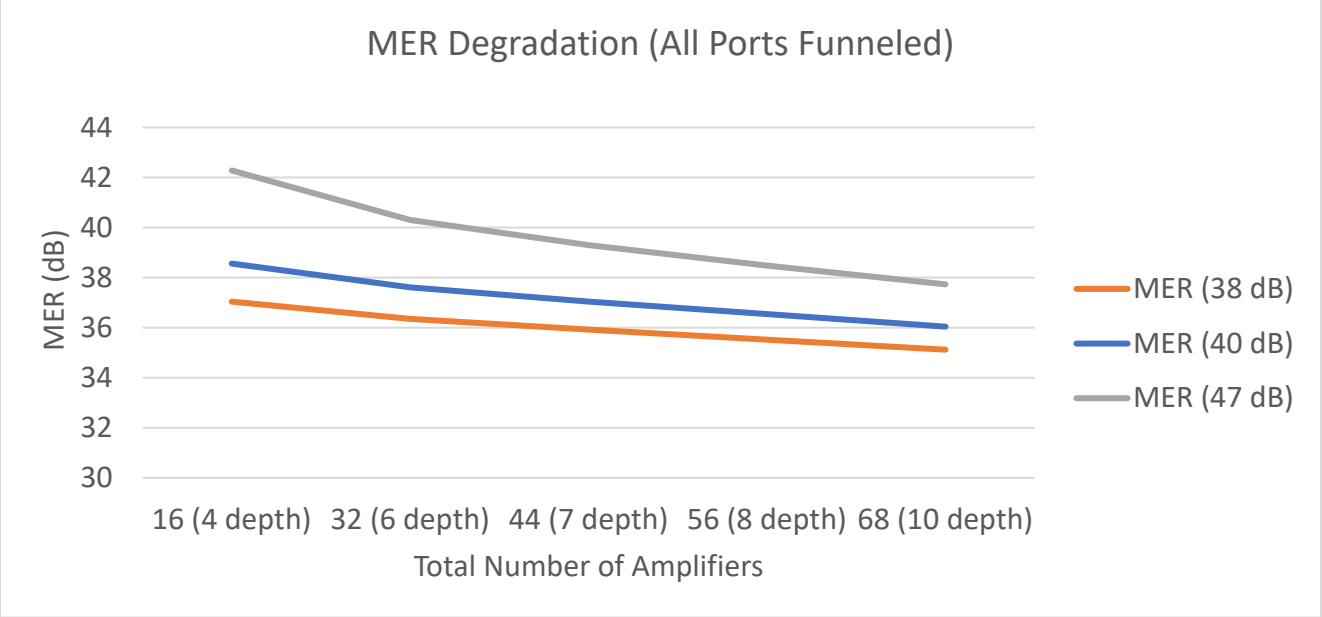


Figure 12 – MER vs. Total Number of Amplifiers – All Ports Funneled

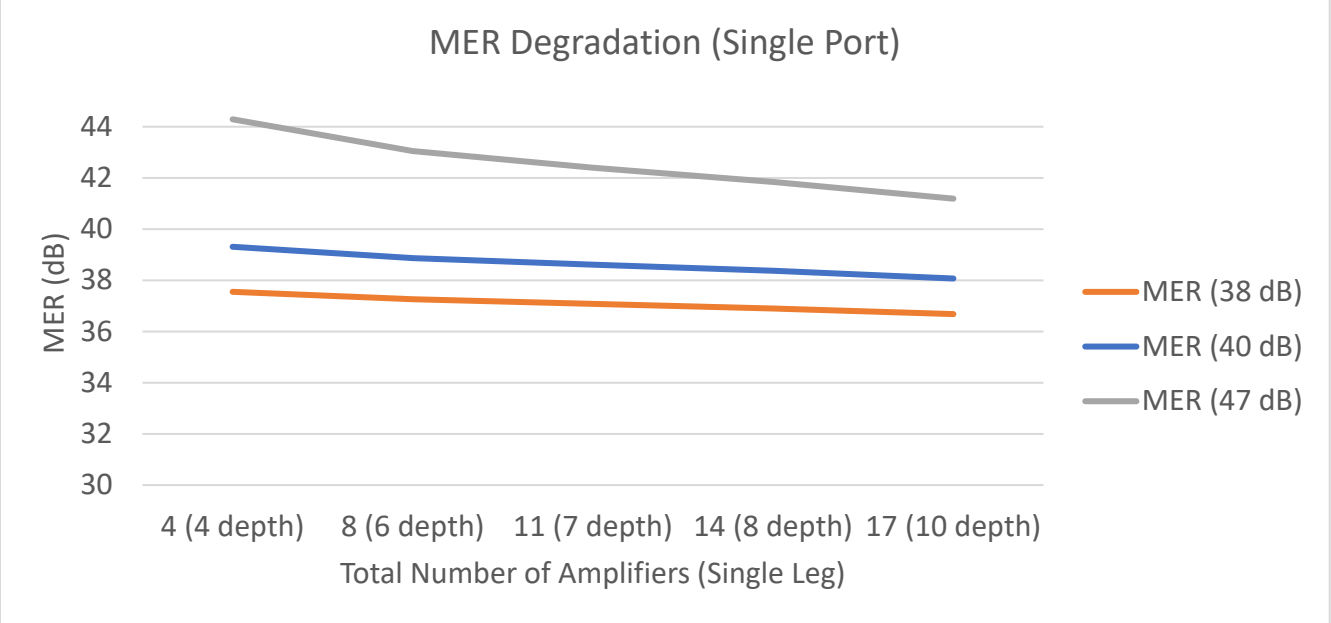


Figure 13 – MER vs. Total Number of Amplifiers – Single Amplifier Leg

When considering Tables 11 and 12, as well as Figures 12 and 13, a few interesting observations can be made. First and foremost, it can be observed that higher starting MERs such as 47 dB are more subject to degradation when exposed to noise and distortions. Figure 13 demonstrates that even with the addition of four amplifiers a ~3 dB MER reduction is realized. This reduction is then somewhat 'flattened' with further additions to noise and distortion products.

The most interesting observation from the figures above is that the 'starting MER' is arguably the most important factor in achieving higher orders of modulation. As an example, let's focus on a source MER of

40 dB as per Figure 12. There is only ~2.5 dB of additional MER reduction when the total number of amplifiers is increased from 16 to 68. Expanding on that further, focusing on the 38 dB MER graph in Figure 12, it can be seen that by reducing the number of amplifiers from 56 to 32 by performing a node split, the MER is increased by roughly1 dB. Alternatively, if the starting MER was increased by 2 dB to 40 dB, the same results can be achieved. Conversely, the total number of amplifiers or reduction of the cascade depth does not yield substantial MER increases in the upstream.

Lastly, it can be observed that isolating each leg of the node will yield more noticeable increases in MER for higher MER values. This is also true for reducing the total number of amplifiers.

It is very important to note that the results above do not tell the whole story of why it is important to be able to isolate each leg of the node from one another. Although this will have a direct positive impact on upstream performance, it also isolates noise from other sources (ingress noise) to one leg of the node. Additionally, node splits have many other benefits such as reducing the number of customers sharing the same data pipe and pushing fibre deeper into the HFC networks. These benefits are extremely important but have not been quantified in this paper.

6. Conclusion

DOCSIS 4.0 technology has been developed to enable HFC networks to provide multi-gigabit services. In this paper, a reasonable worst-case scenario plant was analyzed to estimate the capability of current HFC networks, with no amplifier re-spacing. It should be noted that the 35 dB span loss model was considered a ‘stretched’ plant by many MSOs during the development of DOCSIS 4.0 sets of specifications and plant models.

This study observed that achieving higher orders of modulation such as 1024 QAM is possible for the majority of cases. The only areas of concern for legacy plant design are areas where high flat losses are incurred in the upstream, namely due to splitters and couplers. This is only an issue in ‘stretched’ plant areas where there is already a higher amount of insertion loss from the modems installed at the end of drops and the end of line taps. How a modem can overcome the higher insertion loss with higher transmit powers was also discussed, assuming the operator does not utilize the entire 684 MHz band capability of the CM in the upstream. This should be balanced in conjunction with neighbour interference, since higher transmit levels from the CM can lead to additional neighbour interference cases between DOCSIS 4.0 devices and ‘legacy’ devices. MSOs will have to balance many moving parts, especially CM transmit powers, in order to optimize performance in the upstream to achieve higher source MER and high carrier-to-ingress noise ratio.

MSOs should optimize the TCP and power available in the modem to ensure sufficiently high transmit powers for higher transmit MER. This can assist with achieving higher orders of modulation in the upstream, along with a higher carrier to interference noise ratio. We also noted that funneling and the total number of amplifiers play a part in the overall signal quality in the upstream. Further, isolating each leg of the node from one another in the upstream results in a better overall signal quality. It can also help with isolating ingress noise to a particular leg, rather than it funneling to all ports in the return path. Finally, the source MER from the CM is one of the most—if not *the* most—important factor in the overall signal quality in the upstream. This can help MSOs prioritize efforts for increasing upstream capacity in the most efficient manner.

Abbreviations

ACI	adjacent channel interference
CACIR	carrier to adjacent channel interference ratio
CCN	carrier to composite noise ratio
C/N	carrier to noise ratio
CIN	carrier to intermodulation noise
CINR	carrier to interface noise ratio
CM	cable modem
CMTS	cable modem termination system
CNR	carrier to noise ratio
CPD	common path distortion
CSO	composite second order distortion
CTB	composite triple beat distortion
CTN	carrier to thermal noise
DAA	distributed access architecture
dB	decibels
dBmV	decibels relative to one millivolt
DAA	distributed access architecture
DRW	dynamic range window
DOCSIS	data over cable service interface specification
DS	downstream
ESD	extended spectrum DOCSIS
GHz	gigahertz
HFC	hybrid fibre-coax
ISBE	International Society of Broadband Experts
MER	modulation error ratio
MHz	megahertz
MSO	multiple service operator
NF	noise figure
NPR	noise power ratio
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
OSP	outside plant
PoE	point of entry
PSD	power spectral density
QAM	quadrature amplitude modulation
RF	radio frequency
Rx	receive
SCTE	Society of Cable Telecommunications Engineers
SNR	signal to noise ratio
STB	set top box
TCS	transmit channel set
TCP	total composite power
Tx	transmit
UGHB	under grant hold bandwidth
US	upstream

Bibliography & References

[1] Data-Over-Cable Service Interface Specification DOCSIS 4.0 – *Physical Layer Specification CM-SP-PHYv4.0*

[2] Broadband Cable Access Networks – The HFC Plant, David Lafarge and James Farmer

[3] Optimizing the 10G Transition to Full-Duplex DOCSIS® 4.0, Richard S Prodan

[4] Understanding Real-World MER Measurements. Ron Hranc and Bruce Currivan

Private Wireless Networks And Multi-Access Edge Compute

A Technical Paper prepared for SCTE by

Muhammad J Khan
Principal Engineer, Wireless R&D
Charter Communications
6360 S Fiddlers Green Circle, Greenwood Village, CO 80111
+1 (720) 536-1578
Muhammad.J.Khan@charter.com

Mohamed Daoud
Principal Engineer, Wireless R&D
Charter Communications
6360 S Fiddlers Green Circle, Greenwood Village, CO 80111
+1 720-699-5077
Mohamed.Daoud@charter.com

Joerg Ahrweiler
Director, Wireless R&D
Charter Communications
6360 S Fiddlers Green Circle, Greenwood Village, CO 80111
+1 720-699-3580
Joerg.Ahrweiler@charter.com

Hossam Hmimy
Sr. Director, Wireless R&D
Charter Communications
6360 S Fiddlers Green Circle, Greenwood Village, CO 80111
+1 720-536-9396
Hossam.Hmimy@charter.com

1. Introduction

Enterprises are looking for dedicated wireless networks that offer better control, which can be used to solve operational challenges. Private LTE and 5G networks offer reliability, ubiquitous coverage, high user capacity, and built-in security. Private networks combined with multi-access edge computing (MEC) also give the ability for enterprises to keep data on premises and implement use cases that require low latency.

Over the last year, Charter Communications has continued to work with CBRS spectrum in the private networks space. Charter has evaluated various vendor solutions for private LTE/5G networks. In addition, we have also deployed a private LTE network in a large manufacturing plant in Michigan and have plans to work with customers in other verticals. In this paper, we explain the main motivation behind deploying a private LTE/5G network and how it can be implemented. Second, we present different use cases that can be served by this network in the industrial, healthcare, and education verticals. Third, we show the criteria for vendor selection in our lab evaluation. Fourth, we illustrate the deployment of our first private LTE network in a factory.

1.1. Citizens Broadband Radio Services (CBRS) overview

In April 2015, The Federal Communications Commission (FCC) made the decision to establish Citizens Broadband Radio Services spectrum in the United States. This is a section of 150 MHz spectrum between 3.55 GHz and 3.70 GHz, which was originally used by the government for the military or by wireless internet service providers (WISPS). This decision by the FCC opened up CBRS spectrum to be shared on a dynamic basis depending on the user priority [1].

One of the key components of CBRS is the tiered sharing structure. This is achieved by establishing three different tiers of users, each allowed to use the band at any given time. Below is a brief description of some of the key components of this sharing structure in CBRS.

Spectrum Access System (SAS): The spectrum access system was set up to coordinate between the different users using CBRS spectrum. All Citizens Broadband Service Devices (CBSD) must register with the SAS before being allowed to transmit.

Tiers of users

- Incumbents – These are federal government, fixed satellite users and any grandfathered wireless internet service providers. They are granted interference protection from the other tiers below (PAL and GAA)
- Priority Access License (PAL) – These are licensed users who have access to 70 MHz of spectrum and up to a maximum of 40 MHz per licensee. They are protected from interference from the lower tiers but must cease transmitting to protect the incumbents if determined by SAS.
- General authorized access (GAA): These users can use the entire 150 MHz of spectrum as long as there is no one else using the spectrum within the area. They do not receive interference protection from any of the other tiers of users.

CBSD: CBSDs are defined as fixed radiating antennas and could be small cells, remote radio heads, DAS systems or a combination of all. There are two different kinds of CBSDs:

Table 1 - CBSD categories

CBSD category	EIRP limits	Comments
Category A CBSD	30 dBm / 10 MHz	Limited to indoor deployments. Outdoor deployment maximum height is 6 meters height above average terrain (HAAT)
Category B CBSD	47 dBm / 10 MHz	Limited to outdoor deployments and must be installed by CPIs

Certified professional installers (CPI): All CBSDs must be registered with the SAS by a certified professional installer. CPIs are authorized to give detailed information about each CBSD to the SAS before it is allowed to transmit.

End user device (EUD): End user devices are generally user devices like LTE or 5G capable smart phones, tablets, or routers, which communicate via CBSDs to the network. They are limited to a maximum of 23 dBm EIRP.

1.2. Motivation and benefits of private LTE/5G networks

Current connectivity options for enterprises are limited to either public cellular networks or enterprise Wi-Fi. Commercial networks from mobile network operators (MNOs) are designed for public use and do not cater to the specific use cases for enterprises. The advantages described below are a key reason for enterprises to deploy their own private LTE/5G networks.

One of the key benefits of private LTE networks is its ability to cover larger areas as compared to Wi-Fi. Private LTE/5G networks can provide superior coverage in challenging environments like factories with a lot of metallic structures, or in outdoor campus-style environments like universities. They can be used to complement existing Wi-Fi so that end users can benefit from the best of both networks.

Mobility is another use case where private LTE/5G networks can offer a lot of benefit. Seamless handovers and access point neighbor relations are built into LTE/5G networks. This mobility is ideal for use cases like asset tracking, autonomous vehicles, remote guided vehicles, and critical communications.

Private networks also offer a level of customization for the specific use cases for enterprises. Overall bandwidth in a LTE or 5G network can be customized depending on the requirements. The uplink and downlink ratios in the channel can be modified using different TDD frame configurations. For example, a frame configuration of one makes more sense in uplink heavy applications like video streaming while a frame configuration of two can be used for downlink heavy uses. The bandwidth can further be modified by combining channels together using features like carrier aggregation. Frequency re-use can further be used to control interference between different sectors in a venue. Customization can further be achieved by varying Quality of Service (QoS) parameters and assigning them to different groups of end user devices. For example, devices that require high bandwidth can be assigned a QoS that allows for higher throughput and guaranteed bit rates.

Private LTE/5G networks can also be leveraged to bring the benefits of MEC environments. MEC, as defined by ETSI (European Telecommunications Standards Institute,), “offers application developers and content providers cloud-computing capabilities and an IT service environment at the edge of the network” [4]. Thus, this allows for high bandwidth and low latency applications like video analytics, augmented reality (AR), and other use cases, which benefit from edge processing.

2. Private networks architecture

2.1. Components of a CBRS private wireless network

A private wireless network consists of many components which are either managed by the operator, a third-party system integrator or the end user. A CBRS specific private wireless network consists of the following key components:

A. Wireless spectrum
A private wireless network can be deployed on licensed, unlicensed, or shared spectrum like CBRS. This is granted by a governing body such as the FCC in the United States.

B. Spectrum Access System
The FCC has authorized several spectrum access systems for CBRS. These are deployed in the cloud and CBSDs or the domain proxy must have continuous access to this for regular heartbeats.

C. Radio Access Network (RAN)
Deployed on premises for a private LTE network, the RAN consists of the radios, antennas and baseband functions.

D. 4G or 5G core network
The EPC or 5G core can be located on premises, partially on-premises or entirely on the cloud. This is discussed in more detail in the architecture section of this paper.

E. Multi Access Edge Compute (MEC)
Located on premises to handle data analytics, caching, and other processing functions that would otherwise be deployed in the cloud. This allows for a host of low latency and high bandwidth applications.

F. User devices
The end user devices can be smart phones, tablets, cameras, LTE/5G routers, sensors, etc. and can vary depending on the use case.

G. Management systems
These consist of the portal(s) that help manage the RAN, network core, devices, and orchestration.

2.2. Architecture

The architecture for private networks depends on the deployment venue, use case, and technology (LTE or 5G). Presented below are some common deployments for private LTE, with their respective pros, cons, and usage.

2.2.1. Cloud-based core network

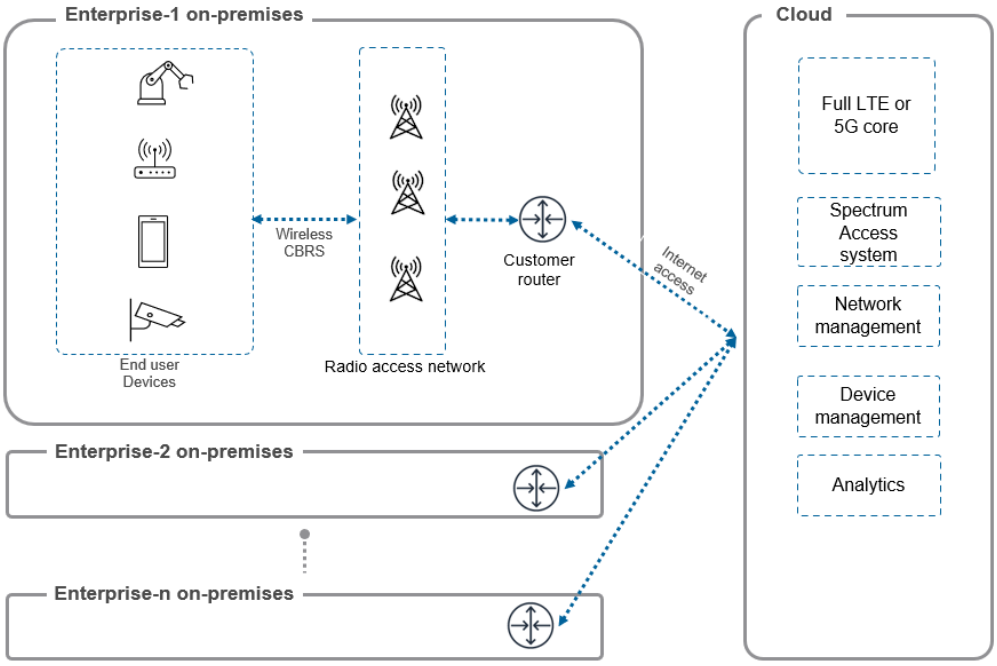


Figure 1 - Network core hosted on cloud

In this deployment, all the network core functions are hosted on the cloud with a dedicated radio access network on premises. This implementation is easy to deploy since there is little infrastructure needed on premises to host the core functions. Further, it can support multiple geographically distant enterprises if they have connectivity to the core.

However, this implementation does not support applications which require low latency since the data has to traverse multiple hops over the internet and back. It is also not efficient for higher bandwidth applications like high-definition video processing or real time streaming. For cameras and video, the backhaul requirements scale very quickly since all processing is done on the cloud. The enterprise still gains the benefit of having dedicated coverage in their venue, which is better than using a public commercial network.

2.2.2. Full network on premises

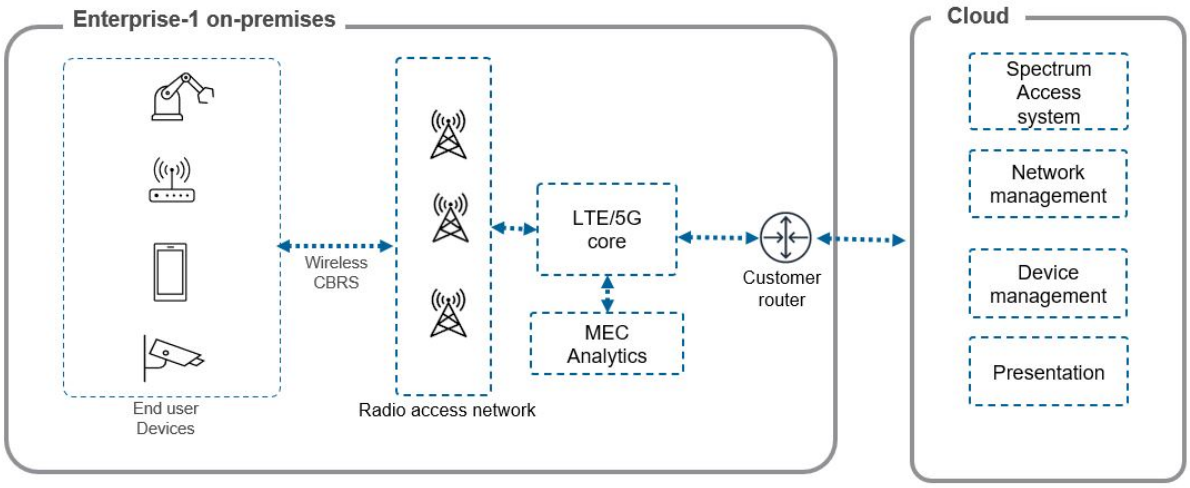


Figure 2 - Full network on premises

In this implementation, all the network core functions are located on premises. This is ideal for low latency applications or applications that require real time processing and high bandwidth like video streaming. For example, this would support a worker safety use case where life feeds from cameras are sent to the edge for further processing and analytics on the edge. While this offers advantages to the enterprise with a dedicated core on premise, it does not scale that well for an operator since separate infrastructure is required on premise for every enterprise.

The management functions can be hosted in the cloud for network operations and orchestration. As in all deployments, the SAS is also hosted separately in the cloud.

2.2.3. Network core at edge data center

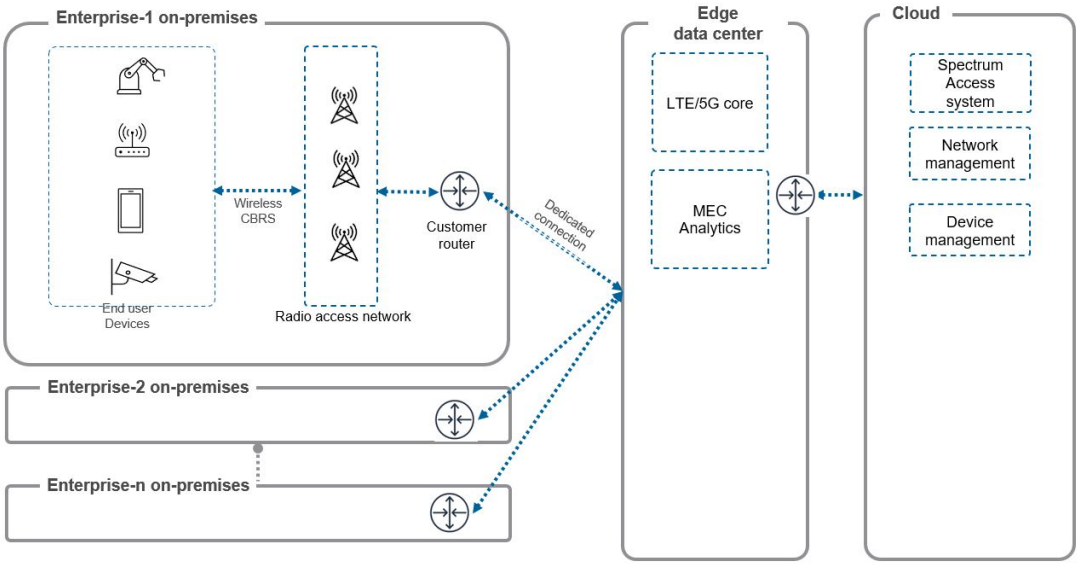


Figure 3 - Network core at edge data center

An edge data center is defined here as a data center with a dedicated connection to the enterprise. This can be owned by either the operator or managed by a cloud service provider as part of their edge offering. This implementation offers a balance between the full cloud and on-premises deployments, while still offering the ability to serve lower latency and high bandwidth applications.

2.2.4. Hybrid architecture

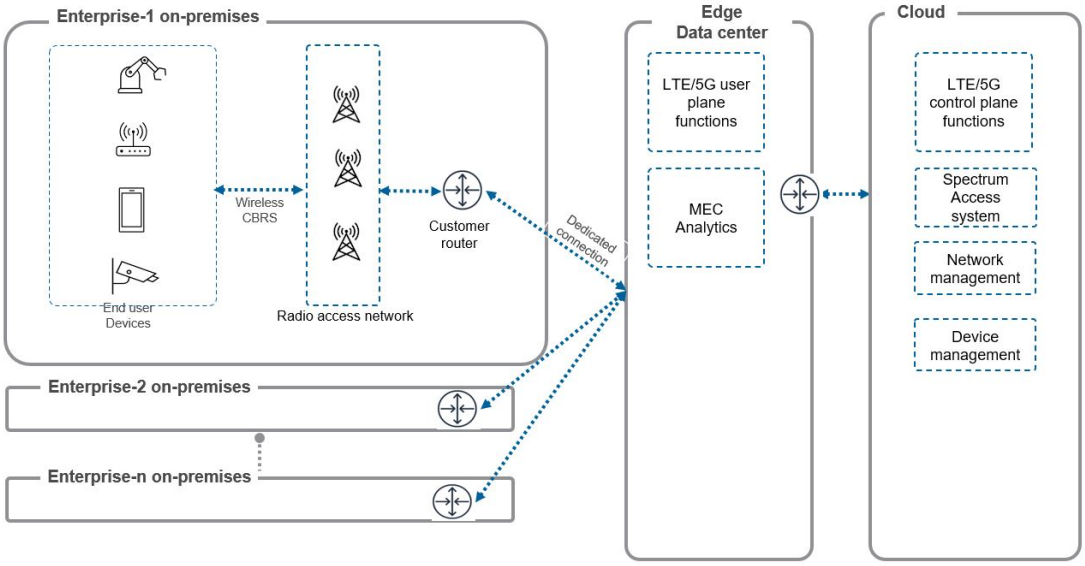


Figure 4 - Hybrid deployment

In this implementation, the use plane functions are located at the edge data center with the rest of the functions hosted in the cloud. For 5G, this means the user plane function (UPF) is hosted at the edge, while the access management functions (AMF), session management function (SMF), policy control function(PCF), network slice selection function(NSSF), authentication server function(AUSF) and unified data management(UDM) functions are hosted in the cloud. This architecture allows multiple enterprises within proximity of the edge data center to be served, while still leveraging the cloud for management.

3. Use cases in different industry verticals

As discussed earlier, private LTE/5G allows for a variety of different applications. Presented below are some of them split by vertical, illustrating how a private LTE network enables them.

3.1. Industrial

a. Worker safety

Safe and accident-free environments are a priority for plant managers in a factory. This can be done is two ways – by enforcing safety standards and then tracking compliance of those standards. Charter, along with its partners, has demonstrated a video analytics worker safety use case. A camera was used to capture video of workers operating a machine in the factory, and sent to the MEC on premises. An algorithm identifies whether the workers are wearing safety gloves, helmets, jackets, and other equipment required for safe operation of the machines. The data is anonymized and recorded for compliance purposes, and automatic notification is sent out if a particular worker is not wearing safety equipment.

b. Autonomous guided vehicles (AGV) and remote-controlled robots

For large warehouses or factories, there is a constant need to move large items from one area to another. Boxes, finished products, and other items must be loaded onto pallets and moved manually by employees dedicated for this task. This uses valuable employee time and resources. Remote controlled robots or AGVs can be used to transport materials within and around the factory to save time. A private LTE network provides this reliable coverage, mobility, and edge processing to realize this use case

c. Push to talk and push to video

In large industrial environments, cellular service is poor because of high penetration loss due to many large metallic structures. In the absence of good public cellular service, private LTE/5G networks can fill the gap. Charter has demonstrated this application to its pilot customer. The push-to-talk server is hosted on premises along with the network core, and factory employees can install a push to talk application on their existing smart phones if it supports CBRS.

d. Predictive maintenance

Factories and warehouses can collect valuable data from their machines with wireless connected sensors. Most mechanical systems have a vibration signature when operating correctly. Sensors can be used to detect any irregular vibrations, which would otherwise go unnoticed. The data is used to train algorithms, which can then predict an upcoming failure. The result is minimizing downtime and gaining insights into machine behavior [3]. Private LTE networks provide the connectivity and reliability to connect hundreds of sensors and process the data at the edge.

3.2. Healthcare

a. Secure communications

Healthcare workers in hospitals and other medical facilities require secure communication with high reliability. Additionally, they may require sensitive patient information to remain on premise. Indoor coverage in hospitals can vary depending on the building material and various medical equipment located indoors. A private LTE network can solve both these problems by providing reliable connectivity with a dedicated RAN, and edge compute to process the data on premises.

b. Augmented reality (AR) applications

Medical personnel equipped with AR headsets can have instructions and information overlaid on medical devices to understand how to properly use them. AR can also be used in a training setting for healthcare professionals to learn to use equipment.

3.3. Education

a. Remote learning

The worldwide pandemic has shown us the importance of having good connectivity at work, universities/schools, and at home. Private networks can be used to extend school connectivity to residential locations around it. Students can access school resources securely and safely from their homes.

b. Campus connectivity

Universities and other campus environments are ideal for private LTE networks since they provide seamless coverage both indoor and outdoor. Students can have access to local applications securely anywhere on the campus. The network can be further used to serve other

use cases like providing connectivity for security personnel, faculty communication and interactive learning via VR or AR headsets.

4. Charter Communications lab evaluation

Charter Communications has conducted an evaluation of private network offerings from several vendors to assess the best fit for its customers. The vendor selection depends on customer, type of use case, cost sensitivity and other factors. Our evaluation attempted to take a holistic look at the different criteria that are important for private networks for enterprises.

4.1. Selection criteria

4.1.1. Network

Private LTE and 5G networks can have varying deployment scenarios - indoor industrial, outdoor/indoor campus, hospitals, schools etc.). Selection of the radio access network determines which of these environments can be served efficiently. Some of the different types of RAN options are:

a. Indoor Pico or Femto radios varying in output power from less than 20 dBm to 30 dBm (maximum allowed power for indoor CBRS deployment)
b. Outdoor rated radios with integrated antennas
c. Outdoor rated radios with an option for external antennas
d. Remote radio head which connects to a central baseband

We chose to go with a vendor that has a wide variety of radio options to serve the different deployment scenarios of private LTE/5G. For example, in a factory which has existing Ethernet cabling, we would choose to go with the indoor small cells which backhaul via copper. However, in another scenario for outdoor campus coverage where the antennas are deployed on a tower, we would go with a RRH and external antenna connected via fiber to a central baseband.

To serve the varying bandwidth requirements within a private network, LTE provides the flexibility to modify different TDD frame configurations. For example, a TDD frame configuration of 2 allows for a 3:1 downlink to uplink ratio. Alternatively, a frame configuration of 0 allows for a 1:3 downlink to uplink ratio. The latter could be used in primarily uplink use cases like video security or worker safety based on video analytics. Another factor we looked at was the total aggregated bandwidth. This could range from 10 MHz to the entire 150 MHz with carrier aggregation. This has a big effect on throughput so was an important consideration in selection of RAN.

The total capacity of the system depends on the number of eNodeBs supported by the baseband and the total capacity of the EPC i.e. how many S1 links it supported. Our evaluation looked at how many of each link the network could support, and if there was room for scaling. Lastly, we looked at the network core configuration from the different vendors. For LTE, the evolved packet core (EPC) could either come in different configurations (small, medium, large) each with a different number of supported. Other vendors did not have an upper limit and scaled the core as the network grew.

4.1.2. Management systems

In a traditional network, there are different user interfaces to manage the different functions from RAN and core. There can be a network management system (NMS) for RAN, a separate command line interface (CLI) or graphical user interface (GUI) for the core, another portal for home subscriber server

(HSS) function, etc. For a private network, it is important to have a single pane to manage most functions. Table 2 describes some of these functions and the intended end user.

Table 2 - Functions of management portals and primary user

Function	Primary user
Adding and removing eNodeB	Operator
Modifying cell frequencies	Operator
Modifying / optimizing network parameters	Operator
Creating different QoS groups	Operator
Viewing network status and alarms	Operator / Enterprise
Adding / removing UEs and sims	Operator / Enterprise
Changing aggregate maximum bit rates	Operator / Enterprise
Creating IMSI groups	Operator / Enterprise

4.1.3. Ease of deployment and operations

Private wireless networks need to be easy to deploy and manage for enterprises. The complexity of LTE and 5G networks needs to be abstracted so that the enterprise can easily integrate the private LTE network into their existing infrastructure. In our evaluation we looked at several aspects of the ease of deployment

a. Ease of physical installation: use of existing IT infrastructure in venues is ideal because of obvious cost benefits. In an industrial environment, the wiring must be done either after hours or between shifts. Moreover, exact location of antennas depends on specific venue restrictions.
b. Ease of operations: after deployment, the private network must be managed remotely, with roles split between the operator and the enterprise customer. This is discussed in more detail in section 4.1.2
c. Support: Vendors need to have 24/7 support teams for any troubleshooting or escalations if needed, depending on the specific service level agreements. Our evaluation also considered this level of support from the network vendors.

4.2. Overview of results

Table 3 - Summary of network evaluation

Category	Vendor 1	Vendor 2	Vendor 3	Vendor 3
RAN – available options	a. 17 dBm Indoor radio b. Outdoor CAT-B radio with integrated antenna	a. 24 dBm small cell b. 24 dBm small cell with option for external antenna c. Outdoor radio with option for external antenna d. Outdoor radio with integrated omni antenna e. Baseband with RRH + external antenna	One radio with adjustable power for both indoor and outdoor. Option for external antenna	a. 24 dBm indoor radio b. 30 dBm indoor radio c. 30 dBm outdoor radio

TDD frame configuration options	1 & 2	0,1,2	0-7	1 & 2
Evolve packet core (EPC)	Full on-premises EPC with redundancy	Full on-premises EPC with redundancy	3rd Party EPC on premises with HSS on cloud	Cloud EPC only. Option for on-premises EPC with 3PP
Max allowable bandwidth (depends on SAS grant)	4 carriers 80 MHZ (40+40 MHZ)	3 carriers. 60 MHz (20+20+20)	8 carriers 150 MHz	2 carriers 40 MHz (20 + 20)
Max modulation	256 QAM	256 QAM	256 QAM	256 QAM
MIMO	2x2 or 4x4	2x2	2x2 or 4x4 (extra radio)	2x2
Maximum capacity	Maximum 24 radios	Scalable	Scalable	Depending on 3rd party EPC provider

5. Private LTE deployment in factory

Charter Communications has deployed a private LTE network in a 600,000 sq. km manufacturing plant in Michigan, USA. The following sections describe the planning, design, deployment, and lessons learned from this pilot.

5.1. Planning and design

Like with any network deployment, our deployment started with customer inputs and a discussion of how they would use the network i.e. primary use cases.

a. Size of the factory: The coverage area is important since exclusion zones need to be defined when doing the RF propagation study. In our scenario, the factory was a little over 600,000 sq. km. The main operations of the factory were on the first floor, while the IT and certain other offices were on the second floor.
b. Interior of the factory: We requested a map of all the equipment in the factory (machines, racks, pallets etc.). This is not always available, so site surveys are valuable when mapping out the venue and any design considerations that need to be made
c. Material properties: Industrial environments have a range of different items which affect the overall propagation. In our factory, the walls were concrete, most of the machines were metallic, there were large wooden structures, and 15–20-foot storage racks in many areas. In addition, there were large aluminum ventilation ducts and metal rails along the ceiling. These may be missing on the drawings, but they do affect coverage so should be accounted in the design as best as possible.
d. Acceptable locations to mount antennas: the factory had restrictions on where we could mount the radios and antennas. This was due to either some high equipment, fans on the ceilings, ventilation ducts, or existing antennas and access points from other technologies. We had to make a tradeoff between desired locations for coverage and these customer restrictions.

e. Networking requirements: IT personnel needed to specify what firewalls they have in place, what IP addressing scheme they have, whether they use a public or private domain name server (DNS), switch configuration, and anything else to integrate the private LTE network into their system.

Considering the customer requirements, we performed a RF design of the factory. Figure 5 shows the coverage, while figure six shows the best server plot.

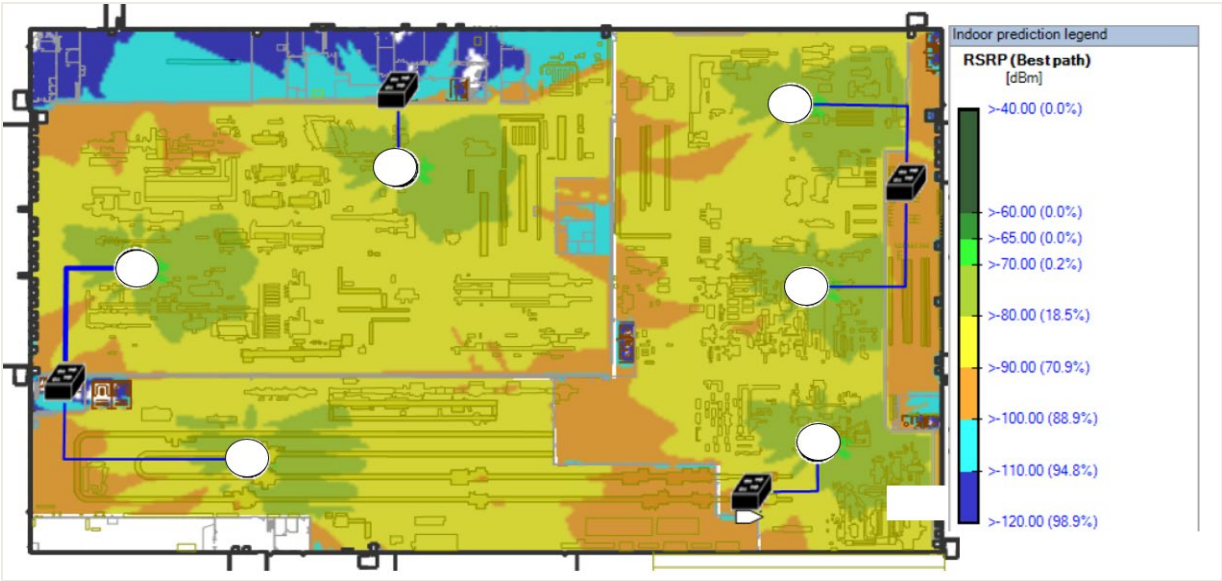


Figure 5 - RSRP plot from RF design

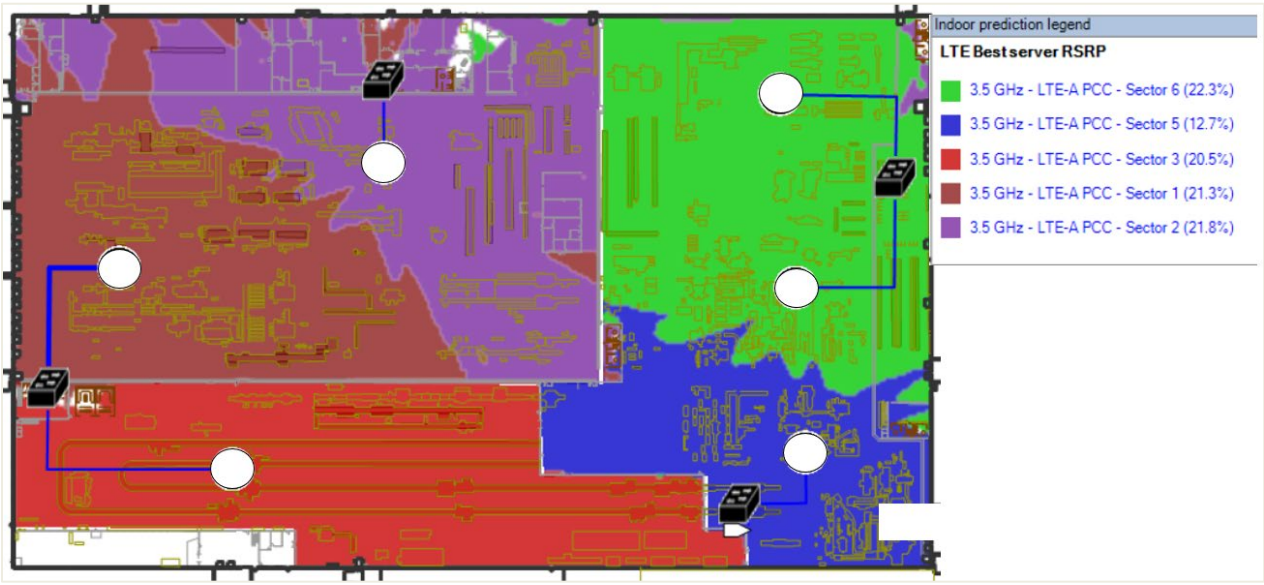


Figure 6 - Best server plot from RF design

The design shows 95% of the plant is covered by a signal level of -110 dBm or better. The intent of the design was to cover the entire factory, and we had to balance between desired coverage and interference between neighboring sectors. While the areas to the north of the factory (cafeteria and copy rooms) had

limited coverage, we can add additional radios there if required. We also implemented frequency re-use to limit inter-cell interference.

5.2. Deployment

The installation was done by the factory's approved installers who were familiar with the specific limitations and approvals needed to install. Custom mounts were created to mount the radios and antennas, which were then hung from vertical rails on the factory ceiling. The EPC and MEC servers were in the factory main distribution frame (MDF) room. Figure 7 shows the deployment architecture.

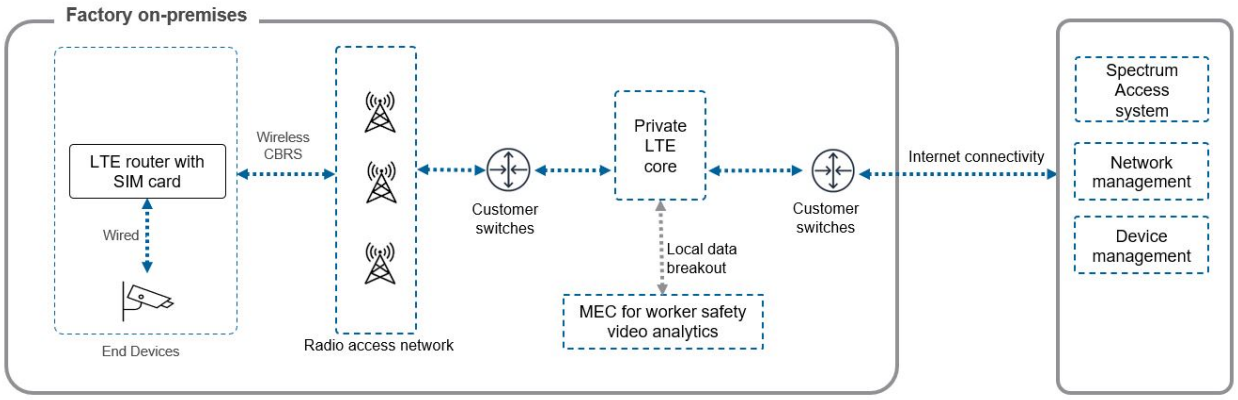


Figure 7 - Deployment architecture

5.3. Post launch testing

After the network was live, Charter conducted performance testing at pre-determined test points to assess coverage, quality, and throughput across the factory. The table below has the results of the testing, and comments where the results deviated from the expected performance.

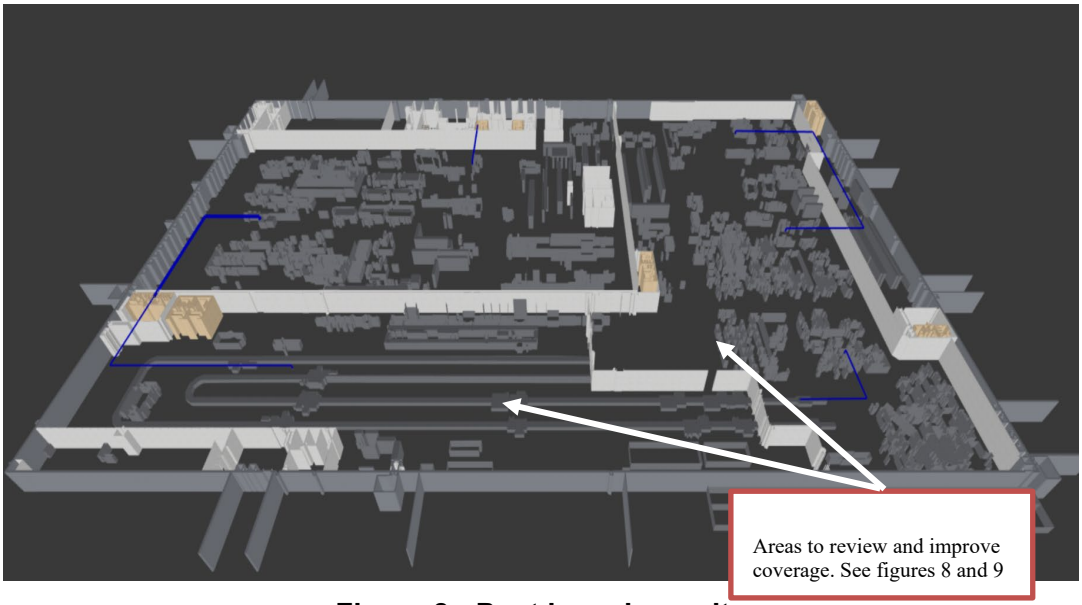


Figure 8 - Post launch results

Table 4 - Performance measurements from live network

	Measured Values				
Test point ID	**RSRP**	**SINR**	**DL Throughput (Mbps)**	**UL Throughput (Mbps)**	**Comments**
1	-81.9	29			
2	-78.9	30	170	22	
3	-97.5	24	105	17.6	
4	-104	20.1	90	17	Marked as exclusion in design
5	-105	18	70	5.8	Marked as exclusion in design
6	-75	30	172	17.6	
7	-94	25	100.5	17	
8	-94.5	28	115	17.5	
9	-95		114	17.1	
10	-110	24	95	17	See figure 7
11	-75	30	170	22	
12	-105	15	75	11.2	See figure 7
13	-96	25.9	80	17.5	
14	-108		40	12	
15	-82	27.9	170	17	
16	-95	23	105	17	
17	-72	30/30	170	17	
18	-108	12	90	11	
19	-80	29	125	22	
20	-107	19	105	17	See figure 7
21	-110	10	70	11	See figure 7
22	-102	19	105	17	See figure 8 - Antenna mount needs to be lower or moved
23	-106	23	120	17	See figure 8 - Antenna mount needs to be lower or moved
24	-87	30	123	17	
25	-95	27	112	17	

Based on the results, we assessed where the network performance could be improved and reasons for lower-than-expected performance. Figure 8 shows an example of an area within the factory. Large metallic structures like this can cause fluctuations in signal strength due to diffraction and are difficult to account for in the design. Directional antennas can better accommodate hard to reach areas or can be excluded altogether if the customer does not wish to cover them.

Figure 9 - Metallic and large structures effect coverage

Figure 9 shows the effect of antenna placement on coverage. The antenna in the picture had to be moved because of factory restrictions. However, it is not located between a metal duct and a rail which limits coverage in one direction. Our solution here is to either lower it to avoid the blocking, or to relocate it to a different area.

Figure 10 - Antenna blocked by metal duct and rail

5.4. Use case

Charter has demonstrated a worker safety use case based on video analytics. Cameras were set up to stream video over CBRS to the MEC located in the factory MDF. An algorithm was trained to detect if employees were wearing helmets, safety gloves, jackets, and boots. The plant manager then collected metrics for compliance purposes, to see if there were any areas of the factory which needed better safety protocols or signs in place.

5.5. Lessons and best practices from deployment

First, the operator needs to have a detailed discussion of the existing network, firewalls, and any other policies that the IT team has in place. We faced some initial issues with default DNS being blocked, but were able to overcome them by working with IT. This can be resolved by including networking requirements in the initial customer information questionnaire.

Second, the RF designer should factor in movable structures in the design, especially in an industrial environment where this can happen often. This could be accomplished by running several scenarios of propagation with equipment located at different areas. At the same time, defining exclusion zones is important since not every area can be covered. Third, the network installers and designers need to work closely to ensure proper installation. A ten foot change in antenna location may not make a big difference in an office or outdoor environment, but in a factory this could potentially impact coverage if the new location is behind obstructions. Fourth, there should be a discussion on the primary use cases so that coverage and capacity can be targeted to serve those regions as best as possible.

6. Conclusion

In this paper, we presented an overview of CBRS and the motivation for private wireless networks on CBRS. Second, we presented the main use cases in different verticals and how a private LTE/5G network enables those use cases. Third, we presented the evaluation of different network vendors in Charter's lab in Denver, Colorado. Fourth, we presented the planning and results from an actual deployment in a large factory in Michigan.

Private LTE and 5G networks have the potential to streamline operations, provide greater insight into operations and help solve critical challenges for enterprises across many different verticals. Charter communications is committed to investing in private LTE/5G networks and serving the needs of enterprises across the nation.

Abbreviations

AP	access point
bps	bits per second
FEC	forward error correction
HD	high definition
Hz	hertz
SCTE	Society of Cable Telecommunications Engineers
5G	5th generation Cellular technology
LTE	Long Term Evolution
MEC	Multi access Edge compute
CBRS	Citizen Broadband Radio Services
SAS	Spectrum Access System
WISP	Wireless Internet Service provider
PAL	Priority Access License
GAA	General Authorized Access
FCC	Federal Communication Commission
CBSD	CBRS Device
CPI	Certified Professional Installer

HAAT	Height above average Terrain
EUD	End user device
MNO	Mobile network operator
QoS	Quality of Service
TDD	Time division duplex
IT	Information technology
ETSI	European telecommunication standard institute
GSMA	Global system for mobile communications association
RAN	Radio access network
UPF	User plane function
AMF	Access management function
SMF	Session management function
EPC	Evolved packet core
MIMO	Multiple in multiple out
UDM	Unified data management
PCF	Policy control function
NSSF	Network slice selection function
AUSF	Authentication server function
RF	Radio frequency
AGV	Autonomous guided vehicle
VR	Virtual reality
NMS	Network management system
AR	Augmented reality
eNodeB	E-UTRAN NodeB
CLI	Command line interface
GUI	Graphical user interface
QAM	Quadrature amplitude modulation
DNS	Domain name server
RSRP	Reference signal received power
MDF	Main distribution frame
UL	Uplink
DL	Downlink
SINR	Signal - interference noise ratio

Bibliography & References

1. Kaelble, Steve. *CBRS & OnGo. 1.* https://ongoalliance.org/wp-content/uploads/2021/04/CBRS-OnGo%C2%AE-For-Dummies%C2%AE-OnGo-Alliance-Special-Edition.pdf.

2. OnGo Private LTE Deployment Guide;CBRS Alliance. Link: https://ongoalliance.org/wp-content/uploads/2021/05/OnGo-Private-LTE-Deployment-Guide-2.0.pdf

3. *Vibration analysis for predictive maintenance solutions*. Radio Bridge. (2021, April 28). https://radiobridge.com/solutions/vibration-analysis-for-predictive-maintenance-and-condition-monitoring#:~:text=Radio%20Bridge%20provides%20high%2Dquality,while%20delivering%20exceptional%20customer%20service.

4. Dahmen-Lhuissier, S. (2019). ETSI. ETSI. https://www.etsi.org/technologies/multi-access-edge-computing

5. GSMA. (2020). *5G IoT Private & Dedicated Networks for Industry 4.0 A guide to private and dedicated 5G networks for manufacturing, production and supply chains*. https://www.gsma.com/iot/wp-content/uploads/2020/10/2020-10-GSMA-5G-IoT-Private-and-Dedicated-Networks-for-Industry-4.0.pdf

Proactive Asset Decommissioning in Critical Facilities to Accelerate Energy Management and Sustainability on the Road to 10G

A Case Study of How Comcast is Embracing Sustainable Asset Lifecycle Management to Power the "Future of Awesome"

A Technical Paper prepared for SCTE by

Mike Gala
Executive Director
Comcast Cable
1800 Arch Street Philadelphia, PA 19103
(215)286-8937
Mulchand_Gala@comcast.com

Sikander Chatha, Comcast
Sr. Manager
Comcast Cable
Sikander_Chatha@cable.comcast.com

1. Introduction

In-rack assets in datacenters and other critical facilities continue to contribute significantly to the energy usage at cable companies. SCTE•ISBE has championed Energy 2020 as an Energy Management Program to provide cable system operators with the new standards, technology innovation, organizational solutions, and training desired to drive continued cable telecommunications network growth, availability, and reliability. While there is a lot of excitement for a new product/application launch and deployment of assets, decommissioning has been an afterthought for many organizations. Instead, companies have been focused on delivery of new projects to optimize service delivery and revenue. With advancements in technology many compute, storage, and networking assets are able to increase their useful lifespan to 5-10 years for certain non-demanding or legacy workloads. This results in little desire or incentive for the planning of asset retirement. New assets and infrastructure have always been priority, and organizations have been squeezing every last drop of productivity out of their assets, extending the asset lifecycle well beyond the manufacturer's End of Service Life (EOSL) dates.

Cable system operators can significantly and immediately impact their bottom lines by deploying solutions that can reduce data center power needs and increase energy efficiency, according to a presentation at the Smart Energy Management Initiative (SEMI) forum by the Society of Cable Telecommunications Engineers (SCTE).

At Comcast NBCUniversal, teams hold themselves accountable to be environmentally responsible. Kicking off a proactive asset decommissioning program was a step in the right direction to help achieve our asset retirement and sustainability goals. This technical paper will highlight how we were able to apply a renewed focus on asset retirement and how this helped with keeping our utility bill flat in 2019/2020.

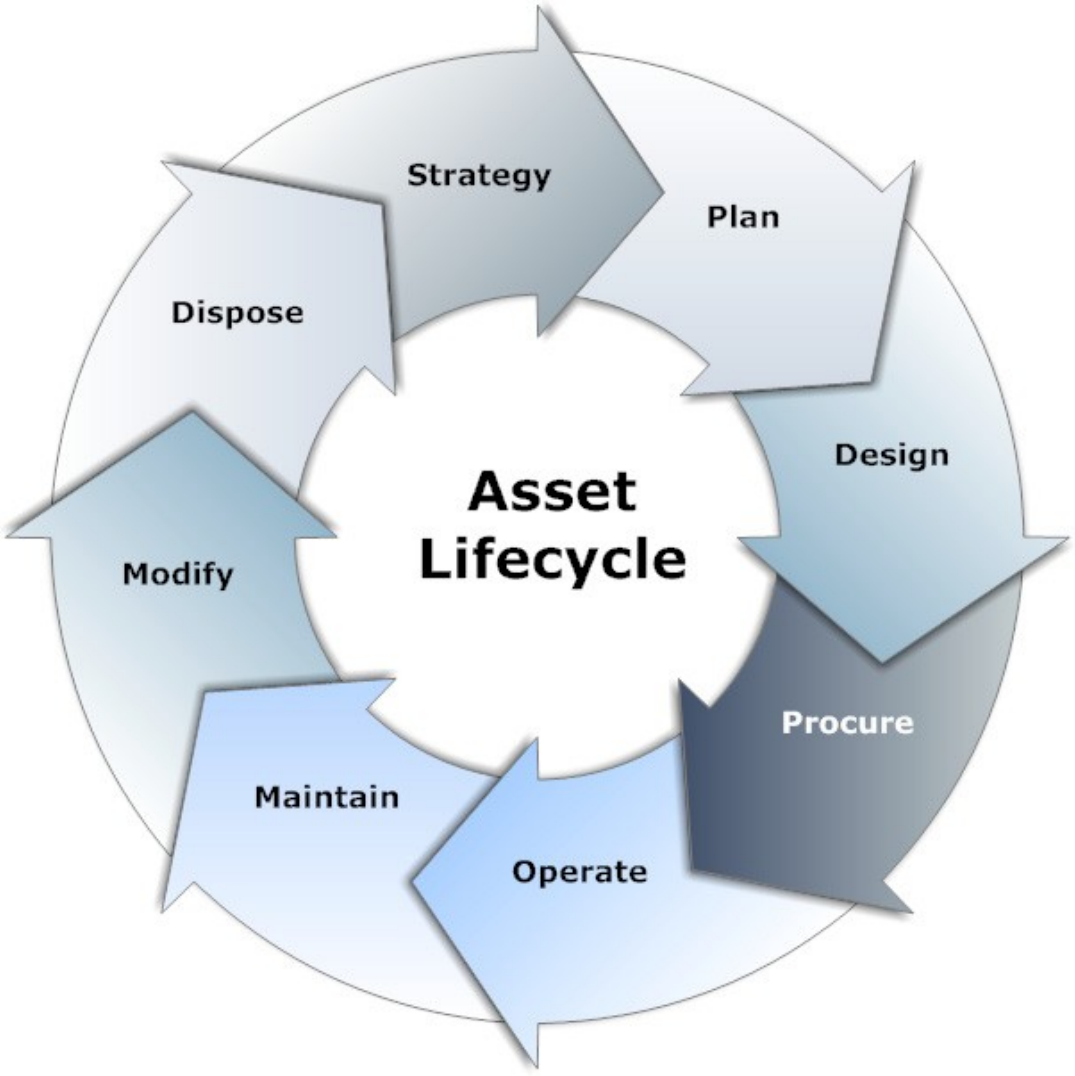


Figure 1 - Typical Asset Lifecycle (NIST SP 1800-5)

2. Benefits

The benefits of proactively running a Decom Program are numerous and include energy savings, reducing security risks, better inventory management, sustainability, reduced capital, and software spend. These are described below:

Energy Savings – Getting unused devices off your raised floor can have a significant impact on your organization's energy bill. Furthermore, newer servers are usually more energy efficient which should be an incentive to get aged devices out the door to help with your organization's goals for overall energy conservation.

Security Risk mitigation – Unused equipment and other aged devices may not be able to receive the latest patches and updates, and this can be cause for major concern due to lack of awareness of these vulnerable systems. Unpatched servers lead to persistent backdoors for hackers and increasingly unpatched servers are becoming a target for cyber-attacks, exploiting vulnerabilities in our organization. Keeping up and patching these systems when we know the ownership and applications being hosted on them is hard enough for any organization. The added burden of patching older abandoned and unused servers can turn into an impossible task.The benefits of retiring these systems early not only reduce risk but also reduce potential business intrusions caused by downtime. This is where a Decom Program can assist and get these unused vulnerable systems off your organization's network and help reduce unnecessary risk.

Help institute better Inventory Management – Getting your organization on board to focus resources on the proactive planning and retirement of unused and aged servers also has a direct impact on your organizations ability to track and inventory assets. Performing a book to floor and floor to book audit of all facilities housing your company's assets is step one of standing up a decom program. Knowing where all your assets are and who owns them is critical to success and this will force your organization to not only get a better understanding of what you own but will also open up the opportunity to better the process in managing these assets going forward. According to a recent Mercer report, the average employee turnover rate was 22% in 2018 which is a nightmare for asset management and keeping application ownership aligned. Getting ahead of the next re-org by understanding who owns and supports your organization's business applications and underlying assets and putting in a process to maintain this information will save your IT department countless hours down the road. This exercise will also help identify orphaned assets which would be ideal decommission candidates and lead the way to costing savings.

Sustainability – As your organization's compute and storage requirements grow, you will want to find ways to not only find alternative sources to power your data centers but also figure out ways to reduce the overall energy footprint. This is where retiring older assets and buying more energy efficient devices combined with the ability to do more with less will empower your organization to meet your individual sustainability goals.

Reduced Capital Spend – Many legacy systems are not cloud-ready, which emphasizes the need to stay lean and avoid the need to expand your physical footprint as eventually most if not all of your organization's workloads will be headed to the private or public cloud. Removing unused and aged devices and leveraging capabilities to optimize workload efficiency will ensure your organization is on the right side of change with shrinking and not growing their physical footprint.

Optimized Software Licensing – Software assets are a major cost in an organization's overall IT spend. The first step in optimizing these costs is knowing what you are consuming and then identifying opportunities to reduce your footprint. Reduction in software licensing and maintenance costs will be an indirect benefit your organization will receive by proactively decommissioning your physical assets. Not only will this result in huge savings, but this may also help your organization stay on side with your licensing obligations to your vendors.

Optimize overall Costs of IT - With the rising cost of IT including capital and operational spend, organizations are trying to find creative ways to avoid costs and save money. According to Gartner, worldwide IT spending will grow 4% in 2021. A major part of this increase and overall spend is associated with Data Center Systems and Enterprise Software. Standing up an Asset Decommissioning Program will directly attack both of these costs contributors and lead your organization not only into substantial cost savings but also provide an optimal setup for your organization to start optimizing IT

spend through virtualization and consolidation of workloads. Knowing what you own and what is and is not being used is step one to any organization's journey into standing up a Decom Program.

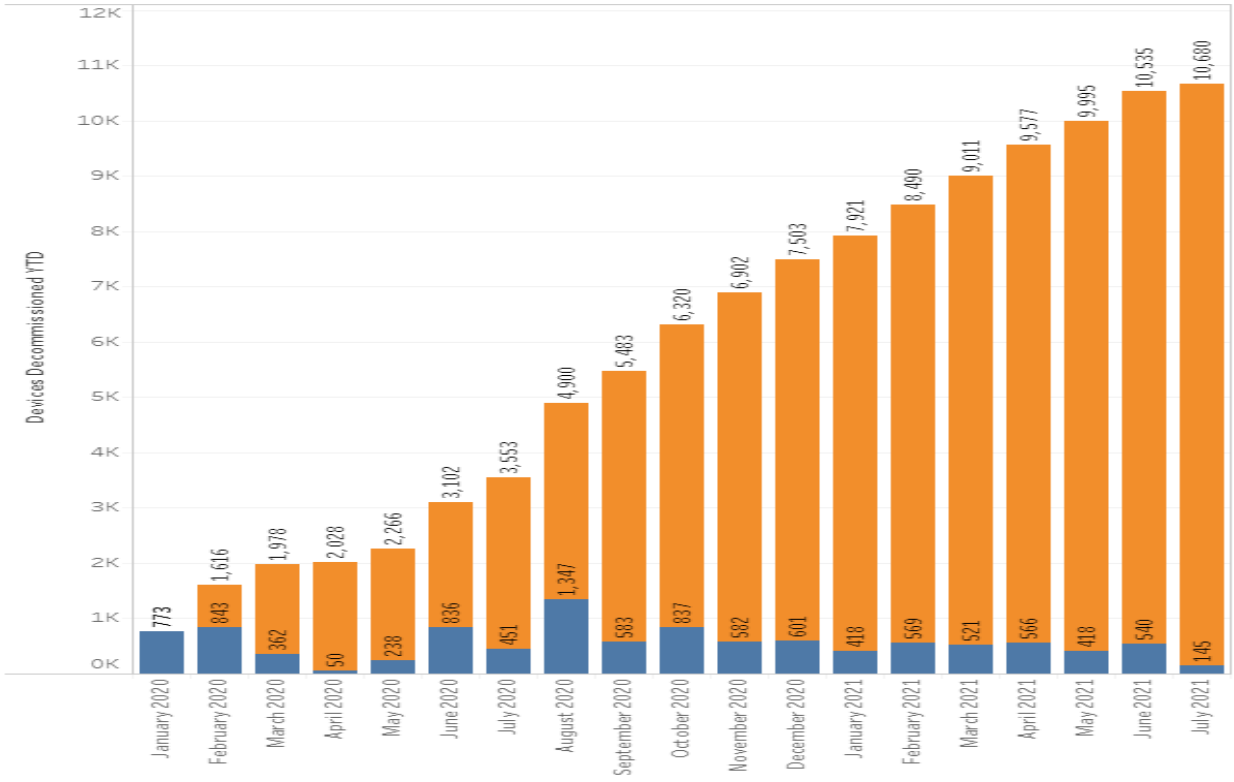


Figure 2 - Decommissions 2020-2021 (YTD)

3. Why is this not already a priority at many organizations?

Focus on new deployments – Generally speaking, there is a lot of excitement and related prioritization for effort involving new product launch and related new asset deployment. There is not as much focus on decommissioning those applications and assets as time passes.

Silos and reorganizations – The life of in-rack assets could be as long as 5-7 years. Many large companies go through multiple reorganizations of teams in such timeframes. This causes lack of ownership of these assets for the new teams. Also, organizational silos impact the ability to see an application retirement, asset decommissioning, and its sustainable disposal end-to-end.

Lack of proper inventory management – A well-maintained inventory with right CDEs (Critical Data Elements) and ownership is a necessary ingredient for a successful decommissioning program. Many companies are not able to create or manage this inventory as new assets come in and teams are reorganized.

4. Comcast Case Study

Why we did it?

Historically, the decommissioning of data center assets has been an afterthought for most organizations. This often results in abandoned physical workloads sitting collecting dust in critical facilities. These abandoned assets not only consume valuable resources but also pose a security threat to the organization. Our journey at Comcast begun in late 2018 as we started to cleanup and get a better handle on our assets. We worked to proactively optimize our decommissioning process to avoid the common pitfalls that a lot of organizations find themselves in, where unused assets are left in the racks, collecting dust, consuming power, and occupying space.

After realizing these potential impacts, Comcast embarked on the journey to shift from a reactive state to proactively managing decoms, similar to how assets are planned for and deployed during project delivery. The need for this was long overdue and some of the sure tell-tale signs were:

- Finding powered off or non-functional assets in our data centers during Data Center audits.
- Finding mismatches in asset status in various asset inventories.
- Finding sudden power drop offs in devices through our DCIM system.
- Finding gaps in our current Decom process, resulting in failed asset decoms.

Inventory Management

With the focus on physical device decommissions, it was very important that we did a floor-to-book and book-to-floor inventory of all our physical devices across all sites. This inventory was updated in our DCIM tool and reconciled with our CMDB to help identify ownership for each device. The team assessed various sources of inventory that can help identify the Critical Data Elements needed for each asset that would drive the identification of the initial assets to target:

Table 1 - List of Critical Data Elements needed for each asset to smoothly run through Decom process

Critical Data Element (CDE)	Note
Make/ Manufacturer	E.g. Hp/Dell/Cisco/Juniper to identify the OEM
Model Number	E.g. DL-360 to match up with EOSL list
Serial Number	E.g. 1234567 to uniquely identify each asset
FQDN/IP Address	E.g. billing-server1@comcast.com to identify network name and IP address
Application	E.g. Billing System
Application Owner	E.g. Consumer_Billing_Ops_team@comcast.com to work with for decommissioning process

Gaining buy-in from all stakeholders

The ability to proactively stand up a program around an area that has been historically ignored for years was an easy sell once the business case was developed and presented to leadership. The opportunities as discussed earlier were a clear and the thought of not moving forward with this was highlighted as a risk our organization was not willing to take.

Besides gaining senior leadership buy in, it was critical that we communicated the importance of this to all stakeholders, including both those submitting request to decommission and those fulfilling the downstream process to get decommissioned assets powered down and unracked. This was especially important because, like most organizations, the resources doing the deployments were the same resources that would be helping with the decommissioning of assets.

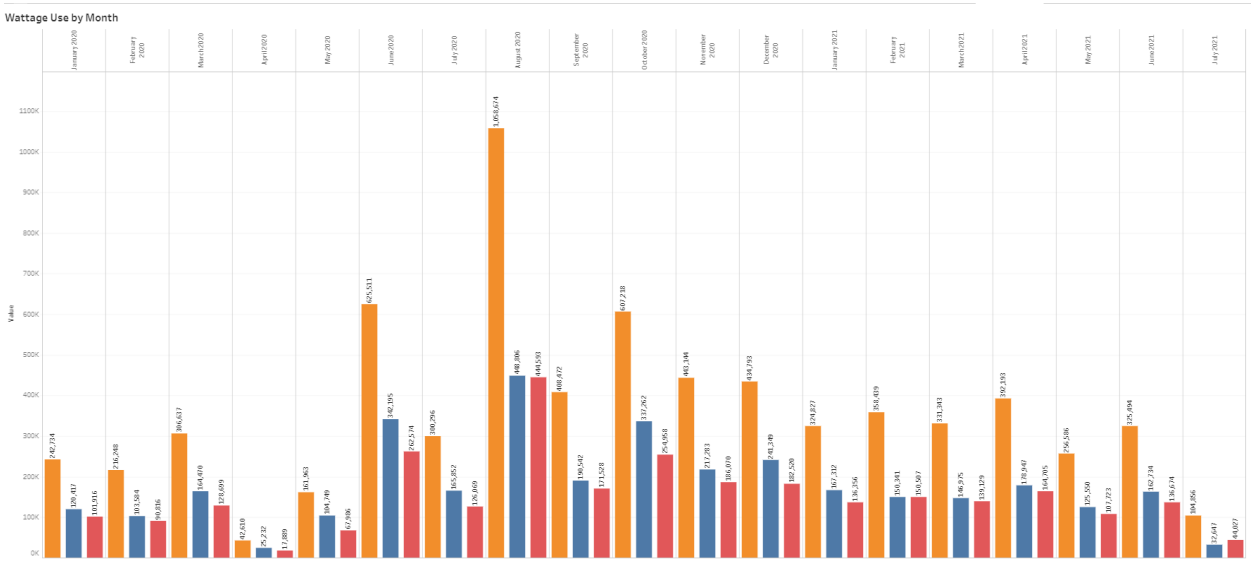


Figure 3: Power watts saving per month 2020-2021 (YTD)

4.3 Evolving the process

Evolving the process involved reviewing the current process and looking for ways to optimize. Looking for opportunities to streamline the process across different areas of the organization and cover various decom scenarios was a critical part of getting started. This involved removing unnecessary steps and introducing critical controls and checks and balances to avoid potential fallout.

4.3.1 Decom request submission

Without having the time or development resources to overhaul the current process, we put checks and balances in place to be notified of any decommission request that was being submitted. This simple step gave our team the power to see decoms through to the end, catching any fallout from decom rejections and potential failed changes.

4.3.2 Decom Governance & Control

The addition of governance tasks along with critical controls, really helped give us get a good handle on our decommissioned assets while we refined the existing process.

A few examples are listed below:

- We made sure new deployment projects highlighted hardware slated for decommissioning upfront so it could be flagged in our CMDB as a potential decom candidate. This really gave us a good handle on decoms associated with tech refreshes and data center migrations.
- We implemented a tool change that no longer allowed teams to diassociate their devices from their respective applications in the CMDB without a valid request to decommission the device. This ensured that we had a valid owner group associated with all devices until they were successfully decommissioned.Along with the above, we also ensured that the device state could not be updated to "decommissioned" without a valid change request to decommission the system. This would help address the concern we found where states were being flipped to "decommissioned," but the actual devices never got decommissioned from our critical facilities.
- Lastly, it was very important to align our decom program with our hardware maintenance process for our physical compute and storage devices. This helped us to identify potential decom candidates when devices are taken off maintenance and vice-versa.

4.3.3 Decom Fulfillment

The decom program showed early signs that the workload was putting a strain on fulfiller teams, and as they tried to balance the challenging task of giving decoms the same prioritization as new deployments, the queue backlogs showed that something was clearly not working.

The backlog of queues showed that tasks near the backend of the process, including port and IP reclamations, firewall entry removal, and unracking of the devices, were getting squeezed for resources, as these activities were involved the most risk and time to execute.

To address this issue head on we met with the teams to find alternative solution for these risky and time-consuming steps as the benefits of this program were much less appealing if assets just sat awaiting decom for months, especially when one use case we were trying to address was to minimize security threats.

After meeting with the teams, we were quickly able to enhance the current process by removing some of the more time-consuming steps that did not need to be in the actual decom flow. The teams flagged their work through the tasks and executed the work through a pre-approved change outside of the current decom process.

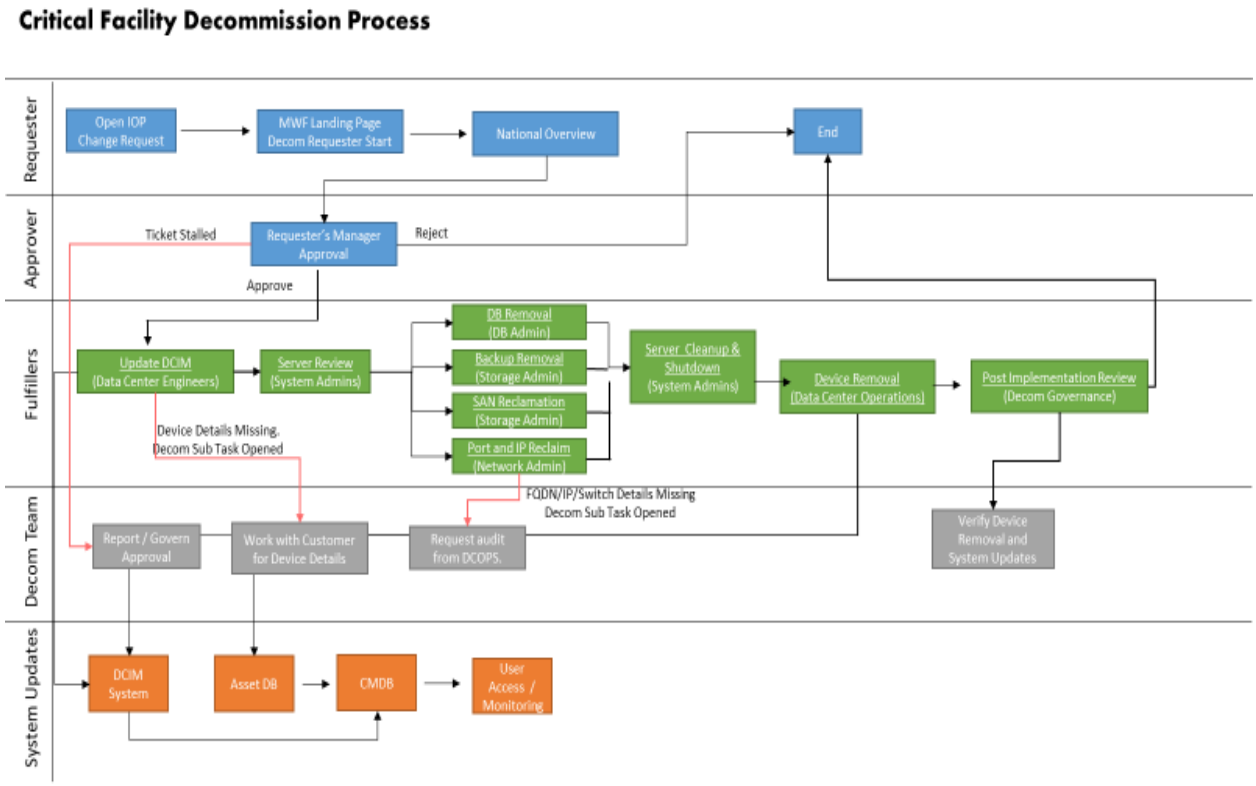


Figure 4 - Example Decommissioning Process

4.3.4 Orphaned Devices

Shortly after standing up the decom program and aligning with our CMDB as the source for ownership relating information, we launched our "Abandoned Server Decom Process" to address devices without ownership or other critical data elements. A few identifiers that we used to determine potential abandoned devices were:

- No ownership in our CMDB - This was the primary indicator.
- State mis-matches between our data sources.
- Powered down or off-network devices.
- Aging devices that were in service longer than 10 years or off maintenance.

In a very controlled manner we went through a 7 step process to turn down orphaned devices and ultimately decommission them if no one raised their hand within 30 days of shutdown. The steps were:

- Try to identify ownership through various sources and resources.
- Perform a physical audit to see if anything can be determined from the device itself. Check if it's even still racked, plugged in or on the network.
- Check for patterns in power usages and network traffic.
- Open a Change Ticket and get approvel for disconnecting or diabling the network port.
- Send out mass communication to Change Management ditro and all hardware service desks and monitoring teams.

- Disconnect or disable the network port.
- Wait 30 days and un-rack the physical device.

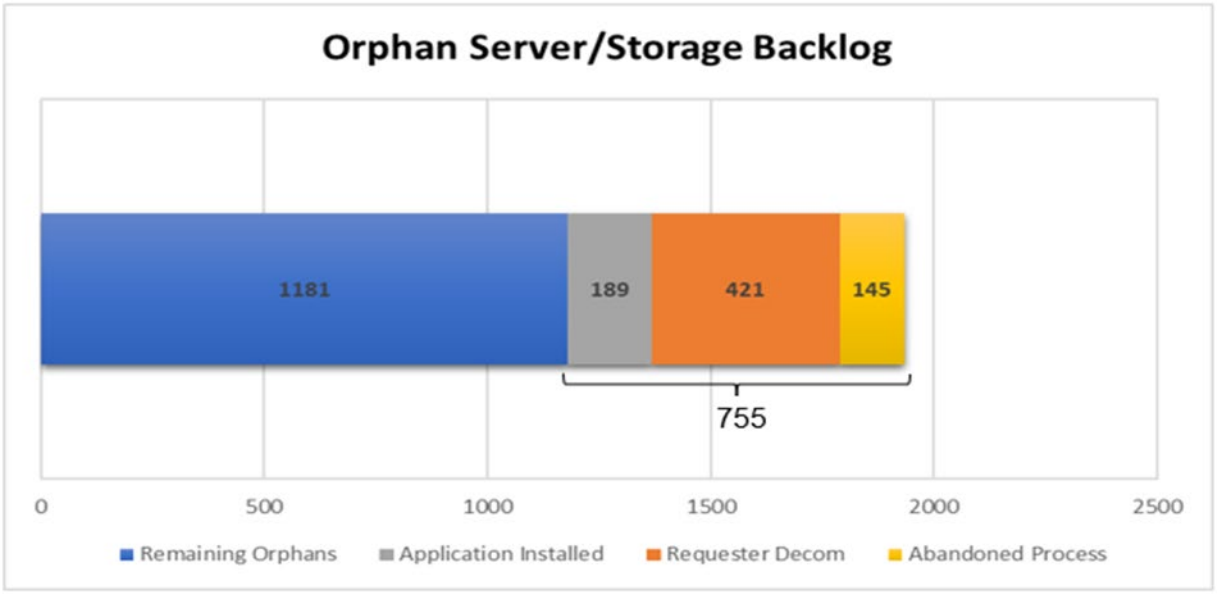


Figure 5 - Orphaned Device Progress Tracking

4.3.5 Remote Sites

The more labor-intensive tasks related to decommissioning of physical servers were still having trouble prioritizing the decom of the physical devices, especially at our non-staffed remote sites. To address this, we were able to generate dashboards and reports highlighting and grouping server decoms by site. This helped our data center teams prioritize and group deployments and decommissions at our remote sites, avoiding additional operating expenses that would have resulted in separate visits to our sites by our vendors

Since inception, we have successfully identified and decommissioned over 12,000 servers, including hundreds of abandoned servers.

5. Next Steps

Building on the success of the program, the team has further evolved the process to manage end-of-life/end-of-service makes/models to drive them thru the decommissioning process with advanced planning. The team is also enhancing the portfolio of the assets to include other appliance-like devices to help expand the scope. Teams are also discussing expanding the portfolio to other sites not in the program scope at this time, including all critical-facilities types, like our Headend & Hubs.

The evolution of the decom program will align with the One Network Program and will require partnership and collaboration across the entire organization. As we move forward we will continue to onboard teams and rationalize tools and processes so that the benefits of this program can be realized across the company.

6. Conclusion

If your organization has not already shifted from a reactive to a proactive approach to handle asset decommissioning, you may find that obsolete assets are taking up valuable resources and possibly draining your IT budget.

Since inception, Comcast has successfully identified and decommissioned over 12K servers including hundreds of abandoned servers and realizing a significant amount of power savings. This would not have been possible without the great work and prioritization given to this program from our Data Center partners and other fullfilment teams at Comcast.

Abbreviations

EOSL	End of Service Life
Decom	Decommission
CDE	Critical Data Element
DCIM	Data Center Infrastructure Management
CMDB	Configuration Management Database

Bibliography & References

https://www.gartner.com/en/newsroom/press-releases/2020-10-20-gartner-says-worldwide-it-spending-to-grow-4-percent-in-2021

https://resources.infosecinstitute.com/topic/linux-vulnerabilities-how-unpatched-servers-lead-to-persistent-backdoors/

https://www.imercer.com/articleinsights/North-American-Employee-Turnover-Trends-and-Effects

https://www.scte.org/criticalfacilities/

https://searchdatacenter.techtarget.com/tip/IT-asset-retirement-in-the-data-center

https://www.blancco.com/resources/bp-data-center-decommissioning-gap-analysis/

https://www.generatorsource.com/We-Buy-Generators/Generator-Decommissioning-and-Removal/Data-Center-Decommissioning-Checklist.aspx

https://www.energy-manager.ca/scte-spotlights-data-center-powering-efficiency-914/

https://www.nccoe.nist.gov/publication/1800-5/VolB/ (Figure 1 – Typical Asset Lifecycle)

Rapid and Automated Production Scale Activation of Expanded Upstream Bandwidth

A Technical Paper prepared for SCTE by

Rob Thompson
Director, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
robert_thompson6@cable.comcast.com

Rob Howald
Fellow, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
robert_howald@cable.comcast.com

John Chrostowski
Executive Director, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
john_chrostowski@comcast.com

Dan Rice
Vice President, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
daniel_rice4@comcast.com

Amarildo Vieira
Engineering, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
amarildo_vieira@comcast.com

Rohini Vugumudi
Director, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
rohini_vugumudi@comcast.com

Zhen Lu
Manager, Next Generation Access Networks
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
zhen_lu2@comcast.com

1. Introduction

Enabling the upstream portion of the access network for Mid-Split or High-Split takes more effort than just configuring the cable modem termination system (CMTS) to activate new DOCSIS carriers. Without considering the potential consequences of in-home networks and their effects on DOCSIS and video services, it may not be as seamless to customers or field operations. Most of the characterization required can be remotely measured and observed through telemetry available in the network. Operators can use algorithms to predetermine if there is work to be done in the field or home to enable new spectrum.

While today's tools are helpful in keeping technicians out of homes to make measurements, they were designed for raising flags, enabling swivel chair dashboards, or troubleshooting specific cases. By contrast, turn-up of network spectrum must initiate action instantaneously at large scale, through multiple scenario permutations, and publish essential data and actions cross-functionally to stakeholder organizations such as Sales, Technical Operations, Care, and even Warehouse Ops.

Instead of reacting to consequences or relying on manual vetting of homes, software running on modernized cloud infrastructure can proactively identify blockers, on an individual customer basis, but in massive scale with machine robustness that is immune to human error. The cloud software creates connective tissue between the upgraded network infrastructure and the incumbent suite of tools currently used by operations. This approach facilitates seamless interaction with the customer experience by maximally enabling new services and identifying blockers to frontline teams to optimize efficiency of support.

The work and development in upgrading the plant and services to mid-split is directly related to future upstream expansion including High-Split. Many of the workstreams will continue and will build off the mid-split effort. Moving the downstream spectrum to allow room for the extended upstream spectrum, the process of upgrading the physical plant with nodes and amplifiers that support different frequency splits, the deployment of Orthogonal Frequency Division Multiple Access (OFDMA) and adjacent channel interference (ACI) -- all correlate to high split deployments. The tools and processes to validate mid-split in-home ACI can be used to evaluate neighbor interference in a high split system. ACI is dependent on in-home splitter isolation and neighbor interference is dependent on tap-to-tap isolation. In both cases, the tools and processes are similar. The mid or high split device can be exercised and the level of interference or effect on the adjacent or neighbor device can be measured prior to permanently moving the customer's device into mid or high split operation. In both cases methods have been developed which are non-service and non-customer impacting.

This paper will introduce the reader to Comcast's cloud software architecture, the high-level algorithms used to "score" a home, and the downstream systems fed by the system that will enable valuable upstream services in scale while minimizing impact to customers and maximizing efficiency of its delivery for the operator.

2. Background

Options for upstream capacity enhancements arrived in DOCSIS 3.0 around the 2006 timeframe [1]. The 42 MHz to 85 MHz or mid-split band could become the first step toward supporting new upstream DOCSIS channels for U.S.-based operators. Since then, new DOCSIS 3.1 and 4.0 standards have been developed to support even more channels, in even larger return bands, like high-split, up to 204 MHz, and Full Duplex DOCSIS (FDX), up to 684 MHz. Those new upstream channels would become part of the larger channel set, enabling operators to increase upstream service capacity appreciably beyond what has been historically constrained to 5 – 42 MHz, known as the standard-split or low-split operating band. During the mid-split introduction, diplexer switching within the cable modem (CM) would also be introduced, to remotely control a new and more flexible diplexer function and thus, the channels over which a CM could communicate. Switchable diplexer functionality enabled operators to seed their network footprint with mid-split-capable CPE, while the rest of the network's technology could catch up to support the activation of enhanced upstream services. What wasn't well understood at that time was the new work needed to enable new upstream spectrum activation -- a topic that will be introduced in this section.

Managing the outside plant to support mid-split has become well understood through ongoing upgrade activities. For example, preparations including vacating downstream spectrum, and relocating the downstream out-of-band (OOB) signal to a higher frequency have become standard operating procedures. Because these outside plant pieces were easily implemented, many operators have activated or at least tested mid-split services, and that experience has proven to be appreciably more challenging for some operators. Two of the challenges, standard-split drop amplifiers (SSDAs) and ACI susceptibility, were recently documented in [4] as home network-based challenges that will block the successful activation of mid-split services.

Mid-split tests at Comcast have helped us learn how to remediate these challenges. SSDAs are ideally removed from the home network for many reasons including that they are a more costly, higher-rate point-of-failure than their passive splitter counterparts. Furthermore, SSDAs used in a home network will have to be removed because of its standard-split diplexer, to make way for mid-split upstream services. Therefore, limiting drop amplifier use, wherever possible, may emerge as one of the preferred remediation approaches.

Remediating ACI susceptibility is accomplished by increasing isolation between the mid-split CPE (MS-CPE) and standard-split (SS-CPE) [4]. The following recommendations have emerged based on our learnings during early mid-split testing, to ensure mid-split connectivity between the home network and the outside plant, while minimizing ACI:

1. If there are few (≤ 4) devices within the home requiring outside plant connectivity, receiving video and/or DOCSIS quadrature amplitude modulation (QAM) signaling, use a higher isolation splitter
2. If there are many (> 4) devices within the home requiring outside plant (DOCSIS) connectivity, use a higher isolation passive network (POE DOCSIS-MOCA Passive or PDMP)
3. If 1 and 2 are not feasible due to severe drop loss, consider the following

a. Reduce devices requiring outside plant (DOCSIS) connectivity, and subsequent drop loss, by replacing them with Wi-Fi capable devices
b. Use a drop amplifier sparingly, which actively adds gain to overcome drop loss

4. If 1 through 3 still do not provide adequate isolation for ACI susceptible devices, use a notch filter

The remediation strategy above can be distilled down to the practice of removing drop amplifiers and swapping splitters --something installers are well-versed in dealing with in their day-to-day. The new work is focusing on when and where to perform the above remediation tasks, since we've already learned that doing them for 100% of the mid-split customers at the time of spectrum activation may not be feasible for some operators [4]. Even though it wasn't called this at the time of publication, in-home Health Assessment Tool (iHAT), as it has since become known, has demonstrated that Remote Feature Control (RFC) and Remote Health Monitoring (RHM) can be used to orchestrate remote identifications of mid-split-capable devices that can be activated without remediation. That's opposed to those that need remediation, which can be placed in a queue for activation when that work completes at some point in the future [4]. Activating mid-split upstream will include performing iHAT tests, and distributing results to teams, including installers, who can perform the remediation tasks described above.

3. New Mid-split Activation Tools

3.1. Mid and High split spectrum activation launch (MUSL/HUSL)

As described, the innovative iHAT tool provides a relatively non-intrusive view into a customer's home and, on a home-by-home basis, making basically a Go-No Go declaration with respect to activating spectrum in the Mid-Split band for that customer's device. The iHAT results score a home's DOCSIS readiness for passing spectrum to 85 MHz, and/or its likelihood of creating video interference.

Of course, the home cannot be assessed, nor the spectrum activated efficiently on a home-by-home basis, by relying on manual processes. The information iHAT needs to run, and the information needed by other systems to act on the iHAT outcome, must be automated and the interfaces to these other functions built for production scale.

On the input side, iHAT needs information from external systems to identify if a house is eligible for Mid-Split from an equipment standpoint, and thereby worth running iHAT at all. It needs the inventory of Mid-Split capable CPE on a per node basis identified so that it can target those devices to run the test. The output of iHAT – the scoring of a home's DOCSIS readiness or risk of degrading the video experience – is information that multiple other external systems need. For example, technicians in the home need to know what to fix if a problem related to Mid-Split is identified. Care agents need to know how to diagnose a possible Mid-Split related issue, guide a trouble call, and dispatch the proper support. Salespersons need to know if an upstream speed that requires Mid-Split capability is possible. Data sciences team need iHAT results to populate databases to analyze trends to adjust and optimize roll-out processes and operation support. We will look at these in more detail in a subsequent section.

3.2. Manual vs. Automation Processes – Cloud Infrastructure Choices

It is undeniable that a successful business must find ways to scale in order to stay relevant. This is extremely true when applied to the telecommunications industry. The bandwidth requirements of people and businesses, especially in the upstream, have been increasing dramatically year over year due to advancements in technology. To keep up with the high demand for broadband services, innovation is required. There are some problems that can be solved by throwing people at the problem; however, when you have customer devices in the tens and hundreds of millions, this becomes untenable.

In the field of software, the most basic reasons to build an automated solution have always been the need for smart investment on resources, and finding budget efficiencies, instead of spending to do manual work year after year. The manual method is undeniably easy and quicker to build but has its own downfall. Questions raised are 1) What can be automated? 2) What kind of automated solution would work? 3) What defines the "right" solution? 4) What platform is optimal to architect the solution? 5)What factors play a part in choosing and creating the solution? 6) What are the dependencies and integrations? 7) Perhaps the most important question of all is, how much of an automated solution can sustain the business growth year after year, without re-investing and redoing a lot of the work? A companion question: What about component usability, for other parts of the business?

The downside of the automated solution is that its machines must be built and trained in a way that mimics the thought processes for human troubleshooting skills. Outcome planning involves thinking through as many of the business's use cases as possible, considering regulations and validating as many inputs and outputs as possible. Also, implementing ways to listen and respond back to alerts, and handling notifications to support the operations and avoid potentially impacting customers.

Let's talk about what factors will make automation an answer for any business, which include but aren't limited to: Scalability, availability, maintainability, reusability and budget. It is imperative that incoming data feeds are well sanitized before entering an automation solution, so the data that comes out has integrity for better analysis and reporting. Effective automated solutions have a proven record of reducing the cost overtime -- emphasis is on "effective automation solution."

Taking some time to answer the questions raised above will help to decide upon the right solution. In fact, that's what led us to choose the solution for iHAT. Like any other software solution, we started with the version 1.0 build, described in [4], to research, implement, test and analyze. This pushed us to build iHAT version 2.0 from the lessons learned. Some of the downfalls of version 1.0 included:

- Manual process
- Scalability
- Speed test failures
- SNMP V3 key reset after every reboot made manual process difficult
- Availability of system depended on human resources
- Time consuming and not easily maintainable
- Handling unexpected exceptions
- Automatic rollbacks
- Having to wait until next day to see the runs, if resources are not available

That list of "cons," along with other elements we learned with the manual implementation of version 1.0, gave us clarity on next steps. It also elicited a new list of questions and requirements, including: 1) Developing a list of features to be automated 2) Cloud-based automated solution is the preferred option 3) Reduce human dependencies 4) Ensure that the platform is highly available 5) Automation, manageability, high availability, scalability, cost effective 6) Not possible to reduce dependencies, which will make the solution to be focused on understanding data collection and distribution and build reliable integration to pull right data from source of truth (SOT) and be the most reliable source of distribution (SOD). By taking these lessons learned into consideration, and after answering all the questions, a solution can be envisioned and built that encompasses the qualities that sustainably support business.

3.3. In-Home Health Assessment (iHAT) 2.0

iHAT is a method, or in programming terms, a script, whose inputs must include the Mid-Split-capable devices and any set top boxes (STBs) that share a home network connection with them [4] on an operator's Mid-Split-capable network. iHAT outputs provide operators with results indicating which Mid-Split-capable devices have successfully switched over to Mid-Split and which will require remediation to do so [4]. Remediation associated with Mid-Split was discussed in Section 2. The iHAT PoC from [4] has proven it will work well for small tests with up to 1,000 devices per instance. Performing iHAT for a larger population (millions of devices) revealed challenges in three key areas:

- Automation – Automation of iHAT, specifically input gathering and output distribution wasn't described in [4] but is a critical step at the beginning and ending of such a test. Taking steps to ensure the accuracy of the deployment data, which is constantly changing due to business as usual (BAU) activities, including device swaps, can prevent aborted iHAT tests attributable to incorrect device information. Minimizing the time between gathering inputs and executing iHAT will ensure integrity of the inputs and is best accomplished via automated process.
- Service Impact – [4] described the importance of the speed test application for iHAT testing, essentially behaving as a catalyst for determining whether a mid-split based service is going to degrade other customer services, like video. Using a speed test application in this manner is service affecting because it blocks a customer's use of their CM during speed testing, and therefore is limited to maintenance window activities. Another example of an iHAT service affecting feature is associated with rebooting a CM, which is required when CM switchable diplexer is changed from standard-split to mid-split. Eliminating service affecting components from iHAT process, like the speed test application and reboots, could enable iHAT testing to coexist with customer use.
- Application Dependencies – for many operators, RFC and RHM described in [4] likely translates to multiple application interactions. A sunny day scenario is that all the applications used to support iHAT are responsive, meaning they complete their function in a timely manner. When the sunny day scenario doesn't occur, iHAT is forced to retry, perhaps multiple times, and ultimately abort a test when dependent application functions fail. One example is when RHM doesn't return STB signal-to-noise (SNR) data necessary to perform a pass or fail decision on ACI susceptibility. Other examples of application dependencies include the speed test application in [4] and the associated class of service (COS) change application to test at the higher mid-split service rate. Most standard-split COS rates support up to 35 Mbps, but to fully utilize a DOCSIS 3.1 upstream [2], a new COS supporting a much higher rate, like ~300 Mbps, would have to be available for use during iHAT testing. Hardening iHAT for scale operation

requires a careful review of the function calls it makes, an assessment of their ability to fulfill their function in a timely manner, and a consolidation of all functions wherever possible.

As it turns out, there was an opportunity to consolidate application dependencies of the iHAT process involving the assessment of ACI susceptibility. The original version of iHAT described in [4] was successful in testing during worst-case conditions experienced on a home network, which were simultaneous occurrences of (1) fully utilized mid-split upstream and (2) ACI susceptible service, including a customer watching video. There were at least five dependent functions in the original version of iHAT that were required to assess ACI susceptibility:

1. CM COS change to support higher mid-split rate
2. Switchable diplexer state change from standard-split to mid-split
3. CM reset associated with both (1) and (2)
4. Speed testing application to simulate customer activity
5. Telemetry polls to assess against pass/fail criteria

Upstream carriers energized by an upstream speed test can be replaced with a OFDMA Upstream Data Profile (OUDP) burst signal. An OUDP request and response are shown in Figure 1. Using OUDP bursts can be facilitated without originally required COS changes of (1) but more importantly they give greater flexibility than speed test applications for estimating ACI susceptibility. OUDP bursts can be quicker, managed via CMTS scheduling functions, and more benign because their total transmission power can be reduced appreciably using smaller bandwidth, for example 1.6 MHz. This change may set iHAT up for no longer being restricted to maintenance window activities, especially if other service affecting aspects of iHAT, like reboots associated with switchable diplexer changes, are also eliminated.

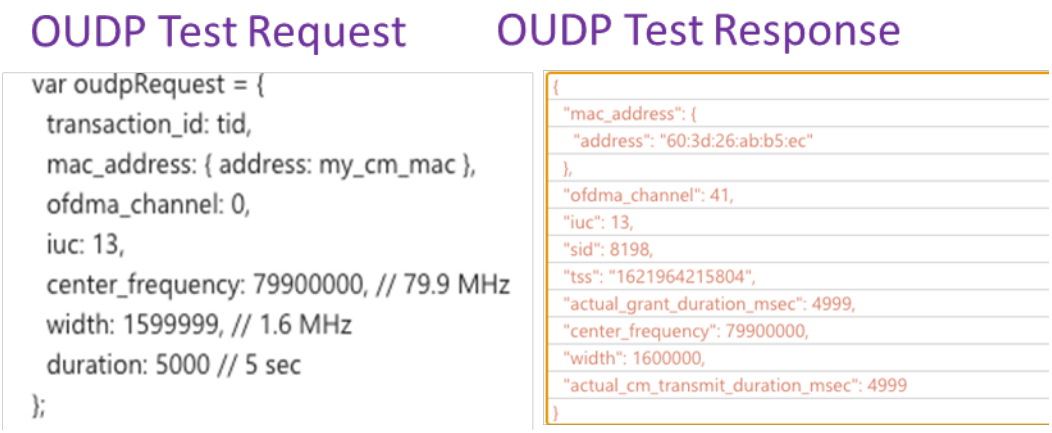


Figure 1 - OUDP Burst Request and Response Parameter Set

With these changes, iHAT will produce a different view of the ACI susceptibility problem, which will become a spectrum-based view from the SS-CPE point of view. While the OUDP signal is bursting in Figure 2, spectrum captures are performed by the SS-CPE at the highest rate possible, providing a multiple-sample spectral view of both the desired downstream signals and the upstream OUDP signal leakage into the SS-CPE, per Figure 3. From Figure 3, ACI assessment becomes more about visualizing OUDP signal leakage into the SS-CPE receiver, than about driving the SS-CPE into a failure state, like the original iHAT v1.0 using the upstream speed test application in [4].

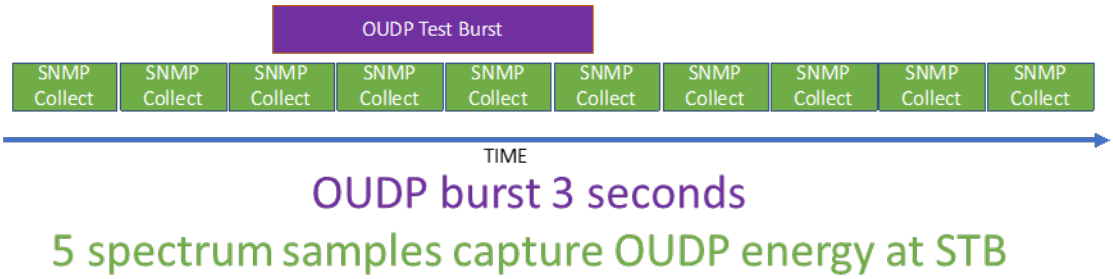


Figure 2 - Simultaneous MS-CPE OUDP Burst and SS-CPE SNMP Data Collection (incl. spectrum capture)

From this new spectrum view, two new parameters can be estimated: MS-CPE to SS-CPE (1) isolation, in decibel units (dB) and (2) interference level, in decibels relative carrier (dBc). When combined with the original metric collection described in [4], the isolation is a simple calculation based upon the telemetry polling of the MS-CPE transmit power in dBmV and the SS-CPE receive power in dBmV. Equation 1 shows this simple relationship, based on decibels relative to one millivolt (dBmV) per 1.6 MHz channels.

$$Isolation\ (dB) = P_{MS-CPE\ OUDP\ TX} - P_{SS-CPE\ DS\ RX} \qquad \text{Equation 1}$$

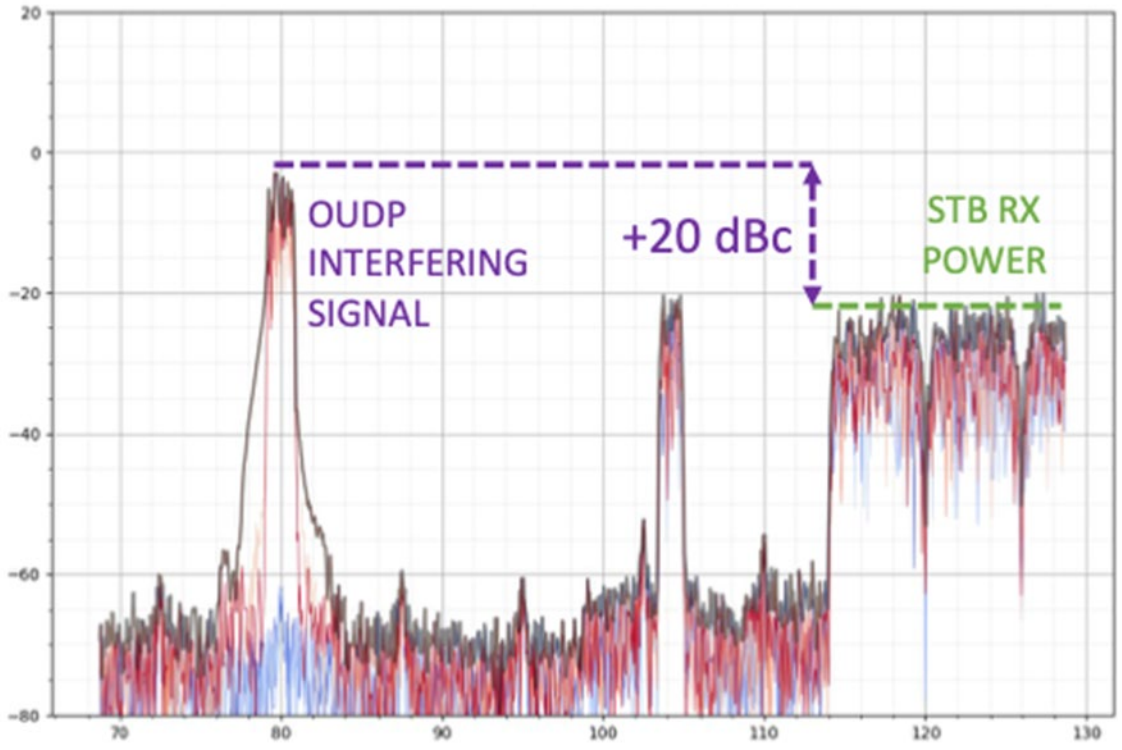


Figure 3 - Spectral View of MS-CPE OUDP Burst, and SS-CPE OOB and Low Frequency Downstream Signals

Adding SS-CPE spectrum data to dependent function 5, "telemetry polls" above and coordinating their capture to occur during the scheduled MS-CPE OUDP burst transmission will enable the estimation of interference level in dBc. Interference level can be estimated via spectrum capture of the SS-CPE per

Equation 2. Multiple captures are required to ensure good estimates can be made for both the OUDP burst and the downstream signal powers. Therefore, the power levels will be a statistical representation of multiple spectrum samples. In particular, the maximum hold, black trace of Figure 3, of the available traces is used to estimate channel power on a 1.6 MHz basis.

$$Interference\ (dBc) = P_{SS-CPE\ OUDP\ RX}\left(\frac{dBmV}{1.6MHz}\right) - P_{SS-CPE\ DS\ RX}\left(\frac{dBmV}{1.6MHz}\right) \qquad \text{Equation 2}$$

Deciding pass or fail will be different than what was originally shared in [4], where failure assessments were made when SS-CPE MER became worse than a threshold value of 28 dB for example, a point past which degraded video would likely be observed for most STBs. New thresholds targets, like 25 dB for isolation and 20 dBc for interference levels, would be revised as data is aggregated from many iHAT tests. Interference targets are based upon laboratory and other investigations, where greater than 20 dBc interference power would result in a degraded MER and FEC metrics. In other words, the undesired OUDP leakage is 20 dB higher than the desired downstream power observed at the SS-CPE input. This level of interference would put certain makes, models, and vintages of SS-CPE at risk of failure due to ACI susceptibility issues, based on internal testing performed in Comcast labs. Interference thresholds would also consider total power level differences of the OUDP signal (1.6 MHz) versus the overlapping mid-split power of OFDMA transmission in the SS-CPE receive window, see Figure 4. Lastly, process accuracy would be proven over full range of SS-CPE downstream receive and MS-CPE upstream transmit power. The 25 dB isolation target is based upon minimum output-port-to-output-port (OP2OP) isolation requirements specified for Comcast passive components.

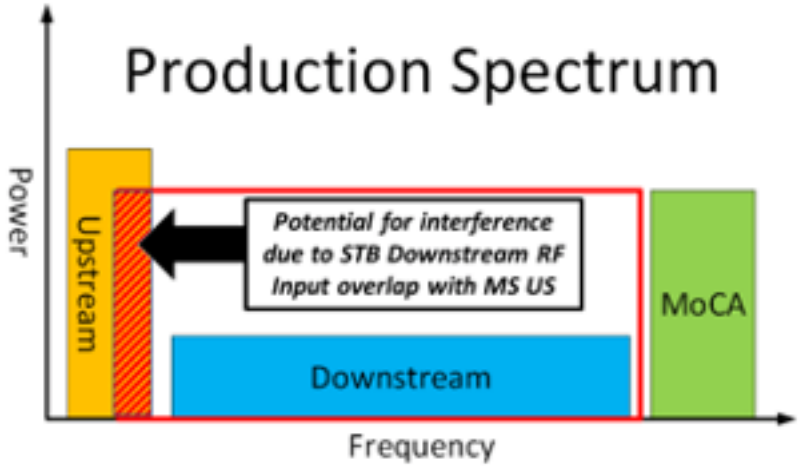


Figure 4 - 30 MHz MS-CPE, SS-CPE Receive Window Overlap

The last piece of scaling mid-split involves configuration of the MS-CPE switchable diplexer and as it turns out, the SNMP set approach discussed in [4] is works well for network utilization improvements and optimizing serving group capacity. Another way operators can manage the MS-CPE's switchable diplexer is using bootfiles. However, if there are no COS changes then redundant bootfiles would need to be deployed to have individualized control over a MS-CPE diplexer state. This could be challenging if there are many mid-split-capable platforms with many different COSs. For operators wanting to introduce a higher service tier, like within their commercial service markets, then co-managing the diplexer state with the COS should coexist nicely, and make more operational sense.

3.4. iHAT as the Engine for Mid-Split Upstream Spectrum Launch (MUSL)

As described, the innovative iHAT tool provides a relatively non-intrusive view into a customer's home and, on a home-by-home basis, will make a Go-No Go declaration with respect to activating spectrum in the Mid-Split band. iHAT scores a home's DOCSIS readiness for passing spectrum to 85 MHz, and its likelihood of creating video interference.

Of course, the home cannot be assessed, nor the spectrum activated efficiently on a home-by-home basis via human interaction. The information iHAT needs to run and the information needed by other systems to act on the iHAT outcome must be automated, and the interfaces to these other functions built for production scale. A logical flow diagram for the overarching Mid-Split Upstream Spectrum Launch (MUSL) ecosystem is shown in Figure 5. As shown, within the MUSL framework, iHAT is the engine. Note that "Tier 1" and "Tier 2" are arbitrary labels meant to represent individualized operator upgrade strategies and illustrate these how these new processes can fit together.

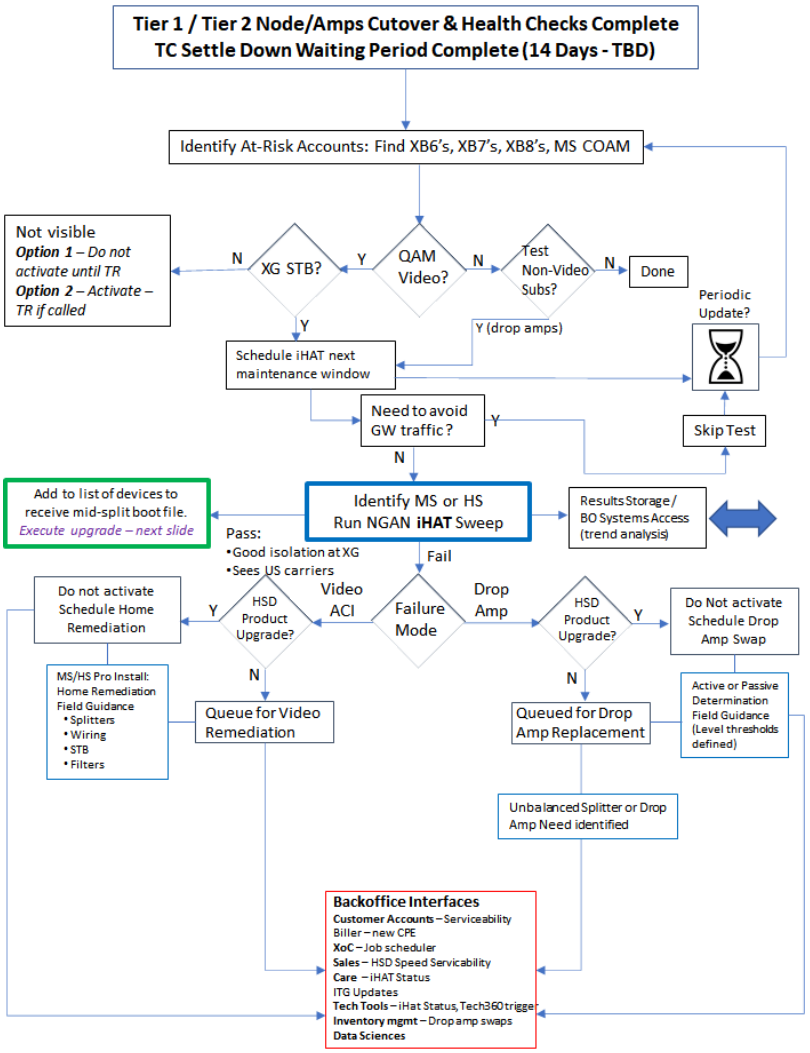


Figure 5 - iHAT is the Engine of the Mid-Split Upstream Spectrum Launch (MUSL) Framework

The interfaces for iHAT for its use in production are highlighted in the red box at the bottom of Figure 5 and briefly described below. These represent interfaces to consider for MUSL to distribute this important information to stakeholders in the success of Mid-Split activation.

Customer Accounts – Serviceability: It is important when there are new upstream speeds that only Mid-Split can provide that the systems to upgrade a customer, whether online or through a service call, recognize if the iHAT status of the home verifies this can be done safely and meet the service speed the customer expects. Alternatively, these tools can trigger an instant iHAT test for an updated result.

Biller – new CPE: When a customer changes CPE, possible iHAT variables that are affected are the device's DOCSIS capabilities, the sensitivity to interference of a new video CPE, and the possibility of a wiring change in the home. It is prudent given these potential risks to the iHAT state as recorded to test (or re-test) the home.

Operations – Job Scheduler: When a home "fails" iHAT, it goes into a remediation queue, with a flag for what needs to be remediated (video or HSD). This allows Tech Ops to plan proactive remediations, occurring routinely and not waiting for a house call to take care of iHAT-known issues.

Sales – Serviceability – Similar to Customer accounts, sales representatives should be able to quickly assess whether a customer, such as an MDU property, is eligible for MS speeds by accessing iHAT status in existing sales tools

Care – iHAT status, ITG Updates: When a TC arrives at an agent, after some amount of ITG-led questioning, the possibility of the issue being MS-related should be considered. A check on the iHAT status of that home, or an instant iHAT test, can help the triage process.

Tech Tools – Tech360: Like Care, when a tech is enroute or onsite to a customer home, part of the awareness the Tech should have is the MS status of the node, and the iHAT status of the customer. Further in the tools, the sequence of steps to diagnose and fix a MS-related issue should also be available.

Inventory Management – Drop Amp swaps: As remediations are made at relatively large scale to remove old drop amps, procurement awareness to the deployment of alternative solutions can ensure the supply pipeline is tracked and cared for

Data Sciences – As iHAT data is accumulated, new information about the home RF environment in dB performance, trends over time, and correlations across neighborhoods can be stored and processed for future optimizations and to estimate future process implications and costs.

3.4.1. Applying MUSL End-to-End

Referring to Figure 6, we can show how iHAT fits within the broader perspective operationally, going from the trigger of a distributed access architecture (DAA) node cutover on the far left to the completion of activation on the far right.

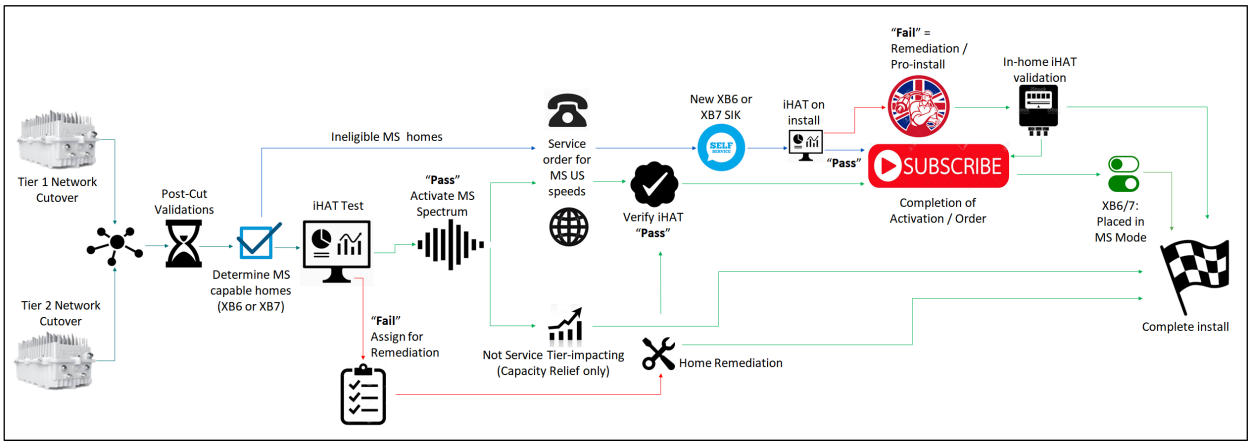


Figure 6 - Simplified Mid-Split Activation Flow: Cutover through to Activation

Beginning on the left, when a mid-split network upgrade occurs, the arbitrarily labeled "Tier 1" or "Tier 2" network cutovers can trigger internal tools, which can notify systems when construction is complete, and the cutover officially closed out. This can then trigger the spectrum activation process. Two things must happen prior to letting iHAT sweep across the node an validate homes where Mid-Split can be turned on. These are

1) *Post-Cut Validation* – Make sure that the network has settled to BAU metrics after the cutover. It is not uncommon to have a short period of elevated trouble calls shortly after a cutover, and it is desired to have any residual cutover issues resolved prior to moving to Mid-Split. This can be time-based, or it can be directly associated with trouble call metrics pre-cut vs post cut.

2) *Determine which homes are eligible for activation* – This boils down to whether the DOCSIS CPE is capable of Mid-Split. At present in Comcast, all DOCSIS 3.1 Gateways are Mid-Split capable.

On item 2) above, if a home is ineligible, then iHAT does not run. Following this arrow to the top path in Figure 6, there is no immediate required step to get that customer a Mid-Split capable modem. There is an effective loss of capacity for every CM that cannot access the DOCSIS 3.1 spectrum, because it forces utilization in the Low-Split band, rather than access the more efficient OFDMA spectrum above 40 MHz.

There is guidance in the field on what triggers a DOCSIS 3.1 upgrade for a customer – a particular speed tier for example. Over time, DOCSIS 3.0 CMs will organically disappear from the field, and it is likely at some point there will need to be a proactive effort to remove the stragglers still in the network to maximize the DOCSIS 3.1 capacity.

Now, as shown in Figure 6, if the customer decides to upgrade their speed tier to one that requires Mid-Split, then of course getting them a gateway capable of that becomes a priority. Also, of course, this customer's home needs to be evaluated for its ability to be placed in Mid-Split mode. So, as a new Mid-Split capable gateway is brought onboard, one of the first things it needs to do is call on iHAT and determine the state of the home for Mid-Split. If iHAT "PASS" is recorded, then the activation process continues, and iHAT sets the device into Mid-Split mode and it becomes capable of using the OFDMA

spectrum, in this case between approximately 40-85 MHz. If iHAT records a "FAIL," then the customer is notified that a technician must come to the home to complete their install, and that their new speed tier will not be available until this "Pro Install" step happens (nor will they be billed for it!). When the remediation is complete, the technician will validate onsite with iHAT triggered locally from the Performance health test (PHT) application.

If the eligibility conditions are in place – Mid-Split capable CM, and a STB model with the necessary Telemetry capability – we move to the right of the blue checkmark of Figure 6: "iHAT Test." Let's now follow the lower path – "iHAT FAIL."

3.4.2. The Remediation Queue

As noted, unless there is a speed upgrade required by a customer, there is not necessarily an immediate need to provide them with a Mid-Split capable gateway. However, it is still important that the iHAT score be logged, and the fact that the home needs to be remediated is documented and populated into tools used by agents and technicians. Homes in this category are placed into a Remediation Queue. iHAT will identify what the failure mode is so that Technicians know what needs to be done. In general, remediation tasks are well-understood and known tasks to technicians – changing out home amplifiers for alternative devices, checking the splitter configuration, model, and wiring to Comcast compliance. After remediation is performed, the iHAT test is run to validate readiness for Mid-Split spectrum, and the activation then completed.

When a home is scheduled for remediation, assuming there is no speed tier ask requiring it, is a business decision with a number of variables having to do with capacity, efficiency, and proactive expense. Ultimately, however, all homes in the remediation queue will need to get serviced to extract the full DOCSIS 3.1 capacity and maximize the upstream runway these architectures are made to deliver.

Also note that a customer's iHAT "score" is not necessarily static. Changes to the coaxial network in the home made by the customer, or new CPE brought into the home can both affect the iHAT score. These events are "On demand triggers" that will force iHAT to run off-cycle even after the initial iHAT sweep of the node at cutover.

3.4.3. iHAT PASS

The most straightforward flow in Figure 6 is right now the center, left to right. Both branches are logical and easily understood. An iHAT "PASS" means that the DOCSIS signals up to 85 MHz are ably received at vCMTS receiver, indicating that there is no home amplifier or filter blocking this transmission. AND it means that the home has been checked for RF isolation between the CM and the STB and determined not to be at risk.

Going to the lower green flow down the center of Figure 6, this is the case where there is no speed upgrade involved. The spectrum is being turned on to maximize efficient use of upstream capacity. The mid split plant upgrade plans, arbitrarily labeled as "Tier 1" and "Tier 2", are counting on use of this capacity to defer any future network augmentation by many years. So, while it may not be noticeably service impacting to a customer, it is network impacting and indirectly service impacting by lowering the congestion on that node overall.

The upper green flow is the case when a speed tier upgrade request is made, and there is already a Mid-Split capable device present. Because of the fact that an iHAT score is not static, a new iHAT score may make sense to obtain prior to upgrading the customer. The customer expectations for the new service will

be higher, and the awareness acute to service impacting issues, so it is prudent to be certain that the home is still in a "ready" condition. In addition, because the customer now has, for example, a speed tier of 200 Mbps, they will have bursts of energy more likely to utilize a wide chunk of the mid-split band at once, a condition that more aggressively exposes the STB to energy that can cause video degradation. If this "updated" iHAT result is still "PASS," then activation is completed. If not (this is not shown), this home reverts to a Remediation state, and because of the desire for a new service tier, it is a Remediation Queue with a higher priority.

3.5 iHAT Software Solution Implementation

An overview of the innovative iHAT tool process flow is described above. This section contains a detailed deep dive into the automated software solution that was built for iHAT version 2.0. The new version of iHAT has been built exclusively for the devices that exist on a distributed access architecture (DAA) platform only.

There are several aspects of the software that have been prioritized when upgrading to version 2 from version 1.

Invasiveness

The major purpose of iHAT is to test whether a customer can successfully be switched to mid split without impacting their service. To do so, some service test must be initiated and compared to a baseline. Previously, in iHAT version 1, a speed test was run on the customer device. While this provided impressive results and proved to be a robust option, it left something to be desired in terms of efficiency.

The end goal of iHAT, and really any type of automated software that has the possibility of impacting the customer experience, is to lessen that impact as much as possible. To this end, iHAT version 2 has switched to a lightweight OUDP burst in lieu of the more heavy-handed speed test that version 1 relied upon. While running a speed test on customer hardware during peak hours would be out of the question, scheduling a short OUDP burst is not. This opens the possibility of running iHAT tests outside of maintenance windows without impacting the customer experience. The only current barrier to a truly unobtrusive experience is the need for a device reset after the split type has changed on the gateway's diplexer – a limitation that will hopefully be lifted with a future software upgrade.

Scale

To properly scale out software improvements that can cover the entire network footprint, agile software is needed. Historically, network configurations and optimizations were carried out in an ad-hoc method by network technicians. While there may always be a need for humans to make tough decisions about how to go about solving some problems, there are other problems that are easily solved by simple logical automation.

One case study of this can be seen in the Octave project, where cable modem level data was gathered by the genome service and used to regularly optimize the DOCSIS configurations of CMTSs [6] [7]. Comcast as a company is striving to upgrade the network at a pace that cannot be matched by cherry picking small parts of the network one at a time. Therefore, the system that facilitates such sweeping network changes must be able to operate in a seamless manner.

In the case of iHAT, this automation comes at several stages. Version 1 had to manually run; this type of model can only be tenable for a handful of node segments/accounts/devices. In Version 2, node segments on the DAA will be automatically registered for the iHAT test pool when they become capable candidates based on network software and hardware criteria. Once a node segment is flagged as capable, automation is set in motion to filter the CPE devices that are serviced by the node segment to those that are iHAT eligible and testing will commence. MUSL will then periodically run iHAT as necessary based on network and account level changes.

Version 2 of iHAT is built to be flat and scalable. All tests are run at a gateway device level leveraging Amazon Web Services (AWS) step functions and lambdas. This type of architecture should theoretically make the iHAT service scale as needed until it can encompass the entirety of the network footprint. The only bottleneck comes from a DOCSIS limitation that only allows one OUDP burst to be scheduled at a time on a remote PHY device (RPD) level basis. This means that the theoretical maximum number of gateways that can be run per RPD is defined by the (maintenance window time)/(OUDP burst duration) with some padding on either end given for device resets.

Speed

The actual test only takes a few seconds of gathering empirical data and a second or so of computing. The bulk of the time during any given maintenance window is taken up by waiting for device resets, which will ideally be eliminated with hitless split changes sometime in the future.

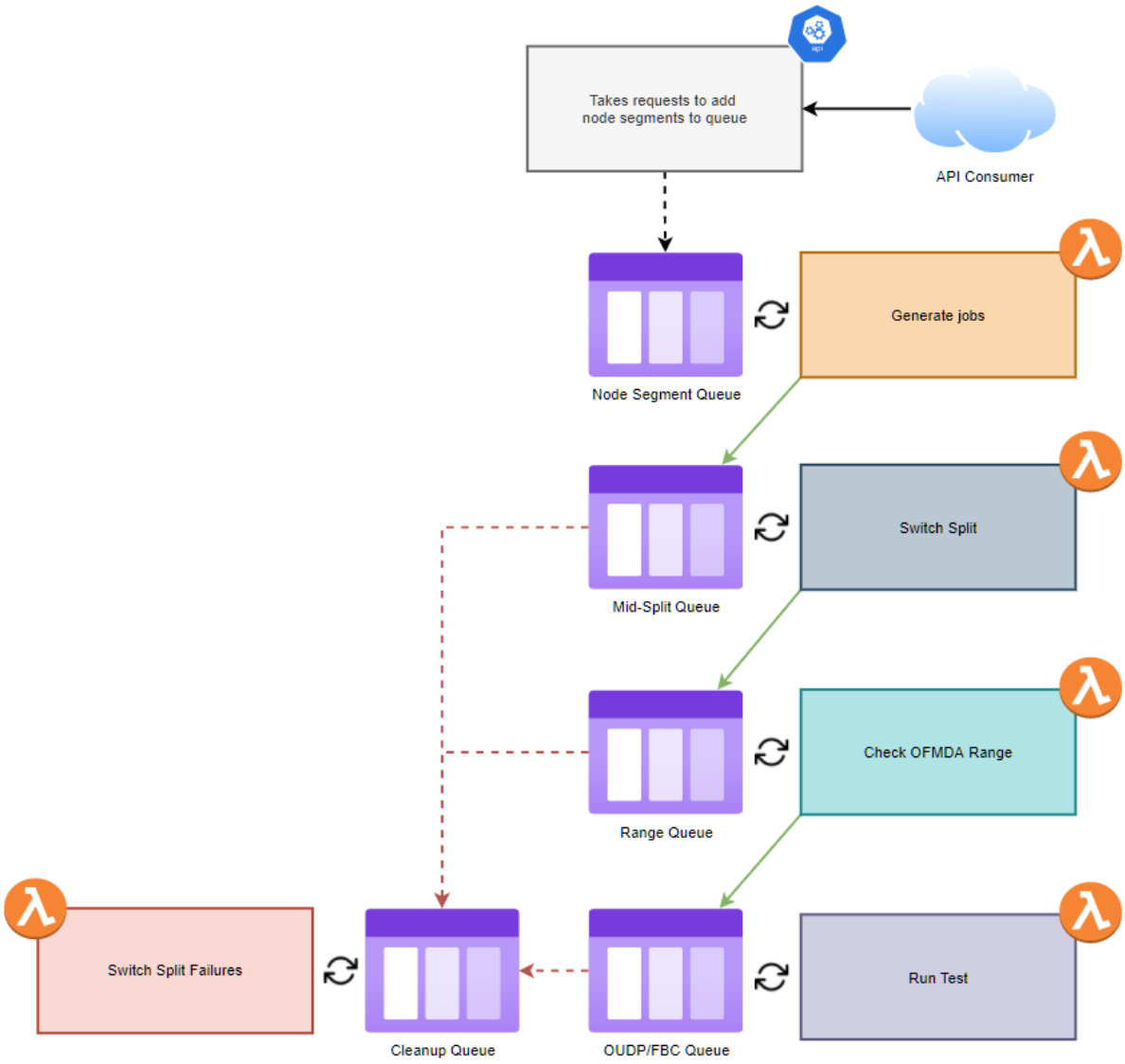


Figure 7 - High-Level iHAT Architecture

Error! Reference source not found. illustrates the high-level architecture that iHATV2 employs to orchestrate mid-split testing at scale. There are 6 distinct phases that are shown here.

Phase 1

Phase 1 is the acceptance phase. This is facilitated by having an application program interface (API) layer that accepts requests for node segment/RPD pairs to be added into the iHAT testing queue to be run in the next maintenance window. This API's sole purpose is to provide an authentication/authorization abstraction layer that wraps the node segment queue.

Phase 2

Phase 2 is the prep phase. This phase starts with spinning up lambda functions using CloudWatch events at the beginning on a maintenance window. Two integrations come into play in this phase. First, the federated data service (FDS) API is called with each node segment name. Node segment names are swapped for MAC addresses that are grouped by service location. At this point, the MAC addresses are opaque, and the genome service must be queried for each MAC address to get device metadata. Genome provides important data such as the model number and current diplexer switch type. The starting switch type is stored so that each device can be placed back in its original state in the case of inconclusive test results.

The model number is then used to classify device types and group devices on an account. Accounts can then be further classified as mid-split eligible based on the kinds and configurations of the devices. Each eligible gateway on an account will generate one "device-level" job. Each device-level job will include one gateway and all set-top-box devices on the account that can produce full band captures (FBCs). Jobs are placed in the switch-split queue for further processing.

Phase 3

Phase 3 is the switch split phase. The phase 3 lambda is responsible for making sure that all modems are switched to mid split mode prior to data collection. If the gateway is already in mid-split mode, then the lambda is responsible for moving the job from the switch-split queue to the range queue, so the cable modem can be immediately processed. If, however, the gateway is not in mid-split mode, then the lambda is responsible for first calling genome to initiate the switch-split into mid-split, and then calling genome again to initiate a modem reset. Once the modem reset has been initiated, the job can be moved from the switch-split queue into the range queue with a 2-minute message timer. This message timer will ensure that minimal resources are spent on waiting for the modem to reset. It is important to note that the rest and delay portion of this lambda may still be required when iHAT switches to a hitless mode of operation.

Phase 4

Phase 4 is the ranging phase. It imperative that the modem is in the correct state before iHAT data collection begins. The range queue that serves phase 4 is the only queue that is not a first-in, first-out (FIFO) queue as it must house two different types of messages. The first type of message represents cable modems that were already in mid split and did not need a modem reset. These messages should appear immediately in the queue as they have no message timer associated with them. The second type of message represents those cable modems that required a modem reset. These messages will only appear after the 2-minute message timer has elapsed. Once a lambda consumes either type of message, it polls genome for a device online status. Once the device is online, then the lambda confirms that the device can range on OFDMA. If the device fails to range, then the modem is sent to the cleanup queue, where it can be switched back into sub-split mode. This type of failure is most likely due to blockage by a drop amp. If, however, everything looks good at the device, --it's online and ranging, the job can be moved to the OUDP/FBC queue with a message ID corresponding to the RPD that the customer gateway is connected to.

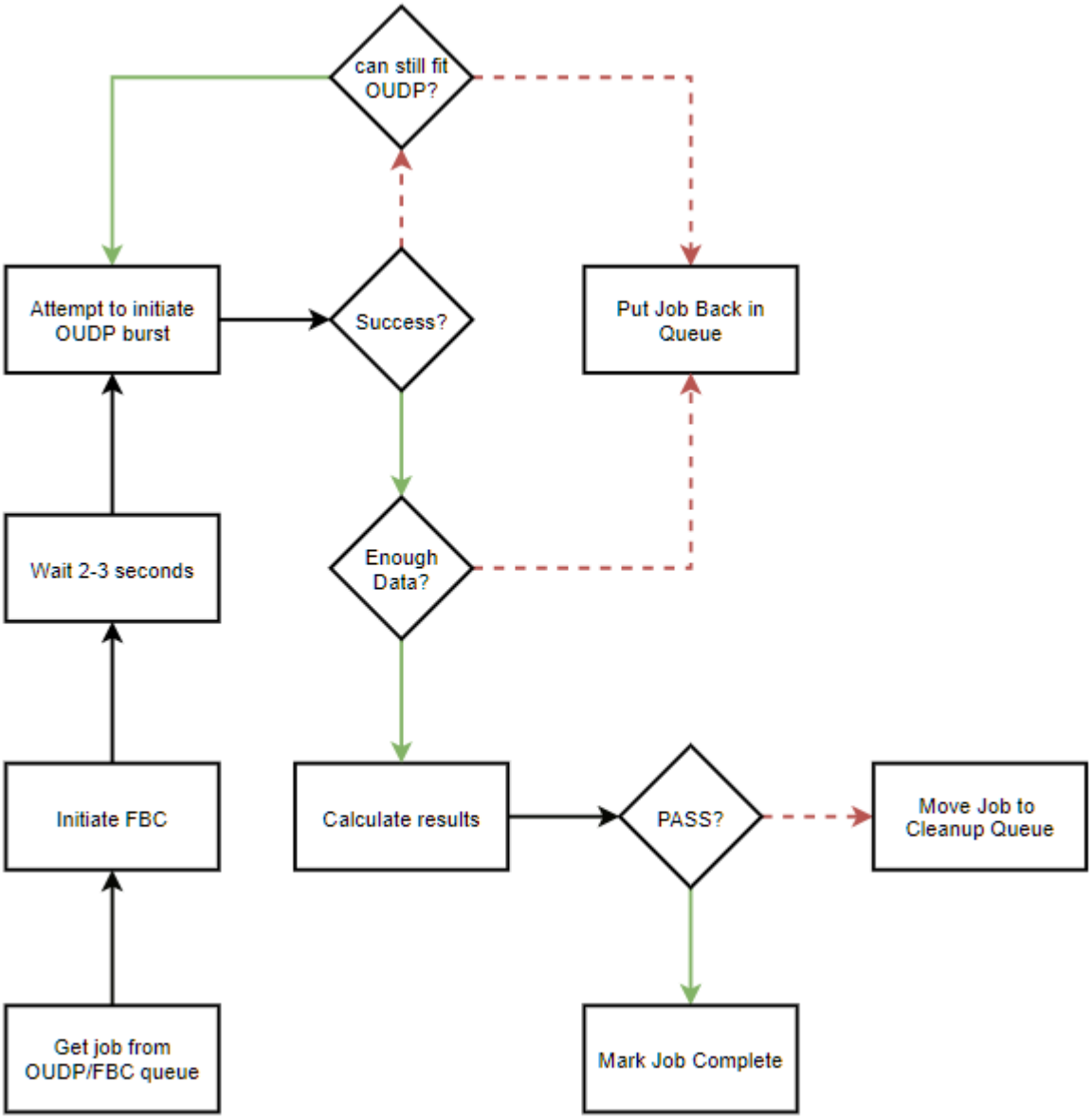


Figure 8 – OUDP/FBC Orchestration Workflow

Phase 5

Phase 5 is where the meat of the iHAT test is performed. A visualization of the workflow is depicted in Figure 8. This phase is slightly more involved than the rest of the phases because careful orchestration of the OUDP burst resource is required. As per the DOCSIS spec, only one OUDP burst on a certain frequency range can occur on a given RPD at one time. This means that one must be very careful when attempting to scale a process that heavily depends on the availability of the OUDP burst service.

The concurrency of this process is controlled by making sure that RPD names are used as message group IDs in the OUDP/FBC queue. In all other queues that exist in the iHAT V2 workflow, the message group ID is the unique value given by the job ID. Amazon simple queue service (SQS) enforces that only one message from a given message group ID is consumed at a given time if the message batch is set to 1. This

means that if all jobs that are associated with a given RPD are associated with the same message group ID, they should execute in serial.

The maximum theoretical concurrency for the data collection process is equal to the number of unique RPDs. Earlier iterations of the architecture tended towards forcing a concurrency equal to the number of unique RPDs in the maintenance window. The fact that a full range of scaling options is available is important, as the iHAT workflow relies on the availability and scale of many other systems to function properly. Being able to calibrate scale to successfully integrate with other systems is a must.

The workflow begins with a lambda firing from the SQS event trigger. First, genome is called to initiate FBC on each associated set top box in the job. This takes 2-3 seconds to initiate, so the lambda sleeps for 3 seconds before attempting to initiate the OUDP burst through the gateway. If the OUDP burst fails, then it is attempted up to 2-3 additional times depending on the FBC duration that was chosen. If it still fails, or if the FBC process does not capture enough data for any of the set top boxes, then the lambda places the job at the end of the queue to be retried later.

If the FBC and OUDP burst orchestration is a success, then the iHAT calculation can commence. This part of the code only takes on the order of a second to complete. If the calculation yields a success, then the OUDP/FBC job is marked complete, and the job has reached the end of the iHAT workflow. If the calculation yields a failure, then the job is moved to the cleanup queue.

Phase 6

Phase 6's purpose is to clean up all failed jobs. It used as the dead letter queue for all queues except the node segment queue. It can also accept jobs from the phase 5 lambda function in the case that a successful test is run, but a failure is noted by the telemetry indicating video service degradation. Categorized failures include, but are not limited to: failure of the device to switch split types, failure of the device to reset properly, failure of the device to range on OFDMA, failure of the OUDP burst to execute properly, and failure due to video noise. It is also possible that the iHAT workflow fails due to underlying software issues, which could be due to network failures or failures of external systems to provide required data.

There are three outcomes of the iHAT workflow. Case 1 is the case where everything is green and the iHAT calculations yield a PASS. In this case, the customer gateway is left in mid-split. Case 2 is the case where the iHAT calculations yield a clear FAIL. In this case, the customer gateway is switched back to sub-split. Case 3 is the case where the iHAT workflow failed and is inconclusive. In this case, the customer device is switched back to its original setting that was recorded in phase 2, whether that is mid or sub-split. Case 2 and 3 may require manual remediation.

The workflow outlined above serves the BAU process of moving customer devices in bulk to mid split and routinely testing them in bulk to ensure network health. There is, however, a separate use case for on-demand testing for customers with a technician in the home. Since the above workflow uses highly modularized components, serving this case is trivial as the individual lambdas can be exposed via API and used by customer care tools. The current use cases are for a technician to move a customer modem from mid to sub-split or vice versa and for an on-demand OUDP/FBC data collection and calculation to yield a score.

The iHAT project requires several integrations with external systems. The one that is used most heavily is the integration with the genome service. The iHAT tool would be much more complicated if it had to deal with authentication and SNMP, but the genome service abstracts those technical details away, allowing for a much cleaner and elegant solution. The next section describes in detail how the genome service achieves this level of abstraction at scale.

3.5.1 Data Collection & Device Interface Platform (Genome)

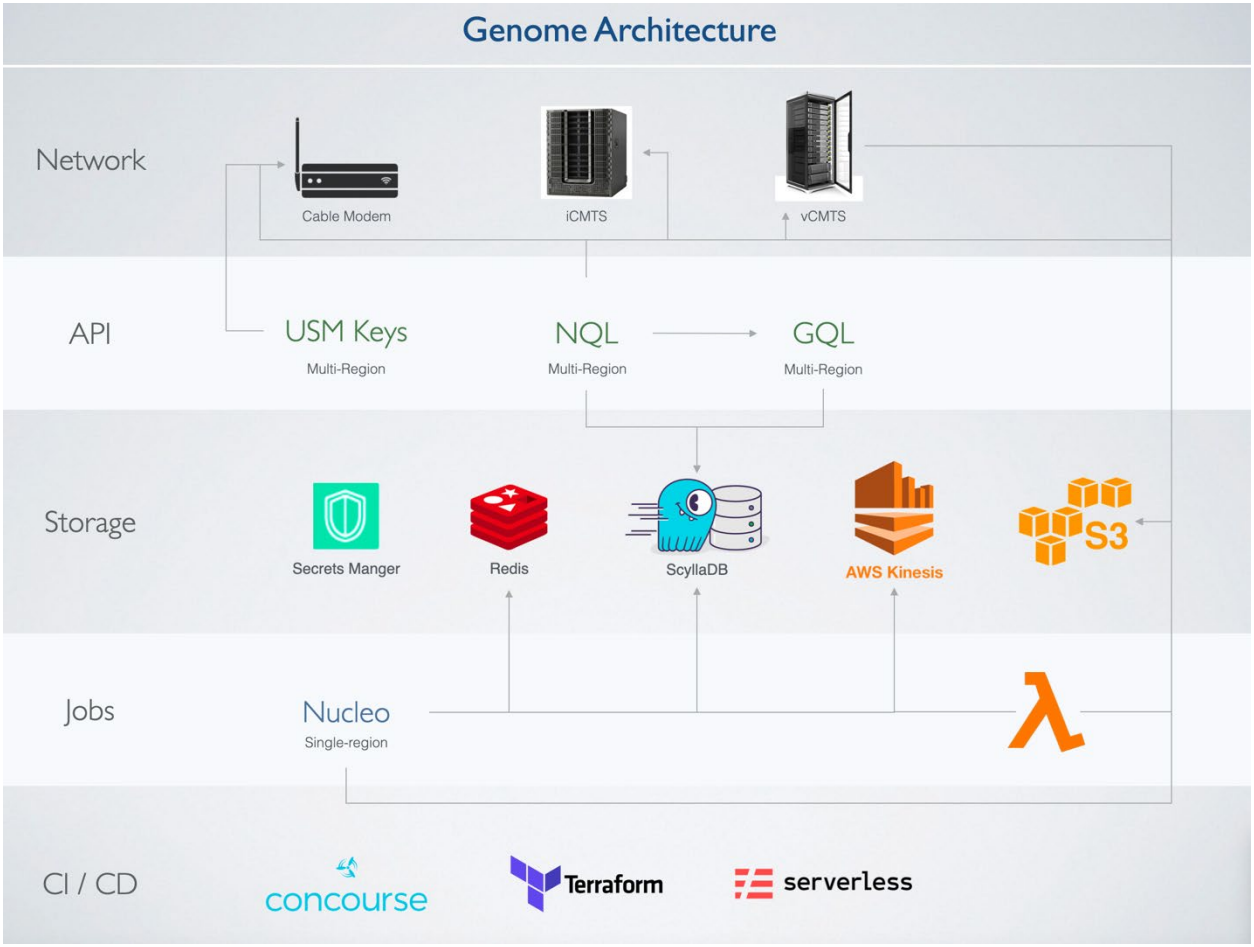


Figure 9 – Genome Architecture

The Genome platform in Figure 9 was originally built to serve the Octave Profile Management Application (PMA) initiative but has grown into a platform of its own since inception [6] [7]. The main purpose of Genome is to actively poll and cache data from devices such as cable modems and CMTSs in a scalable and configurable way while offering consumers the ability to analyze the cached data or get live data in a seamless fashion (on-demand data polling). This positions Genome to be the sole data provider, SNMPV3 secure key collection, data validator and device interface layer for cable modem and CMTS data, eliminating inefficiencies around oversaturating CMTSs or cable modems with multiple connections and over fetching duplicative data. In the context of iHAT, Genome is responsible for verification of eligible accounts and the devices that has been provided by business and their current status, running spectrum analysis while the OUDP has been triggered.

Genome is made up of two layers, the poller layer, and the query layer. A "poller" is simply a piece of software that is responsible for scheduling, collecting, and standardizing a set of data from devices. Pollers are built to be lean, modular, and extensible. The query layer for Genome is called Network Query Language (NQL), which exposes a declarative API service through which consumers request live or cached data. All pollers are primary consumers of NQL to collect live data, while also offering external consumers to do the same. NQL's goal is to offer a declarative abstraction layer for edge network devices, allowing

consumers to query using standard hypertext transfer protocol secure (HTTPS) instead of simple network management protocol (SNMP), trivial file transfer protocol (TFTP), and various other network communication protocols.

Genome requires a list of eligible devices to poll the necessary data for iHAT. iHAT integration layer works as the mediator between business system which holds the list of eligible Nodes/Accounts and verifies the OFDMA being active, along with mid-split capability and eligibility, then feeds Genome with this eligible device list to check the device status to make sure they are in the right state to run iHAT. The polling could be done at any configured frequency or on-demand.

In addition, Genome must ensure that data collection can scale when devices are added over time such that data collection and aggregation can be achieved in a given polling window. It must also ensure that the ingested data is validated and cleaned before it is cached. As the amount of data is large, especially in the case of cable modem data, Genome must also manage data retention policies in order to reduce cost.

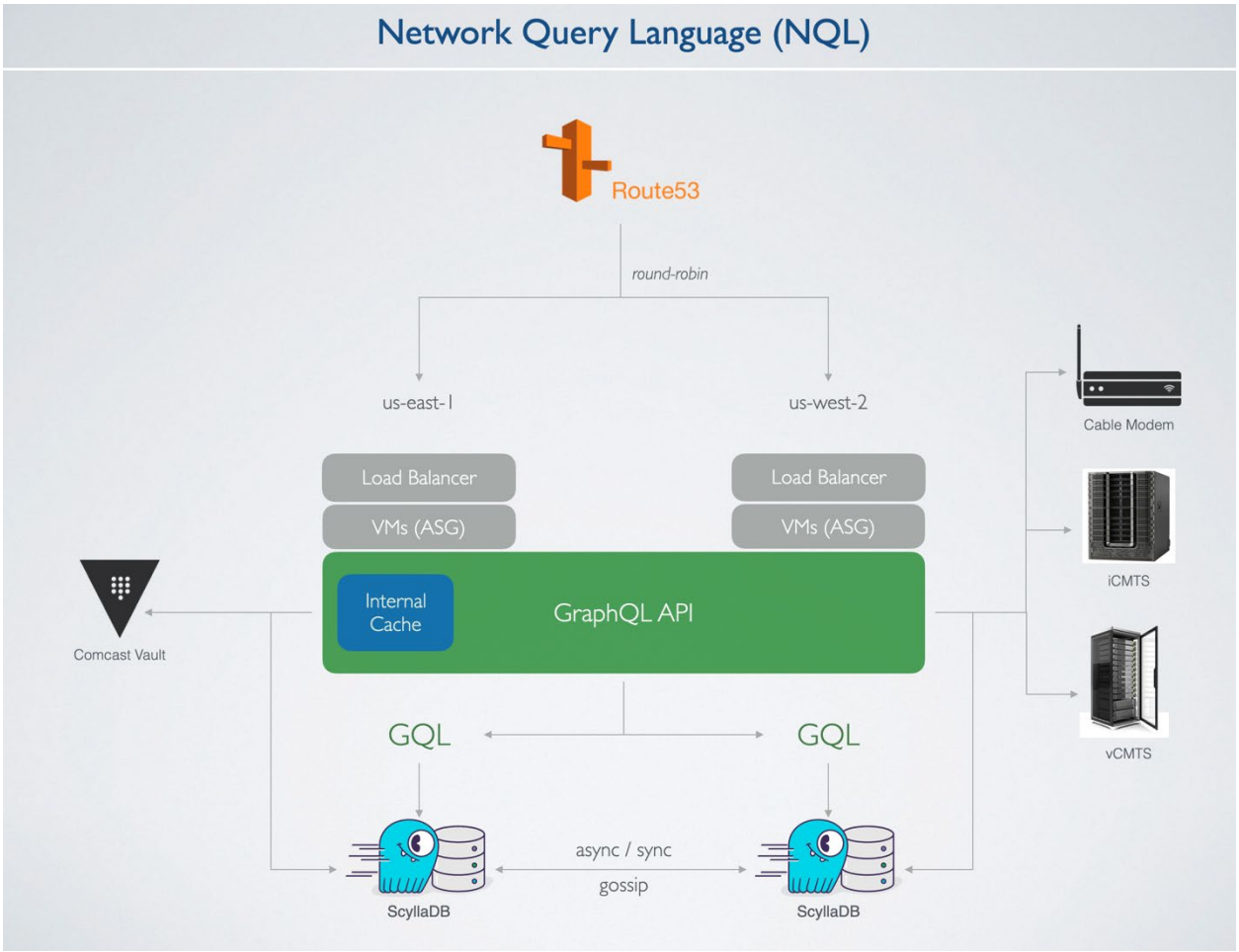


Figure 10 – Genome API – NQL Architecture

NQL's journey began in the early days of Octave development as well. The goal was to create an API service which abstracts away all different protocols around networking and let end users interact through HTTPS. The first iteration came in the form of a representational state transfer (REST) API which accepted object identifier (OIDs) parameters and returned the output through HTTPS. While this iteration ran in

production for several months, several shortcomings were brought to light. For example, use cases involving many data points per CMTS or CMs, necessitated making multiple requests to the API. Each request would connect to the CMTS, which resulted in opening and closing sockets many times over multiple requests. Furthermore, the work of encoding and decoding large amounts of javascript object notation (JSON) data was repeated for each request, degrading the overall performance of the service through repetition of work. Graph query language (GraphQL) was a natural option that allows data schemas to easily evolve and can handle complex relationships between data sets. NQL architecture in Figure 10 was born.

As NQL offers an abstraction layer for connectivity to both cable modems and CMTSs, it must therefore manage all writes as well as reads to each network device. In order to accomplish such a feat, it supports both IPv4 and IPv6 communication protocols. NQL is built primarily using the GraphQL specification and has been optimized at node level. In some cases, consumers may require a request-response type handshake, where the consumer keeps a socket open until the requested data is returned. However, many queries may be long running, and it may not be practical or possible for a consumer to keep a connection open for the duration of the process.

Historically, the team used Nodejs as a declarative jack-of-all languages. However, in this use case, performance is king. As Golang is routinely touted as a performance-oriented language geared towards networking applications, it was a simple choice after orchestrating some internal benchmarking. Genome as of today supports 3+ billion transactions a day with ability to scale up as we need. Simple GraphQL APIs were built in both languages to test a small fraction of our total scale.

NQL's secure shell (SSH) feature allows consumers to log into a supported remote device through HTTPS rather than SSH. This allows NQL to abstract away details around authentication, authorization, and managing the underlying SSH connection. Users are able to connect to a host, run multiple commands, and get output back for each command. NQL uses GraphQL to define the API, which allows developers to develop powerful features and evolve our API without breaking the world. NQLs SSH is a simple layer around SSH, all the logic around what commands to run in what order and what to do with the output is handled by the client itself. For example, configuration management can apply configurations to a CMTS and then analyze the output to see if everything went well. All this is done through HTTPS using NQL. NQL has evolved from a standalone service running on an elastic compute cloud (EC2) virtual machine (VM) to where it could be leveraged anywhere within our codebase and able to run in any container or lambda alongside our codebase.

SNMP v2/v3 capability and the user-based security model (USM) are shown in Figure 11. Genome by default uses SNMP V3, and if the device is not V3 enabled, the fallback option is V2. The role of USM is as follows:

- Get public and manager key from cable modem
- Use manager key and V3 security name as input to get vault secret
- Use cryptographic hash function on vault secret and public key and compute auth priv keys.

With authentication in place, collection of the SNMP data can be performed on an as needed basis and streamed to Kinesis, which will be used by Analytical Engine from PMA and other consumers of the data including iHAT. iHAT uses NQL layer as the interface to do the SNMP set for moving the device from sub to mid-split.

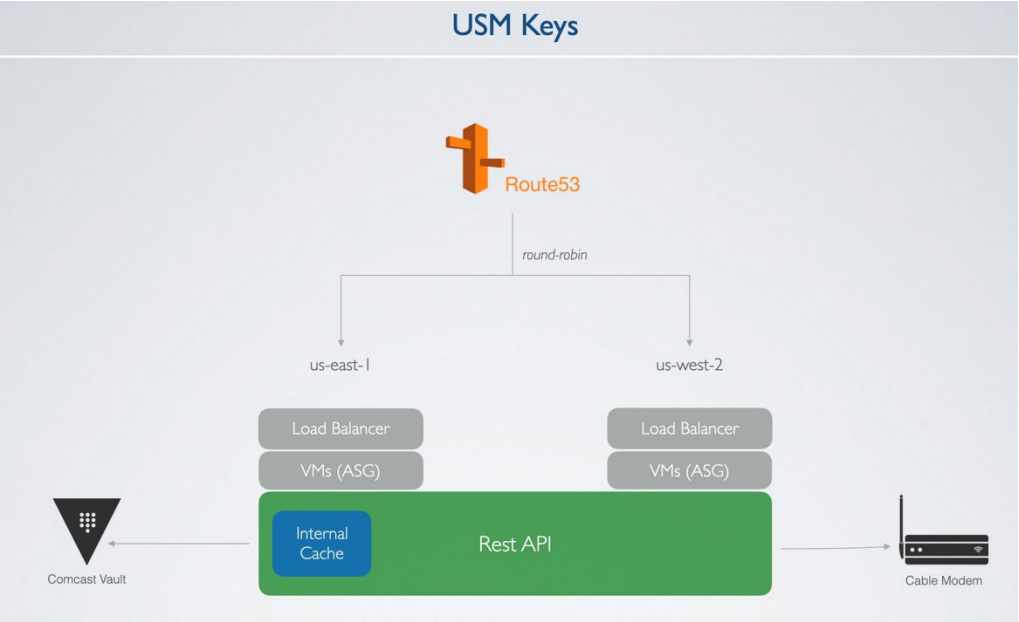


Figure 11 – SNMPv3 USM key architecture

This section has provided a detailed view of the iHAT software architecture covering genome, NQL and USM. Overall, the architecture is like previously presented PMA architecture [6] [7], with many software components being reused to support iHAT functionality. With an understanding of the iHAT operation for mid-split, we would like to now discuss future enhancement considerations for expanding upstream path to high-split and what new functionality would need to be considered for future generations of iHAT software.

4. High-split Future

Plant upgrades with mid-split are just the beginning of extending the upstream spectrum and increasing the upstream throughput capability. Mid-split spectrum activation adds needed capacity in the upstream and extends the life of node splits. To achieve 1 Gbps symmetrical services significantly more upstream spectrum is needed. High-split development and deployments are in process and FDX is right around the corner to achieve one Gbps and greater symmetrical services.

Throughout the development of mid-split, there were many aspects of the access network which were evaluated and many challenges which were overcome. Many of the workstreams and lessons learned in mid-split extend to high-split development and deployment.

OFDMA activation, node and amplifier upgrades, the shift of the downstream spectrum to allow room for the extended upstream spectrum, adjacent channel interference, plant design levels, CPE devices, in home amplifiers and splitting networks were all extensively evaluated and validated for mid-split deployments. The learnings from the development were and continue to be used for high split.

Comcast's network upgrades have predominately been N+0 and recently include amplifiers (N+X) for mid-split deployments. The process of splitting and upgrading a standard-split analog node to mid-split digital nodes is well understood and completed many times a day throughout our networks. Creation and push of the downstream channel map, upstream channel configuration and upgrade of plant equipment; nodes and amplifiers all happen simultaneously and with minimal disruption to our customers. This same

process can be used for mid to high split upgrades. Mid split nodes amplifiers are 1.2 GHz, so there's plenty of spectrum available to move the downstream spectrum above 258 MHz prior to high split equipment installation. Tools, like Comcast's Scout Monitoring, have been developed to monitor and validate many physical layer parameters, including OFDMA in the upstream, and these have been extended to 204 MHz in the upstream. Examples of these tools are shown below. Figure 12 and Figure 13 show the upstream DOCSIS 3.0 and 3.1 channel status via Scout.

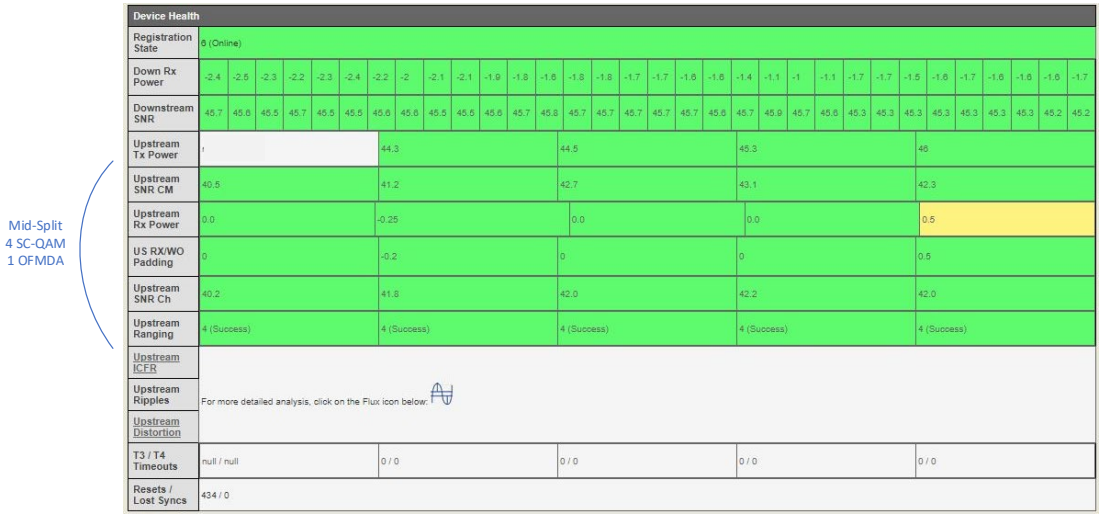


Figure 12 - Mid Split Scout Screen Capture with 4 SC-QAMs and 1 OFDMA

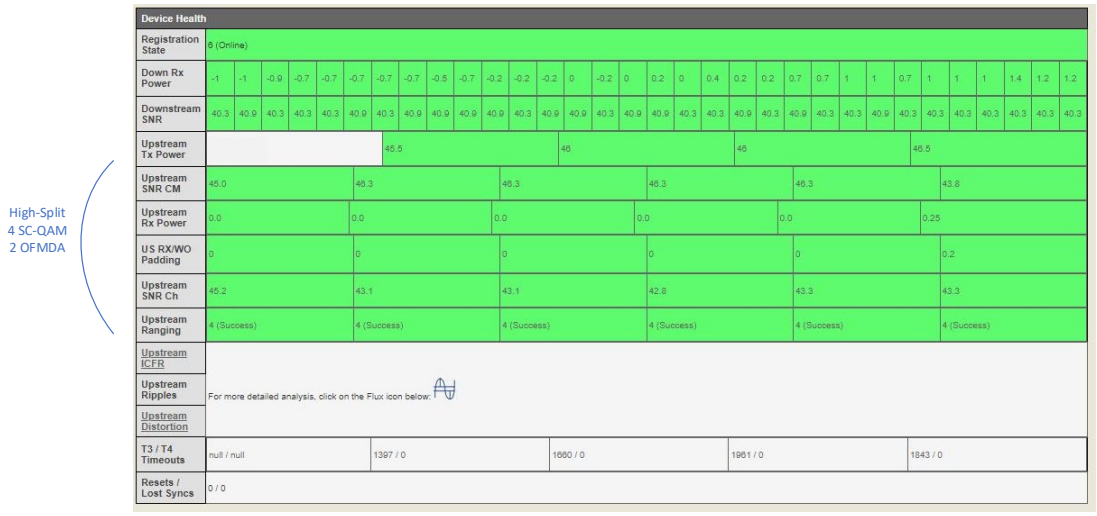


Figure 13 - High Split Scout Screen Capture with 4 SC-QAMs and 1 OFDMA

Figure 14 and Figure 15 show the upstream spectrum as received at the RPD via Yeti for mid split systems to 85 MHz and high split systems to 204 MHz. Yeti is a "real time" upstream spectrum analyzer tool used to evaluate the upstream channels and also the underlying upstream noise.

Figure 14 - High Split Yeti Screen Capture at RPD with 4 SC-QAMs and 1 OFDMA Channels

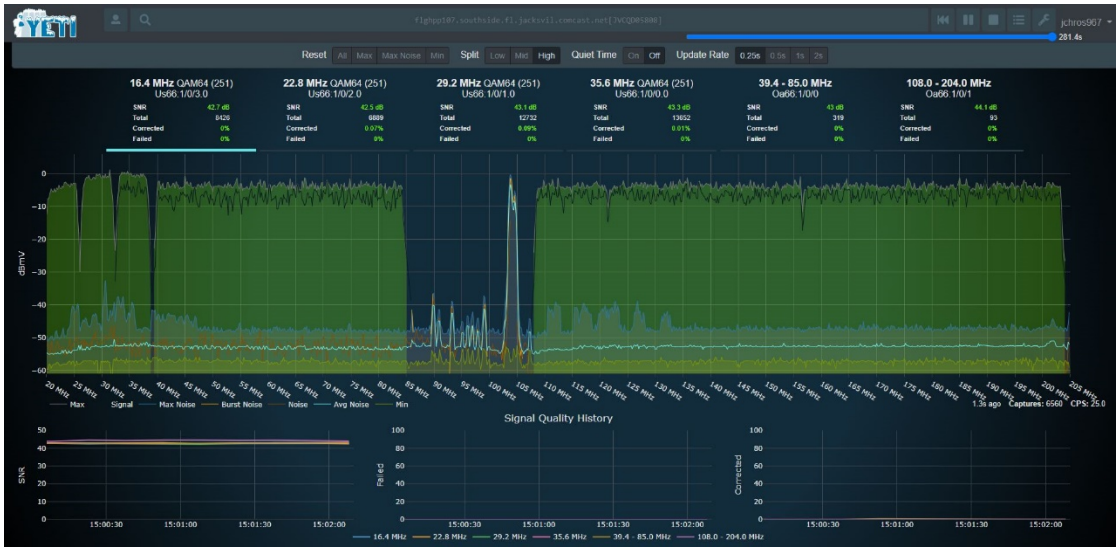


Figure 15 - High Split Yeti Screen Capture showing 2 OFDMA Channels

Note that the lower spectrum edge starts at 20 MHz due to current FFT limitations

4.1. Extending iHAT for Neighbor Interference Case

ACI has been studied extensively for mid-split plants. Thresholds of interference for our set-top boxes are well understood and as discussed iHAT development is well under way for mid-split deployment to analyze ACI with no customer impact. iHAT first started out exercising the full mid-split upstream band with a speed test and the adjacent in-home set-top box performance was monitored during the speed test. This was a very effective way to measure the impact of the mid-split spectrum on the adjacent in-home set-top boxes. The speed test could interfere with customer high speed data usage and interference was possible on the set-top box, if the in-home isolation between the gateway and set-top box was inadequate.

During the development of high-split, a method of evaluating leakage was needed where the high split gateway in the home creates a leakage tone. DOCSIS 3.1 supports OUDP signals and APIs have been developed to have the vCMTS command a cable modem to create an OUDP burst to be used for leakage detection. These same OUDP bursts are being used to evaluate adjacent channel interference. The OUDP burst is created by the mid-split gateway and a simultaneous full-band capture is taken on the adjacent in-home set-top box. Through extensive testing the thresholds for the set-top boxes for interference are well known and the measurement of the delta between the OUDP burst as measured on the adjacent set-top box and the downstream signal is measured to evaluate the potential for ACI. If the delta is less than the threshold the account can support mid-split spectrum activation. If the delta is more than the threshold, the account is targeted for remediation.

This same process can be used for high split systems to evaluate the potential for high split neighbor interference. In the mid-split plant, there is enough tap-to-tap isolation and the interfering bandwidth is smaller, so neighbor interference is not a concern. In a high-split system, the extended upstream bandwidth coupled with lower tap-to-tap isolation with higher frequencies creates the potential for neighbor interference. See Figure 16.

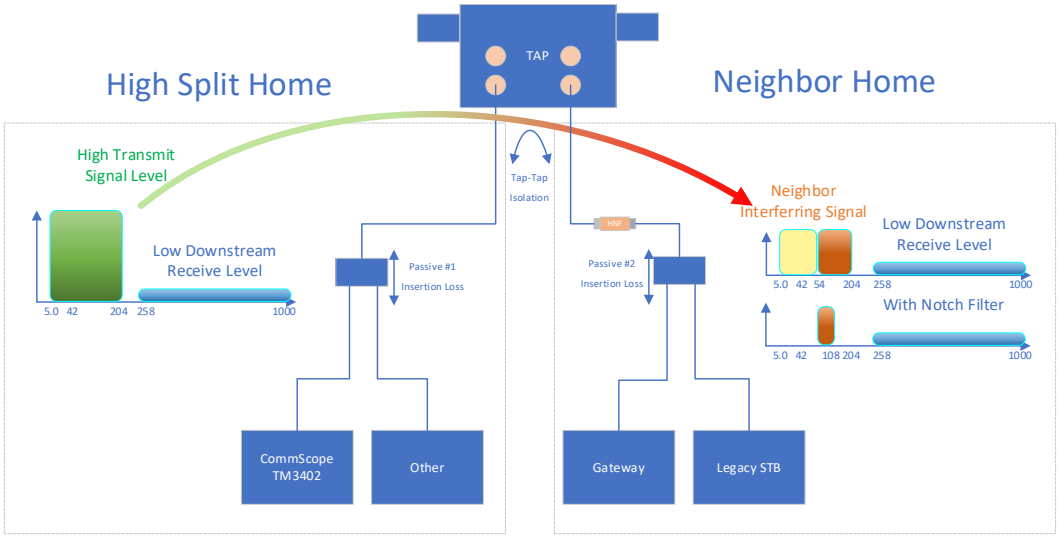


Figure 16 - Neighbor Interference in a High-Split System

The same process the iHAT tool uses with the OUDP burst can be used to evaluate high-split neighbor interference. When high split modems are first deployed and before service is activated, iHAT can be used to create the OUDP burst on the high split device and the neighboring standard and mid-split devices can be evaluated using the same full band capture technique used for mid-split ACI. Additional system and plant information is used to determine which devices are connected to the same tap as the high split device and these devices are monitored.

In a Mid-Split system, ACI and iHAT are focused on adjacent set-top boxes in the same home as the mid-split gateway or cable modem. In a high-split system, neighbor interference can occur both on neighboring set-top boxes and neighboring gateways and cable modems. As in the mid-split case, extensive testing has been completed to understand the interference and threshold levels for high split interference both on the set-tops and gateways.

Testing was completed on both set-tops and gateways with a “Neighbor” high split cable modem passing traffic under various upstream throughput scenarios. This setup for neighbor interference testing is shown in Figure 17.

For neighbor video interference testing there are two main test scenarios

1. Video: system connected to DAC for the STB testing
2. Data: connected to CMTS for CM testing

The same set-up was used for testing both Data and Video (QAM) CPEs. The first vCMTS is connected to the HS node to generate the upstream load to the HS modem is configured for interference as follows

- 4 x 6.4MHz SC-QAM channels (for the SS CMs only)
- 48 MHz OFDMA (up to 85MHz, MS)
- 96 MHz OFDMA (108-204MHz, HS)
- US Traffic and Level adjusted to produce the desired interference signal

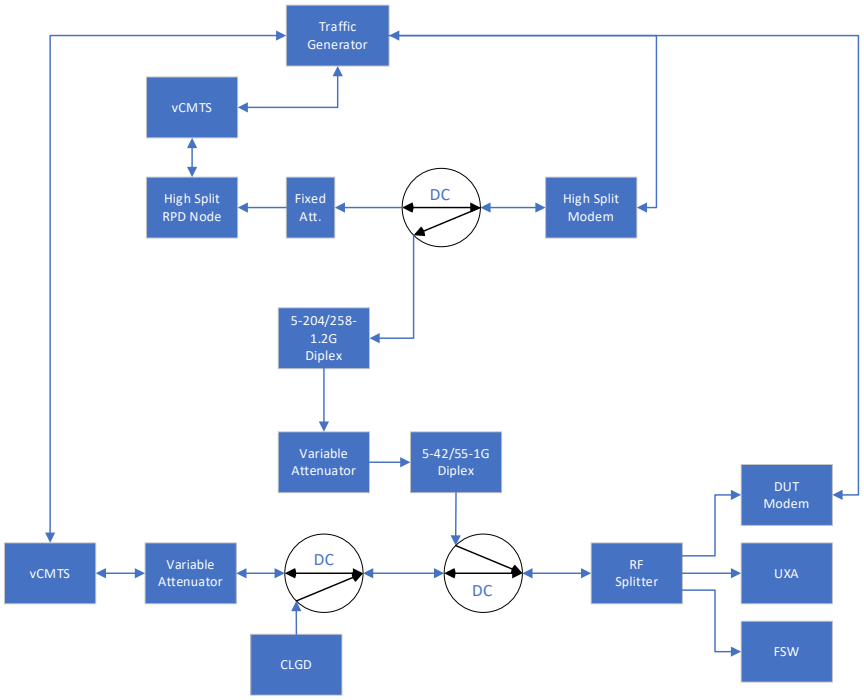


Figure 17 - Neighbor Interference for STB Test Configuration

The DUT modems are configured as follows:

- CM DS Load (258 – 1002MHz):
- OFDM (96MHz, 792 to 888 MHz)
- SC-QAM (603 to 789 MHz, Symbol Rate: 5.360537MSym/s, 256QAM)
- SC-QAM above and below those frequency ranges from CLGD
- Add noise to the DS to decrease the received MER to 38 dB
- A traffic generator is used to produce a fixed DS traffic for the DUT (300Mbps data rate) while the HS data rate was varied from 1 to 980 Mbps

- STB DS Load (258 – 1002MHz)

The Interference Level is defined as the ratio of the high-split Interference power spectral density (PSD) (dBmV/6MHz) with respect to the device under test (DUT) downstream PSD (dBmV/6MHz). It is important to note that the high-split Interference ratio takes in consideration the occupied bandwidth for each high-split interfering traffic rate which is CMTS scheduler dependent. The methodology for determining the High-Split threshold interference was based on a measured DUT DS codeword error rate (CER) above a threshold of $1e^{-6}$ for the Data scenario while an observer noticing video artifacts was used for the Video scenario. The variable attenuator controlling the high-split Interference level applied to the DUT is adjusted for each particular interference high split traffic rate until the CER threshold is reached or video artifacts are observed.

When testing interference low upstream throughputs correspond to low OFDMA utilization and/or a low duty cycle for the OFDMA channels occupying the band from 39.4-204 MHz. As upstream traffic rates are increased, the OFDMA channels are utilized more and the duty cycle is reduced, up until the point that the full spectrum from 39.4-204 MHz is being utilized 100% of the time. The threshold of interference under this testing shows that under very low utilization the interfering threshold is high. As traffic increases, there is a point where the interference threshold is very low and then as the traffic is maximum, the upstream spectrum is at a steady state and the interference threshold rises. This can be seen in both the interference of set-tops and in gateways, as shows in Figure 18.

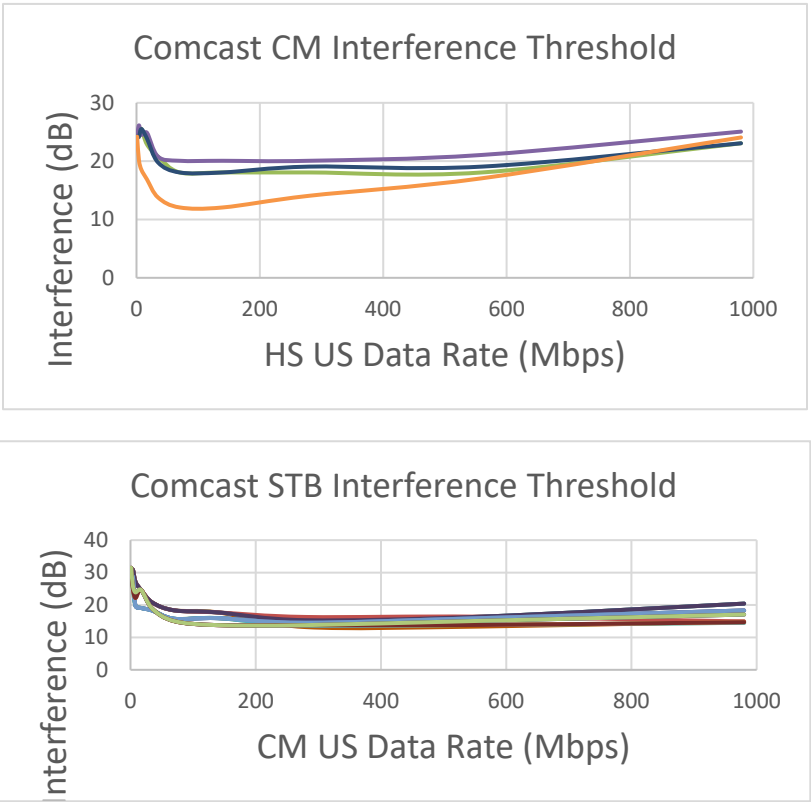


Figure 18 - Cable-Modem & Set-top Box Interference: Threshold vs Neighboring High Split Throughput for Several Different Models of Set-tops

A mathematical model has been developed to better understand the DUT front-end behavior and identify opportunities to improve the CM/STB performance. The goal is to capture the impact on the CM front end including front end overhead, automatic gain control (AGC) behavior, Downstream/Upstream power ratio, high-split upstream occupied bandwidth, and duty cycle. A typical CM front-end behavior is well described in reference [5].

Figure 19 shows the comparison of the model with data measured for a Comcast CM. It also shows the main regions of a typical CM or STB behavior.

- Low Energy Region: range with low interfering signal utilization (occupied BW and duty cycle) requiring higher levels of interference signal to impact the victim CM front end
- Maximum Interference Region: range where the increased occupied BW and duty cycle produces the highest levels of interference and corresponds to minimum interference thresholds
- Quasi Steady-State Region: range where the high interfering signal utilization gets close to a continuous mode requiring higher levels of interference signal to impact the DUT CM front end

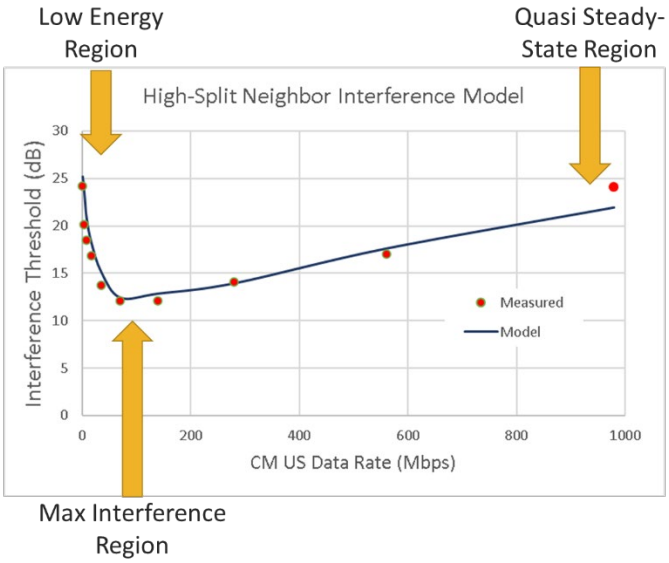


Figure 19 - High Split Threshold vs Neighboring High Split Throughput Model

It is important to note that the maximum interference frequency range is dependent of the CM front-end and CMTS scheduler characteristics. In the case of the CMTS scheduler, grant allocation implementation, which translates into a certain upstream occupied bandwidth and duty cycle, defines the interaction with the CPE front end.

Additionally, for system design purposes, the minimum threshold level is used since the upstream data demand varies all the time. Comcast measurements have indicated a minimum threshold levels in the order of 12 dB which is a little better than the “Maximum average power of carrier input to CM, within any 6 MHz channel from 54 MHz up to 1002 MHz” and “256-QAM Image Rejection Performance” PHYv3.0 DOCSIS specs, which limit the interference level to 10 dB. New case scenarios could be developed and proposed to the CPE chip-set manufacturers to further improve the CPE performance, particularly as the industry starts to get ready for FDX. However, tools like iHAT are critical for a smoother transition.

If high OUDP signal to downstream levels are measured and potential neighbor interference is identified, several remediation techniques can be implemented. One of the potential solutions is the use of notch filters that can be added to neighbor drops and suppress the interfering signal. Although not operationally desirable, these filters can be designed to notch out the complete spectrum from 54 to 204 MHz or to be compatible with mid-split systems, notch filters have also been designed and tested to pass up to 85 MHz and notch the spectrum between 85 and 204 MHz. Figure 20 shows the frequency response of a notch filter used to prevent the impact of the High-Split neighbor interference. Filter suppression higher than 20 dB would suffice for most case scenarios.

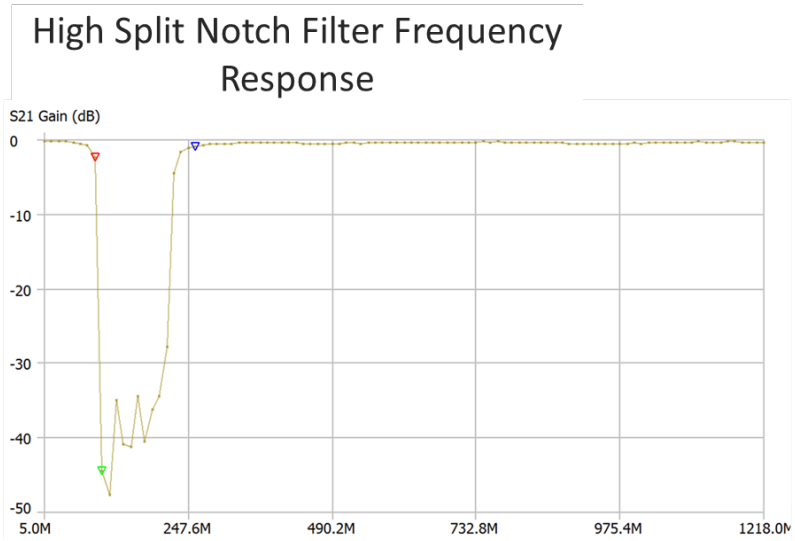


Figure 20 - Neighbor Interference Notch Filter 85-204 MHz)

This allows mid-split and high split devices to co-exist and minimizes device types deployed in the field. New tap plates can also be installed. Newer taps are specified with better tap to tap port isolation to help minimize neighbor interference.

With minor updates to the existing iHAT tools for ACI, iHAT can be adapted to solve the high-split neighbor interference. The threshold of the OUDP to downstream signal can be adjusted to match the measured interference levels for high-split interference, and the back-office tools modified to run/measure OUDP on different accounts. The frequency of the OUDP burst and duration can be optimized for certain areas of the band based on the isolation measurements and the duration may also be optimized.

iHAT and OUDP measurements are an ideal tool to validate the high-split system prior to activation to optimize the customer experience.

5. Conclusion

New work and learning have continued since the introduction and subsequent rebranding to iHAT this past year. Aside from the cool name, improvements upon the process of evaluating ACI using spectrum-based methods have been made that will one day enable OUPD leakage detection to occur in a non-service affecting manner, harmoniously with regularly scheduled traffic. As we work toward scaling iHAT, our intent is to focus on making this tool work as quickly as possible, while offering robustness of carrier grade software at scale using proven tools readily available in the marketplace. A well-defined

object model acting as the glue between iHAT process and downstream APIs can lead to the scale operations needed at Comcast, and reliable distribution of results into the hands of frontline teams who can act in a more proactive manner to address remediation issues. With an eye on the future, including high-split, iHAT has room to grow into a toolset that will one day make it easier for operators to operationalize 1 Gbps upstream speeds. It is our hope that these steps in upstream evolution will help us to embrace new DOCSIS 4.0 FDX sounding technology when the time is ready, and work to harmonize new technologies while always maintaining support for legacy service.

Abbreviations

ACI	Adjacent Channel Interference
AGC	Automatic Gain Control
API	Application Program Interface
AWS	Amazon Web Services
BAU	Business as Usual
BW	Bandwidth
CER	Codeword Error Rate
CM	Cable Modem
CMTS	Cable Modem Termination System
COS	Class of Service
COVID-19	Corona Virus Disease 2019
CPE	Customer Premise Equipment
DAA	Distributed Access Architecture
dB	Decibels
dBc	Decibels Relative to a Carrier
dBmV	Decibels Relative to one Millivolt
CER	Codeword Error Ratio
DAA	Distributed Access Architecture
DAC	Digital Addressable Controller
DOCSIS	Data Over Cable System Interface Specification
DS	Downstream
DUT	Device Under Test
EC2	Elastic Compute Cloud
FBC	Full Band Capture
FEC	Forward Error Correction
FDS	Federated Data Service
FDX	Full Duplex DOCSIS
FIFO	First-In, First-Out
Gbps	Gigabit per Second
GraphQL	Graph Query Language
HS	High Split
HSD	High Speed Data
HTTPS	Hypertext Transfer Protocol Secure
HUSL	High-Split Upstream Spectrum Launch
iHAT	in-home Health Assessment Tool

ITG	Interactive Troubleshooting Guide
JSON	Javascript Object Notation
MER	Modulation Error Ratio
MDU	Multiple Dwelling Unit
MHz	Mega Hertz
MoCA	Multimedia over Coax Alliance
MS	Mid Split
MS-CPE	Mid-Split Customer Premise Equipment
MUSL	Mid-Split Upstream Spectrum Launch
NQL	Network Query Language
OFDMA	Orthogonal Frequency Division Multiple Access
OID	Object Identifier
OOB	Out-of-Band
OP2OP	Output Port to Output Port
OUDP	OFDMA Upstream Data Profile
PDMP	POE DOCSIS MoCA Passive
PHT	Performance Health Test
PMA	Profile Management Application
PoC	Proof of Concept
POE	Point of Entry
QAM	Quadrature Amplitude Modulation
PSD	Power Spectral Density
REST	representational State Transfer
RF	Radio Frequency
RFC	Remote Feature Control
RHM	Remote Health Monitoring
RPD	Remote PHY (DOCSIS physical layer) Device
SC-QAM	Single Carrier Quadrature Amplitude Modulation
SIK	Self Install Kit
SNMP	Simple Network Management Protocol
SNR	Signal-to-Noise Ratio
SS-CPE	Standard-Split Customer Premise Equipment
SSDA	Standard Split Drop Amplifier
SSH	Secure Shell
SOT	Source of Truth
SOD	Source of Distribution
SQS	Simple Query Service
STB	Set Top Box
TC	Trouble Call
TFTP	Trivial File Transfer Protocol
US	Upstream
vCMTS	Virtual Cable Modem Termination System
VM	Virtual Machine
XOC	eXcellence Operations Centers

Bibliography & References

[1] Cable Television Laboratories, Inc., Data Over Cable System Interface Specifications, DOCSIS© 3.0, Physical Layer Specification, https://specification-search.cablelabs.com/CM-SP-PHYv3.0, December 7th, 2017

[2] Cable Television Laboratories, Inc., Data Over Cable System Interface Specifications, DOCSIS© 3.0, Physical Layer Specification, https://specification-search.cablelabs.com/CM-SP-PHYv3.0, December 7th, 2017

[3] Cable Television Laboratories, Inc., Data Over Cable System Interface Specifications, DOCSIS© 4.0, Physical Layer Specification, https://www.cablelabs.com/specifications/CM-SP-PHYv4.0, April 29th, 2020

[4] L. Zhou, "A Proactive Network Management Scheme for Mid-Split Deployment", SCTE Virtual Expo, 2020

[5] S. Shulman, "Operating Legacy Cable Modems in an FDX Environment", 2019 Fall Technical Forum, SCTE-ISBE CableTec Expo, 2019

[6] R. Vugumudi, "Dynamic Data Collection and Configuration Management", SCTE Virtual Expo, 2020

[7] M. Harb, "Lessons Learned from Deploying the Profile Management Application System at Scale and Considerations for Expanding the System Beyond OFDM", SCTE Virtual Expo, 2020

Reducing the Cost of Network Traffic Monitoring with AI

Petar Djukic
Director AI & Analytics
Ciena Canada
Ottawa ON, Canada
pdjukic@ciena.com

Maryam Amiri
Lead AI Engineer
Ciena Canada
Ottawa ON, Canada
maamiri@ciena.com

Wade Cherrington
Software Engineer
Ciena Canada
Ottawa ON, Canada
wcherrin@ciena.com

1. Introduction

This paper describes three visionary approaches to reduce the amount of network telemetry. The problem we discuss is that internet protocol (IP) network monitoring requires collection and storage of a large volume of data. This is a big issue for service providers (SPs), but the solution is usually not thought of outside of conventional approaches. With the advent of artificial intelligence (AI) approaches, especially in estimation, interpolation and imputation, new solution avenues are becoming available.

IP network monitoring often implies the use of NetFlow or Internet Protocol Flow Information Export (IPFIX). However, as we show shortly the two alone are not enough to cover all use cases. NetFlow and IPFIX can be used to turn the network into a large collection of sensors collecting information about IP traffic, which can be used to monitor network usage, identify misconfigured network elements, identify compromised network end points and detect network attacks (Santos, 2016). However, high fidelity network monitoring with technologies such as IPFIX comes with many challenges due to the amount of data that is collected by network elements, then transmitted to where it can be stored and finally processed to get the insights that the operator is looking for. IPFIX also only gives a partial view of the network state and additional technologies are needed to build a complete view of the network.

1.1. Overview of IP Network Monitoring

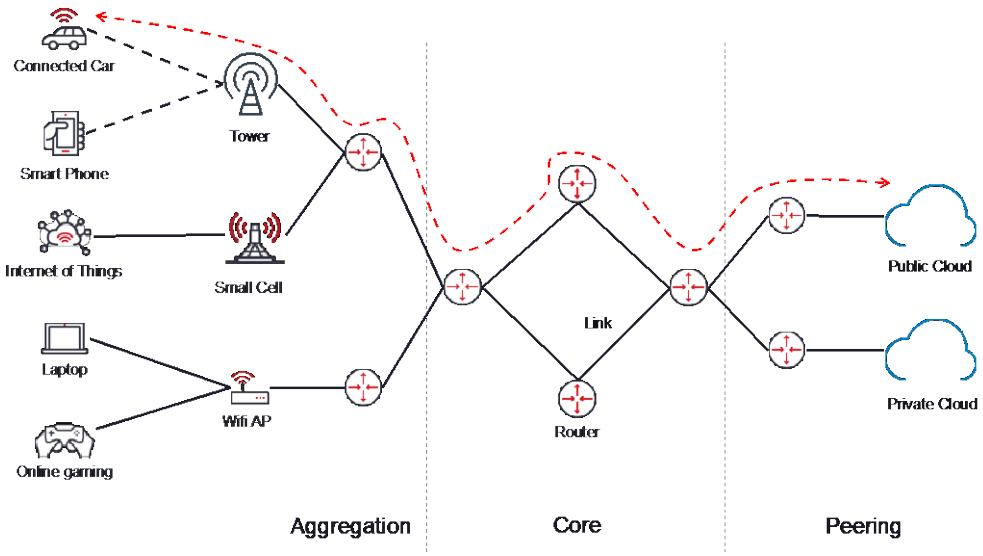


Figure 1 - IP Network

Figure 1 shows a simplified monitored IP network architecture. An IP flow is shown in red and refers to a distinct set of packets occurring in the same period that share a 5-tuple IP header (destination IP address, destination port, source IP address, source port and type-of-service). Aggregate flows, which combine traffic between points in the network may be used in core networks where there are too many flows to track individually. Flows traverse many network

elements, but as the focus of this paper is the IP domain, routers and links are the important part of the figure. Routers and links are both physical elements, however they can have appropriate "digital twins" in network monitoring/inventory software, which is used to create a logical network topology.

There two types of network monitoring: active and passive (Network monitoring, n.d.). In this paper, we are considering passive approaches, which are based on measuring network information without disturbing existing network traffic. Passive IP monitoring can be done at an IP flow level, link level, network level or router level. Active approaches include injecting network packets to probe the network. Some of the approaches for network probing are MTR (Wikipedia, n.d.), based on Internet Control Message Protocol (ICMP) (Wikipedia, n.d.), Iperf (Wikipedia, n.d.), based on a proprietary packet generation and application level protocols, or others, based on RFC 2544 (Wikipedia, n.d.). Active network monitoring can benefit from the applications of AI shown here, but due to space and time limitations these are not discussed.

The goal of IP network monitoring is to get an accurate enough picture of the network state, so that network operators can implement their operational use cases (Quittek, Zseby, Claise, & Zander, 2004). IP network monitoring addresses many use cases:

- *Usage-based accounting* is a business model for selling IP services. A user or a user group is charged based on how much traffic was transmitted. For example, this traffic usage model is used by Amazon Web Services (AWS) to charge for transmitting data out of the cloud. To enable it, very accurate packet counts for the user are required at an ingress/egress point.
- *Traffic profiling* uses information collected about an IP flow to describe it succinctly so that its statistical profile can be used inside of a traffic model. The model of the IP flow can be used in network planning or network dimensioning. For example, the profile may have the average or peak traffic volume and does not require very precise measurements. Depending on how traffic profiling is used flows may be individual or aggregated.
- *Traffic engineering* uses the information collected about IP flows to control the network with the goal of optimizing network resources and traffic performance. Typical measurements are link utilization, traffic volumes between network nodes and routing information. While precise information may be required about flow routing, the information about link utilization and traffic volumes can be approximate. For example, it may be sufficient to know the largest flow on a "hot" link to initiate its route transfer to another part of the network.
- *Attack/Intrusion detection* uses the information about IP flows to detect unusual situations and suspicious flows and then monitors attacking flows. This use case may require stateful packet flow analysis, which requires deep packet inspection. However, to analyse traffic for anomalies all that may be required is a statistical model of the traffic, which can be learned with machine learning.
- *QoS monitoring* uses quality measurements of IP flows to validate QoS parameters negotiated during service level specification (packet loss, latency). QoS monitoring requires correlation of data from multiple observation points, which requires proper clock synchronization. As QoS monitoring is often specified in statistical terms, precisely

collected information is not required to enforce SLAs, rather it is important to build up a statistical profile of the observed QoS experienced by flows.

Each use case has its own requirements on the veracity and type of data required, meaning that there may be multiple protocols and software required to collect the information for any one of the use cases. We note that while some require precise collection and storage of data (usage-based accounting), others may only require precise data collection, while the stored data does not have to be prices (attack/intrusion detection), and some do not require precise collection or storage of data (traffic profiling, traffic engineering, QoS monitoring).

2. Data Collection for IP Network Monitoring

IP network observation is done by collecting information about the infrastructure (routers and their physical connections), information about IP flows traversing the infrastructure and information about how IP packets are routed through the network (routes). This paper implicitly talks about the network observations of the core network (shown in the middle of Figure 1), but the content of this paper also applies to the aggregation and peering parts of the network.

There is no one standardized method to collect all the information required for each of the network monitoring use cases. Typically, a combination of network monitoring protocols is used, and the information is collected by a software tool, provided by vendor specializing in this. Most networks are multi-vendor networks, presenting a commercial and technical challenges to equipment vendors to collect network information from other vendors' equipment. A third-party vendor would resolve those challenges and provide a data collection software. Collected information can be grouped into device, routing, link and flow information. The monitoring software tool correlates the information and presents it in a "single pane of glass". Users can then get subsets of data required for their use case.

2.1. Router information

Router information can be obtained with direct queries using their command-line interface (CLI) (Wikipedia, n.d.). CLI commands typically provide network element information (infrastructure information), which is not otherwise available through other means. For example, configuration, CPU utilization, or debug logs can be obtained in this way. Other information such as routing tables and link utilization are also available, however routers are typically not optimized to access this information through their CLI interfaces and that should be avoided. CLIs can change at the whim of the network device vendor, making it difficult to track changes for a 3rd party vendor providing the monitoring solution.

While not strictly required for IP flow monitoring, some of the information collected directly from the routers can be used for root cause analysis (RCA) of undesirable network behaviours. Particularly, system logs may be useful for this purpose.

2.1.1. Route information

IP routing information can be collected with routing protocol sniffing. Interior gateway protocols (IGPs) such as Open Shortest Path First (OSPF) (Wikipedia, n.d.) and Intermediate System to

Intermediate System (IS-IS) (Wikipedia, n.d.) are used by routers to exchange topology information in an autonomous system (AS), which enables calculation of forwarding tables on routers. A monitoring system sniffs the inter-router traffic for route update packets and builds the known network topology as a router would.

Using the sniffed network topology and shortest path routing, the monitoring system can in real-time build the routing matrix for the network, which is a "digital twin" representation of the network forwarding tables. The routing matrix relates flows to the paths they traverse in the network. Note that since IP uses dynamically changing hop-by-hop routing, the routing matrix is a delayed version of the real-time routing matrix and is only a close approximation of how packets traverse the network.

2.1.2. Link Information

Link information includes volume of packets carried on a link and link utilization. This information can be obtained by subscribing or polling Simple Network Management Protocol (SNMP) (Wikipedia, n.d.) or Network Configuration Protocol (NETCONF) (Wikipedia, n.d.) router's management interfaces. SNMP is an older protocol with many limitations. For example, it has a polling interval of 5-10 minutes, while NETCONF can be configured to stream messages, which are sent on change detected by the router. NETCONF can use the Yet Another Next Generation (YANG) (Wikipedia, n.d.) format to collect information from the network devices. Internet Engineering Task Force (IETF) is actively working on defining new data sources and formatting of their data across router equipment (IETF).

2.1.3. Flow information

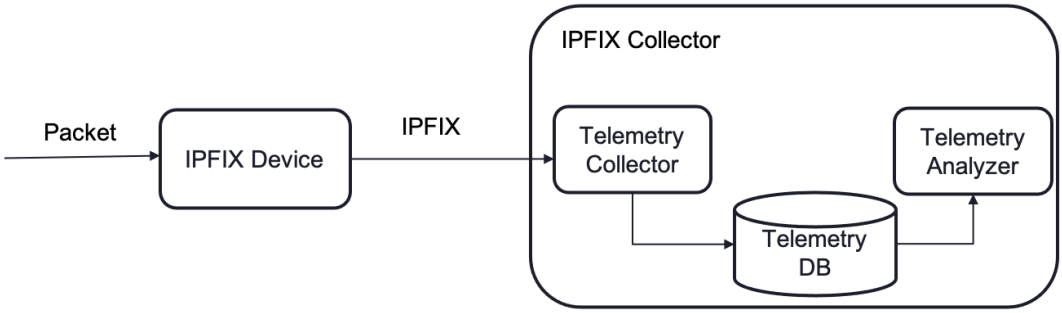


Figure 2 IPFIX Measurement Architecture

Flow information can be collected with NetFlow or IPFIX (Wikipedia), which are equivalent in functionality. IPFIX is the standardized method of doing it and we will limit our discussion to it, keeping in mind that problems with IPFIX also exist in NetFlow. Figure 2 illustrates key components of the IPFIX traffic measurement architecture. IPFIX device and IPFIX collector are two major components of the protocol.

- An IPFIX device is typically a part of routers or switches. It reports information about flows. A dedicated IPFIX device could also be installed to capture packets from the fiber

tap or the mirrored port at a switch. The IPFIX device could be hardware based, or virtualized, so it could also be software installed on a datacenter server.

- The IPFIX collector gathers and analyzes IPFIX flows from multiple IPFIX devices through reliable transport protocols. A typical IPFIX telemetry data record consists of 5-tuple of IP/TCP/UDP header fields, the number of bytes, the number of packets, the flow start time, and the flow end time. In IPFIX, communication between the IPFIX device and the flow collector is done through reliable transport protocols such as Stream Control Transport Protocol (SCTP) or TCP2.

As we just showed, IPFIX way of collecting flow information requires installation of a specialized monitoring system with intermediate collectors, large volume data storage and an enormous number of computational resources to analyze the collected data. The data volume problems start at the network element where the software and hardware are typically not able to track all flows passing the element. Still, the aggregate volume of information may be such that the multiple collectors, distributed across the network, may have to be used. Finally, the volume of collected data makes it impractical to keep IPFIX collected data for long periods of time.

Due to hardware limitations, the monitoring system is usually unable to track a majority of the traffic. One practical approach to mitigate the collection overhead in IPFIX is a technique called threshold compression. In this technique the router reports only flows above a threshold to the collection station. The main disadvantage of this method is that the information of flows below the threshold are not sent. It is quite possible that up to 90% of the flows don't cross the threshold and are not reported.

An alternative method of measuring traffic is use direct counters on traffic tunnels (aggregated flows). However, here we focus on standardized solutions such as IPFIX.

2.2. Overview of the paper

The rest of the paper is organized as follows. We start with a short description of how AI is implemented using deep neural networks (DNNs). Then we describe three ways to use DNNs to decrease the volume of telemetry and stored data, while keeping the fidelity of the data above what is required by traffic monitoring use cases. These are solutions to the problems with IPFIX that we just outlined.

Throughout the paper we cite Wikipedia and AI blogs for various AI concepts. While this may appear to not be the most scientifically sound, we found these articles easy to follow and they always link to the more complete computer science papers for the keen reader. There is much DNN jargon used in the paper. We introduce DNN-specific terms in quotes "" to emphasize their jargon origin.

3. Foundations of AI technologies

This section is intended as a general overview of AI technologies. Readers familiar with concepts of deep neural networks (DNNs) and machine learning (ML) can skip it. More information about DNNs can be found in (Goodfellow, Bengio, & Courville, 2016).

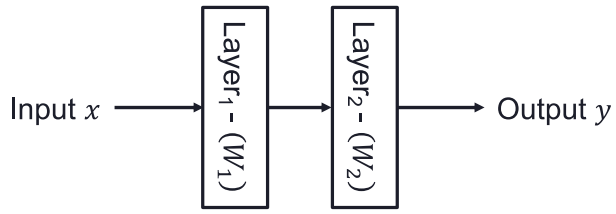
AI technologies are based on the use deep neural networks for machine learning. Machine learning is a computer science concept in which functional software blocks are created by showing the computer examples of correct outputs from inputs, instead of explicitly writing functions that instruct the computer on how to produce outputs from inputs in structured programming (Wikipedia, n.d.). Instead of coding the algorithm, a generic machine learning algorithm is "trained" with examples of what the correct outputs are for given inputs. In recent years, DNN technology has elevated machine learning to the level of human capability in some cognitive tasks. For example, it is now possible to train a DNN-based machine learning algorithm to read a paragraph of text and answer questions about more accurately than humans, or to categorize X-ray images better than a radiologist (Zhang, et al.).

DNN technology is based on basic linear algebra components – matrix multiplication and addition and basic calculus – derivatives. Most of the knowledge required for DNNs has been around for hundreds of years since the time of Carl Friedrich Gauss (Wikipedia, n.d.) and Issac Newton (Wikipedia, n.d.). What is new at this time is that the advances in parallel computing have made it possible to deal effectively with large matrices and train very large DNNs. The most common computing platform are the Graphic Processing Units (GPUs) (Wikipedia, n.d.), which can be used even beyond DNNs, and they are now being complemented with Tensor Processing Units (TPUs) (Wikipedia, n.d.), which are specialized computing units for DNNs. AI accelerators such as TPUs are now found almost anywhere from being embedded in laptops, cellphones, and dedicated data center servers.

For completeness and interest of the reader we now give a simplified overview if how DNNs make predictions and how they are trained.

3.1. How DNNs make prediction

A DNN is a set of algebraic equations that describes how outputs are determined from inputs using matrix operations. A graphical version of the DNN representation is shown Figure 3a, which shows the most basic type of building block for DNNs, known as "dense blocks". The 2-layer DNN shown in the figure is shallow. A typical may have dozens of layers (blocks).



a) Graphical Description of a 2-layer neural network

$$y = \max(0, W_2 \times \max(0, W_1 \times x + b_1) + b_2)$$

b) Mathematical Description of the 2-layer neural network

Figure 3 An example 2-layer DNN

The network in Figure 3a evaluates an equation involving linear algebra shown in Figure 3b. In this example equation, $W_1 \times x$ is a matrix multiplication (Wikipedia, n.d.) and the max function ensures that the result of all operations is positive. Terms b_1 and b_2 are called bias for the layer. The input to the network is x, while the output is y, so the equation describes the functional blocks used in the DNN. The output y is also called a prediction. The input x is a mathematical vector whose components is called a "features". Each feature is a separate input variable contributing to the output of the DNN.

The simple set of algebraic transformations in this example is very powerful as it can be shown mathematically that a neural network with enough layers – depth – can approximate any function. For this reason, DNNs are known as "universal approximators" (Hanin & Sellke, 2018).

A pictorial description of a DNN shown in Figure 3a can be translated into the above equation in Figure 3b by an AI engineer and then made into a software program that performs the set of algebraic equations. In practice, the software piece is simple to write using libraries such as TensorFlow (Abadi, et al., 2015) and PyTorch (Paszke, et al., 2019). The DNN can also be exported into the Open Neural Network Exchange (ONNX) (Open Neural Network Exchange, n.d.), which describes the equations and can be loaded into many DNN software frameworks.

3.2. How DNNs learn

So far, we have described the prediction or inference part of a DNN. If we know the weights of the DNN (e.g. matrices W_1 and W_2 in Figure 3) then upon receiving the inputs, the set of matrix calculations described by the DNN is performed to determine the outputs. The outputs of the DNN are called the "predictions". This process of making prediction is sometimes called "inference". Weights are determined during a process of training.

For example, if we have the function

$$y = 2x^2 + 3x,$$

we can generate a dataset of training samples shown in Table 1 in the two left-most columns. With the dataset, we can use a training function provided in open-source software such as TensorFlow (Abadi, et al., 2015) to determine a set of matrices W_1 and W_2 that result in the best approximation of the function. The training function is called "fit" as it fits the weights of the DNN to the dataset during the training. The fitting function minimizes the error of the predictions $\hat{y}$ for the dataset compared to actual values in the dataset y. The error can be measured with the Mean Absolute Percentage Error (MAPE) shown in the right-most column of Table 1.

Table 1 Example training dataset for $y = 2x^2 + 3x$

Input x	Actual output y	Predicted output $\hat{y}$	MAPE $\frac{\|\hat{y}-y\|}{y}$
1	5	4.5	10 %
2	14	15.6	11.5 %
3	27	25.4	5.9 %

The great DNN research achievement in recent years has been to devise an efficient training procedure that finds the set of internal weights W to minimize the error of the DNN. The training procedure uses "stochastic optimization" who's understanding essentially requires a PhD in mathematics or computer sciences. However, this understanding is not necessary to use DNNs as almost anyone who understands software development can write approximately 10 lines of code to create the DNN and train the function.

The training procedure is iterative and takes in a set of examples of inputs and outputs in batches. For each batch, the training procedure takes in the inputs and generates predictions using the current weights. The predictions from the DNN are compared with the known outputs to find the error in the predictions (the "loss" function) and this error is used to calculate the adjustment to the current weights. The adjustment is usually done using a gradient descent that takes a "learning rate" as an input and calculates the error of the predictions from the weights and number of predictions. The learning rate determines how quickly the descent happens and how closely to the best fit the training gets. The gradient of the whole DNN is calculated using "backpropagation" (Wikipedia, n.d.), which is an algorithm applied backwards through the DNN to differentiate it. Backpropagation is an example of automatic differentiation using the chain rule (Wikipedia, n.d.).

3.3. DNN Models

A DNN model is a trained DNN. A single DNN may have multiple models for different versions of the training data, or different versions of the training algorithms. Each version may have the same structure (number of matrices, size of matrices, and flow through the matrices), but the values in the internal matrices may be different. In a parallel to software development, DNN models have different versions, which presumably improve with higher version numbers. Unlike software, a DNN model is not guaranteed to work well over time, as the inputs may have significant changes in their statistical properties. An example would be traffic demands changing if a new data center peering point is added to the network.

Automating re-training of DNNs due to changes in input data is beyond the scope of this paper.

4. Reducing network monitoring data with AI

Now, we take a deep dive into three approaches for reducing network monitoring data volumes collected with IPFIX. Currently, this problem is resolved in unsatisfying ways. For example, IPFIX may not be active in the network at all, so end-to-end flow information is not collected. If IPFIX is turned on, it may not monitor all the flows – to reduce the pressure on the routers, some flows are simply reported. The collected data is often stored temporarily, or if it stored for longer time, it is reduced by averaging across days, weeks, or months.

We propose three new approaches to deal with the volume of data in the context of IPFIX. First, we consider how to reduce the volume of IPFIX measurements by using link volume information, without a substantial loss of end-to-end flow data quality. Second, we consider how to actively reduce the volume of IPFIX telemetry by taking advantage of correlations in the data. Third, we show how to use AI to compress stored network measurements without major loss of fidelity.

4.1. Reducing the amount of collected information

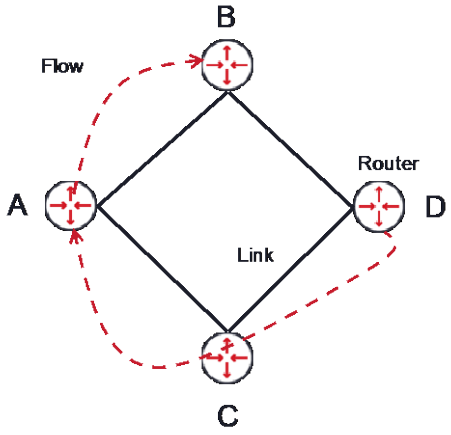


Figure 4 Relationship between links and end-to-end flows

Traffic profiling and traffic engineering are two major use cases for IP network monitoring. Each requires an estimate of the Origin-Destination (OD) traffic matrix, which describes the amount of traffic between each OD pair in the network. Figure 4 shows the relationship between OD pairs (end-to-end flows) and links. In the figure we have 4 routers A, B, C, D and 4 links AB, AC, BD and CD. In case of a core network, we would want to know the aggregate traffic between each OD pair. There are 12 relevant OD pairs, for example for router A there are AB, AC, AD. The premise of the idea in this section is that if we can deduce OD flows from link measurements, we can limit data collection to links. In the case of Figure 4, the amount of reduction would be 75 % (instead of collecting information on 12 OD pairs, we can collect information on 4 links).

Currently, IPFIX is a go to method to obtain OD pair information. IPFIX samples packets transiting through a given router and infers their origin and destination from packet headers. We now show that finding the entire traffic matrix can also be done by utilizing the available link counts and routing information.

Obtaining end-to-end traffic volumes from link measurements is a mathematical problem, which requires a matrix inversion. The instantaneous traffic matrix can be related to link measurements and the routing matrix with

$$\boldsymbol{y} \approx R\boldsymbol{x}$$

where $\boldsymbol{y}$ is the vector of measured link loads over links in the network, R is the routing matrix, and $\boldsymbol{x}$ is OD traffic matrix with one row corresponding to the demand of each OD pair. A flow in the matrix is denoted with row $x_i \in \boldsymbol{x}$ in the OD matrix x. The routing matrix is structured in a way that the link measurements $\boldsymbol{y}$ correspond to the sum of OD flows that traverse the link. Due to packet latencies and the random nature of OD pair traffic, the equation is approximate.

For the mathematically inclined, it may be obvious that the instantaneous traffic matrix can be estimated with

$$\boldsymbol{x} \approx R^{-1}\boldsymbol{y}$$

where R^{-1} is the "inverse" of the routing matrix. Alas, the routing matrix undetermined and is typically not invertible, so this solution is not possible.

One way to solve the undetermined equation is to find the OD traffic matrix $\hat{\boldsymbol{x}}$, which minimizes a distance between the true end-to-end flow measurements and their estimate:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|R\boldsymbol{x} - \boldsymbol{y}\|.$$

which is a problem solvable using AI technology. The equation reads: find an approximation of end-to-end flow volume, $\hat{\boldsymbol{x}}$, which has the smallest error, compared to the true matrix $\boldsymbol{x}$. If we have many link measurements taken during different times of the day, the estimate becomes the maximum likelihood estimate of the end-to-end flows.

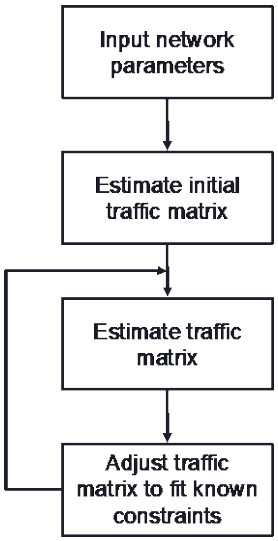


Figure 5 Iterative estimation of the traffic matrix

The optimization approach alone does not provide the best results. We have found it necessary to also combine that approach with stochastic approaches to deal with the randomness in network measurements. The details of how that is done go beyond the scope of this paper due to the volume of mathematics involved.

The approach used above ignores many of the known network constraints since it is defined as an unconstrained optimization. To get around this, the AI method can be incorporated in an iterative procedure shown in Figure 5. The method takes an initial traffic matrix estimate and then uses an estimation procedure followed by an adjustment procedure. As the procedure goes on, it produces a sequence of the traffic matrix estimates $\hat{\boldsymbol{x}}_0, \ldots, \hat{\boldsymbol{x}}_n$, each of which is expected to be closer to the true traffic matrix $\boldsymbol{x}$. As the initial traffic matrix estimate and the estimate in the iterative step may produce a traffic matrix which may not match information known about the traffic (e.g. ingress/egress aggregate counts), an adjustment procedure that projects the estimate into the known constraints is used to fix this.

We evaluate the performance our methodology using real traffic traces from a backbone network. Our source of data is the IP-level traffic flow measurements collected form every point of presence (PoP) in the Abilene Internet2 back bone network (Roughan). Abilene is the major backbone network, connecting over 200 universities in the US, and peering with other research networks in Europe and Asia. At the time the data was collected, the Abilene network had 11 routers resulting in 121 origin–destination flows; there were 15 links in the network. We note that Abilene was collecting OD pair information using flow sampling technology and it was also simultaneously collecting link information, which allows us to estimate OD flows from link measurements and then use true flow measurements to evaluate the performance of the AI-based approach.

Table 2 summarizes the results we obtained on the Abilene dataset. We make several observations. First, the AI approach reduces the error in the estimate significantly (by 42%).

Second, the estimates are now in the range of what is acceptable for the traffic engineering task. Third, the amount of data collection has been reduced by 78%, as we only need packet counters from links.

Table 2 Summary of OD flow estimation from link data performance results

	MAPE	Reduction in MAPE
Non-AI approach	29.6%	-
AI Approach	17 %	42%

We think that this may be a promising approach to reduce operator dependence on IPFIX in estimating OD pair demands for the purposes of network planning and traffic engineering.

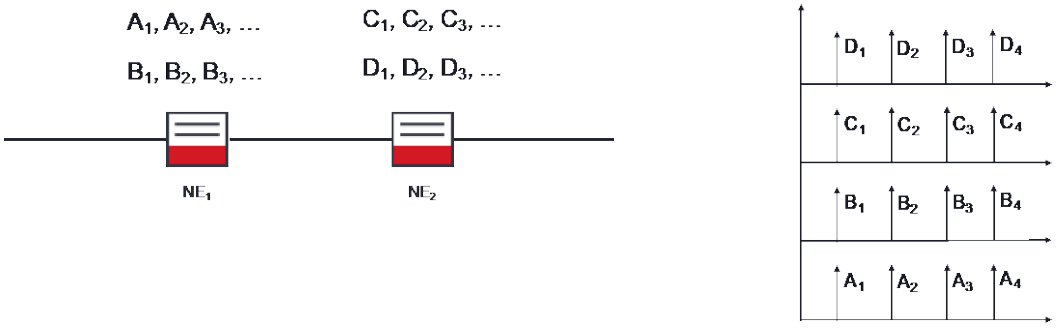
4.2. Reducing the amount of telemetry

Another way to reduce the amount of IPFIX data is in network telemetry. Here we look at ways to reduce number of samples of the collected data. Unlike the previous method, which reduces the number of data sources, where the data is collected, this method reduces the amount of data collected by each data source.

The main idea is to reduce collection of data by decreasing the amount of information collected about some flow, and then to use the information collected from other sources (flows) and AI to infer what data that was not collect would be.

There are two uses for the method described in this section. First, it could be used to reduce the amount of telemetry data. Second, it could be used to increase the frequency of some measurements, while keeping the volume of telemetry about the same. The way this works is that we increase the frequency of measurements on some of the sources, while reducing the frequency of measurements of other sources. We then use AI to impute (estimate) the missing values in the sources with reduced measurement interval.

Network data samples are taken at a prescribed measurement interval (typically in the order of minutes). When picking the sampling interval, the network operator is typically not concerned about sampling interval from the point of view of the Nyquist criterion (Wikipedia), which is required to reconstruct the sampled data without the loss of information. The operators are not trying to reconstruct the data at the point of the collection and processing; data is typically collected for other purposes (forecasting, anomaly detection), so the precise reconstruction of the underlying random process is not important.



Spatially correlated time-series Reported time-series

Figure 6 Spatially and time correlated multi-variate time-series

As an example of time-series collection, we show two network elements and 4 time-series in Figure 6. NE_1 collects time-series A and B, while network element NE_2 collects time-series C and D. Time-series are on the same element are correlated, which means that information of both flows exists in either flow. For example, A(C) could be link utilization and packet latency could be series B(D). If link utilization is high, packet latency is also high, meaning that the two are correlated. Similarly, time-series on the same path are correlated. So, if A is link utilization on NE_1 and C is link utilization on NE_2, they are correlated due to the shared flows on those links. For example, if link utilization is A is high, this could be due to a large flow traversing NE_1, which is also traversing NE_2, so C is also high.

Data pruning

To reduce the information generated and transmitted by NEs we drop (prune) some of the samples. This is called "measurement sampling". Measurement sampling is a well-known technique used in both packet and flow-based measurement to reduce the data volumes required to report. The main idea in this technique is to take only a subset of packets or flows out of all packets or flows to obtain reasonable result for the measurement. For example, we could prune every k^{th} sample of each time-series. This is called subsampling and can be undone for each time-series using a low-pass filter, if the Nyquist criteria is satisfied for the subsampled time-series. *However, this is not what we do.*

Given measurement sampling, we propose a system architecture shown in Figure 7. In the network element, (1) the data pruning module receives the data from the data gathering module and removes some portion of the data before (2) transmitting it to the network analytics. In the network analytics module, a data recovery module (3) reconstructs (imputes) the data and passes it to the recovery evaluation module to (4) determine if the recovery was of high enough quality. Finally, the pruning logic module (5) instructs the data pruning module on how to prune data to improve performance.

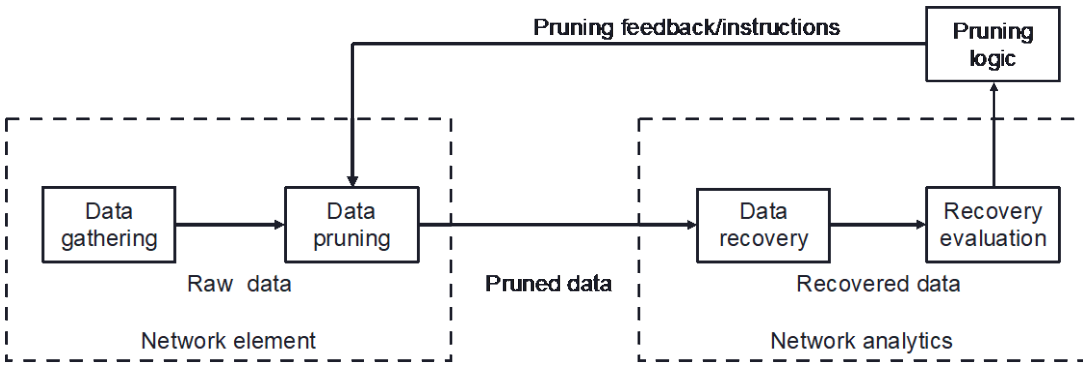


Figure 7 Pruning and reconstruction architecture

Correlation means that information about one time-series is available in another time-series. This is the fact we use when pruning and imputing missing information. We remove parts of a time-series so that some information is always available in another, correlated time-series. The information can be removed in many ways, but the easiest is to use an offset between the time-series when the k^{th} sample is removed. For the example in Figure 6, the samples can be reduced by only sending A_1, A_3, B_2, B_4, and C_1, C_3, D_2, D_4.

Data reconstruction

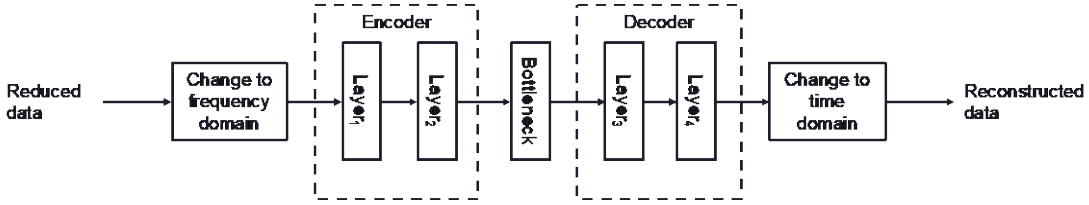


Figure 8 Reconstruction architecture

There are many ways of imputing the missing values received from the network elements. Here we discuss one way of doing it as a way of an example. The DNN treats the missing values as noise in the data and can work with any pruning strategy - the method does not require any information about how the data was pruned.

The AI reconstruction is shown in Figure 8. We use a multistage DNN. The pruned data is used as the input, while the reconstructed data is determined as the output. The method treats missing values as noise, so it is using a denoising method using an autoencoder (Wikipedia) (shown as the encode-bottleneck-decoder architecture). As the values are missing, the reduced data is first transformed in the frequency domain (using inverse Fourier transform or wavelet domain using wavelet transform). The transform into the frequency domain interlaces the missing and present values and so the structure of the data is pronounced even if some of the values are missing. The autoencoder structure denoises the frequency representation of the data thus emphasizing the structures in the data. The inverse Fourier transform then returns the denoised frequency domain data into the time domain.

Table 3 Summary of pruning/reconstruction performance results

Data reduction	5%	10%	20%
MAPE	4%	5%	9%

To evaluate this approach, we use the Abilene backbone dataset, which was discussed earlier. This is a relatively small dataset for this purpose, so learning to reconstruct the data is harder. Nevertheless, our analysis shows that data reduction is possible. For example, if we can tolerate a 10% reduction in data precision in data sources, which are being pruned, we can reduce the amount of transmitted data by 20%.

We think that this may be a promising approach to reduce the amount of telemetry used in IPFIX at a small loss of data precision.

4.3. Reducing the amount of stored information

Useful network data is high in volume, making it difficult to store for extended periods. Here we discuss compression of the data to make storage cheaper. For example, a network with a million data sources may generate 46Gb of 1-minute sampled data in one day, which translates to 16Tb of data per year. While the number of data sources may seem high, even in a traditional IP network this number would be inside the realm of possibility when considering each flow to be a data source. The number of data sources would be much higher when considering cloud services, IoT networks with billions of devices, multiple network layers, or high sampling network measurements (e.g., state of polarization, or wireless SNR measurements).

Using the example of 16Tb of data per year, the Amazon Web Services (AWS) S3 (AWS) cost to store it would be around $5000 in the first year and accumulating to $25,000 in year 5 of storing this data (AWS). Here, we describe a DNN-based lossy compression scheme to compress network data, which has a compression ratio of 100x-200x (Wikipedia). The one-time cost of compressing 16Tb using this compression scheme would be approximately $2.50 and the cost of storing the data for a year would be approximately $150-$250 for the one-year period, depending on the compression ratio. Over a 5-year period, the savings from compressing the data would be around $24,000, or 96%. These are significant savings that could be used in other business areas, instead of for simply storing data in the cloud.

Today's solution is not to store the data, reducing the operator's ability to make data-based decisions. Due to cost, network measurement data is not stored for extended periods, or it is aggregated in larger periods of time (daily, weekly, monthly, yearly), thus losing fidelity in an important historical record of what has happened in the network. The process of aggregation/averaging is a very crude way of lossy compression for time-series. For example, averaging represents a time-series with a single number over a period, so its accuracy is not good. The compression of averaging is not that good either - compressing 15-minute into a daily bin only has a compression ratio of 96, which we show is easily attained with neural networks.

Time-series compression with DNNs

We propose a method of storing network data in a deep-neural network, which greatly reduces the volume of information that needs to be stored. Figure 9 shows an example architecture for a network using the compression scheme. Data is collected from network elements (NEs), stored in temporary storage, compressed, and deployed as a deep neural network. Data is deleted after compression. Instead of using a database to store the data we use a server which hosts the DNN and allows querying of it to retrieve the data.

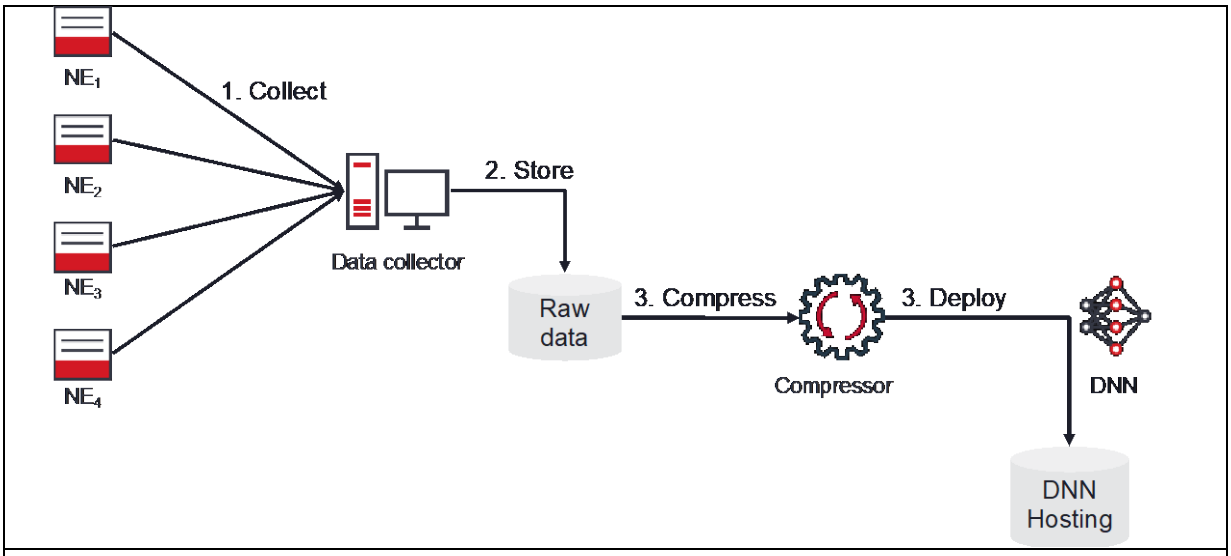


Figure 9 Network data collection infrastructure

The compression of the time-series is accomplished by using the ability of DNNs to memorize the mapping for any function. For example, we can train a DNN to memorize a relationship between a random index and a sequence of measurements. Suppose that the network data consists of sampled samples $\mathrm{ts}_1, \ldots, \mathrm{ts}_n$. The samples are windowed into fragments $\mathrm{w}_i = (x_1, \ldots, x_p)$ of p samples. Each fragment is associated with an index i.

Now we have a dataset where the features are the bits of the index i and the labels are the values in the fragment. We use a DNN to learn a function which maps a 32-bit integer into a p sample long fragment of a time-series. As an example, a 3-layer DNN is shown in Figure 10. Input to the DNN is the index of the bucket i and the output calculated by the DNN is the approximation of the time-series window $\widehat{\mathrm{w}}_i$

$$\widehat{\mathrm{w}}_i = W_3 \max\{0, W_2 \max\{0, W_1 i + b_1\} + b_2\} + b_3.$$

We note that due to lossy compression $\widehat{\mathrm{w}}_i \neq \mathrm{w}_i$, but their difference can be made arbitrarily small.

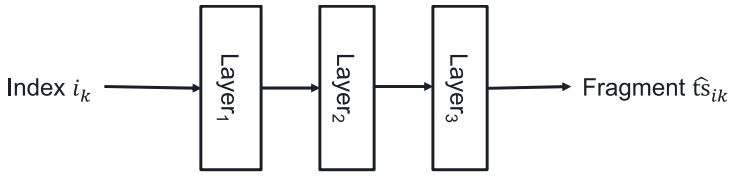
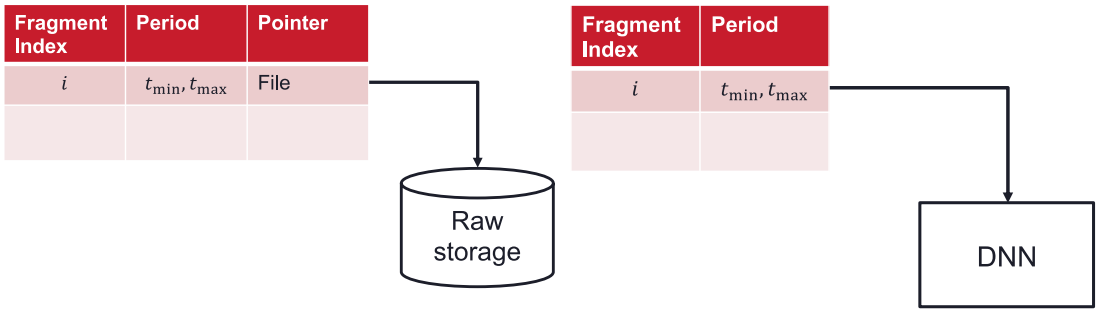


Figure 10 Decompression

Figure 11a compares the DNN approach to how data is stored today. Raw time-series are stored as fragments on disk and a table is used to relate times-series, periods, and the files on disk. Figure 11b shows how data is stored into a DNN. A table is utilized to relate the time-series, periods, and fragments with *indices*. The DNN contains the relationship between the index and the stored time-series. When an index is presented to the DNN, the DNN reconstructs the time-series bucket for that index.



a) Raw storage b) Compressed storage

Figure 11 Time-series storage

Achievable compression

The level of compression is strongly related to the level of achievable precision. In general, the stronger the compression, the less achievable precision is possible. This means that the compression can be used judiciously to save on space and time to train the DNNs. When comparing compression algorithms, an important metric is the compression ratio (Wikipedia). The compression ratio is the ratio of the size of the uncompressed data to the size of its compressed form. For example, if the original size of a dataset is 16 MB and its compressed size (size of the DNN) is 4 MB, the compression ratio would be 4x.

In many cases, the size of values in a time-series may vary significantly. Consider the case of "mice" and "elephant" flows in an IP network. There may be several orders of magnitude difference in the size of these flows. In the case of traffic engineering, it is much more important to know the size of large flows precisely than the size of small flows precisely. For example, if there are 100 small flows that can be compressed at 100x compression ratio and 1 large flow that can be compressed at compression ratio of 10x, to maintain its acceptably high precision for each

set of flows, the average compression ratio compression could be 91x[1]. This should be compared to compressing all flows at the compression ratio of 10x, which the minimum required by elephant flows.

Table 4 Summary of compression performance results

Compression ratio	10x (XOR)	20x	40x	80x	128x	200x
MAPE	0%	0.5%	1%	2%	5%	10%

Table 4 show the compression achievable by DNNs. We used an IP-like dataset that we generated with a simulation. We use this dataset since we can control the size of the time-series, which makes it easier to see the performance of the compression scheme across different sizes. We also tried this on an IP dataset with similar results. The first column is for the XOR compression (Time-series compression algorithms, explained, n.d.), which is a lossless algorithm designed specifically for time-series. The MAPE for this algorithm is 0% since it is lossless. Next columns in the table use the compression we just showed. The compression varies from 20x to 200x, depending on the error in the estimates. We note that at 5% MAPE, which is very reasonable across a range of use cases the compression ratio is well over 100x.

We think that this may be a promising approach to reduce the disk space required to store IPFIX data at a small loss of data precision.

5. Summary

This paper has talked about various methods to reduce the amount of transmitted and stored telemetry data in IP network monitoring. We have described why network monitoring is important in relation to the monetization strategies used by network service providers. In most network monitoring use cases, the precision of the data is less important than the cost of collecting and storing the data. This introduces opportunities to trade off precision in the collected and stored IP telemetry with the cost of collection and storage.

We have reviewed 3 approaches that can reduce network telemetry 20% to 75% and the amount of stored IPFIX data by orders of magnitude. These are visionary applications of DNNs in network monitoring and the authors would appreciate feedback on their potential usefulness to service providers.

[1] We get this by calculating the compression ratio of large flows and small flows and then averaging out across all flows.

Abbreviations

AI	Artificial Intelligence
CLI	Command-line interface
CPU	Central Processing Unit
DNN	Deep neural network
GPU	Graphic Processing Units
IGP	Interior gateway protocols
IP	Internet Protocol
IPFIX	IP Flow Information Export
MAPE	Mean Absolute Percentage Error
ONNX	Open Neural Network Exchange
PoP	Point of Presence
RCA	Root Cause Analysis
S3	Simple Storage Service
SP	Service Providers
TPU	Tensor Processing Unit

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . Jozefowicz, R. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* Retrieved from Software available from tensorflow.org: https://www.tensorflow.org/

AWS. (n.d.). *Amazon S3: Object storage built to retrieve any amount of data from anywhere.* Retrieved 07 21, 2021, from https://aws.amazon.com/s3/

AWS. (n.d.). *AWS Simple Monthly Calculator.* Retrieved from https://calculator.s3.amazonaws.com/index.html

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.

Hanin, B., & Sellke, M. (2018). *Approximating Continuous Functions by ReLU Nets of Minimal Width.* Retrieved from https://arxiv.org/abs/1710.11278

IETF. (n.d.). *Operations and Management Area Working Group (opsawg).* Retrieved 07 13, 2021, from https://datatracker.ietf.org/wg/opsawg/documents/

Network monitoring. (n.d.). Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Network_monitoring

Open Neural Network Exchange. (n.d.). Retrieved from https://onnx.ai/

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., . . . DeVito, Z. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems 32*, (pp. 8024--8035).

Quittek, J., Zseby, T., Claise, B., & Zander, S. (2004). *Requirements for IP Flow Information Export (IPFIX).* Retrieved from https://www.rfc-editor.org/info/rfc3917

Roughan, M. (n.d.). *Internet Traffic Matrices.* Retrieved 07 16, 2021, from https://roughan.info/project/traffic_matrix/

Santos, O. (2016). *Network Security with NetFlow and IPFIX: Big Data Analytics for Information Security.* Cisco Press.

Time-series compression algorithms, explained. (n.d.). Retrieved 07 22, 2021, from https://blog.timescale.com/blog/time-series-compression-algorithms-explained/

Wikipedia. (n.d.). *Autoencoder.* Retrieved from https://en.wikipedia.org/wiki/Autoencoder

Wikipedia. (n.d.). *Automatic Differentiation*. Retrieved June 5, 2021, from https://en.wikipedia.org/wiki/Automatic_differentiation

Wikipedia. (n.d.). *Backpropagation*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Backpropagation

Wikipedia. (n.d.). *Benchmarking Methodology for Network Interconnect Devices*. Retrieved 07 09, 2021, from https://datatracker.ietf.org/doc/html/rfc2544

Wikipedia. (n.d.). *Carl Friedrich Gauss*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Carl_Friedrich_Gauss

Wikipedia. (n.d.). *Command-line interface*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Command-line_interface

Wikipedia. (n.d.). *Data compression ratio.* Retrieved 07 21, 2021, from https://en.wikipedia.org/wiki/Data_compression_ratio

Wikipedia. (n.d.). *Graphics processing unit*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Graphics_processing_unit
Wikipedia. (n.d.). *Internet Control Message Protocol*. Retrieved 08 09, 2021, from https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol
Wikipedia. (n.d.). *IP Flow Information Export.* Retrieved 07 13, 2021, from https://en.wikipedia.org/wiki/IP_Flow_Information_Export
Wikipedia. (n.d.). *Iperf*. Retrieved 07 09, 2021, from https://en.wikipedia.org/wiki/Iperf
Wikipedia. (n.d.). *Isaac Newton*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Isaac_Newton
Wikipedia. (n.d.). *IS-IS*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/IS-IS
Wikipedia. (n.d.). *Matrix Multiplication*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Matrix_multiplication
Wikipedia. (n.d.). *MTR (software)*. Retrieved 07 09, 2021 , from https://en.wikipedia.org/wiki/MTR_(software)
Wikipedia. (n.d.). *NETCONF*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/NETCONF
Wikipedia. (n.d.). *Nyquist frequency.* Retrieved 07 16, 2021, from https://en.wikipedia.org/wiki/Nyquist_frequency
Wikipedia. (n.d.). *Open Shortest Path First*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Open_Shortest_Path_First
Wikipedia. (n.d.). *Simple Network Management Protocol*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Simple_Network_Management_Protocol
Wikipedia. (n.d.). *Structured programming*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Structured_programming
Wikipedia. (n.d.). *Tensor Processing Unit*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Tensor_Processing_Unit
Wikipedia. (n.d.). *YANG*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/YANG
Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., . . . Perrault, R. (n.d.). *The AI Index 2021 Annual Report.* Retrieved from arXiv: https://arxiv.org/abs/2103.06312

Reliable Power Monitoring is Critical to Successful 10G Deployment

A Technical Paper prepared for SCTE by

Tim Cooke
Director of Product Management
Amphenol Broadband Solutions
Chatham, VA
434-489-4713
tcooke@abs-go.com

1. Introduction

As network speeds increase, and their elements move ever closer to end users, the need for good, reliable and steady power has never been greater. With the changing architectures of networks – copper, fiber, wireless, small cells, etc. it is equally important that power solutions be scalable and intelligent.

It is important, therefore, that modern power solutions provide usable data that not only tracks their own health, but also can collect and transmit information that allows an operations team to make better informed decisions on how to best manage the entire network. Voltage, amperage and temperature monitoring, for example, can provide actionable insight in the following ways:

1. Within a co-location facility – understand the actual power usage when access to the metering device or monthly power usage invoice is not available. Such information will prove useful for contract negotiations related to power consumption.
2. Provisioning analytics can determine at what time of day the highest power consumption takes place and offers the ability to tailor tasks for periods when power is less expensive and places a lesser demand on the grid.
3. Mitigation – with advanced monitoring, the ability to avoid disruption of services exists by enabling the ability to recognize and identify issues before a failure occurs.

Each of these areas, and more, may be improved through the effective use of power monitoring.

2. Technology Driving a Shift in Network Architecture

As technology, media, and other sectors push an ever-increasing number of modern applications in such diverse markets as residential, healthcare, entertainment, manufacturing, farming, transportation and hundreds of others, the speed and latency of traditional networks simply cannot keep up. There is a need for all forms of broadband networks – fiber, copper, wireless, etc. – to work together seamlessly to carry the information needed to allow the continued advancement of the promise of the IoT, autonomous vehicles, and hundreds of other applications yet to come.

With the accelerating complexity of available “connected” products, and even those that are only imagined now, it is important to find a common denominator that can bring some level of order or commonality. This may well be the fact that all of the parts, from the tiniest sensor inside a suburban home’s washing machine, to the complex computers that run statewide networks must be powered. This fact provides engineers with a methodology to monitor much of the energy health of 10G networks.

Meanwhile, in the increasingly automated and connected homes of today, power reliability and quality become ever more critical as IoT and connectivity demands increase. According to a recent analysis by Emergen Research, “The global power monitoring market size is expected to reach $5.86 Billion in 2028 and register a steady CAGR over the forecast period.”

Power monitoring equipment is now available for the homeowner to monitor and troubleshoot incoming power and power usage throughout the day. As solar and alternative energy sources become more affordable and utilized, the ability to monitor and ensure the requirements are being met, increases. As power monitoring capabilities move to the level of the homeowner, how much more critical is the system intelligence for the network power, whether HFC, DOCSIS 4.0, PON, EPON, etc. All of these levels related to power availability and quality are essential for post pandemic workers to stay connected and ensure network QoS for the rapidly changing work force.

With advanced monitoring, the ability to avoid disruption of services exists by enabling the ability to recognize and identify issues before a failure occurs. Through constant measurement and comparison against expected power thresholds, deteriorating power or other conditions can trigger the necessary actions to avoid intermittent, or even total loss of power. Such mitigation ensures reliable network service and the highest level of customer satisfaction.

3. 10G Speeds Demand Distributed Architecture

With remote work becoming the norm since the pandemic, the 10G initiative is being brought to the forefront quicker than originally anticipated and along with it, the need for clean, reliable power to service the network and the end-use products in the residence and elsewhere.

Distributed Access Architecture (DAA) migrates the intelligence and functionality that had once existed in head ends and data centers out much nearer to where these things are actually applied. In other words, the data remains closer to where it will actually be needed, decreasing the delay of making a round trip to the head end. (Figure 1)

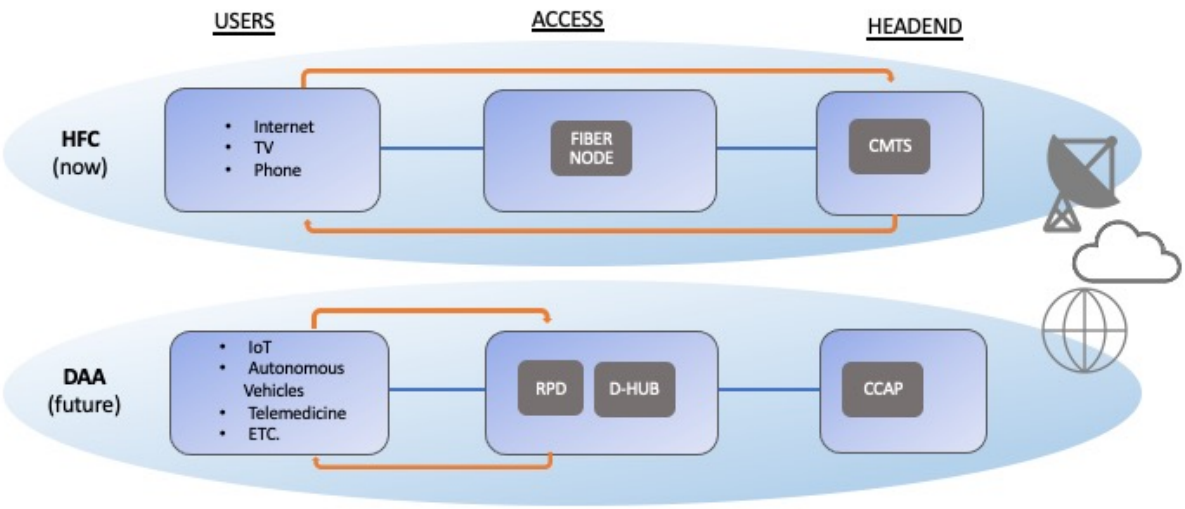


Figure 1 – DAA Migrates Intelligence & Functionality Closer to User

Consider autonomous vehicles. In this example, a future vehicle designed to operate at Level 5 (true driverless) will not only need to rely on the information available on-board the vehicle, but also such data as what traffic may be approaching from around a blind corner. This information is very location-specific,

and therefore it would be inefficient to forward all of this to a central data repository at a head end merely to push it back to the vehicle.

With distributed access “local” information can stay local, lessening the burden on networks to carry every bit of information to and from a central hub. Thus, speed is increased, and latency reduced.

Moving nodes closer to end users will result in a network that is more complex. Equipment with the intelligence to move data within the node, as well as beyond, will be deployed. To support these intelligent nodes, basic infrastructure such as secure shelter, health monitoring, and intelligent power sources must be available.

As equipment is deployed deeper in networks, exponential growth of equipment locations naturally occurs. Each network element must have the ability to sense possible problems, and to transmit the need for support before a failure occurs. Imagine the possible problems in our autonomous vehicle scenario should a network failure occur. Like all other critical equipment, power panels should have such intelligence and capability.

4. Power Panel Design is Evolving

The main function of power panels, of course, is to provide the power needed by each piece of network equipment in a given site. Whether using fuses, circuit breakers, high-, medium, or low-current options, connectorized or lug outputs, or any of dozens of other technology options, the basic functionality of these panels is the provision of power to each piece of equipment that needs it.

The latest generation of power panels makes the integration of power management directly in the power platform easier than ever before. While power panels have provided the ability to monitor their own health in the past through simple alarm closures, the intelligence to report and, more importantly predict problems is now available within the panel itself.

Within the latest generation of power panels, power monitoring has been isolated from the provision of power itself, meaning that, depending on the design of the actual panel, monitoring devices, such as “hot swappable” cards, may go “bad” without interrupting the flow of power to critical equipment. Thus, the card may be removed for service or replacement without affecting services.

Power panels have had the ability to deliver simple alarm information via output contacts for some time. These output contacts have been tied into a network’s alarm systems for decades. With the critical nature of traffic within every distributed node however, intelligent information that serves to predict problems before they occur is imperative.

Within a node, depending on its size, multiple panels may be required to supply the power necessary for all equipment to function at peak performance. In this case, the ability to “daisy chain” multiple panels over a serial cabling scheme which then use a single network connection to transmit critical data back to an operations center will simplify monitoring of multiple bays of equipment.

Intelligent power panels should monitor current at the individual circuit level, providing detailed information that is simply not available with feed-level monitoring or threshold alarming. By measuring parameters such as current per individual circuit, total input bus current, panel input voltage, temperature, and data from external probes, personnel now have access to a much more complete data picture, allowing operations to determine actual current usage, perform trend analysis, and gain critical insight

into equipment performance. This data facilitates proactive action to avert disruptions to service as related to power and power quality.

5. From Alarm Closures to Predictive Analysis

Simple Network Management Protocol (SNMP) is a way to monitor network devices that are on an IP network. Information is requested by an SNMP Manager about the device and connected equipment and status. With baseline measurements and continuous updates, equipment performance can be tracked and controlled. Additionally, SNMP traps sends alerts instantly whenever an event occurs.

The SNMP manager interacts with a network device and its connected equipment. Through the local area network (LAN), an SNMP manager routinely requests information, such as power provisioning, remote site management and circuit threshold data from network devices. This information is recorded and stored via the SNMP manager and can be viewed in a user-friendly table or graph. The data collected can often be displayed by equipment type and location, performance and power usage, and typically monitors threshold levels which allows proactive maintenance with equipment, mitigating downtime by allowing maintenance to be scheduled on a routine basis.

SNMP Traps send instant alerts from the network device when an event occurs. The network device sends these messages without being prompted by a request from the SNMP manager. As soon as an event occurs an alarm is triggered indicating where the event occurred. Immediate access to equipment alarms can prevent unnecessary downtime. During an event, an SNMP manager promptly notifies the local technician who can then make repairs or prevent equipment damage

6. Using data to drive energy improvement in the network

Troubleshooting and prevention alone is a compelling reason for intelligently monitoring power supplies throughout the network. Still, with a growing emphasis on energy efficiency and cost savings, a compelling argument can be made to use data as a way to drive continuous improvement in energy usage in each individual location, as well as across entire networks. Consider each of these scenarios:

Within a co-location facility the ability to understand the actual power usage as measured by the service provider’s own equipment, allowing comparison to readings from the metering device or monthly power usage invoice as provided by the power company. Such information will, along with confirming usage, prove useful for future contract negotiations related to power consumption.

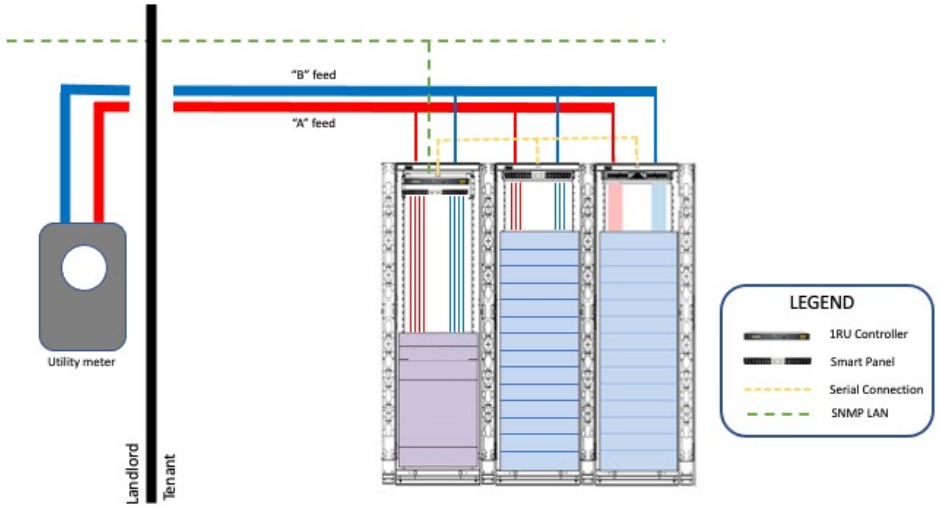


Figure 2 – Co-location Power Scenario

Perhaps even more critically, provisioning analytics will help determine the time of day or even day of the week when the highest power consumption takes place. This knowledge will contribute to better planning and cost savings by offering the ability to tailor tasks to occur during periods when power is less expensive. Further, the network provider will place a lesser demand on the power grid based on this information and planning.

7. Conclusion

With network speeds increasing exponentially, the need for good, reliable and steady power has never been greater, while changing network architectures indicate that it is equally important that power solutions be scalable and intelligent.

Modern power solutions must provide usable data that not only tracks their own health, but also can collect and transmit information that allows an operations team to make better informed decisions on how to best manage the entire network. Voltage, amperage and temperature monitoring can provide actionable insight in both the inside plant (ISP) and the outside plant (OSP), allowing engineers to determine key information such as time of day power consumption or, more importantly the ability to avoid disruption of services due to power loss.

Abbreviations

CAGR	compound annual growth rate
CCAP	converged cable access platform
CMTS	cable modem termination system
DAA	distributed access architecture
EPON	Ethernet passive optical network
HFC	hybrid fiber coaxial
IoT	internet of things
ISP	inside plant
LAN	local area network
OSP	outside plant
PON	passive optical network
QoS	quality of service
RPD	remote PHY device
SNMP	simple network management protocol

Bibliography & References

Power Monitoring Market Size to Reach USD 5.86 Billion in 2028, Emergen Research, June 2021

R. Abbi, et al "*Powering the future 10G access networks - An End-to-End Perspective",* SCTE Paper, 2020

R. Abbi, S. Dharanikota, "Powering the future of 10G Access Networks", SCTE Journal of Energy Management, September 2019

Right Technician at the Right Time
Using Machine Learning to Predict Network Maintenance Issues

A Technical Paper prepared for SCTE by

Anastasia Vishnyakova
Machine Learning Engineer
Comcast Cable
1800 Arch Street, Philadelphia PA 19103
267-260-3277
Anastasia_Vishnyakova@comcast.com

Rama Mahajanam
Director of Machine Learning
Comcast Cable
1800 Arch Street, Philadelphia PA 19103
267-260-4237
Rama_Mahajanam@comcast.com

Mike O'Dell
Director Network Maintenance Comcast Cable
Virtual Location
412-417-0481
Michael_Odell@cable.comcast.com

May Merkle-Tan
Sr. Researcher, Machine Learning
Comcast Cable
4100 E Dry Creek Rd, Centennial, CO 80122
no contact number
Heng-RuMay_Tan@Comcast.com

Catherine Hay
Manager, Process Management
Comcast Cable
676 Island Pond Rd, Manchester, NH
603-222-7420
Catherine_Hay@cable.comcast.com

Lisa Pham
Machine Learning Engineer
Comcast Cable
1800 Arch Street, Philadelphia PA 19103
267-260-4319
Lisa_Pham@comcast.com

1. Introduction

When a customer encounters a service problem that cannot be resolved by regular triage with a call or chat agent, a fulfillment truck roll is scheduled. A standard technician, who specializes in repair issues located inside the premise and/or related to the service drop, performs onsite troubleshooting. If they discover an issue within the broader service network typically impacting multiple customers, they escalate the problem to a line technician. Line technicians expertise is in repairing and maintaining the Outside Plant (OSP) network, such as the nodes, amplifiers, passives, and hardline cables that provide service to multiple customers. This two-step troubleshooting practice incurs the cost of dispatching both types of technicians when an OSP network maintenance impairment is determined to be the root cause of a customer's service issue. More importantly, the follow-up escalation often delays the problem resolution for the customer.

In this paper, we will show that resolution efficiency can be improved by harnessing machine learning (ML) to predict when a customer's service issue has a high likelihood to be escalated to the line technician. Our approach leans on network and device telemetry as well as monitoring processes that assess the integrity of our service network (such as checking for outages, impairments, and performing problem segmentation). A key component of our ML modeling is the integration of graph features derived from network topology, that help identify issues related to equipment within the network that serves multiple customers. This model is in the trial stage and is being tested by selected regions. We also discuss how we can evaluate model precision to incur savings for implementing the model. Finally, we describe how the model will be integrated into the internal troubleshooting software.

2. Acknowledgements

We would like to acknowledge contributions of several Comcast colleagues. Michael Kreisel, who has since left Comcast, developed the initial RTM (Refer to Maintenance) model and the topology data. The extended Applied AI engineering team contributed to the setting up and maintaining of the data pipelines supporting the model. We would like to thank Kevin Bohinski, Ryan March, Likhitha Thebatni, Pat Dwyer, Nick Pinckernell, and Koundinya Venkata Sai Ravulapati for supporting the data structure for the model. We thank Applied AI researchers Yonatan Vaizman, Tianwen Chen, Fan Liu, Navdeep Jain, Scott Rome, Joshua Jackson, and Zhipeng Liu and director of Machine Learning Hongcheng Wang for their expertise and advice on the model development. We are grateful to our intern Vishnu Sharma who has supported model operations and contributed to the feature development. We thank Garret Beatty for his work during the Fall of Code program. We thank our product owners Dave Monnerat and Jason Stevens for being champions of the model among stakeholders. Background on plant operations and tools was obtained from interviews with many Comcast leaders and experts. We want to thank Larry Wolcott, James Sayer, Scott Johnston, Gary Ventriglia, Reid Downey, Scott Shrader, Andreas Lebaudy, Hari Palaiyanur, and Rich Aleong. We thank our field operations colleagues Ray Stewart, Mark Bosteder, Josh Halbrook, Nicole Atherholt, Jason Fletcher, Ed Reece, Brandon Yawn, Karen Washington, Robert Millis, Marcellis Price, Mike Mahaney, Andrew Herr, and Marc Tucker for their support during the model trial. We are grateful to Leslie Ellis, Larry Wolcott, and Jan Neumann for their edits and revisions.

3. Plant Operations

If recent global experiences have taught us anything, it is that connectivity to communications, information and entertainment are essential to modern life. There are many different technologies available to consumers for these needs, with cable TV / broadband networks being among the most common and widely available service delivery network types. During the pandemic of 2020, many service delivery providers, ourselves, and others, experienced significantly increased demand for data, as

well as increased expectations for reliability and availability. Energy is still being invested in building scalable solutions to meet the increasing demand for bandwidth to the end user. For the purposes of this paper, we are going to examine and define a typical system, dividing the network into several segments, primarily the ISP (Inside Plant), OSP (Outside Plant), and Customer Premise.

The ISP, also commonly referred to as the Headend, is the facility where signals are generated, received, and processed from content providers, as well as from Customer Premise Equipment (CPE), like broadband gateways and set-tops. It is typically a climate-controlled facility where the equipment is housed to manage the routing of video, data, voice, and other signals to their intended destinations. The Headend also houses CMTS (Cable Modem Termination System) devices for broadband connectivity.

The actual customer premise – the household, multiple dwelling unit (MDU) or business -- can have slight variations in definition, primarily around where the demarcation point exists between the premise wiring and the Outside Plant. In some cases, the outside drop cable that connects the hardline cable tap to the premise is considered to be customer premise wiring, and in other instances, the customer premise wiring is defined as the customer side of the grounding or bonding block. In the context of the Refer to Maintenance process, this demarcation is an important distinction. For this paper, we will define the premise as any and all wiring from the hardline cable tap port to the Customer Premise equipment (cable modems, WiFi gateway, video set top boxes, etc.), including all coaxial cables, amplifiers and splitters in the customer premise wiring.

We can now define the Outside Plant as the transmission medium(s) transporting signals between the Headend or Hub and the hardline cable tap port. This would include the fiber and coaxial cables, as well as the active components (nodes and amplifiers) and passive components (splitters, couplers, and taps) between the Headend and the customer premise.

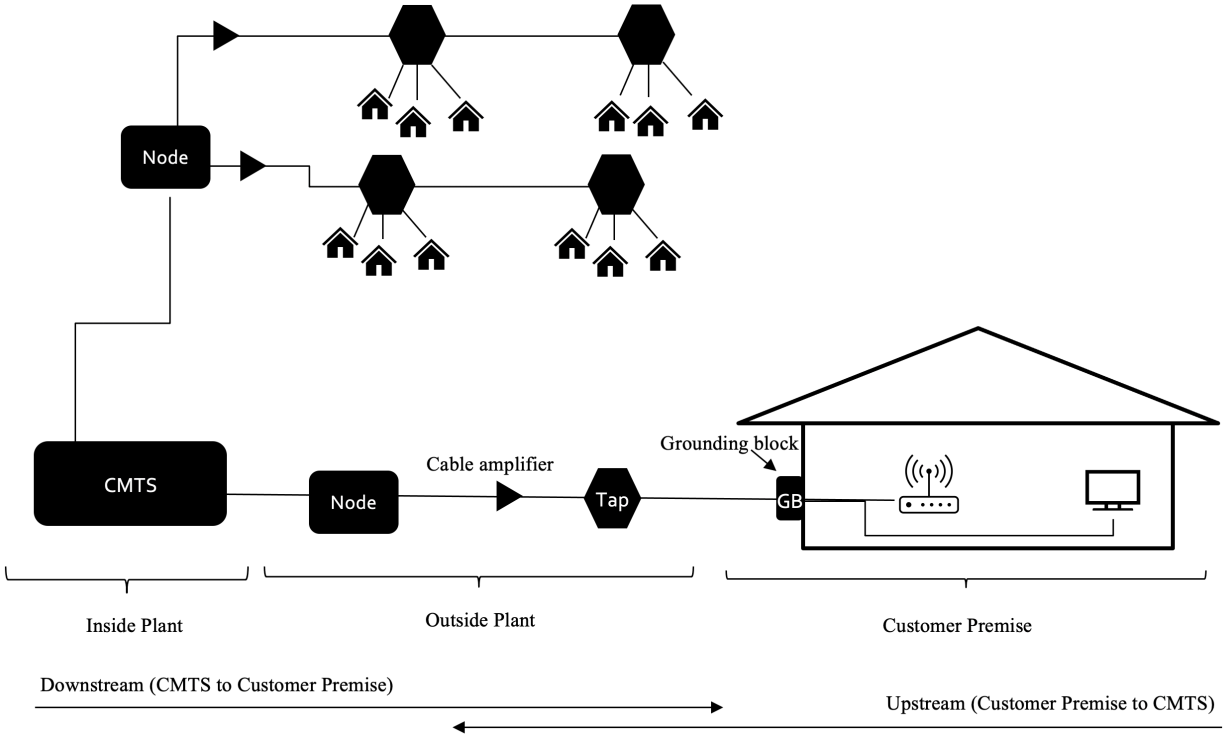


Figure 1 – Plant Diagram

As mentioned previously, this portion of the network is evolving rapidly to meet the increasing demand for bandwidth by the broadband consumers. The Hybrid Fiber Coax (HFC) portion of the network is becoming increasingly diverse as new technologies are developed and deployed, while previous generations of hardware are continually optimized to meet the signal quality and bandwidth needs of the customers. HFC networks built in the late 1990s can still reliably provide service to customers today. Even so, the components that make up the HFC section of the network have remained largely unchanged. To that end, we need to consider the failure points that are common to all generations of HFC networks.

The headend or hub locations that we mentioned have also undergone technology evolutions, but fundamentally, their purpose is to aggregate all the signals designated for a service area, and then modulate them into an optical format for transmission over a fiber optic transport network to a fiber optic node. The node translates that optical signal into an electrical signal and directs that signal into coaxial cable through a series of mechanical connections. The coaxial cable then transports those signals toward the target customer population through another series of mechanical connections via coaxial cable, connectors, couplers, splitters, and finally to the cable tap. The cable tap diverts a portion of that electrical signal to the customer service drop through another mechanical connection at the tap port. The service drop then transports the signal through yet another series of mechanical connections such as bonding or grounding blocks, splitters, amplifiers, connectors, and wall plates, until the intended signals arrive at a cable modem, set top box, or other customer premise equipment. In this example, we outlined many of the connection points that a signal must traverse in a single direction to reach the customer premise equipment, but this example did not include any active devices (amplifiers). This example of a single fiber node feeding a node service area without requiring any amplifiers is often referred to as a Node + Zero (or N+0) architecture, meaning there is a node, but no amplifiers to “boost” the signal to send it further into the network. Amplifiers allow operators to serve a larger area without adding more fiber cables or nodes. It is not uncommon for operators to still maintain HFC networks with N+6, or fiber optic nodes trailed by as many as 6 additional amplifiers, in cascade, required to reach all customers in the node service area.

Thus far we have only discussed the service delivery of signals to the customer in a single, or one-way direction. That is, signals are delivered from the headend or hub to all the customers in the node service area, or “broadcast”. The reality, of course, is that there are signals flowing in both directions, from the headend to the premises, and conversely, from the premises back to the headend, aggregated via the HFC network components. However, all signals, whether upstream or downstream, can be adversely affected by impairments at any point in the HFC network. The degradation of the signal level or quality can be experienced very differently by the customers and our intelligence tools, depending on which component in the transmission medium is damaged or compromised.

The HFC network is a complex yet robust asset. It is arguably one of the most important assets for any operator, and as such, much effort and energy is invested in keeping it healthy and performing at an optimal level. As we described at a very high level, there are a very large number of components that must individually be uncompromised, and collectively maintained to ensure the physical, mechanical, and electrical integrity of the overall network.

Fortunately, the maintenance of these networks has gotten much more efficient over the years. While many of the troubleshooting practices have changed very little, the telemetry we are able to collect from CPE and other assets in the HFC infrastructure have allowed us to better identify those areas which are exhibiting symptoms of impairment. From that additional telemetry data, we can prioritize our response by customer impact potential, and dispatch our Line Maintenance teams to the right area to repair the network and remove the impairments. Previously, plant maintenance was either 100 percent preventative, meaning we would sweep or inspect the network on a scheduled basis, or it was reactive, meaning we

responded to an area when our customers reported an issue. Using the data points that will be discussed later in this paper, we are able to collect metric data points from devices located in different points of the network and infer when those symptoms are caused by an impairment in the physical transmission medium.

4. Remote Network Telemetry Measurements

DOCSIS (Data Over Cable Service Interface Specification) technology is supported by comprehensive data collection systems that offer insights into the state of the access network remotely. Comcast has developed and operates several intelligence tools supported by DOCSIS specifications to help identify network impairments and prioritize repairs. Wolcott et al [1] wrote about the role of the network remote telemetry in the Proactive Network Maintenance (PNM) systems:

… [with] the capability of nearly all CPE [Customer Premise Equipment] to report some level of intelligence back to the tool sets, and [we] have a comprehensive view of all corners of the network. No longer is plant performance informed by a small quantity of DOCSIS channels on a relatively small number of devices. The ability to construct, analyze and match full downstream signatures in multiple portions of each individual premise within a node, in addition to the traditional DOCSIS frequencies, means that an operator can develop an accurate map of all the impairments in the upstream and downstream at virtually every component level within that node. Individual premise issues can be isolated and identified, and with the integration of system design prints, network level impairments can be correlated to active or passive components with a fairly high degree of certainty. This allows for impact scaling, and prioritization of impairment resolution with greater precision, as well as improved task management. Since causality can be attributed to the component level with greater accuracy, dispatch of the proper fix agent is more effective and efficient.

4.1. PNM Software Poller

As the DOCSIS specifications have evolved, so has our ability to poll CPE for meaningful data to analyze the performance of the network. Multiple pollers and polling frequencies can be aggregated to form a comprehensive picture of the performance of the network and premises. Using multiple pollers, or polling timing sequences can be useful in managing the demand on the systems by spreading the requests across time, while optimizing which data points are required at specific intervals. For instance, one PNM software programs could poll devices 3 times per day to report raw measurements describing connectivity between the modem and the CMTS, while others collect different data sets like Full Band Capture (FBC) energy in the FM band at more frequent intervals. We describe some of that data collected by the various pollers in the tables below.

Table 1 – Measures Collected by the PNM Poller

Telemetry	Description
Downstream Receive Power	Power (decibels relative to one millivolt, dBmV) received by the cable modem on the downstream DOCSIS channel. For a DOCSIS cable modem to work within the specification, the downstream power level needs to be in the -15 dBmV to + 15 dBmV range.
Upstream Receive Power	Power (in dBmV) received by the CMTS from a specific modem.
Downstream SNR	Signal to Noise Ratio (SNR) refers to the strength of the signal being received relative to the noise on the line. For the cable modem to

	work within the specification, the SNR needs to be at least 23.5dB; a desirable ratio is 30 dB or higher.
Upstream SNR Individual Modem	Similar to Downstream SNR (see above) but for signal sent upstream
Upstream SNR Channel Average	Similar to Downstream SNR (see above) but for the upstream channel average
Upstream Transmit Power	The strength of signal transmitted by the cable modem. Target levels could be between 42 and 50dB.
T3 Timeouts	Number of instances of T3 Timeout (i.e. ranging request retries exhausted). The cable modem has sent 16 Ranging Request (RNG-REQ) messages without receiving a Ranging Response (RNG-RSP) message in reply from the CMTS within 200ms. The cable modem is therefore resetting its cable interface and restarting the registration process. A T3 timeout is typically associated with an upstream impairment.
T4 Timeouts	Number of instances of T4 Timeout (i.e. when the cable modem did not receive a station maintenance opportunity to transmit a Ranging Request (RNG-REQ) message within the T4 timeout period (30 to 35 seconds) after receiving the previous RNG-RSP). The CMTS must provide each CM a periodic ranging opportunity at least once every T4 (30 to 35) seconds. A T4 timeout usually initiates a reboot of the CPE due to lack of communication: the cable modem resets its cable interface and restarts the registration process, reinitializes its mac after T4 seconds have elapsed without receiving a periodic ranging opportunity. This typically indicates an occasional, temporary loss of service, however, if the problem persists, it may indicate possible service outages or maintenance activity on the CMTS. A T4 timeout is typically a downstream problem; if many modems are affected by T4 errors it may indicate a section of the HFC network being affected. T4 timeouts can also be associated with a CMTS that has extremely high usage (e.g. > 95 percent capacity).
Lost Syncs	Number of times the CM lost synchronization with the downstream channel. Discontinuities in the value of this counter can occur at reinitialization of the managed system.
Resets	Total number of resets happened on the CM
System Uptime	Time in seconds (s) that the device has been ‘up’ and running.
FEC (Forward Error Correction)	Data does not always transmit through cables perfectly, there are often some errors in the bits (1s and 0s); a 0 might be flipped to a 1 or vice versa. To correct those errors, some extra data must be attached to every codeword that goes out (a codeword is just a fixed chunk of data). If it can fix the error, then all is well. However,

	sometimes codewords have too many incorrect bits and the error correction algorithm cannot fix it. If more than 1 percent of the codewords are uncorrectable, we may start to encounter some issues. FECBlks—The total number of FEC blocks (both good and bad) received by all the upstream ports associated with a given downstream. FEC is generally measured using codewords. A codeword is 16 bits with 2 bits for error correction.
Upstream FEC Unerrored	Total number of code word blocks received without any errors. In ideal scenarios it should be equal to total FECBlks
Upstream FEC Corrected	The total number of FEC blocks received by all the upstream ports associated with a given downstream that were slightly corrupted by noise or ingress and that could be corrected and recovered by the FEC algorithm. Number of code words detected as errored, and which have been corrected.
Upstream FEC Uncorrectable	The total number of FEC blocks received by all the upstream ports associated with a given downstream that were so corrupted by noise or ingress that they could not be corrected or recovered by the FEC algorithm. High values imply presence of data loss; this can happen if there is lot of noise in the signal.
Downstream FEC Unerrored	Similar to Upstream FEC (see above)
Downstream FEC Corrected	Similar to Upstream FEC (see above)
Downstream FEC Uncorrectable	Similar to Upstream FEC (see above)
Upstream | Downstream Interface ID	Generally, a device will be connected to the same interface ID over a period of N days; a change to another interface can be an indicator of network-related issues
Upstream ICFR	In-Channel Frequency Response: the max-min variation peak to valley of the response represented in dB. Minimal to no variations are ideal; large variations indicate network integrity issues.
Upstream Ripples	The number of amplitude variations within a 0.75dB threshold.
Upstream TTE	Total Tap Energy upstream from CM expressed in dB
Upstream PMT	Post Main Tap – the Adaptive Equalization (EQ) determined tap that is expending the most energy beyond the main tap. (In the absence of RF impairments, all signals will traverse through the main tap which will have the highest value.)
Upstream PMTE	Post Main Tap Energy – The PMT amplitude expressed in dB. Energy observed in these PMTs is indicative of RF impairments such as micro-reflections being present and the CMTS is activating the equalization gain states in its attempt to compensate for impairments. Energy level is positively correlated to the degree of impairment.
Upstream ETAP1, ETAP2	The EQ tap information for echoes or impedance mismatches beyond the PMT greater than the established threshold. ETAPs provide approximate ‘distance’ to issue (fault/echo). “Distance to issue” can be inferred as the ETAP number post “main”

	tap (outside premise that serves a cluster of residences) with highest 'energy' multiplied by approximately 85 ft per ETAP increment) [2]
CM Phase	Cable modem signal's measured phase angle deviation before adaptive equalization; frequency carrier wave's phase information is a key component in Quadrature Amplitude Modulation (QAM) technique for transmitting digital data as an analog signal.

4.2. PNM Account and Device Network Analysis Degradations Tool

Comcast has a PNM tool that seeks out account and network degradations. The tool consumes data from multiple pollers and runs the impairment algorithms to examine devices six times per day, to report account and device level degradation issues related to disturbances in the RF spectrum. Using comparative analyses, the tool attempts to isolate impairments related to network, drop, in-home wiring, loose connections, and whole home concerns. The table below lists selected metrics used in the impairment detection.

Table 2 – Measures Analyzed by the Algorithmic Impairment Analysis Tool

Metric	Definition
In-Channel Frequency Response (ICFR)	ICFR is the peak to valley measurement of the amplitude flatness. Ideal ICFR is flat and high ICFR can be caused by impedance mismatches due to loose connectors or other cable integrity issues. The algorithmic tool checks for matches between the ICFR signatures and PMT (Post Main Tap) values to identify a common impairment across neighbors.
Carrier to Interference + Noise Ratio (CINR)	CINR is defined is the range between highest and lowest Modulation Error Ratio (MER). CINR help us capture ingress on non-FBC-capable devices.
Full Band Capture (FBC) FM Noise	Full Band Capture -capable devices can identify if there are FM signals present at the RF input. Loose connectors and other cable integrity issues cause FM signal ingress.

Table 3 – Impairment Locations Identified by the Algorithmic Impairment Analysis Tool

Impairment Location	Description
Solo	The tool cannot isolate the fault to specific location in the home or drop but have checked neighbors and excluded a potential network issue. Fault can be at the device, ground block, drop but is not isolated to that single outlet or device.
Outlet	The tool isolated the fault to the specific outlet that feeds a single device.
Home	The tool isolates the fault to a common point in the home that feeds multiple devices but does not impact all devices in the home.

Drop	The fault is isolated to a common point that feeds all devices in the home and is not a network issue.
Network	The tool has correlated at least one other near neighbor with the same impairment on the same channel(s) indicating an Outside Plant impairment.

4.3. PNM Node-Level Network Degradation Impairment Analysis Tool

Alongside the account- and device-level impairments, there is a tool that consumes both account-level data as well as node-level data. This algorithmic tool that scans the network for RF impairments and reports them continuously and creates event logs indicating a problem and a list of accounts impacted. Events move from soaking to confirmed stages and vary by severity. Events include outages, service call alerts, plant faults, spectrum power analysis events, and upstream CMTS port analysis events.

Table 4 – Node-Level Impairments Identified by the Algorithmic Analysis Tool

Event Type	Description
Outage	Event generated when 4 or more devices that share a common problem are offline. Criteria area adjusted for multi-dwelling units.
Power outage	Event is declared if 4 or more devices appear to be a part of a commercial power outage, identified by devices on battery backup.
Service call event alert	Event is generated if the percentage of customers with a scheduled trouble call on the node exceeds a threshold
Plant Fault	Events are generated when a proportion of customers within the node violate the thresholds for the RF Levels. Problems include out of range upstream transmit power, upstream or downstream SNR/MER, downstream receive power.
Spectrum Power Analysis Events	Events are generated by looking for a cluster account matching a specific RF spectrum signature. Signatures include suckouts (a frequency-specific collapse in the RF energy caused by a short distance and high energy reflection), waves (a sinusoidal pattern detecting multiple peaks to nulls in the spectrum 4.5 dB or greater), and tilts (a delta threshold between upstream transmit level in the highest and lowest frequency channels that exceeds 3dB)
Upstream CMTS Port Events	Events are generated by monitoring utilization of upstream channels, failed and correctable FEC errored rate, number of upstream channels in use, and channel width and modulation.

We discuss how the inputs from the remote network telemetry are transformed into the model features in section 6.1.

5. Refer to Maintenace (RTM) Field Operations

When customers encounter service issues, they can reach out for help through various channels, like IVR (Interactive Voice Response), call or chat agents and the Xfinity Assistant interactive chat application. If a chatbot application detects issues that cannot be solved by system refresh or restart, it will route subscribers to a chat or a call agent. Call and chat agents use Interactive Troubleshooting Guides (ITGs) to localize the problem and offer the most effective solution. If an agent identifies an impairment in the network, they will schedule a service technician to visit customer's home and continue troubleshooting on site.

Ventriglia et al [3] describe a workflow for on-site troubleshooting. On site, technicians use field tools to take measurements at different points in the drop network. A technician might start with taking measurements at the ground block closest to the customers' premise. If these measurements are within acceptable specs, the technician will move to investigate impairments inside the customers' home. If a fault is found in the grounding block, the technician will continue to investigate issues at the tap. If no issues are found at the tap, the technician will do the work to replace the drop and perform a premise health test to confirm that customers' issues are resolved. If an impairment is found at the tap, the technician will escalate the problem to the line tech team by creating an RTM request. Issues are triaged by field operations teams and if they are confirmed as a plant impairment, a line technician job is scheduled. Field operations teams monitor non-severe issues, identify neighborhood impact, and send a line technician once a neighborhood impact is confirmed. When the standard technician finds a network issue, a cost is incurred to send both a standard technician and a line technician. Moreover, we estimate that a subscriber might wait up to two days for a standard technician appointment and then possibly another day for the line technician appointment. This historically has been the repair model when a single customer reports a service issue that is ultimately attributed to an impairment in the outside plant network.

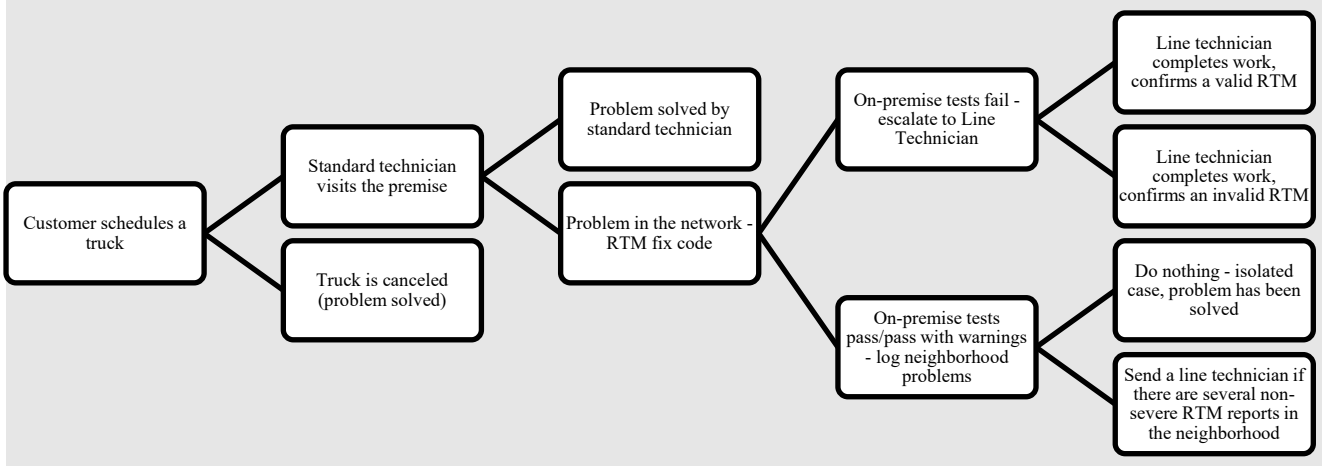


Figure 2 – Refer to Maintenance Field Operations

6. Refer to Maintenance Machine Learning Model

We propose a workflow where a predictive model will recognize that the issue will be an RTM problem, with high probability. In this case, we will not send a standard technician, but schedule a line technician instead. If the model predicts a low RTM probability, we might message the standard technician that the issue is an unlikely RTM.

We have developed a machine learning classification model to identify standard technician visits that will result in the RTM escalation that is also confirmed as "valid". The model is trained on the set of data comprised of standard technician visits. The model learns the target labels derived from the line technicians the "valid RTM" label, whereas all other codes are included as negative labels. We exclude trouble calls where technician could not access the premise, or no one was home.

The model has access to two sets of features: telemetry collected by pollers for all subscribers in the company footprint and information collected ad-hoc for subscribers who reported issues during the call with an agent.

We used a two-step modeling approach, which we refer to as *tier 1* and *tier 2*, to make the final prediction. Our *tier 1* model is trained on all the available telemetry information to determine the probability of RTM among the subscribers' neighbors. This is then used to evaluate the probability of RTM for the subscriber during the day prior to the day of call. Subsequently, our *tier 2* model uses features collected ad hoc during the call with an agent, archived telemetry features, including the *tier 1* derived probability of RTM during the previous days as well as the probability of RTM among neighbors.

Following sections offer details of model development within each stage.

6.1. Feature Engineering

6.1.1. DOCSIS Telemetry Features

The model uses features from PNM telemetry sources that collect and aggregate DOCSIS measurements. These sources poll and analyze the state of the network, looking for outages and impairments. Other systems collect network data and do the fault segmentation analysis.

We batch-process features from the PNM sources once per day. Each source has a varying aggregation window. Data from the pollers reporting connectivity between modem and CMTS and account degradation algorithm tools arrive 3 to 6 times per device per day.

For the tool reporting account-level degradation, we used a one-day lookback window and aggregated the data across devices and polls reporting the proportion of polls with a certain impairment. Aggregated features include a potential filter indicator, number of correlated accounts found by the algorithm, FBC FM Noise and ICFR impairments.

The PNM tool reporting on connectivity between the model and CMTS offers measurement and count data. Devices report cumulative counts of T3 and T4 timeouts, lost syncs, resets, FEC corrected, uncorrectable, and total codewords until they are rebooted. We transform count features to get the incremental increase since the last poll unless there has been a reset, otherwise, we take the feature value as is. Measurement features include receive and transmit power, SNR, TTE, PMT, PMTE, CM phase, upstream ripples, and ETAP1 and ETAP2 measures. We further transform FEC errored and corrected codeword counts as percentage of total errored and corrected codewords. Then, we proceed with data aggregations: we take averages for features across all polls for all devices attached to the account during the day prior to the model evaluation. Finally, we examine data across four days prior to the model evaluation to estimate standard deviations for the features.

We collect data from the node network impairment algorithm for all events created one day prior to the model evaluation tied to the subscriber who requested the truck. We estimate the number of accounts

impacted by the same events and identify events by type: plant issues (ICPF plant faults, upstream and downstream FEC), wave, and suckout events. We count the number of events reported for each account.

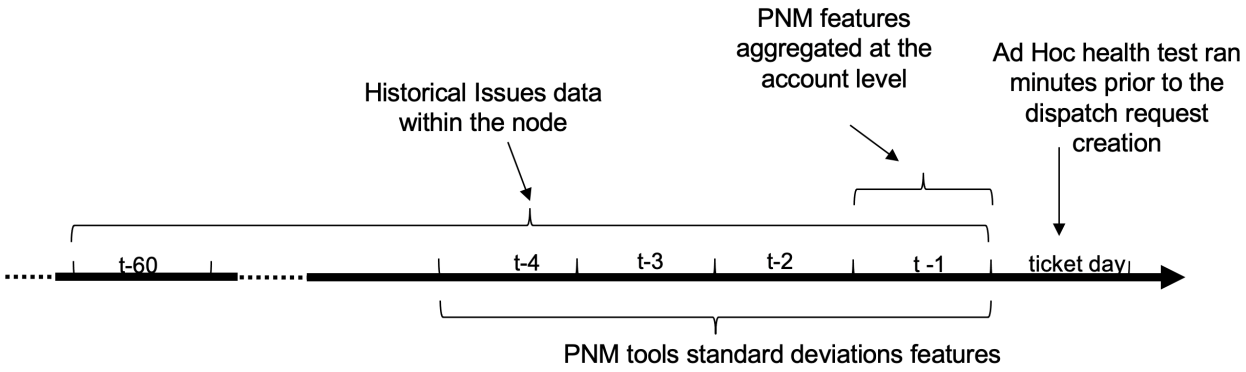


Figure 3 – Feature Extraction Timeline

6.1.2. Historical Issues Reported by Technicians

Another data source added to the model are historical rates of different types of issues reported by the technicians in the same nodes in the previous 60 days. Issues include construction and underground crew escalations, drop, ground block, and prior RTM escalations. We include these features as rolling aggregations for problems 60 days prior to the model evaluation.

6.1.3. Ad Hoc Remote Telemetry Health Checks

Agents use the remote Application Programming Interface (API) to run ad hoc telemetry checks using the capabilities of polling tools to collect the most recent data about the customers' network. Data from PNM pollers are collected 3 to 6 times per day and lag in time due to processing and data availability delays. Ad hoc telemetry health tests offer data from the same pollers but are retrieved on demand. In addition to the raw data, the tests also report pass/fail/pass-with-warnings analysis and offer heuristics for interpreting the raw data ranges (such as labels for failures in each DOCSIS analysis section). We transform raw data with similar aggregations described for PNM tools earlier and add labels and heuristics data as one-hot-encoded or ordinal (based on severity levels) features.

6.1.4. Topology Aggregations

Node topology was instrumental in creating features that detect network impairments. Node topology maps describe how equipment such as amps, taps, splitters, and cables are connected from the node to customers' homes. Data are collected by calling on the geographic information system API with a backend database to execute spatial intersect queries to identify equipment related to the locations supported by the Comcast footprint. Data are stored as a directional graph representing paths from the node to the tap, along with coordinate and equipment details.

Node topology maps were used to aggregate and average measurements, to provide a wider view of the network events surrounding a customer when the service call was being scheduled. One type of aggregation is at the parent level, which averages the telemetry over all customers who are immediate neighbors in the graph. Another aggregation is a comparison of parent measurements to those of parent's

neighbors. The idea is that if a single piece of network equipment is broken, we can observe the impact of this by comparing all the customers who depend on it (and thus have impaired service) with customers who do not depend on it but are otherwise similar because they share the same upstream network components. We also identified locations on the graph where amplifiers are located and estimated the average and standard deviation of measures across amplifiers on the path from customers' home to the node.

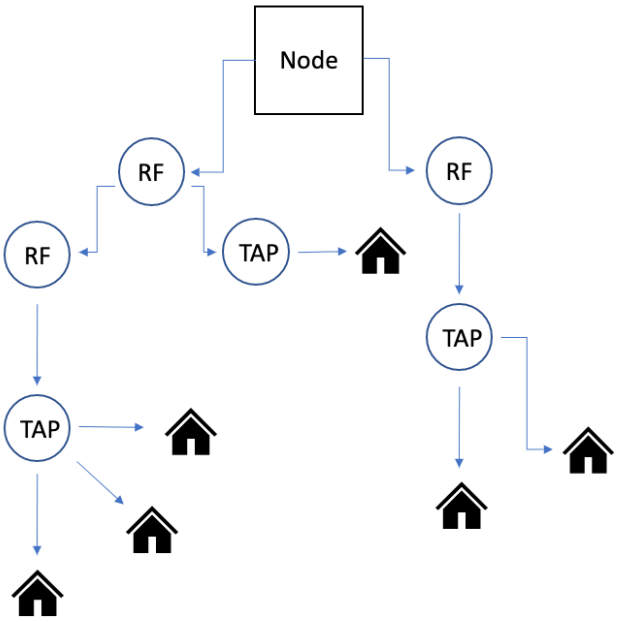


Figure 4 – Node Topology Diagram

Figure 5 provides an overview of the data, feature engineering and integration in our two-step modeling approach.

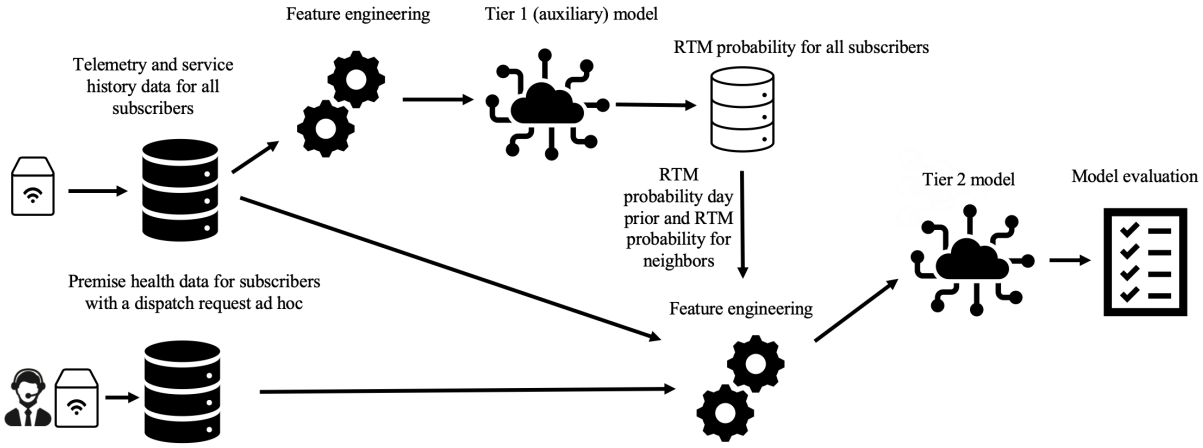


Figure 5 – Model Training Process

6.2. Model Evaluation

6.2.1. Model Performance

Our best-performing models is trained with an open-source XGBoost classifier (Tianqi and Carlos, 2016 [4]). The classifier has been calibrated for best hyper-parameters and uses 500 estimators to build trees with maximum depth 2. 239 features are used as model inputs in *tier 1* model and 375 features are used as model inputs in *tier 2* model. Our target label has a very small prevalence, resulting in the class imbalance problem. To assist the models to learn from a data set with class imbalance, we used the class-weighted version of the model to help remediate the class imbalance.

During the training stage, we separated data into training and testing sets, using the time of ticket creation to separate testing set of observations. Further, *tier 1* and *tier 2* models are trained on separate sets of data. In the latest iteration, tier 1 model was trained on 2,231,591 observations collected during December 2020 through March 2021.Tier 2 model was trained on 1,021,398 observations collected during April through June. During evaluation, tier 1 model is tested on full available out-of-sample data (April through August 2021), however, here we quote performance numbers for both models tested on the same data with 587,975 observations, collected during July through August 2021.

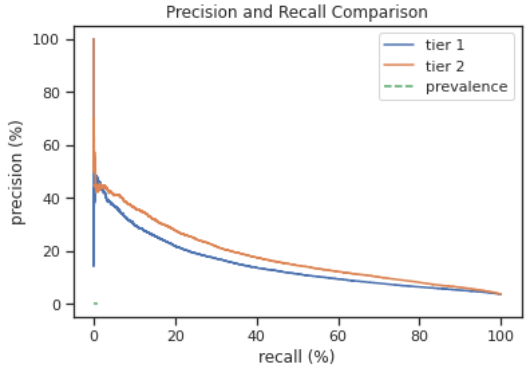


Figure 6 – Precision Recall Curve on Testing Data

The *tier 1* model is trained with telemetry data available for all subscribers. Calibrated at 5 percent recall, the model achieves 37.2 percent precision. Model performs 10 times better compared to the random guess.

The *tier 2* model is trained with combined telemetry, RTM probability of neighbors and subscriber prior day RTM probability, and ad-hoc health test features. Calibrated at 5 percent recall, the model achieves 41 percent precision: this means that the model calibrated to return at least 5 percent of true RTM trucks, can accurately identify 41 percent of cases. Model performs 11.2 times better compared to the average prevalence (random guess).

While Model stacking increases complexity of the model training process, it helps us to increase model precision by 3.5%. Figure below shows the precision and recall curve for the *tier* 1 and tier 2 models evaluated on the test data.

6.2.2. *Exploring Feature Importance*

We used normalized model weight to explore top features contributing to the model importance. Following graphs show features ranked and identified by source.

In *tier 1* model, features extracted from the account degradation analysis tool make up the majority of top features. Many of the features contributing to the model are based on the topology estimates – you can identify them by "parent" (tap aggregations) or "AMP" (amplifier aggregations) flags.

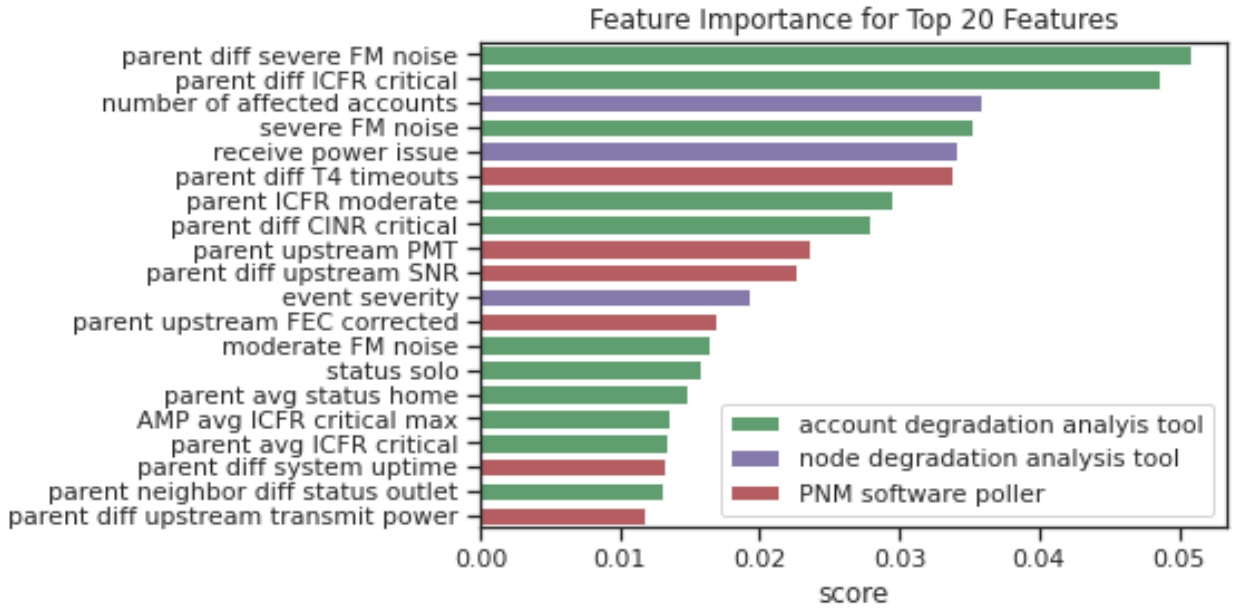


Figure 7 – Feature Importance for Tier 1 Model

In *tier 2* model, features extracted from the premise health check tool make up the majority of top features – it is expected as these features are extracted during the call with the agent and represent the most recent state of the device performance. The top 2 features contributing to the model come from the auxiliary tier one model and indicate probability of RTM 3 days prior to the call. The importance of these features leads us to believe that network problems might impact customers in a matter of days prior to the call. We also explore the directionality of features for *tier 2* model.

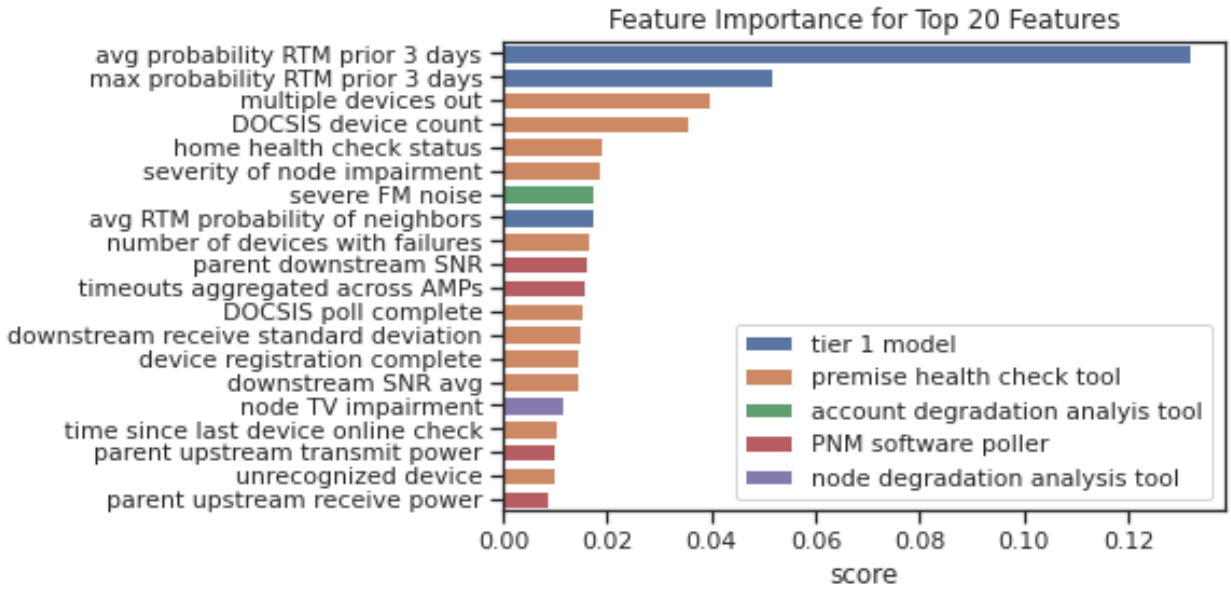


Figure 8 – Feature Importance for Tier 2 Model

Features positively correlated with the RTM outcome are number of device failures identified by health check, probability of RTM days prior, high downstream receive standard deviation, failed (high value of) home health check status, multiple devices out/offline, number of events with a TV-related impairment evident in the node, severity of node impairment, severe FM noise event incidence, and timeouts aggregated across amplifiers. For these features, higher values indicate a higher chance of RTM.

Features negatively correlated with RTM probability are time since last device online check/registration, parent upstream receive power, complete DOCSIS polls, found unrecognized devices, DOCSIS devices on account, complete registrations, downstream SNR, and parent downstream SNR. For these features, lower values indicate a higher chance of RTM.

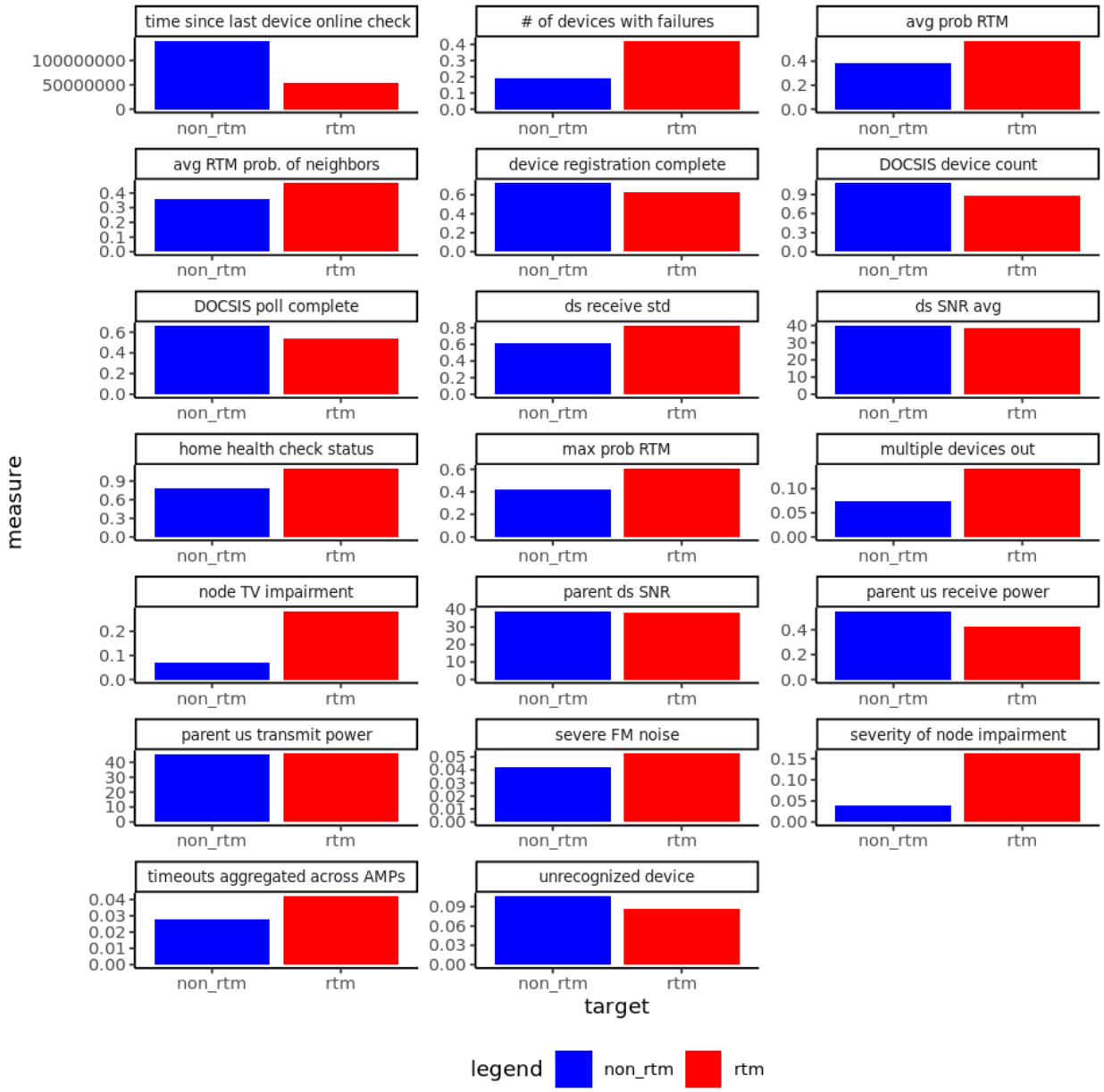


Figure 9 – Feature Exploration for Tier 2 Model

6.3. Cost Calculation

To assess the financial impact of the RTM model, we constructed a cost analysis comparing the business as usual (BAU) process to our model (RTM). The goal of this analysis is to identify the minimum precision required by the model to incur a net zero savings relative to BAU. This allows us to know when our model precision exceeds the break-even point before deploying in a live environment.

We assess the net savings relative to BAU when the RTM model makes a positive prediction (i.e. when the model recommends sending a line technician instead of a fulfillment truck). Note that the net savings for a negative prediction is effectively zero since a negative prediction defaults to the BAU protocol.

The cost analysis presented here also assumes that the model will make recommendations without any subject matter expert (SME) interventions. Asking SMEs to review model predictions, thereby utilizing the "human-in-the-loop" framework, improves the model precision but its impact to the overall cost is yet to be estimated.

Running a machine learning model incurs a fixed cost of computing, data storage, and engineering resources. Additional costs might include if we use a human-in-the loop model and each prediction is reviewed by an expert. The cost to run a model would be a fixed cost per prediction.

There are two approaches we used to assess the net savings incurred by our model. First, we describe in more detail the simpler of the two approaches. *Customer-level approach* estimates the net savings by taking into consideration only the customer with the model prediction in question. The latter approach, referred to as the *neighborhood approach*, additionally takes into consideration the impact of sending a line technician to the customer's neighborhood. Sending out technicians early and preventing more fulfillment trucks from going out into the neighborhood can generate additional savings. Let's first denote the following costs:

Table 5 – Cost Model Definitions

Cost of fulfillment truck	F
Cost of line truck	L, where L > F
Operations and research cost per prediction	R

6.3.1. Customer-Level Cost Calculation

In a true positive scenario where the RTM model correctly predicts that a line technician is required, the RTM model incurs a cost of L and the BAU model incurs a cost of sending both trucks, $L + F$. Therefore, the net savings for a true positive is

$$(L + F) - L - R = F - R$$

In a false positive, the RTM model incorrectly recommends a line technician and we will assume that in this case, we would need to send a fulfillment truck after incorrectly sending a line technician, costing the company $L + F$. We note here that this is a very conservative estimate as we can expect that sometimes a line technician can resolve issues that normally require a fulfillment truck (e.g. when an issue exists outside the home). For a false positive, the BAU's cost would simply be F. Let's also assume that sometimes customers cancel their scheduled trucks, and we will denote this cancellation rate as Y. Then, the net savings for a false positive is

$$F * (1 - Y) - (L + F) * (1 - Y) - R = -L(1 - Y) - R$$

We define precision as the proportion of predictions the model made correctly. Using these two estimates, the net savings, S_{net}, of the RTM model with precision P is given by

$$((F - R) * P) - (L * (1 - Y) - R) * (1 - P) \; = \; S_{net}$$

By setting $S_{net} \; = 0$, we can solve for precision P to get the precision required to break even with BAU's costs, which we denote as P_{even}. In customer-level approach,

$$P_{even} \; = \frac{L(1 - Y) - R}{F - 2R + L(1 - Y)}$$

Graph below illustrates the relationship between savings and the model precision. Actual estimates are omitted to comply with our company's data reporting policies. We offer the direction of the positive savings and show the precision required to achieve the break-even point.

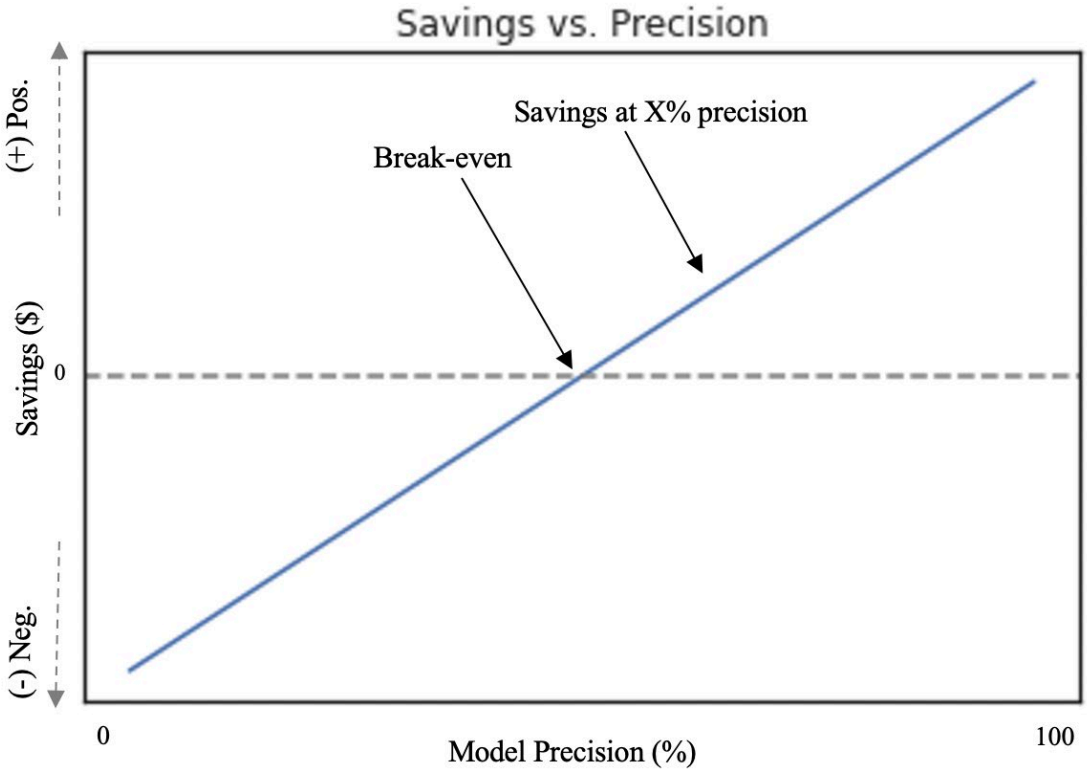


Figure 10 – Precision vs. Savings Curve Illustration

6.3.2. Neighborhood-Level Cost Calculation

In the neighborhood-level analysis we also consider the RTM model's impact to the neighborhood. By sending out line technicians early, there is a positive impact to other customers in the neighborhood suffering from internet issues that are predicted to also be RTM issues. In this scenario, if there exists a customer C_A with an RTM prediction, then we will delay sending out a fulfillment truck to C_A's neighbor, C_B, if the model also recommends a line technician for customer C_B within 2 days of C_A's prediction. In this analysis, we define two people as neighbors if they not only share the same CMTS node but also all the devices between the CMTS node and their respective homes.

In this case, we consider the true positive and false positive costs from customer-level calculation for the original customer C_A and we *additionally* consider the cost of a true positive and false positive to a neighbor C_B. A true positive for neighbor C_B occurs when the model correctly predicts an RTM for the original customer C_A *and* correctly predicts an RTM for customer C_B. In this case the true positive net savings for customer C_B is the cost of sending out a fulfillment truck needlessly, F.

A false positive for customer C_B occurs when we delay the customer from receiving a fulfillment truck when we incorrectly predict an RTM for C_A *or* C_B. The cost here is much less severe than only customer-level approach, as it is mostly a delayed response and possibly a second call to an agent made by the neighbor. If we denote the cost of contacting an agent as A and we assume that Z percent of the time, customer C_B will opt to call an agent instead of using an automated scheduling service, then the cost the model incurs for being wrong with customer C_B is $A * Z$. Because both the neighborhood sizes vary and the number of trouble calls made in a neighborhood vary across customers/regions, we used the test data to generate the net savings of our model as a function of the model's precision. We found that this neighborhood analysis decreases the required precision to break even, suggesting that implementing this delay service will have an additional positive financial impact.

6.4. Model Deployment

6.4.1. Model Deployment for Proof of Concept (POC)

We explored a "soft" deployment of the model for selected markets during our POC stage of development. Our goal was to obtain buy-in from our colleagues and test the model validity without incurring negative costs to the business or impacts to the customer experience.

During the POC we worked with the "human-in-the-loop" strategy. In this process, our program evaluated dispatch requests created by agents every hour and flagged "highly predictive" RTM candidates identified by the model. We then sent emails with the predictions to the regional subject matter experts (SMEs) for triage using AWS SES (Simple Email Service) email automation and posted in the Microsoft Teams channel. Emails were sent right after we received dispatch request logs to simulate a near real-time processing and prevent SMEs access to the truck outcomes (so that their judgement won't be "poisoned" by the truck outcome knowledge).

One of the participating markets explored dispatching a line technician the day a prediction was deemed valid by the SME, with the intent of remediating any customer pain from plant impairments prior to the standard technician visit. The majority of the market SME's evaluated network telemetry via desk top tools, logged notes and then validated the label (whether RTM is needed or not). Of note: the markets opted not to cancel the standard technician visit and schedule a line technician in their place until the predictions provided a higher degree of confidence. We have collected over 4,500 observations from the POC stage. SMEs evaluation of the model predictions resulted in 48 percent model precision. One of reasons for higher precision during the trial could be that evaluations were made on the dispatch request and not on the confirmed service calls. On average, 35 percent of scheduled service calls are canceled.

A debrief with the involved SMEs helped us to discover additional data sources, improve our feature engineering process, and identify false positive cases.

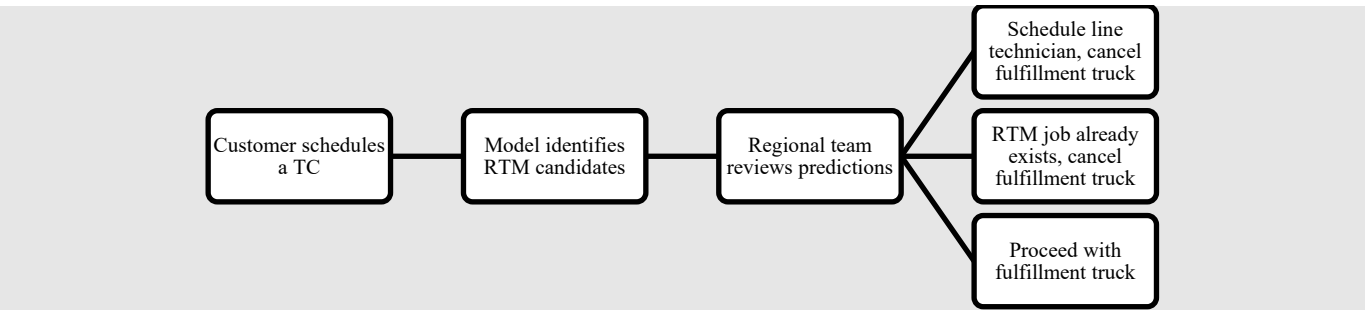


Figure 11 – POC Stage Model Deployment Workflow

6.4.2. Production Model Deployment

Figure 10, below, illustrates how we plan to deploy the RTM model into production. We will rely on the Redis (Remote Dictionary Server) [5], a fast, in-memory, open-source key-value data store to host the data. Our internal platform will host the model API. Model will consume two sets of features described earlier: payload from the ad-hoc premise health check passed by the agent making the API call, and features derived from the archived telemetry and *tier 1* model input stored in Redis feature store. Model engine will process features and output predictions to pass to the call agent. Subsequently features used and model result will be logged in AWS S3 (Amazon Web Services Simple Storage Service).

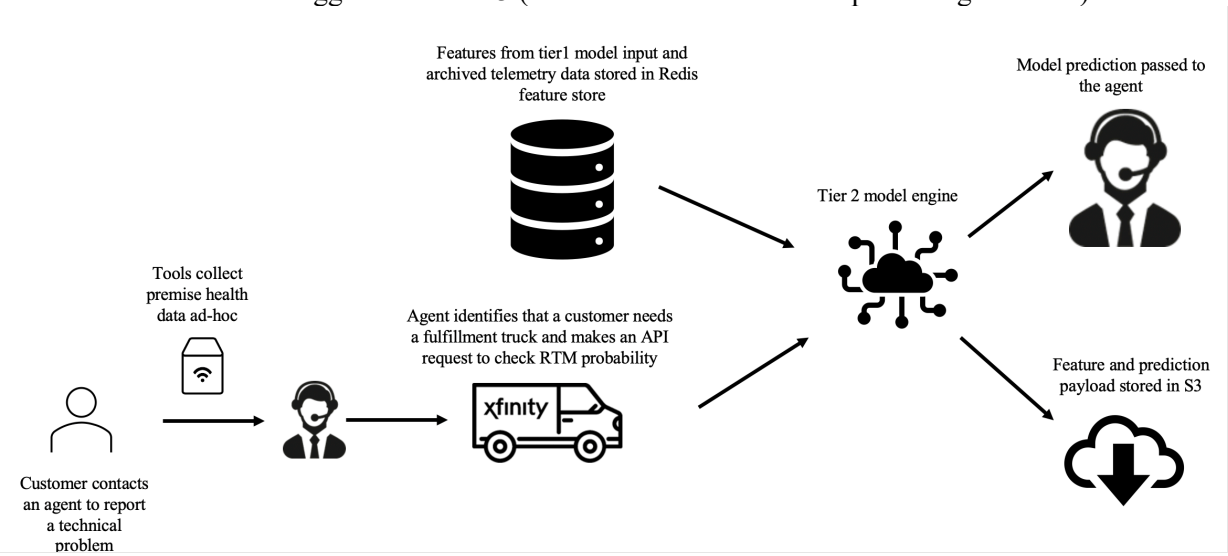


Figure 12 – Model Deployment Diagram

7. History of Model Development and Future Work

Network telemetry and topology features are used across multiple use-cases across the organizations. Originally, a set of features derived from the network telemetry was developed and applied to the classification problem predicting whether technician will only need to do work outside of the home. This model was implemented during COVID-19 business operations and helped us reduce the interactions between customers and technicians. We discuss this model in the Vengriglia et al [2] paper mentioned earlier.

The RTM model has been developed based on the telemetry features and benefited from RTM-specific features derived from topology and RTM probability of neighbors. During the development of the model, we have received feedback from the field operations colleagues and redefined the target label for the model. Originally, we trained our model to recognize the RTM fix code entered by the standard technicians. We then learned that regional teams put the requests through a triage process and only send a line technician for severe service disruptions only, taking other steps to help customers with non-severe service disruptions. Receiving this feedback, we have redefined our target label and updated the model to learn that the standard technician recommended an RTM, field operations triaged the case to send a line technician, and the line technician confirmed that the referral vas valid.

Our model development process does not stop with this paper or the model deployment, which is an early marker in the overall trajectory of ML and network operations. We have worked and will continue to work with experts to identify several telemetry sources that can help us improve model precision. First, we have learned that upstream interface problems can help us detect issues impacting multiple customers. Second, we have also acquired and plan to integrate packet data, hoping that this source can help us detect intermittent issues better.

8. Conclusions

In our paper, we described the complexity of the HFC network and the intelligent systems that collect measurements to monitor the health of the plant. While the telemetry measurements collected from the plant help us detect the symptoms of impairments, the machine learning approach is a predictive approach to plant maintenance and can enable us to dispatch line technicians to the right area in the network and remove problems more efficiently. The implementation of the Refer to Maintenance (RTM) with Machine Learning model would help us move from a reactive to a predictive maintenance approach. In addition, an ML approach can identify and uncover intermittent issues that are not always prioritized (or able to be prioritized) for line technician work, and can deteriorate the customer experience.

Adopting a machine learning approach will help us to continuously deliver an improved customer experience and streamline efficiencies across the service pipeline. Instead of multiple visits by different technicians, our customers could have their issues resolved faster with a single visit by the right technician. Further, we expect that a corrected network issue will have a positive impact on the whole neighborhood, delivering an improved experience to all customers in the area and thus reducing contact rates and unnecessary truck rolls.

We have described how cost analysis can help determine the appropriate model precision to achieve before deployment. Using this approach, we can deploy the model that is tuned to achieve an outcome beneficial to the business. In summary, the Machine Learning approach is a game changer that will help us to achieve continuous operational transformation. By identifying and prioritizing network issues, we can deliver the right technician at the right time and improve customer experience.

9. Disclaimers

We collect, store, and use all data in accordance with our privacy disclosures to users and applicable laws.

Abbreviations

API	Application Programming Interface
AWS	Amazon Web Services
BAU	Business as Usual
CATV	Community Antenna Television
CM	Customer Modem
CMTS	Cable Modem Termination System
CPE	Customer Premise Equipment
dBmV	decibels relative to one millivolt
DOCSIS	Data Over Cable Service Interface Specification
ETAP	Electrical Transient Analyzer Problem
FBC	Full Band Capture
FEC	Forward Error Correction
HFC	Hybrid Fiber Coax
ICFR	In-Channel Frequency Response
ISP	Inside Plant
ITG	Interactive Troubleshooting Guide
IVR	Interactive Voice Response
MDU	Multi-Dwelling Units
ML	Machine Learning
OSP	Outside Plant
PMT	Post Main Tap
PMTE	Post Main Tap Energy
PNM	Proactive Network Maintenance
POC	Proof of Concept
RF	Radio Frequency
RNN	Recurrent Neural Network
RTM	Refer to Maintenance
S3	Simple Storage Service
SES	Simple Email Service
SME	Subject Matter Expert
SNR	Signal to Noise Ratio

Bibliography & References

[1] SCTE Expo 2016 - A Comprehensive Case Study of Proactive Network Maintenance; Wolcott et al

[2] VIAVI 2015 - Understanding Equalizers, Pre-Equalizers, and Their Use in Localizing and Troubleshooting HFC Issues; Fenton, Mike and Jump, Larry

[3] SCTE Expo 2020 – Training Machines to Learn from Signal Meter Readings; Ventriglia et al

[4] XGBoost: A Scalable Tree Boosting System, March 9, 2016; Chen, Tianqi and Guestrin, Carlos

[5] Redis, Documentation retrieved from https://redis.io/

SD-WAN Security and SASE

How to Secure SD-WAN and Role Of SASE

A Technical Paper prepared for SCTE by

Charuhas Ghatge
Senior Manager, Product and Portfolio Marketing
Nuage Networks by Nokia
520 Almanor Ave, Sunnyvale CA
(510) 299-2989
Charuhas.ghatge@nokia.com

1. Introduction

Software-defined wide area networks (SD-WAN), a software approach managing wide-area networks, offer ease of deployment, central manageability, and reduced costs, and can improve connectivity to branch offices and cloud. End users are excited about SD-WAN because it enables them to manage and add network functionality using a cloud-first strategy for application access and delivery.

Compute resources and associated cloud services are exploding, and traditional enterprise network boundaries have expanded into the public cloud, branch locations, and intelligent edges. So, what does this all mean to the branch security?

Cloud computing has created several challenges since networking and security are incompatible with the cloud-centric and mobile-first business models. The network is rigid and static. Security is heavily centered around the data center, fragmented across multiple domains of physical locations, cloud resources, and mobile users. Networking and security have created silos that were built and implemented decades ago, and new functionalities are added and patched in as needed in a haphazard way. Secure Access Service Edge (SASE) is a new paradigm defined by Gartner that combines network and security into a single cloud-based service.

In this whitepaper, we will discuss SD-WAN security features and explain SASE – what is SASE, why is it needed, what does it comprise of and its deployment considerations.

2. SD-WAN Security

Existing security models cannot effectively address the new security requirements driven by move to cloud and the evolving threat landscape.

First, due to SD-WAN allowing the use of broadband internet as a transport mechanism, the internet, which is traditionally not a guaranteed secured link, the access to it needs to be made secure.

Second, current protection model in Enterprise branch is basic and not enough to secure local internet breakout to cloud as all traffic is steered over Multiprotocol Label Switching (MPLS) to Data Center (DC) sites where security is applied. Also, there is not much end to end micro-segmentation between branch and DC/cloud applications across the enterprise.

Third, with the increasing attack sophistication and evolving threat landscape we cannot assume that all attacks can be prevented by protective controls. Currently there is not much visibility to branch user traffic. Visibility and security analytics are key to help detect attacks.

Last, but not least, the current security provisioning model for applications is largely manual and device-centric.

Major security functions for a secure SD-WAN are discussed in the subsections below.

2.1. Securing broadband internet access

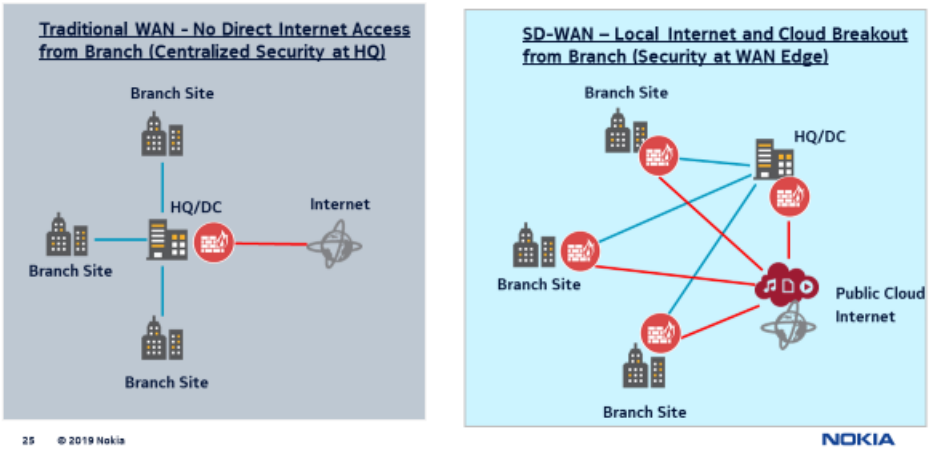


Figure 1 - Local Internet breakout

The Internet is not very secure for enterprise WAN requirements. Hence cloud-based application traffic is often backhauled from the branch to the enterprise Headquarters (HQ) before being handed off to the Internet. This introduces delay and jitter and hence application performance is often compromised because of WAN bandwidth constraints at the branch and added latency from backhauling connections.

The solution is to use direct internet connectivity to the cloud and web applications from the branch. The SD-WAN solution needs to make these internet connections secure and reliable by creating encrypted tunnels between every site in the SD-WAN, while taking advantage of Secure Socket Layer (SSL) security provided by the Software as a Service (SaaS) application for traffic going from the branch to the application directly using the Internet. This makes the Internet access more secure. With such encrypted links and a stateful firewall, an SD-WAN solution can prevent unauthorized outside traffic from entering the branch. The stateful firewall is usually implemented directly on the Customer Premise Equipment (CPE) device of the SD-WAN and no external security hardware or software should be needed for the stateful firewall functionality.

2.2. Threat prevention using IDS and IPS

Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) are important to detect and prevent the known attacks by recognizing the virus signatures. Threat prevention component prevents malware from penetrating the network, regardless of application traffic in which they are hiding. It is important that the IDS/IPS functionality should be implemented natively on the CPE of the SD-WAN.

It uses signatures of known attacks to match traffic that passes through the CPE to prevent attacks and these signatures have been divided into groups of relevant signatures. IPS/IDS policies can be defined and managed centrally via the SD-WAN GUI or APIs. It is recommended that signatures be updated dynamically from cloud and applied on the local CPE device of the SD-WAN. The SD-WAN GUI should display statistics and generate reports of intrusion event details rule hit counts etc.

2.3. Stateful Firewall

A stateful Firewall filters packets based on the state and context of network connections and provides full protocol inspection considering the STATE+ CONTEXT of the flow, thereby eliminating additional attacks surface.

A stateful firewall understands the network flow and can identify data packets of a flow. Since a stateful firewall can look deeper into packet payloads, it can support DDoS (TCP/UDP/ICMP Flood), Port-scans etc.

Step one is to secure the branch within the SD-WAN network and to secure the links for the internet breakout. This is done by implementing a stateful firewall right on the CPE device without any additional external equipment or implementing any third-party solution. This way local branch users can safely access the corporate resources and SaaS applications in the cloud while being protected from both, inside and outside threats.

Enterprises can define access rules and policies to allow or deny traffic to/from the application – for example the administrator can define a policy to deny access to cloud storage app that is not in the corporate IT's management domain. There could be a policy allowing access to an application like say, Office365.

Note that, Stateful Firewall should preferably be validated by 3rd party for PCI-DSS network firewall requirements.

2.4. Threat Intelligenece based on IP Reputation

Threat Intelligence feature detects branch device communication to risky IP addresses and sites. It also generates reports on access to risky IP and sites from SD-WAN branch user devices. Threat Intelligence uses IP reputation database that is updated daily from cloud.

2.5. Web/URL Filtering

URL filtering limits access by comparing web traffic against a database to prevent employees from accessing harmful sites such as phishing pages. Users spend increasing time on the web, surfing their favorite sites, clicking on email links, or utilizing a variety of web-based SaaS applications for both personal and business use. While incredibly useful to drive business productivity, this kind of unfettered web activity exposes organizations to a range of security and business risks, such as propagation of threats, possible data loss, and potential lack of compliance.

2.6. L7 and SaaS Application Control

One of the prime benefits of SD-WAN is its ability to allow a direct access for a branch user to the cloud and SaaS applications. A good secure SD-WAN must have the ability to restrict user access to a specific

application, be able to set application-based policies and monitor and log application usage. For this, it needs to have a layer-7 Deep Packet Inspection (DPI) engine to recognize thousands of application types and pre-defined SaaS services - Office365, WebEx, Salesforce, GitHub, JIRA, Azure, AWS, Google among others for easy access as well as monitoring.

2.7. Automated and rapid response to threats

A rapid and efficient incident response continues to be the biggest challenge facing security teams today. The sheer volume of these signals means that a lot of critical alerts miss getting the timely attention. Security teams need help to scale better, be more efficient, focus on the right issues, and deal with incidents in a timely manner.

By defining automated response action, the users can prevent malware from infected branch device from entering corporate network. For example, leverage network security analytics to identify suspect end-points based on threshold alerts and use service chaining to dynamically insert services (such as NGFW or IPS) for suspect traffic. The suspect traffic could be diverted to cloud-hosted security service.

2.8. Realtime security analytics and automation

Some threats can be quite sophisticated and cannot be prevented by the protective methods. Real time traffic monitoring and security analytics provide fine granular visibility to locate the cause and reprogram the security response. Automated action reduces the time to react to the anomalies and reduce the impact.

With end to end visibility and control for each application, the operator can detect, protect resources at a very granular level, and use automation to respond in real-time to threats. SD-WAN security monitoring should allow you to do a contextual flow visibility of each flow.

Realtime security analytics helps in Threat hunting, Network Forensics as well as troubleshooting. The forensic reports can be used for compliance and security audits – both internal audits as well as external agency audits.

Realtime network security monitoring allows you to generate alerts based on security events - port scan detection, port sweep detection, security policy violations and volumetric DDoS attacks.

3. Third-party Security Technology Ecosystem

Third-party security products should be a part of the overall effective security functionality for an SD-WAN solution. Third-party security solutions should be incorporated in a couple of ways:

Most enterprises have an existing set of security infrastructure and solutions they use. SD-WAN vendor should partner with the best of breed security appliances already present in the IT infrastructure.

Secondly, for a tighter alignment with the security vendors' advanced functionality, the security VNFs (Virtual Network Functions) from those vendors should be integrated onboard the CPE. Both these options are described below.

3.1. Security partners and Cloud Security services

When it comes to security, it's simply not feasible for a single SD-WAN vendor to provide every security functionality on its own. The scope of threats, risks, and corresponding technologies is simply too great. SD-WAN vendor should establish technology partnerships covering several security domains, such as industry-leading next-generation firewalls and secure web gateway and Cloud Security services.

The integration with a Cloud Security vendor allows you to route local branch Internet-destined traffic directly to the security cloud to enable a fast and secure experience This eliminates the need to backhaul local traffic to the internet gateway.

You can route specific traffic to the Cloud Security vendor's security cloud through IPSec tunnels for further security. For example, you can define an action to route the traffic to Cloud Security vendor's cloud as part of the response to a threat.

3.2. Hosted VNF on CPE and Service Chaining

The CPE of the SD-WAN solution also acts as a powerful platform to host VNFs. SD-WAN implantation should have the service chaining functionality. By service chaining the VNFs (many a times, on-demand), a dynamic and advanced security features are provided.

4. Secured Access Service Edge (SASE)

4.1. Overview – What problem is being solved by SASE

As cloud becomes pervasive and driven by digital transformation of enterprises, the networking and security needs of an enterprise are evolving. Traditional enterprise network and security paradigm was centered around applications in private data centers. Although networking was complex, the security risk profile was well defined. Migration of applications to the cloud has redefined the networking and security. Many companies use SD-WAN to securely connect branch offices to their corporate networks instead of relying on traditional and expensive MPLS links. SD-WAN also facilitates direct access by corporate branch offices to the public clouds and SaaS applications. This creates stringent security requirements from the branch to the cloud.

Gartner observed this trend of security and networking requirements and has recently defined a new framework that converges network (SD-WAN) and security into a single cloud-based service: Secure Access Service Edge (SASE).

This section discusses what is SASE, it's advantages, key components, deployment options and recommendations for a successful SASE implementation.

4.2. What is SASE and its key components

SASE is a framework that brings together networking and security services in one unified solution designed to deliver strong security from edge to edge, delivered as a service. It is not an RFC or a static architecture, but rather a recommendation and a framework.

SASE has two major functional blocks – Networking (SD-WAN) and Security, as shown in figure x. SD-WAN is the foundation of SASE and security features are offered on and beyond SD-WAN.

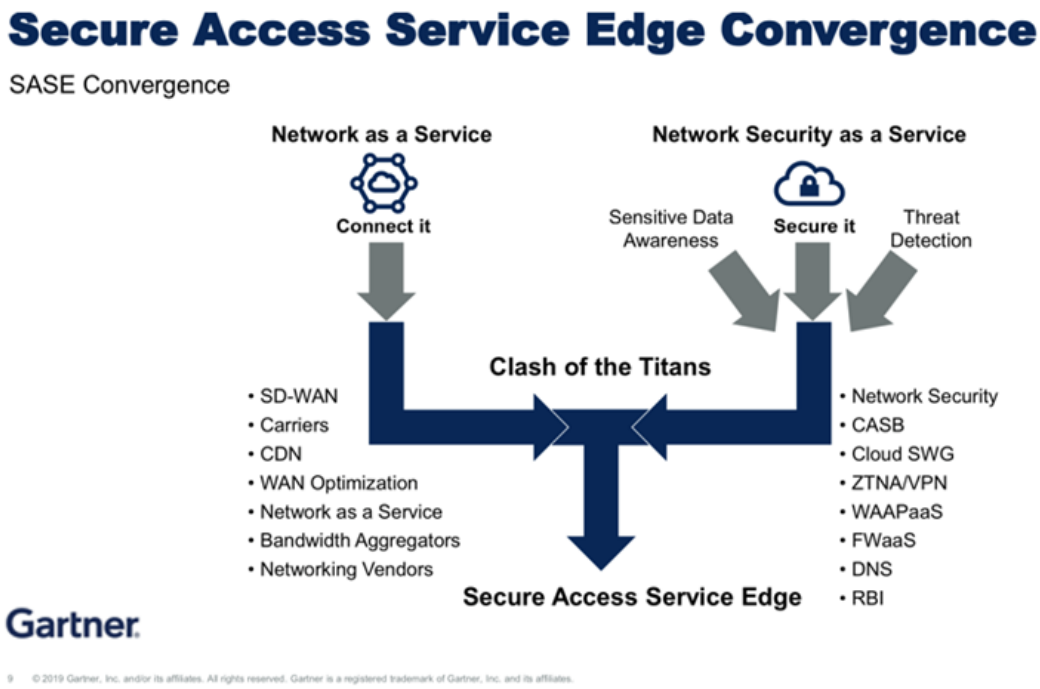


Figure 2 - SASE Components

4.3. SASE Networking (SD-WAN) – features and benefits

SASE networking capabilities offers the following benefits and capabilities:

- Network Agility – Flexibility and choice of MPLS, broadband or LTE.
- Multi-Cloud and SaaS Connectivity without the need for backhauling.
- Network Management and Automation - Real-time network monitoring, analytics, and reports.
- Application Performance Assurance – Business-policy based application prioritization.

Table 1 - SD-WAN Features and descriptions

Feature	Description
Comprehensive Routing capabilities	Full stack of routing protocols to support switching and routing personalities
Access and Connectivity to and from Anywhere	Seamless connectivity and policy management across fixed (internet, L2 and L3) and mobile WANs
Performance based POP selection	Support for multiple paths and POPs and performance-based selection ability
Application aware routing and traffic steering	Providing optimal application experience based on application types

Feature	Description
Hybrid WAN support (Full MPLS/Ethernet) for legacy Datacenter access	Seamless integration of existing networking to access data center and apps
Multi-Cloud & Hybrid Cloud connectivity	Policy based access to and across applications in private cloud and multiple public clouds
Connectivity Security – VPN, IPSec	Embedded encryption and end point security
SD-WAN Service Portal	Multi-tenant SD-WAN portal hosted by CSP for the visibility and control the CSP service and operations teams need to manage multiple SD-WAN services.
WAN Optimization & Bandwidth Aggregation	Optimizing the use of available network for availability and performance

4.4. SASE security components

Table 2 - SASE Components

Component	Description
IPS	Intrusion Prevention system
IDS	Intrusion Detection System
Firewall	Stateful Firewall
Realtime Security Analytics and Automation	With end-to-end visibility and control for each application, the operator can detect, protect resources at a very granular level, and use automation to respond in real-time to threats.
SWG and DNS Filtering	Secure Web Gateway is used to protect users and devices from online security threats by enforcing internet security and compliance policies and filtering out malicious internet traffic
ZTNA	Zero Trust Network Access
CASB	Cloud Access Security Broker - According to Gartner, a cloud access security broker (CASB) is an on-premises or cloud-based security policy enforcement point that is placed between cloud service consumers and cloud service providers to combine and interject enterprise security policies as cloud-based resources are accessed.
DLP	Data Loss Prevention - DLP provides visibility across all sensitive information, everywhere and always, enabling strong protective actions to safeguard data from threats and violations of corporate policies.
FWaaS	Firewall as a Service

4.5. Deployment considerations for SASE

SD-WAN, although a new technology, has been maturing in recent times and the number of deployments is growing rapidly. The security technology including cloud-delivered security is mature and most enterprises have deployed security in some form or other. Given this, a SASE deployment of rip-and-replace is not practical because of the existing investments. The most pragmatic solution will consist of utilizing existing security, especially cloud-delivered security to offer SASE. SD-WAN's flexibility to integrate with existing security vendors is very important. A complete SASE solution from a single vendor would compromise completeness due to vendor's technology limitations, it will reduce flexibility in a very dynamic space of enterprise security, and it will also risk the vendor lock-in. It should also be noted that MEF has specified SD-WAN standards, however, the security architecture is and likely will continuously evolve.

So, a good SASE solution should have the flexibility of:

- Scalable, high performance SD-WAN because SD-WAN forms the foundation of SASE.
- Exhaustive natively embedded security functions within SD-WAN itself.
- Integration with cloud security vendors for advanced security functions, as well as, evolving capabilities in this space.
- Ability to deliver SD-WAN as well as security features via Managed Service Provider or MSP partners' cloud and POPs.

This flexibility will allow the Managed Service Provider to offer cost effective SASE solutions based on its enterprise customer's specific needs rather than one size fits all expensive approach. It will enable MSP to differentiate its offer from other cookie-cutter (me too) solutions.

5. Conclusion

Since, SASE is a framework, rather than a standard, each vendor's implementation is unique to that vendor, depending on their expertise. Also noteworthy is that, like any new technology, there is a hype cycle and then there is reality. Currently, SASE is at the peak of its hype cycle. We believe, and Gartner agrees, that SASE is a 5-10-year journey versus a defined destination.

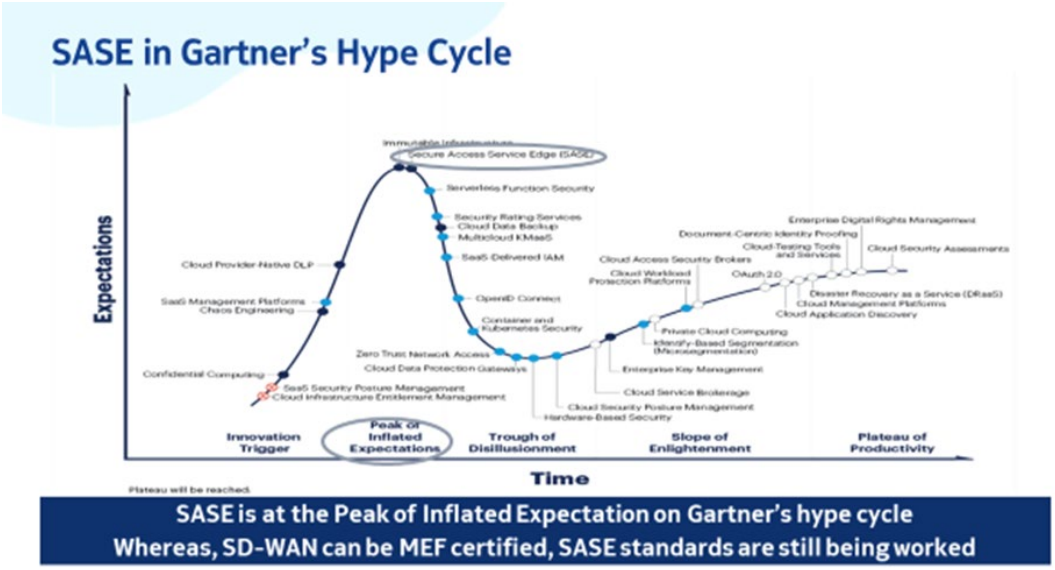


Figure 3 - Gartner Hype Cycle

Considering this, an MSP should opt for a solution that provides strong foundational capabilities and at the same time provides flexibility to evolve the solution in a vendor agnostic manner.

SD-WAN is being widely adopted by enterprises, greatly simplifying the branch network environment by integrating multiple functions (Internet and hybrid WAN connectivity, Advanced Security & NGFW, Cloud on-ramp, Application Experience, Wi-Fi, etc.). The SASE platform provides another strategic advantage for MSPs to offer SASE, a comprehensive networking and security solution, as a Service. With its foundational SD-WAN capabilities, advanced security embedded within the platform, and open vendor-agnostic platform affords MSP a future-proof SASE solution.

Abbreviations

AWS	Amazon Web Service
CASB	Cloud Access Security Broker
CDN	Content Delivery Network
CPE	Customer Premise Equipment
CSP	Communications Service Provider
DC	Data Centre
DLP	Data Loss Prevention
DNS	Domain Name System
E2E	End to End
FWaaS	Firewall as a Service
GCP	Google Cloud Platform
IDS	Intrusion Detection System
IPS	Intrusion Prevention System
IPSec	Internet Protocol Secure
MEF	Metro Ethernet Forum – An Industry consortium defining SD-WAN standard
MPLS	Multi-Protocol Label Switching
POP	Point of Presence
SASE	Secure Access Service Edge
SD-WAN	Software Defined Wide Area Networking
SWG	Secure Web Gateway
WAAPaaS	Web Application and API Protection as a Service
WAN	Wide Area Network
ZTNA/VPN	Zero Trust Network Access / Virtual Private Network

Bibliography & References

Secured Access Service Edge (SASE) – Gartner, 2019

Security Strategies in the Wake of Nation-State Attack Evolution

A Technical Paper prepared for SCTE by

Emma Rochon
Security Architect 2
Comcast Cable
1800 Arch Street, Philadelphia PA 19130
215-262-3275
emma_rochon@comcast.com

Nancy Davoust
VPII, Security Architecture, Identity and Access
Comcast Cable
1800 Arch Street, Philadelphia PA 19130
303-862-0143
nancy_davoust@cable.comcast.com

1. Introduction

Over the past decade, private businesses have increasingly been the victim of nation-state cyberattacks, which are defined as attacks carried out by a hacker, or a group of hackers, working with adversarial government to commit cybercrime against another country. Notably, there has been an 100% increase in nation-state incidents from 2017-2020. Specifically, nation-state attacks rose from 17% of all known attacks to over 40% during the last 3 years [3].

In this paper, we will highlight some of the issues with nation-state attacks and provide some guidance on how to defend against these attacks on a regular basis. There is no one way to defend against attackers, but there are strategies that can be implemented to prevent an attack from being catastrophic.

2. Nation-State Attack Background

The Cybersecurity and Infrastructure Security Agency (CISA), the Federal Bureau of Investigation (FBI) and the National Security Agency (NSA) and other government agencies continue to post warnings about nation-state attackers and their techniques to attack as they continue to advance. Many of these attacks could be applied to the cable industry. The cable industry has evolved significantly in the past 20 years with the rise of Internet and connected services to become a prized attack surface for nation-state actors.

Who are these nation-state attackers? Often these are known as Advanced Persistent Threats (APTs) because they are constantly working to come up with new attacks and they never go away. The largest groups of APTs come from Russia, China, North Korea, Iran and the U.S. [4]. The attacks are becoming more impactful, with posturing around government influence and even threats of physical war. Additionally, economic damage caused by nation-state attacks impact its citizens. The pipeline attack earlier this year in the U.S. caused fuel shortages for cars in the impacted areas which impacted people getting to work and school [10].

Just like all hackers, nation-states take advantage of known issues. Unlike the general hacking community, nation-states have political goals as well as financial goals. These nation-state attacks include theft of political or military data, advancing foreign policy, disinformation campaigns, or financial motivation. The nation-state attackers profile includes larger targets, more financially rewarding targets, and targets that will bring damage to other countries [4].

And most importantly, nation-states are well organized, funded, and very patient. It is not always directly obvious that a nation-state is sponsoring a cyber-attack, and many attacks do not have confirmation on their origin yet are heavily suspected to be a nation-state attack. If nation-states are going to make their goals, they will be persistent. Once attackers get into systems, they try to stay in them. A key activity for nation-states is to perform reconnaissance once in a system and prior to striking, so they have a good inventory of all the attack methods and damage they can cause before they strike. However, some attacks are very much predicated on timing of new zero-day attacks being found. Zero-day attacks are serious software vulnerability exploits, which the developer or vendor may not be aware of. It is always a race to exploit a new zero-day attack before a patch can be applied to fix the issue [5].

To make matters worse, on July 14, 2021, SecurityWeek reported that China has passed a new law that any zero-day vulnerability discovered by anyone in China is required to be shared with the CCP, and is prohibited from being shared with anyone, or any government outside of China. Product manufacturers may learn of vulnerabilities in their products, but this is not completely clear. This will make it more

difficult to gather information. Until then, security researchers around the world shared information about zero-day vulnerabilities they found.

As we look at how to best secure access from people and software to software resources storing and using sensitive data on various types of devices, we need to understand how to best apply our time and resources to protect against these nation-state attacks. Nation-state attacks have evolved from political statements to other types of attacks such as military intelligence, election interference, resource interference and ransomware. Protections include doubling down on security overall, and especially ensuring there are strong authentication, access management and authorization systems to protect against attacks like phishing, and escalated privileges.

There are many exploitable attack surfaces and nation-states understand how to take best advantage of many of these. One well known path is using supply chain weaknesses to plant malware which may propagate into other systems or provide access and visibility once implanted within a company. There have been cases of malware being included in devices or software specifically in products destined for the country they desire to attack. Therefore, when a device is powered on, software is now running inside the country and company they are trying to attack, and it does not appear to be an attack from the outside. Some security experts recommend not to buy devices or software from countries in question.

3. Past and Current Cyberattacks

Analysis of past and current cyberattacks is among the best tools when it comes to understanding where to apply more security to defend your systems. Knowing what occurs in those successful exploits can shape updated guidelines on what to consider when defending your own systems. Often, security requirements present as clinical, general statements and theories that take effort to apply to systems. And while those statements and theories are critical to system security, they fall into the category of security hygiene and overall best practices. Ensuring that data is using NIST-approved encryption mechanisms is a best practice, since any deprecated encryption practices could be broken by attackers, if they were able to breach the data in the first place. Best practices come from the assumption that attackers are already present. And zero trust principles require continuously questioning trust, aligning with the thinking that attackers are already present in systems. That doesn't mean every system has an attacker in it, it just means that we must assume that attackers can break past the first layer of security and are able to gain access to systems.

As was introduced earlier in this paper, nation-state attackers take advantage of undiscovered zero-day software vulnerabilities. Many of us receive notifications daily which ask us to update software. Outdated software is where exploitable vulnerabilities lay, and without updating software often and consistently, systems are more vulnerable to common attacks. When systems choose to release software patches, they often release patch notes on what was changed. Attackers often use these patch notes to create exploits for systems running the vulnerable version and will use these exploits to target unpatched systems. Possibly the most notable ransomware, WannaCry, originated from outdated software. Attackers used the exploit on the SMB protocol to create WannaCry, a ransomware that successfully attacked over 300,000 devices [9]. Ransomware has affected everything from gas prices to meat production and it is expected that it will continue to grow [10].

A recent attack that occurred using zero-day vulnerabilities was to a popular email, calendar, and collaboration tool [2]. In this situation, attackers had discovered four zero-day vulnerabilities within this server. Patches for this vulnerability were released about a month and a half later. The four vulnerabilities, when used in tandem, can lead to remote code execution (RCE). Remote code execution is

how attackers can hijack servers, create backdoors, steal data, and further any malicious code deployment. This attack was traced back to an APT group from China. APT groups, or Advanced Persistent Threats, are threats that originate from nation-state attackers, and often consist of an attacker gaining entry to a system and lying dormant for an amount of time. Although an update and security patch have been released to remediate these vulnerabilities, that does not mean the threat is gone. Anyone using the affected software versions is vulnerable to future attempts of the exploit. Furthermore, current systems may have been compromised even with the patch, and techniques such as an undetected backdoor or a time bomb file that will execute on a certain date may still exist on the patched server. This attack demonstrates the capability of nation-state attackers.

Another malware attack that occurred in 2020 offers insight into how nation-state hackers operate. In this situation, attackers were able to use forged tokens to obtain privileged access in a system used for IT administration throughout many organizations. Attackers then used this elevated privilege to access whatever software they wanted to within the organization by using the forged token. Forged tokens can only work if they are not cryptographically secure tokens, if the keys to secure the tokens were stolen or the applications validating the tokens are susceptible to replay attacks and do not validate the signatures with unique data for each transaction. These types of compromises provide unauthorized access to data and software.

In the past two years, we've seen an increase in ransomware, effectively making some systems completely unusable. From 2019 to 2020, ransomware attacks rose 62% worldwide, and an incredible 158% in North America [10]. Ransomware infects a system and encrypts the system data against the user's will. Then, the user is presented with an option to pay a ransom, which will hypothetically decrypt the user's data. Ransomware is complicated to resolve. In the past, ransomware victims were warned not to pay the ransom, as that would encourage other attackers to go down the ransomware route. There is also never a guarantee that paying the ransom will unlock your system. In some ransomware, it was easier for the program to wipe the system instead of encrypting it, so that by the time the ransom was paid, the data was already gone. Many of the new ransomware infections require multiple payments, the first to unlock your encrypted data, the next to avoid selling data to others. There may be a third payment set up as a monthly fee, to avoid any re-locking of data. In most cases of ransomware, it is unlikely that the targeted system will be able to recover normally, or that the organization will recover financially.

Nation-state attackers, as well as individual hackers, are seeing the current landscape of how organizations respond to ransomware attacks. The popularity of ransomware is partly due to how many organizations are paying out the ransoms. Statista surveyed over more than 600 IT organizations found that most of them had been infected with ransomware in 2020, and 68% of those infected paid the ransom [6].

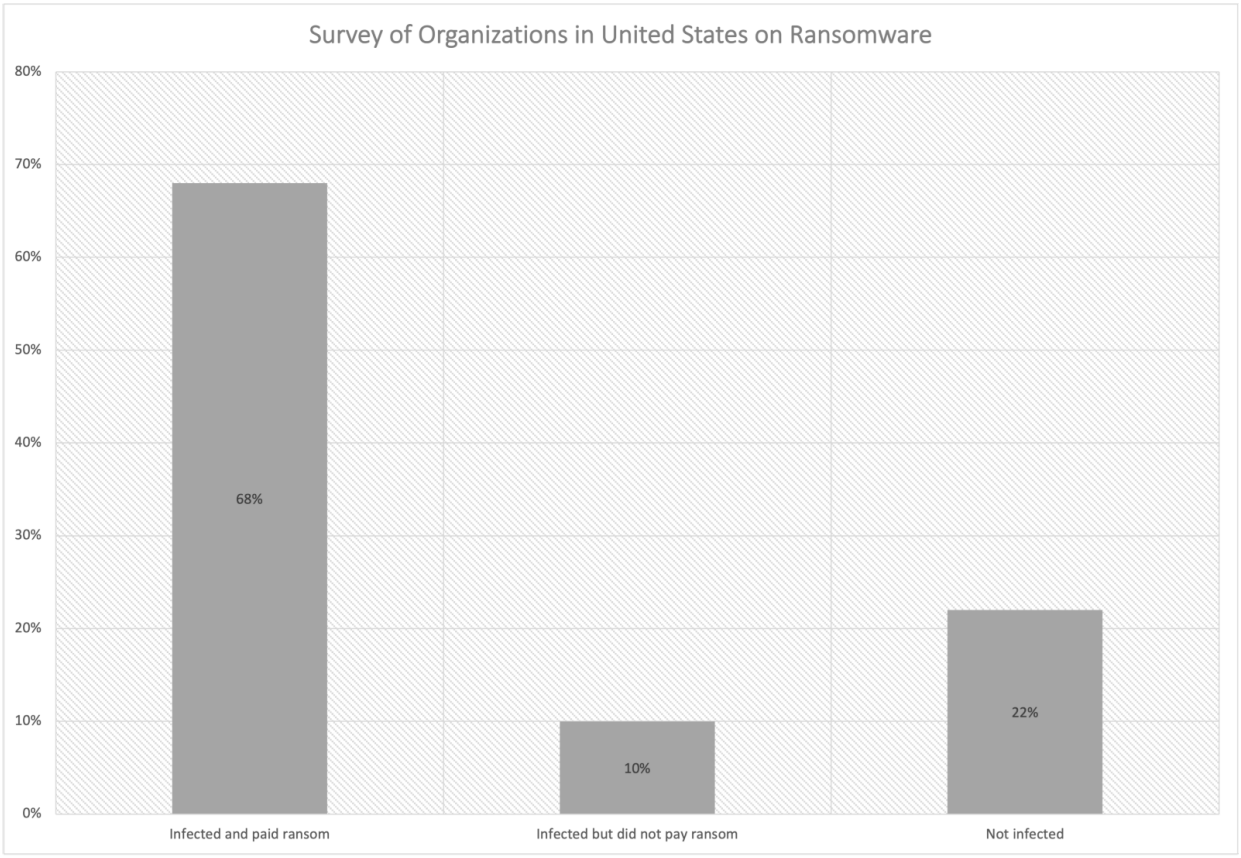


Figure 1 - Organizations on ransomware response in 2020

The effect of ransomware is two-fold. First, the ransom itself generates financial gain for the attackers. Second, the ransomware takes the system offline. Other attacks for taking systems offline include denial-of-service attacks, where a system is flooded with requests, too many to handle. Ransomware can take a system offline, as well as generate significant money for the attackers, which a typical denial-of-service attack cannot do. As more organizations continue to pay out the ransom, more ransomware attacks will happen. Instead of waiting for a ransomware attack to hit a system, it's better to plan.

The best defense against ransomware is to ensure that your system can be re-built, and re-instated, without touching the infected portion of the system. Often, this is done by building a disaster recovery (DR) environment, that replicates the production environment of the system. However, if a disaster recovery system is on the same network, or attached to the production environment in some way, it is useless. The disaster recovery environment should be geographically and logically separated from the production environment. The DR plan needs to include a back-up copy of any needed data, that is located off-site and stored in a separate way than other data. Many ransomware systems attempt to discover automated backup systems first, then ensure to lock both the original and the backup. Ensure your business-critical systems are equipped to deal with ransomware appropriately, to avoid unwanted payments and unwanted outages for your business.

4. IoT Device Security

Nation-state attacks can use any possible method to achieve its goals. One area that has been overlooked in the nation-state conversations, but could become an issue, is IoT devices. According to Juniper

Research, over 46 billion IoT devices will be attached to networks around the world by the end of 2021. That is a lot of potential attack surfaces.

IoT devices, even though small in processing power, can still propagate malware or overwhelm a network with traffic, especially if tens of millions start to chat all at once. Additionally, many IoT devices openly communicate with various back-office systems for software updates, configuration changes and service level monitoring. Many IoT device manufacturers have yet to appropriately implement security, because of resource constraints on the device, yet the device is able to connect to networks using various protocols. If an attacker could compromise the "right" (unprotected) IoT devices, they could be used to deny service, lock or steal data, and more. Some of the connectivity challenges are not only issues with specific protocols, but, just like software, which version of protocol is supported on the devices. Some more popular protocols include WiFi, Bluetooth, Bluetooth Low Energy (BLE), LoRa, RF, NF, Zigbee and others. Each protocol is created by a different organization at different times over the last 10+ years to help connect devices to a network. Even the protocols that start out with some security have a difficult time keeping up with new versions that are secure since IoT software stacks are difficult to update.

The Zigbee organization has changed its name to the Connectivity Standards Alliance (CSA) and now includes Matter (its newest secured protocol), Zigbee, Green Power, Smart Energy, JupiterMesh, RF4CE and Dotdot. We are happy to see the newer protocols evolve security protections. The Professional Service Association (PSA) organization certifies security principles for IoT devices for a unique identifier, security lifecycle, PKI certification, secure boot, secure updates, anti-rollback, secure storage and trusted cryptographic services. Earlier in 2021, they had certified 30 chips/SOCs, 14 software platforms and 10 OEM devices to date. Wi-Fi improved security with WPA3 which includes a QR code with a public key you can scan with your phone instead of using a password and has stronger cryptographic algorithms with forward secrecy and updated keys.

While we are making progress, it is not enough to rely on external organizations to provide all the necessary security for devices attaching to our networks. One thing we can do is participate more in setting standards and enforcing compliance from manufacturers. We can learn, from the many security standards on other products and services that have emerged from International Standards Organization (ISO), National Institute of Standards and Technology (NIST), Payment Card Industry (PCI), Health Insurance Portability and Accountability Ace (HIPAA), various privacy laws (GDPR, CCPA & CPRA, and many others), Financial Industry Regulatory Authority (FINRA) and others to help mold where we are today. We can certainly take many lessons learned from across the spectrum and apply as many as possible. Then where there are gaps, we need to step-up and lead innovative solutions.

To help understand how to best mitigate issues, study and understand past and present attacks. A good source of attack information is available from the U.S. CERT vulnerability advisory, which provides information on discovered vulnerabilities every day. A surprising number of vulnerabilities are published on a regular basis. As you learn about new security vulnerabilities, you can assess your environment, applications, devices, services and resources for issues. Some issues can be mitigated with workarounds or additional protections applied to help protect your company and your customers before formal patches can be released.

For IoT devices, let's apply the normal baseline security practices for larger devices. There are 5 major areas to ensure device security. Most of these are currently inadequate, which means they do not exist at all or are highly immature:

1. Standards to which IoT devices should comply
2. Designs which allow for improvements in features, including security updates over time -- which is more difficult to implement but necessary to resolve

3. Testing and compliance measures to ensure implementations meet laws and standards
4. Ongoing operational support for scanning and patching.

The real trick with an IoT device is to ensure it can be as secure as a server, but with very little resources. Many manufacturers' margins are thin because of tough competition and rising costs for resources and labor. Manufacturers have little incentive to increase spending on security unless other manufacturers are investing as well. Most people do not understand security well enough to know what look for, and have no influence on manufacturers of IoT devices. After all, if an IoT device breaks due to a security issue, the consumer must buy another appliance. That seems counter to the cycle we'd like to see created, where manufacturers are required to build security in to keep devices, people, and networks secure.

Part of a secure IoT device architecture could depend upon being behind a secure gateway in the home, however not all ISP connections and gateway products will be at the same level of security. Additionally, IoT devices may respond to a myriad of protocol announcements to join a network or be discovered and accessible and may never traverse on a path that is secure to the Internet. And then there are exploits which could take advantage of these devices that may not be able to discriminate against the signals from a secured gateway vs a rogue device.

Some basic tenants for device security include but are not limited to having secrets encrypted in storage, and obfuscated when in use. These secrets include a hardened identity that cannot be spoofed. One way to do this is to only put identity into hardware and secure it with an x.509 certificate, where the private key is also only available in hardware. Having a hardened identity prevents spoofing and evil twin attacks as well as serving as the basis for strong authentication against a trusted identity.

Devices also need to include secure boot up sequences and operational status to ensure integrity and authenticity of the hardware and software. Secure boot needs to include an unbroken chain of trust between the different layers of components and software, to ensure hackers cannot introduce new software between layers of different authentication keys that are not tied to a root of trust or authentication processes. Authentication is best when it is a cryptographic series of functions that cannot be interrupted when software is first being loaded for use during the boot process.

Devices must include a way to receive secure configuration updates as well as secure software updates. When devices do not include security on configuration interfaces, this leaves security entirely to the environment in which the device is run to ensure that only authorized configurations are allowed. The device must include protections against malware, which include automated software updates, without customer intervention.

Devices must ensure that unnecessary hardware and software ports are not available, which could be an open hole for attackers to exploit. Having open ports and unnecessary protocols or APIs is a common issue. Attackers will learn how to reach devices using these protocols and open ports. Additionally, ensuring security is turned on for the necessary open ports is critical. Using common or global secrets to protect ports is no longer an acceptable practice. Each port needs to use unit-unique public key cryptography to ensure access only comes from legitimate users in the ecosystem and is only destined for that specific device.

One of the most common issues is the inclusion of a default password in devices. This should be prohibited. Many people do not change default passwords. Attackers count on getting into devices this way. Additionally, passwords are easy to crack and should never be used without including multifactor authentication. People needing access can use mobile apps to manage multifactor authentication, or

a QR code with public key cryptography. Logins could use common sign-in systems such as Google, Microsoft or Apple. These are trusted authentication systems.

With the increased use of zero-trust principles which include micro segmentation, the blast radius for compromises can be reduced. Micro segmentation ensures that each resource should only be reachable by other authorized resources, and is accomplished by isolating network access. Authentication of resources is required, in addition to authorization as will be discussed. Another best security practice is to ensure that only things that need to reach the Internet, do. Many people accidentally configure public cloud accounts to be publicly accessible. Additionally, IPv6 addresses can be a source for common access mistakes, such as when people use a public address rather than a private address for internal applications.

One common question that comes up in security is how to assess if a device is healthy, meaning, not currently compromised. We not only need to know if a device is heathy as a part of normal monitoring, but also whether it is healthy upon first configuration for service – a clean start. For IoT devices today, we generally don't know. For IoT devices in the future, without overburdening the necessary hardware and software for an IoT device, it should be mandatory to have device security state, software versioning and file integrity. Creating a virtual device architecture for monitoring device health could improve our ability to contain attacks.

Using common security frameworks, such as the NIST Cybersecurity Framework, to map out all the components of IoT devices for its entire lifecycle is a good way to cover the necessary security protections. The NIST Cybersecurity Framework provides guidance on identification, protection, detection, response and recovery. For example, identification includes visibility into where and what systems or devices are in your ecosystem and how they are lifecycle-managed and governed. How will you able to respond to an attack if you don't know where or how to reach devices, systems or applications? In conjunction with traditional security frameworks, technological maturation processes need also to be assessed and anticipated. Consistency in security processes and operational monitoring need to be applied with good ways to measure against risk for your company.

Continuous scanning and immediate patching are both critical to keeping down the number of vulnerabilities on your devices and in your software. There are great scanning tools available for things like Windows servers, but when it comes to homegrown applications, IoT devices and industry-unique solutions, scanning leaves much to be desired. Each team responsible for building and operating devices and software can keep up their operational excellence metrics with weekly reporting on any known issues and remediation status. Operational excellence is a requirement to ensure there are no outages. Too often, software teams are focused exclusively on new release cycles at the cost of spending cycles to lower technical debt -- which includes security vulnerabilities.

Additional scanning techniques can look for static passwords, which should be immediately replaced with strong authentication mechanisms, managed by automation where possible. Also very important is scanning for certificate expiration dates and replacing certificates before they expire. For example, if a certificate is expired that is used for TLS authentication, TLS leaves a wide-open security hole by failing the connection open with no security. Be sure to configure TLS sessions to prohibit failing open, and monitor and replace certificates before they expire.

Monitoring for the previously mentioned device health, as well as activity such as communications paths are important to understand. If communication should only be happening with a back office system on your corporate network, but you can see communication messaging over APIs exiting your corporate network, that is an alert. It may represent a security issue related to data that may be leaking; access may

be compromised; malware may be propagating, etc. Monitoring should also include using data analytics and machine learning to access patterns for people as well as applications. Understanding what the normal operating model looks like can be key to blocking lateral movement, if access or device software is compromised.

Monitoring file integrity and approved configurations matters. If an attacker changes configurations and plants malware, you want to be able to detect and respond right away. Having a way to roll back or rebuild clean code quickly to remove malware is very important. Often the difference between a small incident and a large compromise is time. If good monitoring techniques are not in place, small issues can go undetected for a long time. Combinations of small issues may fall under the radar of detection and then can be executed together to perform a larger exfiltration, such as a "golden" pile of data being moved, or denying access to services or devices.

Concentrating on all points of network access is also important. That said, monitoring can only be as good as the data it is getting. If unauthorized access is happening in a vendor network or off-shore facilities, is there anything monitoring the network that will trigger unusual activity? This activity could appear to come from a legitimate vendor connection, but the fact that the traffic came at an unusual hour could be a tipoff.

Examining source IP addresses (if you can get them) can tell you if there are connections coming from geolocations that are unacceptable. Collecting context-rich data, containing information such as source IP address, can enable a system to separate legitimate traffic from unacceptable traffic. More systems and services are using location, connecting to compromised data sets for comparison and getting information from government agencies to help identify issues, such as known malicious source IP addresses. Also, security level protections for connections with outside businesses or facilities should be verified to ensure malware protections are in place before accepting traffic from that new connection. DDOS and malware-sourced verification is common.

Using good Secure Development Lifecycle (SDL) management practices is also mandatory. Ensuring every device make and model has gone through a threat model to identify attack surfaces and identify how to best protect them is important to do as a part of the design phase. When threat modeling, review not only the architecture, features and functionality of the product, but also the code build pipeline. Additionally, ensure all APIs and protocols are being designed for secure use. Data encryption and privacy also need to be reviewed as a part of a good threat model. Then ensure pen testing is conducted, in an environment like where the device will actually operate, to provide the right level of operation security configurations. There are many other aspects to the SDL, including scanning for bugs in code, and the previously mentioned monitoring aspects.

5. Authentication and Authorization

So how can we secure access? By using strong authentication, access controls and authorization for users, software, and devices. Authentication and authorization are often confused with one another, but in the scope of system security, it is critical for the two to be implemented and assessed separately but designed together to ensure no security gaps. Authentication is the process of confirming a subject's identity, and authorization is only allowing that identity access to information and systems they are allowed to access. Authentication and authorization are most associated with human users, but modern practices include authentication and authorization of machine access as well.

Secure your front and back doors as well as your windows into your systems, applications, resources, devices and services. Secure your workforce users, workforce admin accounts, workforce

service accounts, guest and vendor accounts, and all customer accounts. If nation-state attackers get their hands on any kind of credentials, they will use them. Work within the zero-trust model, to create checkpoints throughout the system to ensure users, and machines, are not only who they say they are, but that they also have the correct access. Following this model can prevent attackers from taking advantage of authentication and authorization in a system.

Evaluating the system's environment is the first step in considering authentication and authorization. A common solution to insecure connections is to require the use of a VPN, or a Virtual Private Network. The use of a VPN verifies that the user is on a private, usually encrypted, connection. However, before the connection to a VPN can be established, a user must authenticate to the VPN and prove their identity. Access to the VPN may be needed to access sensitive or confidential data, so it is important to use a strong method of authenticating users to the VPN. VPNs generally don't take care of any authorizations beyond hooking into authentication systems such as SSO to ensure they have active user credentials that are authenticated.

Ten years ago, a simple username and password combination would be enough to access sensitive data and critical systems. Today, multi-factor authentication is the industry standard. In fact, multi-factor authentication has a 99.9% success rate in protecting against compromised credentials [5]. The purpose of multi-factor authentication (MFA) is to first, enter your username and password, but as a second step, to prove user presence through a second device.

However, MFA is not invincible, and as it becomes widely adopted, we will see more and more attacks on systems protected by multi-factor. Older MFA solutions such as SMS-based MFA have shown to be easily hacked, which is why SMS-based MFA is becoming less popular and is being replaced by multi-factor that depends on an authenticator application, such as Duo Mobile or Microsoft Authenticator. These authenticator applications provide another level of security that SMS-based MFA does not. The difference between SMS-based MFA and application-based MFA is that it is easier for an attacker to compromise a user's text messages, but a significantly more difficult task at hand if they need to get to an application in the user's device and compromise secure APIs. However, there are even potential vulnerabilities with these authenticator apps, such as user error. If a user gets too comfortable accepting authenticator requests, they have the chance of accepting a false authentication request and approving an attacker's request. Therefore, many authenticator apps have begun implementing multi-factor requests that force the user to engage with the request, and in process, confirming that they are the one who made the actual request.

Passwordless authentication is making its way through the technology world [5]. At the root of authentication, there are three ways a user can prove their identity: something they know, something they have, or something they are. Something they know would be a password, or a security question. Basing user authentication on something they know is no longer considered secure. Multi-factor authentication, as well as SMS-based two factor authentication, depends on something the user has. This is usually a phone but can also be a hardware token. With multi-factor authentication, the first step of authenticating is to enter a username and password combination, and then approve a request on another device. The password is still in use during this authentication mechanism. With something that the user is, however, passwordless authentication is making a rise. This is authentication that is based on user presence, as well as user validation in the form of biometrics.

Passwordless authentication does not only consist of user biometrics. In fact, there are a handful of ways that users can authenticate into a system without using a password. However, biometrics have become one of the most popular ways a user can authenticate, without a password. The different technologies

behind passwordless authentication include, but are not limited to: FIDO2, Windows Hello, and Microsoft Authenticator passwordless authentication.

These forms of multi-factor authentication, app-based or passwordless, are acceptable options for gaining access to a network. However, once a user is authenticated through single sign-on using a specific device, that authentication can carry over to more resources, including other applications, without requesting to sign back in. This is accomplished by using technologies like OpenIDConnect (OIDC), SAML, and OAuth. These technologies rely on cryptographic tokens and keys to be passed from machine to machine. These technologies can be used to identify and authenticate human users as well.
Creating a secured identity for machine access may not be an obvious solution, but attackers have been able to use machine identities to gain entry to a system. How do we prevent attackers from spoofing machine identities? One solution is PKI, with a signature over unique data in the transaction which shows current proof of possession of the private key associated with the public key inside the certificate. The certificate contains the device identity, and the certificate is also signed by the trusted certificate authority.

Authorization ultimately arises from the principles of least privilege, which essentially states that users should have access only to the resources they need, and no more. Authorization is often overlooked, as it requires specifying explicit access to specific resources and tasks as named in an attribute or role-based access control policy. However, it is critical for access to be defined per user, such as a general user role with least permissions and then only giving specific users elevated privileges. Abusing excess privileges is an integral part of how attackers navigate through a network, and access control is how one can prevent an attacker from leveraging them.

While an attacker can certainly leverage excess privileges, an employee can as well. Insider threats are a real concern and make up a significant portion of privileged access attacks. This is especially concerning if an individual has unfiltered access to systems, and their credentials are not revoked after termination. Real-time ability to revoke credentials is crucial in preventing abuse and compromises. If an individual leaves the organization, or does not need access to a resource anymore, the immediate revocation of their credentials or access is critical. This includes not only employees, but also contractors and business partners. Vendor access is another area of user access that often is overlooked. Vendor access must follow the same policies as regular users within the environment, potentially with less access to base resources.

Following a zero-trust model, as well as authenticating and authorizing users for resources by using modern technologies will assist in preventing stolen, forged, or abused credentials from causing a catastrophic incident. Nation-state attackers are resourceful and will take what they can. Minimizing their blast radius from compromised credentials is essential.

6. Conclusion

There is no guaranteed way to prevent attackers from exploiting your systems. However, best practices can be applied, as can a bit of innovativeness about how to thwart unauthorized access into systems, and prevent unwanted events. As described in this paper, nation-state attackers will try any method possible to circumvent system security. Defending against potential attackers requires time, energy, and money. However, it is no longer possible to ignore security when building systems and networks. Using the methods outlined in this paper can help prepare a system to defend against attackers.

Attackers, especially those representing nation-states, will continue to evolve their attack methods as technology evolves. Underestimating these attackers will result in a security posture unprepared to deal

with evolving attacks. As we know, there is no one magic fix for defending against hackers. They are resourceful, unpredictable, and relentless. Instead of a one-time fix, system security involves an ongoing process of evaluating, building, testing, re-evaluating, re-building, re-testing, and repeating. The purpose of this continuous evaluation is to keep up with the attackers, as they are continually and constantly attempting to gain access to cable provider systems. Following the practices in this paper will help create a security posture for cable providers in the modern age.

Abbreviations

API	application programming interface
APTs	advanced persistent threats
BLE	bluetooth low energy
CISA	Cybersecurity and Infrastructure Security Agency
CCP	Chinese Communist Part
CCPA	California Consumer Privacy Act
CPRA	California Privacy Rights Act
CSA	Connectivity Standards Alliance
DR	disaster recovery
FBI	Federal Bureau of Investigations
FINRA	Financial Industry Regulatory Authority
GDPR	Global Data Protection Regulation
HIPAA	Health Insurance Portability and Accountability
IoT	internet of things
ISO	International Organization for Standardization
MFA	multi-factor authentication
NF	near-field
NIST	National Institute of Standards and Technology
NSA	National Security Agency
OIDC	OpenIDConnect
PCI DSS	Payment Card Industry Data Security Standard
PKI	public key infrastructure
RCE	remote code execution
RF	radio frequency
SDL	secure development lifecycle
SMB	server message block
TLS	transport layer security
VPN	virtual private network

Bibliography & References

[1] "*Cybersecurity Framework*", NIST, https://www.nist.gov/cyberframework

[2] *"Microsoft: Multiple Exchange Server Zero-Days Under Attack by Chinese Hacking Group"*, Ryan Naraine, SecurityWeek, https://www.securityweek.com/microsoft-4-exchange-server-zero-days-under-attack-chinese-apt-group, March 2 2021

[2] "*Nation-state cyberattacks see huge rise in 2020*", Sead Fadilpašić, TechRadar, https://www.techradar.com/news/nation-state-cyberattacks-see-huge-rise-in-2020, April 8, 2021

[3] "*Nation States, Cyberconflict and the Web of Profit*", HP Wolf Security, HP, https://threatresearch.ext.hp.com/web-of-profit-nation-state-report/, April 8, 2021

[4] "*New Law Will Help Chinese Government Stockpile Zero-Days*", Kevin Townsend, SecurityWeek, https://www.securityweek.com/new-law-will-help-chinese-government-stockpile-zero-days, July 14, 2021

[5] "*One simple action you can take to prevent 99.9 percent of attacks on your accounts*", Melanie Maynes, Microsoft, https://www.microsoft.com/security/blog/2019/08/20/one-simple-action-you-can-take-to-prevent-99-9-percent-of-account-attacks/, August 20, 2019

[6] "*Share of organizations in the United States that experienced a ransomware attack and paid the ransom in 2020*", Joseph Johnson, Statista, https://www.statista.com/statistics/701282/ransomware-experience-of-companies/, March 5, 2021

[7] "*Technical Deep Dive Into SolarWinds Breach*", Parmanand Mishra, Qualys, https://blog.qualys.com/vulnerabilities-threat-research/2021/01/04/technical-deep-dive-into-solarwinds-breach, January 4, 2021

[8] *"WannaCry Ransomware Campaign: Threat Details and Risk Management"*, John Miller and David Mainor, FireEye, https://www.fireeye.com/blog/products-and-services/2017/05/wannacry-ransomware-campaign.html, May 15, 2017

[9] "*Why ransomware attacks are on the rise – and what can be done to stop them*", Lynsey Jeffery and Vignesh Ramachandran, PBS, https://www.pbs.org/newshour/nation/why-ransomware-attacks-are-on-the-rise-and-what-can-be-done-to-stop-them, July 8, 2021

Seeing Double:
Network Digital Twin

A Technical Paper prepared for SCTE by

Guy Meador III
Senior Solutions Architect
Cox Communications, Inc.
6305 Peachtree Dunwoody Road
Atlanta, GA 30328
404-269-5625
Guy.Meador@cox.com

George Cave
Principal Data Center IP Architect
Cox Communications, Inc.
George.Cave@cox.com

1. Introduction

Operators are facing disruptive change to their networks driven by technological and business forces. The business seeks to increase innovation, reduce costs, and manage risk while network technologies including virtualization, software defined networking, streaming telemetry, and automation drive additional complexity into the network. Typically, an operator's functional disciplines such as engineering, operations, analytics, and planning each have separate, uncoordinated views of the network with sometimes overlapping, sometimes disjoint, data, leading to inefficiencies and uncertainty. Network Digital Twin provides a comprehensive view of the network, combining disparate data sources such as telemetry, monitoring, inventory, provisioning, analytics, planning, and network automation with variable-fidelity models of equipment, network functions, network services, network operations, and analytics, creating a separate digital representation of the live network. The network digital twin is an integrated, consistent, comprehensive, lively, and accurate model of the network and its constituents, connected to and representing the state and behaviors of the actual entities of the operating network. Operators use the network digital twin through the accompanying dynamic views and visualizations to better understand the state and behaviors of the network and, through manipulation of the digital twin, affect changes in the actual network. Functional disciplines within the operator's organization use the network digital twin to break down data silos and leverage the consistent model of the network as the basis for their individual missions. The paper addresses the formation, concepts, and implications of digital twins and their application to the operator's network, including an example from Cox's application of aspects of digital twin to our data center network.

2. System Architecture

The three top-level functional building blocks of the network digital twin functional ecosystem are depicted in Figure 1. The functional building blocks have significant internal composition (other functional blocks) and form an enduring basis for multiple logical and physical architectures. This architecture description, therefore, provides discussion, primarily, of key aspects of a functional architecture for network digital twin.

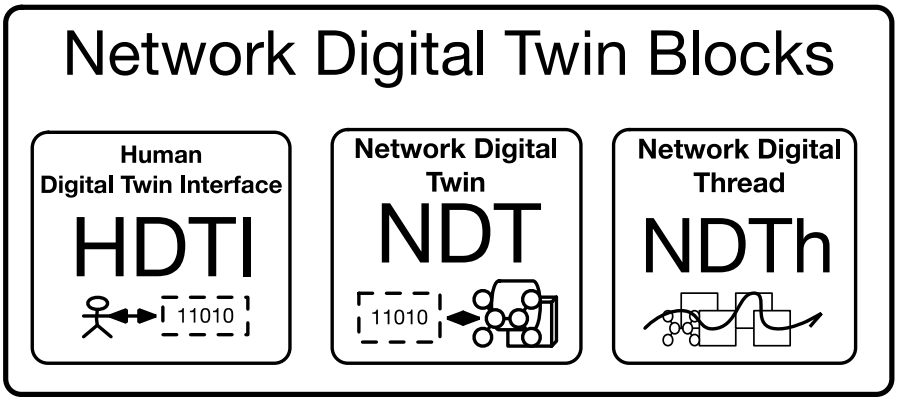


Figure 1 – Top Level Functional Blocks of the NDTE

The top-level functional building blocks of the Network Digital Twin Ecosystem (NDTE) are the Network Digital Thread (NDTh), the Network Digital Twin (NDT) itself, and the Human Digital Twin Interface (HDTI). These three top level building blocks work in concert to provide a transformative approach to the creation and operation of the network.

Network Digital Thread: The network digital thread contains engineering models for the network and its constituent subsystems and interfaces. Also included are any document-based information, specifications, drawings, images, and non-operational data relevant for the network that are created across the system development life cycle (SDLC), such as design and planning information.

Network Digital Twin: A network digital twin is a unified executing engineering model of an operating network entity coupled with normalized data pertaining to the actual operating network entity with which it is associated. The network digital twin can be queried and manipulated separately from the actual operating network entity, with the option for those manipulations to affect changes at the operating network entity. Network digital twins can be created to represent network entities at the logical, virtual, or service layer, as well as composites of any of these (including an entire network).

Human Digital Twin Interface: The human digital twin interface is a human-machine interface (HMI) providing the capability for a person to interact with one or more digital twin instances. It provides dynamic visualization of the digital twin instance and related data and supports human-driven query and intent-based manipulation.

2.1. Theory of Operation

The prototypical top-level functional connections and information flow between the network digital twin functional building blocks and the operating network entity are shown in Figure 2, adding an additional functional block to the mix: the Actual Twin (AT).

Actual Twin: The actual twin is the physical, virtual, or logical instance that is a counterpart to a digital twin instance. Examples of actual twins for the network include a network router appliance, a router network function in a virtual machine or container, and a layer 3 virtual private network (VPN) service.

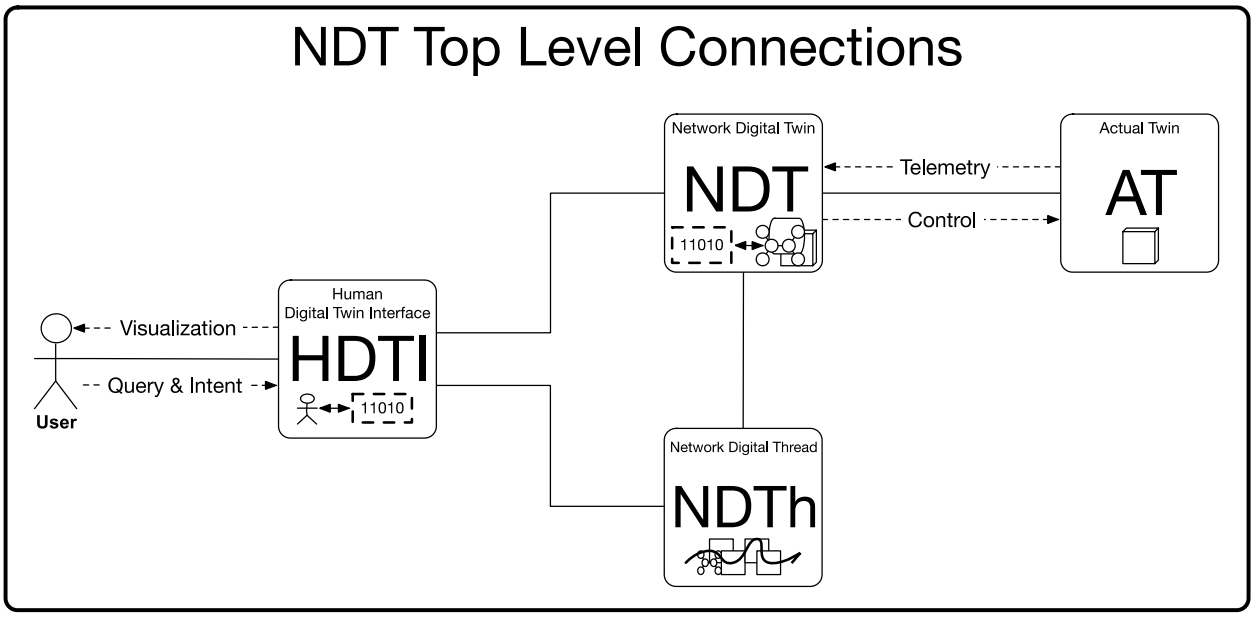


Figure 2 Top Level Functional Blocks and Connections

An overview of the theory of operation for a prototypical network operations situation using network digital twin is as follows.

An executable NDT instance is created with an integrated set of models pertaining to a specific type of network entity, e.g., the type of AT, in a given role in the operating network. The models are drawn from the Network Digital Thread, which is the authoritative source for this information. The NDT is also pre-initialized with data pertaining to operational intent for the instance.

The NDT instance is associated with an actual entity instance in the operating network, e.g. the AT. Telemetry data pertaining to the AT flows from the AT and other systems into the NDT. The telemetry data could be from streaming telemetry, monitoring systems, element managers, and other such intermediaries. The telemetry data is used to update the state of the NDT, causing the dynamic executing model to be evaluated, updated, and to react to the changing conditions of the AT. For example, the NDT may detect model constraint violations, a variance between operational intent and current state, or other such issues.

If the NDT models are constructed to provide control capabilities for the AT, then control signals are sent as needed toward the AT. The control signals could take many forms, depending on the specific implementation. The control signals could be of types such as configuration changes, commands, signals to intervening systems, and events. Incorporating this kind of model into the NDT provides intent-based closed loop control for the network.

The HDTI provides the user/operator experience for the NDT and the Network Digital Thread. It allows the user to see various static and dynamic views of various kinds pertaining to the network digital twin that are updated during operations. It also enables queries to be performed and the ability to express operational intent and other control input from the user perspective.

Ideally, any technical information from the Network Digital Thread is linked with the NDT and can be discovered and viewed using the HDTI environment.

2.2. Network Digital Thread

Network Digital Thread encompasses the engineering and analytical models for the network created across the entire system development life cycle (SDLC). The network digital thread building block is the digital thread created through digital engineering, with the subject being the operator's network.

The INCOSE Systems Engineering Body of Knowledge (SEBoK) states, in part, that digital engineering "…is the creation of computer readable models to represent all aspects of the system and to support all the activities for the design, development, manufacture, and operation of the system throughout its lifecycle. These computer models would have to be based on shared data schemata so that in effect a digital thread integrates all the diverse stakeholders involved…"; and further, "Everything from documenting requirements, technical reviews, architecture design, and so forth would be based on the models in a digital engineering environment (Vaneman and Carlson, 2019). The digital thread would be the authoritative source of truth concerning the system data." - [SEBoK/Digital Engineering].

The network digital thread contains engineering models for the network and its constituent subsystems, properties, interfaces, protocols, behaviors, and data. These models range across multiple disciplines, including network engineering, facilities engineering, reliability engineering, systems security engineering, software engineering, network planning, and operations engineering, to name but a few. The various discipline-specific models are all aligned and connected through the common context and consistency established and maintained in systems engineering models.

Within the engineering disciplines, a major class of model in the network digital thread are those models addressing forward engineering concerns. These models capture the "as-specified", and "as-designed" aspects for the network, addressing, for example, structural, behavioral, functional, non-functional, and operational concerns. In addition to the models that capture specifications and design decisions, this class of model also includes the accompanying engineering analysis models and engineering/test data sets.

Models of the forward engineering class are necessary to establish a correct model of the network as it is intended to be. They establish the interfaces, protocols, behaviors, data formats, and other characteristics of the network and its constituent subsystems, separate and independent of any hardware and software implementation. They are integral to the simulation, emulation, operational characterization, and operational control capabilities of NDT, forming the rational basis with which the live operational data is normalized and paired with the NDT.

One example of a forward engineering model relevant to NDT is a network engineering model for network layer 2-4 network functions, connections, configurations, and behaviors. A formal model is prepared for the functionality, interfaces, and behaviors of a give type of network function (ex. router forwarding), with the model providing state variable and state transition definitions, configuration parameter definitions, and the effect of configuration on those

behaviors. Other instance-independent characteristics of the network function can be included in the model. Given a sufficiently elaborated model, the designed layer 2-4 behavior of the network function can be determined given any set of internal state and external stimulus, as can an arbitrary network of many interconnected network function instances.

Another example of a forward engineering model relevant to NDT is a systems engineering model for a state machine-based autonomous operations closed loop controller. This type of controller is used at the network resource level (ex. device) or multi-resource network service level (ex. layer 3 VPN) to actively establish and maintain the operational state of the controlled entity according to operational intent. A formal model is prepared to specify the intent & configuration parameter definitions, states, state transitions, triggers, and actions of the controller, along with the input and output signals (ex. telemetry/events, control, configuration). Given a sufficiently elaborated model, the designed behavior of a closed loop controller can be determined given any set of internal state and external stimulus, as can an arbitrary set of many interconnected closed loop controller instances in communication with one another. Taken together, this type of model provides a key part of the executable specification for closed-loop controllers for network autonomous operations.

A third example of a model type relevant to NDT is a reliability engineering model for a manufacturer's network equipment model (in the last usage "model" means the manufacturer's model identifier rather than a formal engineering model of the equipment). A type of reliability engineering model is the Failure Mode and Effects Analysis (FMEA) model. The FMEA model identifies the inherent reliability characteristics of the equipment design, including the tree of relationships between possible failure modes and the resulting system performance degradation or failure. Given an FMEA model and a set of one or more fault conditions, the effect of the components failure modes to the system performance or safety can be determined and characterized. With relevant equipment test results and measurements from the population of a given equipment model in operation in the field, a predictive model for fault and failure probabilities can be created to live alongside the FMEA model. With a sufficiently elaborated predictive model of this type and the relevant time series state measurements for an instance of equipment, the probability of failure of the equipment can be predicted, indicating a need for proactive maintenance ahead of the failure. This second type of model may be developed using machine learning techniques and is closely associated with the area of predictive network maintenance.

Network digital thread entails more than models. For example, any document-based information or drawings and images relevant for the network created across the SDLC are part of the network digital thread. These non-model-based artifacts can be linked with the models, but do not play an active part in the executable network digital twin.

The preceding examples are but a few of the many types of models that are associated with network digital thread. Other examples include models for system security, network planning, network design (physical and virtual), and mechanical/physical design, among others. Some of these models are design models specifying correct characteristics of the described atomic or composite entities; others are analytical models of the resulting emergent characteristics of those

designs (ex. fault tree) or are predictive in nature based on the combination of the prior models with aggregated data from tests and measurements from the field.

All model types described above can be used for the purposes of engineering analysis. Some can be used for engineering discipline-specific model execution and/or simulation. Multi-discipline model execution and simulation is possible when coordinated through appropriate systems engineering models and co-simulation environments. As networks and network operations complexity rises, these techniques become more important to employ for the design and analysis of networks, their data, and their operation, even without the full capabilities of network digital twin.. They are used to support exploration of the network and operations design space and trade studies, for example, without connection to an AT.

2.3. Network Digital Twin

The models of the network digital thread designed for execution and simulation in the context of engineering design and analysis, along with the model execution platforms, can be developed beyond the scope identified in the previous section. Doing so enables a spectrum of model execution capabilities directly associated with an actual operating network and its constituent parts. A key characteristic of the approach is that the disparate models are linked together, coordinated, and updated with identical, or at least consistent, normalized, data during their execution.

For example, executable models from the network digital thread (at the needed level of fidelity and from the necessary engineering disciplines) pertaining to an instance of network equipment are brought together, connected, and co-executed in a coordinated fashion on a platform environment designed for that purpose. These executing models are then provided with data from and about the network equipment instance, forming a kind of digital copy of the equipment instance and its state. The digital copy consisting of these models bound with the network equipment instance's data forms the basis of an instance of a network digital twin for the actual network equipment instance.

The dynamic behavior, state changes, and other characteristics of the actual network equipment instance can be simulated, emulated, analyzed, represented, and predicted based on the network digital twin instance. The network digital twin execution platform provides additional data, from or about the actual network equipment instance, to the network digital twin instance as it becomes available.

The data pertaining to the actual network equipment falls into categories that include operational intent, target configuration and operational state, actual operational configuration and state, sensor readings (ex. temperature), network function-specific data (ex. firewall state, traffic counters, routing tables, etc.), and alarm conditions. Additional data about the actual network equipment instance can also include historical information, time-series data, failure data, scheduled maintenance actions, and other instance-related data.

The above data may be obtained for the network digital twin from a variety of sources, including telemetry and monitoring systems, element management systems, operations analytics systems, service provider data systems, or directly from the network equipment instance. The rate and

latency at which changing data from and about the actual network equipment is provided to the network digital twin will affect the extent to which the network digital twin instance reflects its twin's state.

Binding multiple, coordinated, executable models with a consistent set of instance data from an actual network equipment instance into the network digital twin instance enables the ability to achieve useful analytical and operational objectives.

The functional/behavioral models in the network digital twin instance combine with the actual twin data to enable analysis and simulation. For example, given a specification of network traffic arriving at one of the interfaces, the resulting forwarding plane and packet transformation functions and behaviors can be determined and/or simulated without further interaction with the actual network equipment instance. The case study later in this paper employs this type of network digital twin analytical model.

Anomalous or incorrect behavior or conditions in network equipment in the operational environment can be identified through the network digital twin's explicit and automatic checking of invariant conditions, constraints, policies, and state encoded in its models. Examples include the presence of disallowed configurations, variance between intent and operational configuration or operational configuration and relevant established policies, detectable functional behavior variances, and fault conditions.

With a network digital twin approach, the often-brittle workflow-oriented techniques using "golden config" templates and scripts for pre-checks and post-checks are eliminated in favor of an approach whereby the multi-discipline specification of correctness in context is expressed in the network digital twin models with continuous evaluation during model execution, and whenever new data is available from the operating environment.

Further, by employing state machine-based closed loop controller models as a part of network digital twins, the entire operational lifecycle of each actual twin in the operating network can be addressed, achieving intent-based network management with autonomous operations.

2.4. Human Digital Twin Interface

The human digital twin interface is a human-machine interface (HMI) providing the capability for a person to interact with one or more digital twin instances. It provides dynamic visualization of the digital twin instance and related data, and supports human-driven query and intent-based manipulation.

The HTDI is separate from the other functional blocks of the NTDE for several reasons. First, the pace of innovation between the NDT execution environments and HMI variety and innovation pace will be different; de-coupling these two facilitates the ability to take advantage of the pace of each independently. Second, the type of HDTI used for the same NDT is dependent on specific goals for the interaction, roles of the user, and so forth. Multiple HDTI experiences may be in simultaneous operation connected to the same NDT instance.

For example, an HDTI based on web technologies with user access via a web browser matches the typical network operations scenario of today. Static and dynamic views, queries, and controls appropriate to the user role and web browser environment are provided.

An application running on a mobile device in the field can also be created as an HDTI implementation, providing field technicians a new and improved way to view and interact with the autonomously operated network.

Another example of an HDTI is that provided by the class of tools, either web-based or native applications, classified as engineering design, simulation, and analysis tools. These tools are not used for operations but benefit from connection to simulation-only NDT instances, or connection to hybrid test environments containing a mix of simulation-only NDT instances and NDT instances associated to AT instances.

A final example with exciting promise for both operations environments and engineering exploration is that of virtual reality HDTI implementations. These will have a transformative effect on network operations centers enabling users to see and interact with the operating network in significantly new and effective ways.

In the operations environment, it may seem that HDTI is just another typical support portal, but that is not the case. The HDTI approach does not present an order/workflow-based view of the operating network, and user-initiated changes for the NDT are not user-initiated by manipulating orders. Rather, HDTI presents an active, intent-based way for the user to manipulate the NDT in a manner that is more natural and direct. With closed loop autonomous operations at work in the NDT layer, the user wields the HDTI to see the live state of the network and directly express intent for changes to the NDT, seeing those changes progressing and taking effect in the same view of the NDT. The overall approach embraces autonomous operations while hiding the complexities of its implementation, and, at the same time, provides improved visibility and control to the user.

2.5. Observations and Implications for Network Digital Twin

Applying network digital twin in the service provider network operations environment can take several forms ranging from passive uses for visibility, analytics, and verification/constraint checking on one end of the spectrum, to a full-up intent-based network with closed loop autonomous operations on the other.

The types and fidelity of models used in the NDT should be tailored to the specific purposes for which NDT is being used. An ideal network digital twin ecosystem should facilitate the flexible mixing, matching, coordination, and harmonization of the various models and uses.

Employing network digital twin effectively for many useful purposes does not require direct use of AI/ML at all. This is particularly true when using network digital twin for autonomous operations. Intent based network operation built upon executable models for state machine-based closed loop control, policy and constraint checking, service-to-resource configuration mapping, and other autonomous operations capabilities, along with integrated, consistent, and improved visibility, are achievable and valuable without direct reliance on AI/ML. It is preferable for the

automated operation of the network to be explicitly designed, implemented, and verified for desired characteristics through forward engineering disciplines. This approach ensures autonomous network operations remain consistent, explainable, and predictable.

On the other hand, AI/ML plays a vital role for the network digital twin ecosystem. It has already been discussed how machine learning plays an important role in engineering analysis and the formation of predictive models, playing a key role in the design of the executable models of the NDT. In connection with autonomous operations, machine learning has a key role to play in operations analytics (OA) within the network digital twin ecosystem, providing informational/advisory signals back into the closed-loop controller portion of the NDT. For example, the employment of OA functions based on machine learning and connected to the NDT used for evolving intrusion detection, identifying unusual network operational patterns, predicting equipment failure, and providing other operational recommendations. These OA advisory signals flow into the autonomous operations closed-loop controller layer of NDT; the controller layer is designed and configured to have a predictable and tested response, including the option of ensuring any appropriate human-in-the-loop decision making.

Of particular importance for success, and a challenge to be overcome, is the standardization of information across the NDT models and the normalization of network data (ex. from the AT) presented to those models. One of the benefits of moving to a network digital twin approach to network autonomous operations is that it will be a nexus forcing function to that normalization. It is a systems engineering challenge that begins early in the SDLC and affects multiple operator systems and organizational silos.

The network digital twin approach will transform the way provider networks are designed and operated, but it will not arrive overnight and all at once. Organizational inertia of existing tools, techniques, skills, and processes across the SDLC will need to be overcome. Here, too, it will take years to fully realize the complete architecture described in this first section.

However, network digital twin implementation has begun. The next section provides a case study of how one group at Cox has embarked on the journey employing a network digital twin operations analytics model for the data center network.

3. Case Study: Network Digital Twin in Cox's Data Center Network

With the goal of promoting the long-term health of Data Center Network operations within Cox Communications, Inc. (CCI), the Data Center Network Engineering (DCNE) team has embraced Intent-Based Networking (IBN) as outlined in "Intent-Based Networking – Concepts and Definitions (draft-irtf-nmrg-ibn-concepts-definitions-02), A. Clemm, *et. al.*" and in "Intent-Based Networking – Concepts and Overview (draft-clemm-nmrg-dist-intent-03), A. Clemm, *et. al.*", as a guiding framework in the implementation of a holistic automation strategy within CCI data centers. These documents, taken together, do not constitute a detailed blueprint for IBN, rather they define IBN and the structural components necessary to implement an autonomic network at a conceptual level. While investigating the promise and feasibility of IBN, it became

readily apparent that implementing automated change would require a high degree of visibility into the configuration and operation of network devices.

3.1. Background and Discussion of IBN Strategy

In the summer of 2020, the CCI Data Center network team began work developing an Intent-Based Networking strategy that could be implemented in successive stages with each stage providing value on its own and subsequent stages building upon the value of previous stages. Six stages were identified:

1. Visibility
2. Validation
3. Change Automation
4. Event Management
5. Auto-Remediation
6. Autonomic Network Operation

Visibility – Raw data about the network needs to be accessible at a *human* level. Considering the massive amount of network data that can be collected on a tier-3 Internet service provider network, it is important to distill that data into a digestible source usable by both humans and machines. This data must include both device and topology data as well as statistical traffic data. The refinement and analysis of this data into a usable, published dataset is the goal of this stage.

Validation – Testing using the data collected during the visibility stage provides a basis to ensure that the network is operating according to designed intent. In this stage, suites of tests are built to validate configurations, topologies, traffic flows, etc. These tests can then be used to ensure that changes to the network do not violate the original design intent of the engineer.

Change Automation – The risk in performing automated changes to the network is less stressful with a high degree of visibility into the operation of the network and the ability to validate that those changes had the desired result and did not break any existing network functions. This stage is often built in parallel to the previous two with the goal of automating simple, repetitive tasks. As Visibility and Validation mature, Change Automation can be trusted with riskier changes. Part of this stage is the identification of component configurations (config snippets) used to implement specific functions, i.e., implementing an access control list (ACL) to control access to the routing engine.

Event Management – Leveraging Visibility and Validation, AI and Machine Learning can be employed to detect anomalies in expected behavior. Analysis of these anomalies leads to the identification of events, which can be investigated, and appropriate action taken through human interaction.

Auto-Remediation – As the Change Automation and Event Management stages mature, trust in allowing the network to heal itself grows. Once a specific course of action has proven to resolve

an event that commonly occurs, Change Automation can be used to automatically fix or handle those events.

Autonomic Network Operation – This final stage requires significant investment in the previous five stages to the point where AI and Machine Learning algorithms can review high-level intents provided by engineers, validate them against the operating network, review the component configuration libraries of the Change Management stage to make determinations regarding how the network should be configured to realize the intended design of the network and finally schedule and implement the changes identified.

With this strategy in mind, the Data Center Network Design team began developing use cases focused on developing the Visibility and Validation stages.

3.2. Definition of Use Cases

Like most traditional network teams, the Data Center Network team at CCI is staffed primarily with engineers trained mainly on networking and security disciplines with a smaller number of individuals that have cross-training in programming and automation. Staff are comfortable working in a device command line interface (CLI), but less so writing code or leveraging tools like Ansible to make inquiries or changes to devices. To encourage the adoption of a 'NetDevOps mindset', the DC Network team began investigating various tools to facilitate movement *away from the CLI*.

A set of use cases was identified, and various network tools were researched for their feasibility in meeting those uses cases, as well as the part they could play within the broader context of the team's IBN strategy.

3.2.1. Functional validation of the network intent.

As discussed regarding Visibility and Validation, a fundamental component of our IBN Strategy is the ability to collect network data and test it based on an expected outcome. While many might consider it sufficient to validate that device configurations were matching some pre-defined standard template, it is also important to validate those devices were operating as expected and were implementing the intent of the design. The latter validation capability required a considerably more sophisticated testing framework to enable validating the output of various commands.

Examples of intents identified for the proof-of-concept testing were:

1. Route Propagation
2. Expected Path Analysis
3. FWs blocking/passing traffic as expected
4. VLAN Propagation within the Data Center
5. EVPN Functionality
6. Physical link connectivity and bi-directional traffic validation

3.2.2. Network database – searchable codex of devices and linkages.

Another aspect of visibility that was researched is the ability to perform a search and find various conceptual entities and where they might exist within the network. The goal was to quickly locate and isolate things like:

1. Devices and Device Interfaces
2. MAC and IP Addresses
3. VLANs
4. Routes
5. ACLs
6. L3VPNs
7. Text within a device's configuration, such as a neighbor hostname in a physical interface description.

3.2.3. Reporting – Automated Querying of Network Database

A final identified use case is the ability to perform ad-hoc queries on network data for the purpose of reporting. This use case could in turn facilitate other use cases by giving users direct access to search device data, create reports for management, or use search results as inputs to some other process.

3.3. Decision to implement a PoC with a Specific Vendor

Based on research and the use cases identified, an application was selected for implementation in a POC environment. One of the key factors in this decision was the fact that the application chosen builds a mathematical model of the network based on the transformations made to a packet as it passes through the network. Because the data center network at CCI implements EVPN with L3VPNs, it can be confusing to troubleshoot packets from the CLI. The application gives CCI engineers the ability to visualize how complex network configurations affect the path an IP packet takes through the network on a hop-by-hop basis.

3.3.1. Discussion of Implementation

After reviewing the possible security issues with the CCI internal security team, it was decided to implement the application in a SaaS model. In this model, an internal *collector* is installed within the security perimeter of the organization with the ability to make SSH connections to all devices participating in the network model. The collector logs into each device, runs a battery of commands and collects the resultant data generated by those commands. It then bundles up the data and encrypts it for shipping to the cloud environment, which processes the data, builds a mathematical model of the network, and provides a searchable map through the cloud-based GUI.

Depending on the number of devices requiring collection, the sizing of resources dedicated to the collector may vary. Appropriate sizing for the collector is something that needs to be determined

as part of the implementation process. It is important to note that the function of the collector is solely as a means of gathering information, not processing it. If the 'on-prem' application model is used, resource requirements are significantly higher.

Management of the collector is handled through automation provided by the vendor and upgrades to the cloud environment are all managed internally by the vendor.

The ease with which the application can be managed within CCI was an important factor in the decision to eventually move forward with purchasing the fully licensed version of the application.

3.3.2. Discussion of adoption and usage

The CCI DCNE team manages network connectivity for multiple data centers with the organization. There are two primary data centers, one in Alpharetta, GA and one in Phoenix, AZ as well as a disaster recovery data center in the Metro Atlanta area. In addition, there are distributed regional data centers servicing edge applications in the various CCI markets. To quickly gain benefit from the application, maps of each data center were built to assess if there were possible issues at those locations by running basic checks to validate connectivity between devices. After this initial stage and discovery process, a full map of the network including the network backbone was created. This map allows for the analysis of path data between the various data centers and ensures that traffic follows the expected paths based on where it originates and where it is destined.

Adoption of the new application has not happened as quickly as hoped. The initial collection of data in the primary data center included many devices. Because of this, the first maps which were drawn by the application were highly complex and difficult to understand. Though the search features worked as expected and gave accurate results, the complicated nature of the maps discouraged some users from engaging with the tool. To counteract this complexity, the application allows devices to be aggregated into 'clusters' represented by a single node on the map. This has been effective at making the map more digestible but requires a significant investment in effort to curate the maps. Leveraging the application's API, automated tools for map curation are being developed.

3.3.2.1. Device selection and curation of map

As previously discussed, having an accurate map is vital to developing trust in the tool by network personnel. While it may seem obvious on the surface for highly complex networks, determining the set of devices to collect requires some forethought. While the process of collecting and mapping the network can provide significant insights, a haphazard approach to adding devices can lead to a map that requires significant effort to curate. Built within the tool itself are the capabilities to automatically curate connectivity between devices, but this requires LLDP/CDP on devices interfaces, something that for security reasons is not viable on all devices (such as FWs and some LBs). Consequently, manually linking those devices together can become a time-consuming part of the curation process.

A recommendation to make the process smoother is to start from a central point or 'core' network device and then expand the collection of devices towards the edge of the network. By doing this, devices can be added to the network in a manageable way and the location and placement on the map more easily be curated.

3.3.2.2. Usage/adoption among front line staff

A key aspect of adoption is the use of the application as a troubleshooting tool. The search capability allows users to do things like locate hosts within the network or search paths through the network to ensure that end-to-end traffic is following the expected path. During the information gathering phase of the troubleshooting process, being able to quickly collect this data can have a significant positive impact on MTTI (mean time to innocence) and MTTR (mean time to resolution).

While the application has shown significant promise, there has been some difficulty in getting some network staff to adopt a 'new' tool when they are used to using other tools or using the CLI. A four-session training course was presented to show the efficacy of the tool as a troubleshooting resource. Three sessions were presented by internal staff and a fourth was presented by the vendor. Sessions were recorded and made available for the broader network community at CCI. Although this did increase the user base slightly, it did not reach the expected goal for number of users in the allocated time frame.

While some staff have seen value right away and have begun adopting the tool, these staff tend to be higher-level engineers. We are seeking further engagement and feedback from the potential user community to increase future adoption.

3.3.2.3. Expected growth in adoption

Recognizing the need to integrate this new application into the typical troubleshooting workflow for front line network support, we have undertaken to analyze the typical use cases seen by operations staff. More training on how to solve specific troubleshooting problems will be developed. During this process, operations staff will be engaged to provide feedback to the development process.

Another point of attack in growing adoption is the development of network queries to provide value to users by using the application's query engine to gather data, generate reports, etc. Queries that provide detailed validation of proper device configuration, for example, proper MTU sizing on an interface, can be adopted as part of a workflow process to remediate devices that are out of standards compliance. While this may initially be implemented as a manual process, the application provides a facility to make testing an reporting an automated process on all collections.

3.4. Discussion of PoC Results

Based on the use cases described above, the POC was very successful. In fact, very early on in the POC, issues with the network topology requiring immediate attention to prevent possible outages were discovered.

3.4.1. Efficacy of network entity searches.

One of the most useful toolsets is the ability to search the network for various network elements, such as IP addresses, MAC addresses or VLANs. This search capability has proved very useful and provided great insight into how VLANs are distributed in the various data centers. After searches are entered valid results are shown both as tabular data and as graphical representations on the map. When searching for a VLAN, all points that the VLAN is distributed to are shown on the map.

3.4.2. Efficacy of network path searches

Probably the most powerful of all the tools in the application's arsenal is the network path search: given two IP addresses on the network a graphical representation of the path traffic takes between those points is given. That path information includes detailed information from each device in the flow, including how the packet changes from entry to exit as it moves in and out of the device. The detailed information provided is extremely useful because it allows us to view the relationship between underlay and overlay parts of the EVPN configuration.

3.4.3. Use of network queries to identify misconfigured network devices

Another useful tool of the application is the query engine. This tool allows us to generate queries against all the devices in a map, extract data from those queries, and run tests against that data. During the POC we used this facility to validate the configuration of MTU size on device interfaces. Manually validating the MTU size on thousands of device interfaces is a tedious and error prone process. By writing a properly formed query, we can ensure that all device interfaces that are administratively and operationally 'UP' have an MTU of a certain size or greater. This function works very quickly and gives us the information needed to hand off to our operations teams in csv-formatted report so that the incorrect interfaces can be remediated.

3.5. Realities of Digital Twin maintenance

As described in this case study, the digital twin shows great promise in helping validate the proper operation of the network based on the intent of the network design. There are, however, some considerations that affect the construction of a network digital twin as it is being created and managed.

3.5.1. Time requirements for staff to manage.

The initial time commitment for building a digital twin is very high but should taper off as the number of devices in the collection reaches the full set of devices in the network. This is especially true if accommodations can be made to device configurations that make automated mapping of the network easier, such as the implementation of low-level device discovery protocols is enabled, such as LLDP or CDP. While the application was able to infer connections between devices without direct LLDP/CDP data by analyzing MAC addresses and ARP tables on the device, inference of connections can only work if the application can collect the data necessary to make those inferences. In some cases, network configuration can prevent the collection of that data.

Once the topology of the network is reviewed, and any curation of the twin required is completed, the application will maintain an accurate assessment of the topology unless some change requires an update to the map. It is important to note that the time needed to manage map curation at this point relates to arranging the placement of devices/cluster nodes on the map in such a way that it is readily understandable by all network staff.

3.5.2. Assigning overall responsibility/technical advocacy for map(s)

It is important that a RACI matrix be developed for ownership and management of digital twins. Responsibility and accountability for the accuracy of each map must be vested in individuals who have an interest in the value provided by the map. In a highly complex data center environment, especially a service provider environment, it is likely there is no one person who has a full understanding of the entire data center network. Being able to ensure that the map is accurate is an effort that should be shared between design and operations staff and should enlist enough staff to ensure expertise throughout the environment.

An example of the need for this effort would be the case where a connection between two devices was not properly discovered, requiring manual configuration in the digital twin model. In this case, path searches may not resolve as expected, which would be an obvious indication of an issue to someone familiar with the expected path that traffic should take.

3.5.3. Cost/Benefit discussion.

While a full ROI analysis has not been performed on the application, there have already been several instances where the manpower required to analyze and remediate issues in the network discovered by the tool would have been orders of magnitude greater if done manually (if they would have been identified and fixed at all). It is important to note though, that fully realizing the value of the tool requires investment in manpower to customize the tool to the specific environment within which it is deployed, which is something that must be done on case-by-case basis. In addition, a fair accounting of the fully realized ROI would require a significant analysis in time savings for network staff.

4. Conclusion

We hope that the paper has provided a blended view of network digital twin, including both a forward-looking glimpse of the long-term landscape with some of the exciting transformative possibilities, and a practical account of how one such possibility is being realized today.

In the words of Lao-tzu, “A journey of a thousand miles must begin with a single step.” The introduction, development, and wielding of network digital twin is a journey worth taking. We welcome fellow travelers on this journey as we all take the first steps!

Abbreviations

ACL	access control list
AI	artificial intelligence
AT	Actual Twin
CCI	Cox Communications, Inc.
DCNE	CCI Data Center Network Engineering
FMEA	Failure Mode and Effects Analysis
HDTI	Human Digital Twin Interface
HMI	Human Machine Interface
IBN	Intent-Based Networking
ISP	Internet Service Provider
INCOSE	International Council on Systems Engineering
MAC	Media Access Control
ML	machine learning
NDT	Network Digital Twin
NDTE	Network Digital Twin Ecosystem
NDTh	Network Digital Thread
SCTE	Society of Cable Telecommunications Engineers
SDLC	System Development Life Cycle
SEBoK	Systems Engineering Body of Knowledge
VPN	Virtual private network

Bibliography & References

Intent-Based Networking – Concepts and Definitions (draft-irtf-nmrg-ibn-concepts-definitions-02), A. Clemm, *et. al.*; IRTF

Intent-Based Networking – Concepts and Overview (draft-clemm-nmrg-dist-intent-03), A. Clemm, *et. al.*; IRTF

Header Space Analysis: Static Checking for Networks, Kazemian, Peyman, G. Varghese and N. McKeown; *NSDI* (2012).

SEBoK: *Systems Engineering Body of Knowledge*; International Council on Systems Engineering

SEBoK/Digital Engineering: Article titled *Digital Engineering* in the *Systems Engineering Body of Knowledge*, 2021; International Council on Systems Engineering

The Way of Lao-tzu, 64, Lau-tzu, c. 604-c.531 B.C.E.; Bartlett's Familiar Quotations, 18th edition

Small Cell Deployment Strategies for Cable Broadband

A Technical Paper prepared for SCTE by

Toby Peck
Sr Director Product Management
EnerSys
Toby.Peck@enersys.com

Greg Laughlin
Strategic Marketing Manager
EnerSys
Greg.Laughlin@enersys.com

1. Introduction

Cable broadband and wireless operators are working together to create new revenue-generating partnerships and service offerings. While some cable operators are reselling wireless mobility as part of their bundled service offerings, others are selling mobile operators access to their infrastructure for cost-effective small cell connectivity.

Many broadband operators also plan to deploy their own small cell networks, leveraging recent acquisitions of CBRS spectrum and their existing Hybrid Fiber-Coax (HFC) infrastructure. Regardless of the reason, introducing small cells to the HFC network presents many challenges, from making a workable business case, to product design and development as well as operational practices.

Reliable power and quality backhaul are essential for small cells, and radios must be strategically located across a wide geographical footprint. You must also maintain a tightly coordinated schedule while working with a limited budget. The HFC network is a smart choice for distributed connectivity—it alleviates the cost of building a small cell network from the ground up, providing reliable, cost-effective infrastructure through most populated areas.

When HFC is not available to support small cell devices, alternative powering approaches may be more cost-effective. Remote Line Power (RLP) uses high-voltage DC power to distribute longer distances using copper with fiber, often in a hybrid fiber/copper composite cable. Local utility power can also be used, but often requires expensive construction, permitting and metering for new service to cell site installations. All three powering options have advantages and disadvantages.

This paper will evaluate HFC for small cell power and backhaul capabilities and explore different methods of deployment. We'll also look at local utility power as well as remote DC power solutions for small cell powering. Small cell deployment requirements and challenges will be explored with guidance on selecting connectivity solutions based on the situation and environment.

2. Small Cell Radio Deployment Challenges

In a perfect world, operators would be free to place small cell radios wherever they want, and every location would have a power outlet with reliable backup power and fiber for backhaul. The realities of small cell deployment are different, and it can be quite expensive and can take months to bring power and communications to the radio. Local siting restrictions and costly permits can add even more barriers.

2.1. Cost-Effective Infrastructure Availability

Unlike tower-based cellular radios which can provide several square miles of coverage, mid-band small cell radios output much less power with coverage areas of a few hundred feet maximum. This means radios must be close to the coverage target, and a large quantity is needed to cover a sizeable area. Small Cell Forum recommended best practice is to install small cells within 20-40m of heavy traffic locations.[1] In other words, a good deployment solution needs scalability and precise radio placement.

Getting permission to install a small cell can be a challenging process, with hurdles around zoning, securing siting, permitting and regulations. Jay Brown, Crown Castle CEO reported it can take up to 2

[1] SCF 2021 Market Forecast document 050.10.5; http://smallcellforum.org/scf-market-forecast; viewed 7/28/21

years to deploy a small cell on a wall or pole, due to wading through siting and construction issues.[2] That same deployment model can cost $25,000 or more.[3]

2.2. Reliable Power

More than ever, keeping the mobile network running at full capacity without interruption is essential. Networks must withstand daily power anomalies that plague the United States utility grids and cause damage to sensitive network equipment.

Extreme weather events are becoming more prevalent, resulting in more frequent, longer duration power outages. United States utility customers experienced 1.33 billion outage hours in 2020, up 73% from 2019's 770 million outage hours.[4]

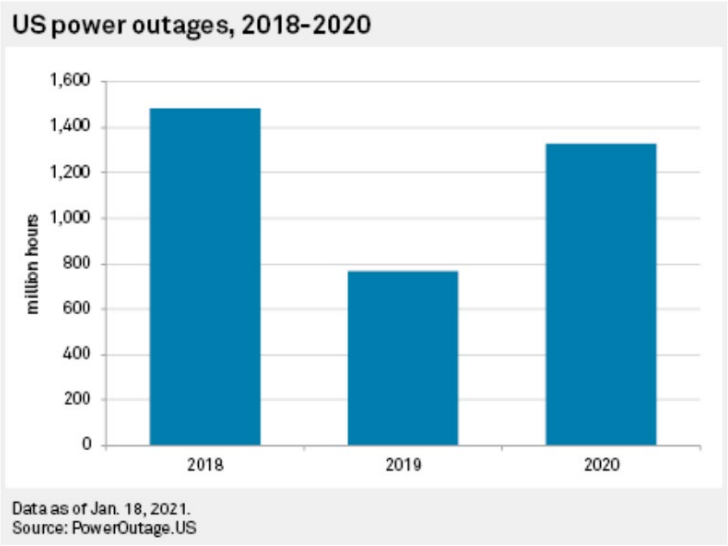


Figure 1 - US Utility Outages 2018 to 2020

A reliable small cell network requires clean conditioned power and, where econmically viable, energy reserves for backup during utility outages.

2.3. High Bandwidth, Low Latency, Reliable Backhaul Connectivity

Cellular networks have traditionally relied on fiber optic backhaul communications due to superior performance in bandwidth and latency. Next generation 5G design parameters specify delivery of up to 10Gbps throughput with less than 1ms latency[5]. These are essential factors for real-time use cases like gaming, video streaming and remote healthcare, where the experience requires seamless communications.

[2] Small cells: Strand-mounted business opportunities; BTR; 9/5/2018; https://www.broadbandtechreport.com/docsis/hybrid-fiber-coax/article/16449380/small-cells-strandmounted-business-opportunities

[3] "Cable's wireless biz 'ready for its star turn' – analyst"; LightReading; J Baumgartner; 6/17/21; < https://www.lightreading.com/cable-tech/cables-wireless-biz-ready-for-its-star-turn---analyst-/d/d-id/770301 >

[4] US power outages jumped 73% in 2020 amid extreme weather events; G. Herring; 1/19/2021; https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/us-power-outages-jumped-73-in-2020-amid-extreme-weather-events-62181994

[5] 5G Low Latency Requirements, McLaughlin, 2019, https://broadbandlibrary.com/5g-low-latency-requirements/

While fiber may be preferred, it is not always cost-effective when construction is needed to bring the fiber to the location where the small cell is installed.

Meeting the goals of cable broadband small cell deployments will require operators to find financially viable ways to provide the fundamental elements for high quality small cell radio connectivity, while maintaining reliability of their existing network. Long term plans should consider product lifecycles and how to stay ahead of evolving radio technologies.

3. HFC as an Advantage for Small Cell Deployment

3.1. Advantage: Ubiquitous Broadband and Power Grid

For many years, the core mission of the cable broadband network has been to deliver fast, reliable and high-quality wired connectivity. What started out as a scattering of coax cable networks intended to provide community access television in the late 1940s[6], has evolved into a vast network of fiber optic and copper coaxial cables delivering not only television, but Internet and phone service around the globe.

In the United States alone, cable broadband networks have over 1.7 million total miles of deployed fiber and coax to make up the HFC network. This vast expanse of physical infrastructure passes 96% of the homes in the U.S., providing service to 78 million customers[7].

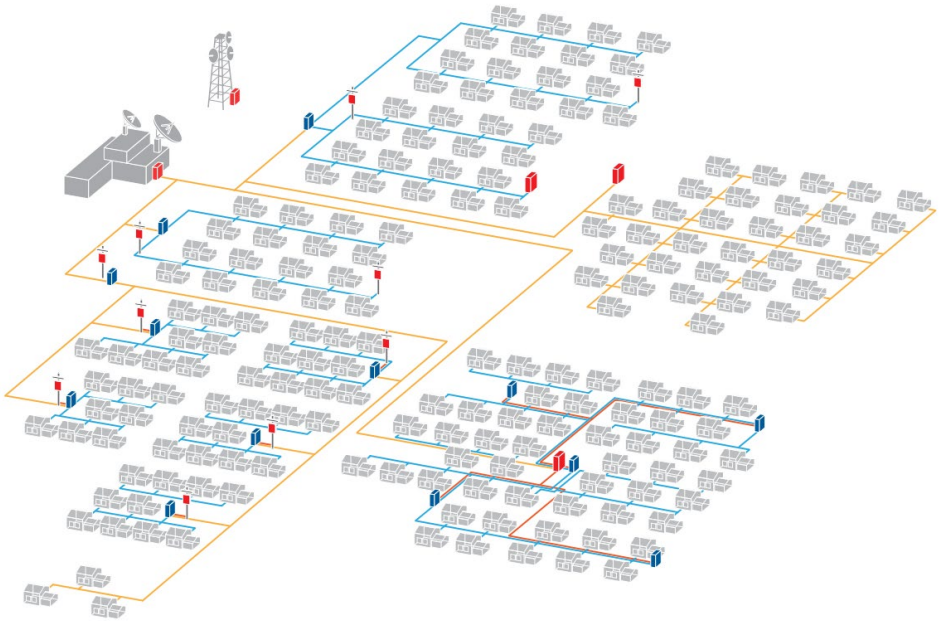


Figure 2 - HFC Network

In the HFC network, fiber optic cables are primarily used as an efficient delivery mechanism to bring long haul network traffic into neighborhood nodes; and in some new construction to end users directly using Fiber to the Home (FTTx). These nodes convert light signals to electrical signals for transmission over coax. The coaxial cable then delivers data for the "last mile" to each subscriber's home through a cable

[6] The Cable History Timeline; Viewed 7/25/2021; https://www.cablecenter.org/images/files/pdf/CableHistory/CableTimelineFall2015.pdf

[7] NCTA Broadband Stats and Facts; Viewed 8/1/2021; https://www.ncta.com/broadband-facts

modem. Remote PHY (CCAP) nodes, traditional fiberoptic nodes, and amplifiers require power, so the practice of using the network coax to carry power is still widely used as illustrated in Figure 3.

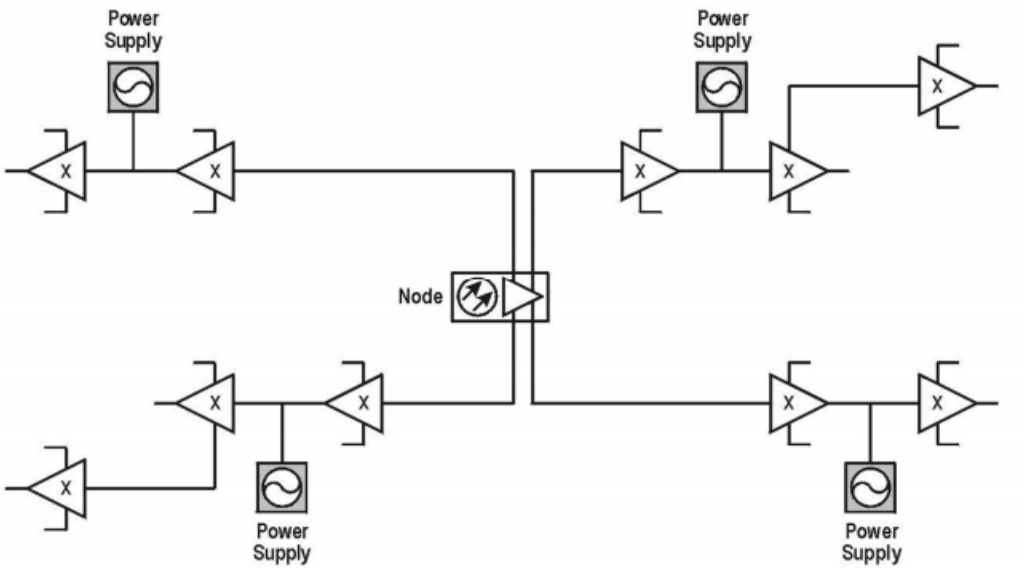


Figure 3 - Typical HFC Network Powering Diagram[8]

Coaxial cable accounts for 20% of the HFC network. This coaxial network can be looked at as 340 thousand miles of existing, accessible and available power lines running through most populated areas.[9] While utility power lines often span the same poles, there are other aspects of the HFC network that make it much more suitable for small cell deployments.

3.2. Advantage: Accessibility and Scalability

The HFC network architecture is built on a distributed access model that is extremely efficient and flexible, enabling Internet connectivity at any point in the network. According to a recent publication from iGR, it is estimated that the cable broadband aerial strand could support small cell installations for over 203 million people in the United States, about a quarter of the total addressable market.[10]

NCTA identifies the flexibility and continuous management as key value props of the HFC network:

> *"The beauty of HFC architecture is it marries the power of gigabit capable networks with an architecture that is extremely efficient and flexible…network engineers don't just plan for today, they build networks that accommodate future growth in consumer demand. That planning allows cable's networks to stay ready and robust in the face of events that may cause internet traffic surges."* [11]

[8] SCTE Standard 238 2017; https://www.scte.org/standards-development/library/standards-catalog/scte-238-2017

[9] NCTA Broadband Stats and Facts; Viewed 8/1/2021; https://www.ncta.com/broadband-facts

[10] iGR, 2020, A Strand of Hope: How strand-mounted small cells can address the demand for 4G and 5G mobile data, iGR Media Center, Viewed 3/15/2021 https://igr-inc.com/media-center/white-papers/strand-mounted-small-cells/strand-mounted-small-cells.asp

[11] *"Why Cable's Broadband Network Is Handling the Pandemic and Ready for the Future", NCTA; 4/29/2020; https://www.ncta.com/whats-new/why-cables-broadband-network-is-handling-the-pandemic-and-ready-for-the-future*

The coax portion of the network is used for broadband delivery, with physical 'taps' cut into the hardline coax to provide individual connections to the broadband network through smaller, more flexible and manageable premise coax. While most taps used for residential customer broadband connectivity prevent power from travelling to the home, a special style of 'power-passing tap' allows for both power and broadband to be extracted. These power-passing taps were initially introduced to support telephony systems and more recently re-introduced to provide power and backhaul communications for Multiple Service Operator (MSO) strand-mount, Wi-Fi access points. With hundreds of thousands of HFC-enabled Wi-Fi access points deployed worldwide, broadband operators have proven the effectivity HFC can play for scalable wireless access.

3.3. Advantage: Reliable Source of Power

HFC power supplies take power from a utility connection and energize the coax, enabling nodes and amplifiers to power directly from the coax, removing the need for utility connections to each device. Since the cable operator's communications channel radio signals (RF content) and alternating current both propagate using sinusoidal waves at different frequencies and can simultaneously travel over the same coaxial cable. This combination of power and communication is unique to cable broadband and a differentiator for small cells.

HFC power supplies utilize ferroresonant transformers to provide a very robust and quality powering platform, ideal for sensitive network equipment. Ferroresonant transformers by nature provide excellent line isolation; 1000:1 ratio for HFC power supplies. This means that any change to the input, small or large will have little effect on the output. A high voltage surge or spike from the utility will be minimized at the power supply before it can affect the HFC network.

Most broadband cable operators co-locate their HFC power supplies in cabinets with batteries to maximize network power availability when utility power fails. Today's HFC power supplies can switch to battery mode within a half duty power cycle, preventing the nodes, amplifiers and radios from dropping power and resetting during the power transfer.

3.4. Advantage: Physical Placement

Furthermore, most cable operators enter into franchise agreements with their local municipality which allows them to connect any network equipment to the HFC without requiring additional site approvals or permits.

New Street research analyst Spencer Kurn noted the value the broadband cable strand played in Sprint's recent 20,000+ small cell deployment in Long Island. "Cable companies already have the public right of way with poles and strands of aerial cables… Historically, cable companies have been reluctant to open their infrastructure to wireless carriers to deploy small cells… Altice is unique in that they did this for Sprint in return for a really attractive wireless Mobile Virtual Network Operator (MVNO)."[12]

Moffet Research estimates that an MSO can install a CBRS small cell for around $2,500 by avoiding the siting costs that are applied to pole and wall installations.[13]

[12] L Hardesty, 2019, Altice's 19,000 small cells in Long Island don't help Sprint's network much, say analysts, Fierce Wireless, Viewed 3/1/2021, https://www.fiercewireless.com/wireless/altice-s-19-000-small-cells-long-island-don-t-help-sprint-s-network-much-say-analysts

[13] "Cable's wireless biz 'ready for its star turn' – analyst"; LightReading; J Baumgartner; 6/17/21; https://www.lightreading.com/cable-tech/cables-wireless-biz-ready-for-its-star-turn---analyst-/d/d-id/770301

3.5. Advantage: High-Capacity Low-Latency Backhaul

MSOs are pushing fiber optics deeper into neighborhoods and moving much of their core network processing from regional headends out to the edge of their network. This "edge computing" application greatly improves processing time, enabling higher capacity with lower latency. A recent CableLabs® analysis showed >50% reduction in total cost of ownership (TCO) for an outdoor use case of backhauling small cells when served by DOCSIS® networks compared to a more traditional deployment served by fiber.[14]

DOCSIS® 3.1 is currently deployed globally in millions of homes and businesses. Today's DOCSIS® 3.1 introduces support for low-latency DOCSIS (LLD), and provides an option to increase upstream spectrum from the traditional 5 to 45MHz range to 5 to 205MHz. This adds significant upstream capacity improvements and sets the groundwork for future improvements coming with DOCSIS® 4.0.

Through DOCSIS® enhancements known as low latency X-Haul over DOCSIS®, 5G backhaul latency can be reduced to approximately 1ms, surpassing 5G vRan targets of 5ms.[15]

Extended frequency division duplex spectrum DOCSIS® 4.0 will take advantage of the extended 1.8GHz of spectrum as well as full duplex to maximize throughput. DOCSIS® technology leaders expect throughput target speeds of 10Gbps downlink and 5Gbps uplink.

3.6. Advantage: Existing Well-Maintained Resource

One major advantage of the broadband HFC network is it's the lifeblood of cable companies and warrants significant, continuous investments and routine maintenance. Cable operators in North America invested $17B in 2020 alone, on infrastructure and networks[16], with total investments of $290 Billion since the year 2020[17]. NCTA credits these cable broadband commitments to investing in their network as key factors leading to industry's ability to meet added demands from the COVID-19 pandemic work-from-home impact.

> *"…hundreds of billions of dollars in investment in infrastructure and technology, smart and responsive network engineering, and a flexible network foundation that is being constantly upgraded with hardware, software and other technical modifications* (NCTA, 2020)[18]*."*

And in addition to capital investments, cable broadband's smart architecture and commitment to localized management provides a very resilient and reliable infrastructure. *"By having eyes on every aspect of network performance, engineers are able to diagnose localized issues and often troubleshoot them quickly."*[19]

[14] DOCSIS® Network vs. Fiber Backhaul for Outdoor Small Cells: How Larger Footprint of DOCSIS Networks Lowers TCO in the Outdoor Use Case, https://www.cablelabs.com/docsis-vs-fiber-backhaul-outdoor-small-cells

[15] CableLabs Low Latency DOCSIS® Technology Launches 10G Broadband into a New Era of Rapid Communication; https://www.cablelabs.com/cablelabs-low-latency-docsis-technology-launches-10g-broadband-into-a-new-era-of-rapid-communication

[16] "Internet Facts and Stats; NCTA; viewed 7/20/2021; <https://www.ncta.com/broadband-facts>

[17] *"Why Cable's Broadband Network Is Handling the Pandemic and Ready for the Future", NCTA; 4/29/2020; < https://www.ncta.com/whats-new/why-cables-broadband-network-is-handling-the-pandemic-and-ready-for-the-future >*

[18] Ibid

[19] Ibid

4. Connecting Small Cells to HFC

4.1. Design and Integration Considerations

The cable broadband HFC network brings a unique set of connectivity considerations for small cell radios. Creating HFC-compatible small cell radios require integration of HFC components and elements into radio design allowing them to mount and receive power from HFC coax, backhaul via DOCSIS® and enable remote system management and reporting.

4.1.1. Housing Size Limitations

Cable broadband operators typically enter into franchise agreements with local municipalities that include rights-of-way. These agreements provide cable operators with the freedom to hang outside network equipment from the steel messenger cable, or "strand" that spans between utility poles and lashes the coaxial network cables. The terms of the agreements typically allow the operator to deploy their networks in this fashion without requiring new permits for each piece of equipment deployed. The equipment deployed on the strand, however, must meet size, weight and power restrictions.

Starting with form-factor, any device designed needs to fit into what the National Electric Safety Code (NESC) defines as the Communications Space. NESC specifies a minimum 12" clearance between cable TV and telephone lines as shown in Figure 4, but some jurisdictions lower that to 11in or even 9in in some cases.

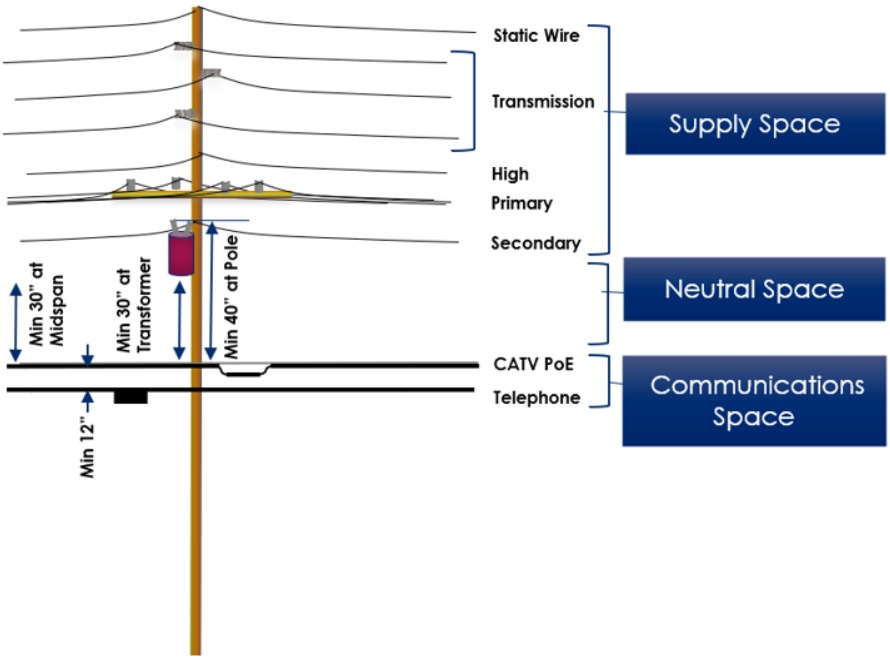


Figure 4 - Communications Space

The coax itself must be accounted for in this space as well which can reduce this space by 2 to 3in, and some broadband operators have individualized stringent space requirements. This means small cell radios should be no taller than 9in (including integrated antennas).

Figure 5 - Small Cell Radio within Communications Space[20]

4.1.2. Physical (Outside) Connections

The main connection points for the HFC small cell are the mounting connections to the strand, the grounding connection, and most importantly the connection to the HFC coax.

There is no industry specification for strand-mount brackets but there are basic guidelines to consider. Messenger strands are typically 0.25in or 0.375in in diameter, so useful designs would include flexibility to support either size. A potentially large cable bundle, up to a 3in diameter may be tethered to the strand, creating an obstacle between the strand and radio. Most strand brackets are designed with a "C" shape to accommodate the coax bundle. Another important consideration for the strand bracket is to check that materials are galvanically compatible with radio housing and the galvanized steel strand. Metals with dissimilar galvanic properties can lead to unwanted corrosion.

Figure 6 - Strand Bracket

The strand messenger, in addition to mounting, also provides the electrical grounding for strand-mounted equipment. While many radio designs use the mounting bracket as the ground connection, a ground lug on the radio housing can facilitate a dedicated cable to secure a ground connection between the radio and strand.

Perhaps the most important physical connector is the HFC coax connection. The coax cable delivers power to the radio, provides the network backhaul and is a potential source of water intrusion and radio noise ingress. There are two standard connection types used for HFC equipment, typically selected based on the size of coax cable connecting the equipment, which is determined by the cable operator based on

[20] Image taken from Airspan website; 8/1/2021; https://www.airspan.com/cbrs/

the amount of power that the equipment draws from the HFC coax. Most strand-mount HFC equipment like nodes and amplifiers connect directly to distribution “hardline” coax, which is much heavier and less flexible than the smaller coax “drop” cable that run from the HFC to a residence. Smaller drop coax, which is more flexible and less expensive than hardline can be used to connect smaller loads like Wi-Fi access points. While the decision is made by each cable operator, general guidelines limit the smaller drop coax for loads under 90W, with hardline coax is suitable for any loads. Smaller coax cables use F-style connectors, the same as home modems and set top boxes specified by ANSI_SCTE_124_2021[21]. Hardline coax uses a standard 0.625in x 24 threaded connector as specified by ANSI_SCTE_91_2015[22]. It is important to follow SCTE 91 specifications closely as there are essential elements like shrink sleeve ridge and galvanic considerations.

Figure 7 - Hardline ‘KS’ connectors

Designing small cell radios around SCTE 91 specifications (hardline connection) provides the cable operator with flexibility to connect using either hardline distribution cable or drop cable by installing a ‘KS male to F’, or “pin-to-F” adapter. The pin-to-F has the standard 0.625in x 24 threads to connect into the radio, with the other end having F-type threads for drop cable.

Figure 8 - Pin-to-F connector

4.1.3. Power

An HFC small cell radio must be able to power from the HFC coax, which delivers an alternating current (AC), quasi-square wave in the range of 45-90VAC, depending on resistive losses created by distance from the local HFC utility power supply. Cable networks work similarly to wireless networks, sending digital data over radio frequency (RF) waves, only the waves are sent within coaxial cables and fiber optical cables instead of openly through the air. A unique and important feature of the HFC network is the

[21] https://www.scte.org/standards-development/library/standards-catalog/ansiscte-124-2011/

[22] https://www.scte.org/standards-development/library/standards-catalog/ansiscte-91-2015/

use of Alternating Current (AC) for power, both power and data can travel through the same coaxial cable at the same time.

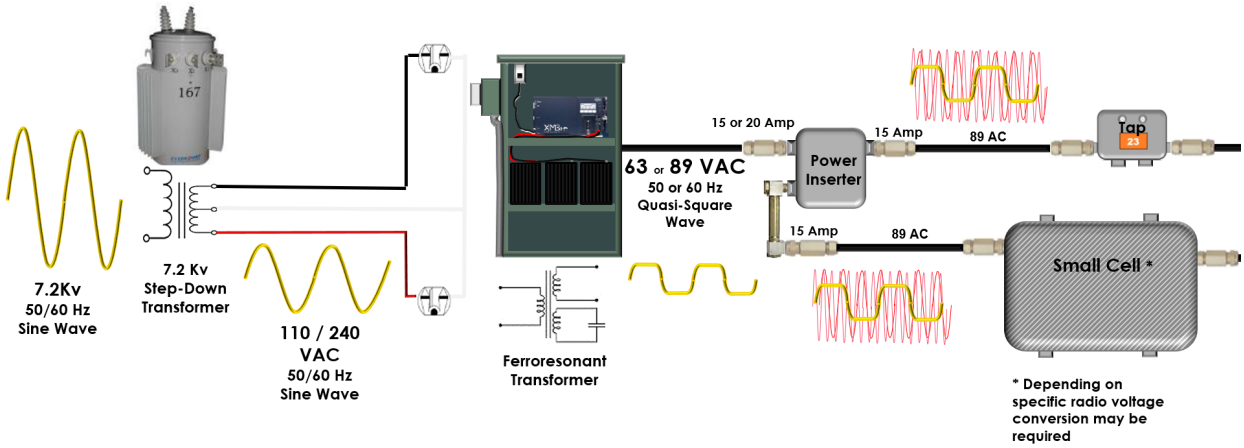


Figure 9 - HFC Power System

The device being powered by the coaxial network, be it a node, amplifier or in this case a small cell radio, will use RF filters to separate the power and DOCSIS® RF. At this point, the device can convert the voltage from AC to DC to power the device, and RF can be used for DOCSIS® backhaul with the integration of a cable modem.

Voltage levels drop over distances per Joule's I2R law, meaning voltage at the small cell will be lower than at the cable power supply. Most cable operators dictate that devices powered by HFC automatically shut down at lower voltages in the 35 to 45VAC range protecting the device from drawing too much current [Current (I) = Power (P)/Voltage(V), therefore a 100W radio will typically draw twice the current at 45VAC compared to 90VAC]. Many cable operators also require the powered-device (small cell radio in this case) sustain operation over a full duty-cycle of sustained power loss, requiring additional capacity in the power circuit.

Another consideration around small cell power is the advantage of using watchdogs and remote power control to reduce outages and costly truck rolls. Most end users have had experiences with home networking equipment locking up, requiring a power cycle to bring the device back online. This is simple when the device is in a home, office, or headend, but problems on the HFC strand require expensive truck rolls to perform the most basic tasks, like pushing a reset button. A good watchdog system should catch any problems with routine operation of the device and will auto-reset or trigger a power cycle to the system without human intervention. It's also valuable to provide remote power control of the radio element independently of the overall radio system. This can be very useful if the radio element gets stuck or if the operator wants to shut down the radio transmitter for a local service call. This value-added feature however requires communications and power control between the DOCSIS® modem and onboard AC/DC power supply.

4.1.4. Backhaul

One of the primary benefits of deploying small cell radios over HFC is the ability to use DOCSIS® for data backhaul. DOCSIS® 3.1 has incorporated features useful for small cells including: low-latency DOCSIS®, mid-split and high-split spectrum options to increase throughput, and additional tools like

business services over DOCSIS® for added management and security. Cable operators deploy millions of DOCSIS® modems to their home and business subscribers. Managing a cable modem inside of a small cell radio should not be much different.

However, there is a significant difference between a home or office modem and an outdoor hardened modem, especially when it resides inside an enclosed housing with an RF transmitter. Outdoor small cell radios are designed to operate in outdoor climates and be exposed to extreme hot and cold temperatures where modems must be built with more expensive, rugged components. In addition to harsh environmental conditions, outdoor HFC modems carry more stringent requirements to withstand power disturbances like surges, line cross and electrical static shock. They are expected to self-recover from network and power outages and can require operator-specific firmware for remote system management.

4.1.5. Safety

Electrical shock hazards and RF radiation are the top safety concerns for strand-mount small cells. The NESC limits products residing within communications spaces to 90VAC which is considered 'low voltage' and safe for technicians to work around without requiring electrician licensing. This however does not mean that HFC power is not dangerous. While it is important for technicians to use safety procedures when working on or near HFC, the products using HFC for power must meet rigorous standards for electrical shock and be proven through agency testing.

Small cell radios fall under the category of 'Information Technology Equipment' for Underwriter Laboratories[23] (UL) and CSA Group[24] (CSA – formerly Canada Safety Agency), the two primary safety certification agencies in North America for cable broadband equipment. Both UL and CSA are Nationally Recognized Testing Laboratories (NRTL) that test to the same harmonized Canadian Electrical Code (CEC) and National Electrical Code (NEC) standards, making their certifications virtually interchangeable.

4.1.6. Environmental

Outdoor small cell radios are subjected to harsh weather conditions, including extreme temperatures, ice, rain and wind. Installations in coastal areas add the element of salt, an enemy to steel and electrical circuitry. Aluminum housings coated with salt-resistant paint, along with designing around dissimilar metals, will go a long way to preserve equipment integrity.

4.1.7. Network Integrity

A challenging but important element of the small cell radio, especially when connected to the cable broadband network, is the mitigation or prevention of electromagnetic interference (EMI) between the cable and wireless networks. Both wired and wireless networks use radio signals for communications, often in the same frequency range. It is essential these networks are isolated from each other and radio signals from one medium do not bleed into the other.

Damaged or improperly installed coaxial cable can act like an antenna and is more susceptible to conducted interference, where RF noise or signals are propagated through the coax. This interference can impact the RF signals from the cable headend, resulting in poor or lost communications for customers

[23] https://www.ul.com/

[24] https://www.csagroup.org/

across the network. Conducted noise can also come from crystal oscillators in tuner circuits, or from AC/DC power supplies within the small cell.

Wireless networks are more susceptible to radiated interference, where RF signals leaking from the cable network via an enclosure or poor connector affect the wireless signal of the radio. The effects are similar to those on a wired network, impacting overall performance and integrity of the wireless network.

A sealed, aluminum small cell housing acts as a Farady cage and provides most of the protection needed to mitigate interference between networks. However, RF inside the sealed housing, especially when housing a cable modem and wireless radio can be exceedingly difficult to overcome. Shielding, grounding, internal wiring, EMI filters and a lot of testing are required to truly protect the cable network.

4.2. HFC Connectivity Methods: Integration and Demarcation

This section will discuss two methods of deploying small cells over the HFC network and provide advantages and disadvantages to both.

4.2.1. Small Cell with Integrated HFC Elements

A DOCSIS® small cell radio is one that has been designed specifically for strand-mount and connecting directly to HFC coax for power and backhaul. The radio manufacturer will design a housing to meet the industry requirements, including but not limited to:

- Strand mounting connections
- Coax connection including seizure screw
- Status indicators (modem, radio)
- Water intrusion
- UV
- Wind loading
- Antennas
- EMI
- Communications space limitations
- Electrical grounding
- RF Protection from external impairments
- Corrosion (water, salt water, dissimilar metals)
- Vibration
- Labelling
- Internal circuitry including RF protection
- Electrical surge

HFC-specific internal components must be designed or purchased to include:

- Coax interface
- RF/Power Filter
- Outdoor hardened DOCSIS® modem including all firmware for management, MSO-specific features

A single box system minimizes the number of cables, connectors and external components. This reduces the number of potential failure points and makes for simple system installation, troubleshooting and management verses a solution with multiple elements for power and backhaul.

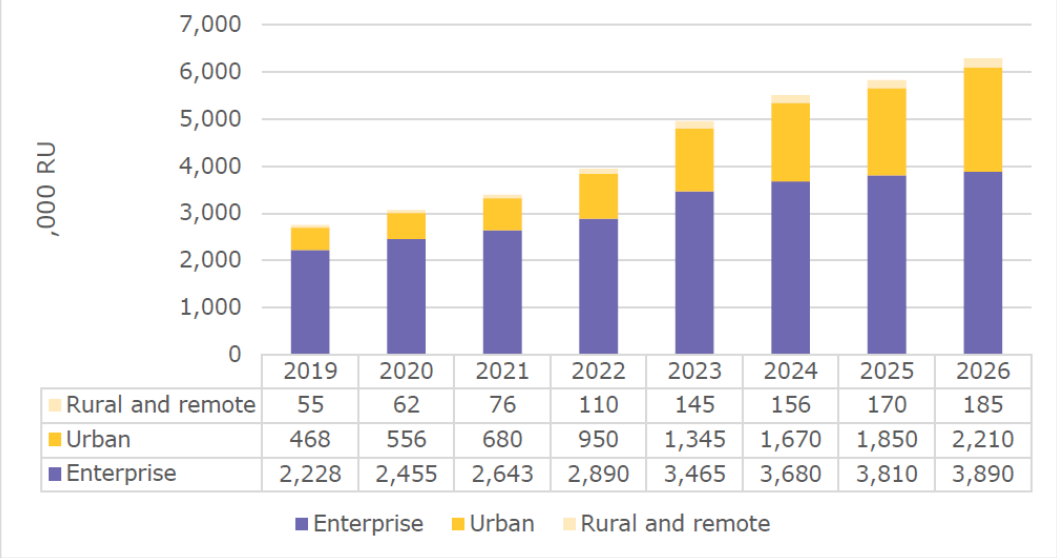
The procurement process is also simplified by having a single manufacturer to coordinate with for purchasing if any problems arise. Instead of spending time to understand which part of the system is causing problems, the radio can be replaced, returning the defective unit to a single vendor.

However, creating a small cell radio specifically for HFC has its own drawbacks. The biggest challenge is with small cell radio manufacturers and their need for a business case to justify the investment needed to bring an HFC-specific radio to market.

Like any manufactured product, one goal is to create products serving multiple markets without additional SKUs, allowing manufacturers to leverage economies of scale to lower their overall component and operational costs.

HFC deployment historically has been a small niche sub-market for outdoor small cells, and outdoor small cell deployments are a fraction of indoor enterprise. For example, according to Small Cell forum around 468,000 outdoor small cells were deployed in 2019, approximately 20% of the total deployments.[25] In that same year Sprint deployed ~20,000 strand mount small cells over Altice's HFC, which would have accounted for ~4% of the total outdoor small cell deployments.

The outdoor small cell market share is expected to increase to 35% by 2026 to over 2 million radios installed, but it is too early to say how many of those radios will be targeted for HFC deployments.



	2019	2020	2021	2022	2023	2024	2025	2026
Rural and remote	55	62	76	110	145	156	170	185
Urban	468	556	680	950	1,345	1,670	1,850	2,210
Enterprise	2,228	2,455	2,643	2,890	3,465	3,680	3,810	3,890

Figure 10 - Small Cell Deployments by Segment 2020 - 2026[26]

Other large MSOs are planning their small cell programs with help from recent allocations of licensed and unlicensed CBRS spectrum by the FCC, and their deployment numbers will likely reach over 100,000, but this is not a proven market and any forecasted numbers must be met with caution. Without solid commitments, radio manufacturers are faced with high-risk investments which pull resources away from other revenue-generating opportunities.

[25] SCF 2021 Market Forecast document 050.10.5; http://smallcellforum.org/scf-market-forecast; viewed 7/28/21

[26] Ibid

Another challenge for the integrated small cell radio is each vendor must go through steep learning curves with regards to cable networks and nuances involved with HFC power, RF mitigation and industry practices. This can add significant delays in development schedules and require radio vendors to become experts in technologies outside of their core competencies. Vendors must invest in costly CMTS and HFC power equipment and may need to build their own cable networks for development and testing.

Certification is another pain point for radio manufacturers. Small cell radios must be FCC certified for each radio band they serve and may require additional certifications for each country where the radio will be deployed. The addition of a cable modem to the radio requires CableLabs DOCSIS® certification at costs intended for high-volume amortization seen in the residential modem market.

Each MSO will further qualify and verify the hardware for operation in their networks. These tests include functional, RF emissions and operational test as well as environmental tests. Test and certification cycles are lengthy and expensive, requiring heavy investments in specialized equipment and expertise. In a best-case scenario, the radio passes all tests on the first attempt but this is rarely seen if ever. Multiply this by several vendors and it can have a significant impact on complexity and schedule.

4.2.2. HFC Demarcations for Small Cell Radios

Another approach to bringing small cell radios to HFC is to create demarcations allowing radio vendors to use off-shelf products with minimal changes. Cable broadband already uses demarcations in the form of taps used to deliver DOCSIS® to residential customers. It is widespread practice within the industry for MSOs to establish standardized RF levels at their taps so technicians can deploy a set length of drop cable knowing it can provide the optimal signal at the user's cable modem.

An HFC demarcation device for small cells would perform the power and media conversions and provide the radio the power and network connections.

Figure 11 - Small Cell with Demarcation Gateway (Author's Photo)

Demarcations allow small cell radio manufacturers to use radios for multiple purposes, eliminating the need for them to design low-volume niche products for HFC. Radio certifications and SKUs are reduced, proving additional economies of scale for the radio vendor. The MSO can then leverage competition between radio vendors which can result in additional cost reductions.

Demarcation gateways are not radio vendor -specific, so operators can source from one or two manufacturers directly, leveraging volume pricing.

Demarcation also provides additional reliability to the HFC network by acting as an isolator between the wired and wireless networks, protecting both from RF ingress. It gives the MSO better control of the cable modem which interfaces directly with the core network. Demarcation gateway manufacturers have expertise in outside HFC plant and DOCSIS® modems, providing additional quality assurance to the operators.

Finally, the overall TCO of the radio network will continue to stay low as next-generation radio products come to market. Wireless products tend to evolve faster than DOCSIS®, and demarcations provide the ability to upgrade your forward-facing technologies without sacrificing the entire system. MSO lab qualification and testing time is also greatly reduced for even better financial results.

A downside of using demarcations is the burden of extra inter-connecting cables between the demarcation device and radio. Extra cables and connectors do add some cost, but more importantly add points of failure to the system. It adds complexity to installations and potential problems when troubleshooting in the field.

When an unexpected problem occurs in a 2-box system (demarcation + radio), a technician must isolate the problem and work with the vendor to troubleshoot. Having two vendors to coordinate with can drag out repairs, adding cost and potential network downtime to resolve. However, that ability to troubleshoot from the demarcation point can provide additional information to speed the troubleshooting process.

Another disadvantage to the demarcation concept is aesthetics. More cities and municipalities are requiring outdoor equipment to be more aesthetically pleasing. Fewer devices hanging from the strand looks better than a multiple boxes with cables in between.

One concept addressing the problem of multiple connected products is a direct connection where the radio and demarcation are physically attached to each other with virtually unseen power and network connections. This model still isolates the radio from the demarcation, providing the benefits stated previously, but provides the benefits of a single-box solution.

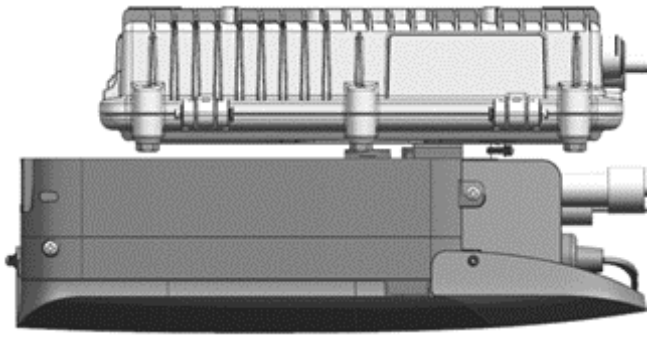


Figure 12 - Demarcation with Radio Attached

5. Local AC Utility Power for Small Cells

While this paper focuses primarily on HFC for small cell power, there are other options broadband operators can leverage.

In rare situations where rooftop power is available from a building where power distribution is managed by an operator, local AC utility power may be the ideal powering solution for small cells if it is the only source available, and if the small cell radio is designed to operate from AC voltage. The benefits are apparent: No time constraints for permits or meter installations, and only incremental monthly power consumption costs. However, the limitations of direct utility power for small cells include limitations around power quality and reliability.

If the availability of the radio is a high priority, or if the radio is designed for DC power, a local power supply should be installed between the utility and radio. A local power supply can condition utility power to remove noise and voltage spikes, and if the radio requires it, convert utility AC voltage to useable DC power. Batteries can also be added to the local solution to facilitate radio availability during utility outages and power disturbance events.

Using local utility power can become a disadvantage if that power is not readily available at or near the radio location.

6. Remote DC Line Power for Small Cells

RLP uses a centralized power hub to deliver DC power long distances over copper pairs to remote locations. The basic design involves converting local utility power to DC, elevating the DC to high voltage for distribution, then lowering the voltage at the delivery point.

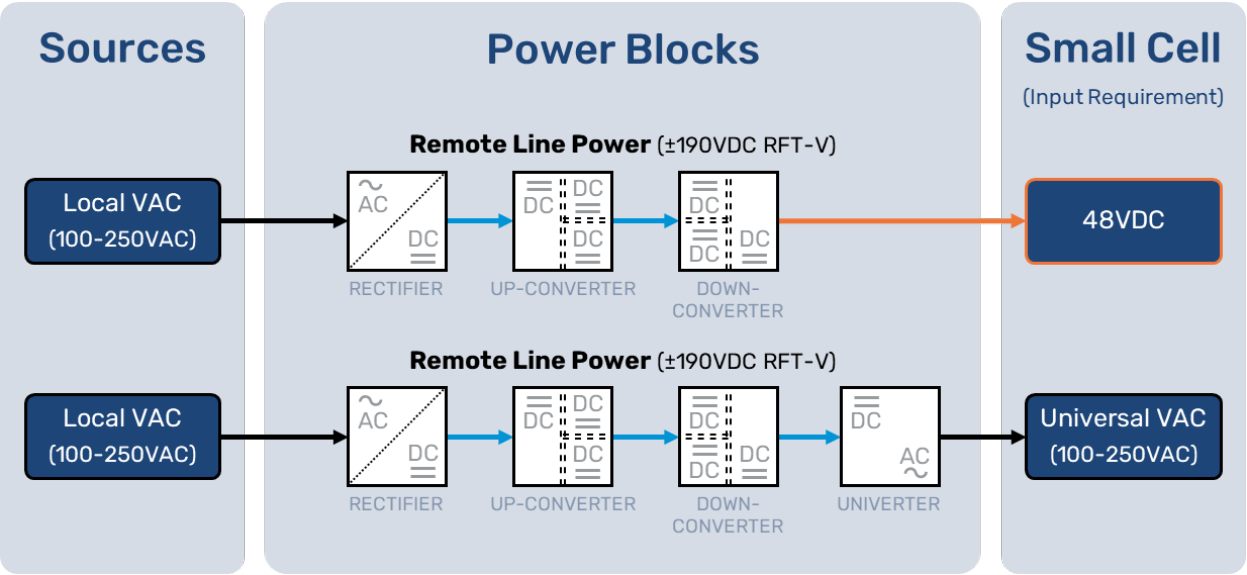


Figure 13 - Remote Line Power for AC or DC Delivery

RLP is a good solution for small cells, especially in green field or locations where other power sources are unavailable. RLP is popular with neutral host providers as it allows them to use existing telephony copper pairs to deliver power in high density areas, eliminating the need for costly construction or meter installations. Advances in this technology supports higher power using fiber and copper pairs, significantly more appealing in cost per node than local AC utility power. The time savings alone can more than justify the use of RLP.

While remote line power is an excellent solution for many use cases there are some limitations and drawbacks. Systems that use existing twisted copper pair may be limited to 100W per circuit. Multiple circuits can be used in parallel but must be aggregated at the small cell site. This aggregation can add more components on the strand for undesirable aesthetics as shown in the figure below.

Figure 14 - Remote Line-Powered Small Cell

7. Conclusion

It's time to bring awareness and identify different solutions to fulfill the objectives around network reliability, wireless connectivity, and corporate management. By highlighting the pain points in deploying small cells we explored three solutions to power small cells: utility power, HFC power and remote line power. We highlighted the benefits of using the HFC infrastructure for low-cost high value connectivity and discussed two methods of deploying small cells over HFC.

HFC certainly provides the cable broadband market with significant advantages in infrastructure availability, not only for small cells, but for any network-enabled product:

- Ubiquitous Broadband & Power Grid
- Accessibility and Scalability
- Reliable source of power
- Physical placement
- High-capacity low-latency backhaul
- Existing well-maintained resource

Hundreds of thousands of outdoor small cells will need to be deployed in the next 5 years to meet 5G demands. Upcoming smart cities, autonomous vehicles and drones will create even more demand for reliable network infrastructure for scaled deployments. Cable broadband operators are well situated to capitalize on these emerging markets.

Abbreviations

AC	alternating current
CBRS	Citizens Broadband Radio Service
CEC	Canadian Electrical Code
CEO	Chief Executive Officer
CSA	Canadian Standards Association
DC	direct current
EMI	electromagnetic interference
FCC	Federal Communications Commission
FTTx	fiber-to-the-home
HFC	hybrid fiber-coax
LLD	low-latency DOCSIS®
MSO	multiple system operator
MVNO	Mobile Virtual Network Operator
NCTA	The Internet & Television Association
NEC	National Electrical Code
NESC	National Electric Safety Code
NRTL	Nationally Recognized Testing Laboratories
RF	radio frequency
RLP	remote line power
SCTE	Society of Cable Telecommunications Engineers
SKU	stock keeping unit
TCO	total cost of ownership
UL	Underwriters Laboratories
UV	ultraviolet

Bibliography & References

[1] SCF 2021 Market Forecast document 050.10.5; http://smallcellforum.org/scf-market-forecast; viewed 7/28/21

[2] Small cells: Strand-mounted business opportunities; BTR; 9/5/2018; https://www.broadbandtechreport.com/docsis/hybrid-fiber-coax/article/16449380/small-cells-strandmounted-business-opportunities

[3] "Cable's wireless biz 'ready for its star turn' – analyst"; LightReading; J Baumgartner; 6/17/21; < https://www.lightreading.com/cable-tech/cables-wireless-biz-ready-for-its-star-turn---analyst-/d/d-id/770301 >

[4] US power outages jumped 73% in 2020 amid extreme weather events; G. Herring; 1/19/2021; https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/us-power-outages-jumped-73-in-2020-amid-extreme-weather-events-62181994

[5] 5G Low Latency Requirements, McLaughlin, 2019, https://broadbandlibrary.com/5g-low-latency-requirements/

[6] The Cable History Timeline; Viewed 7/25/2021; https://www.cablecenter.org/images/files/pdf/CableHistory/CableTimelineFall2015.pdf

[7] NCTA Broadband Stats and Facts; Viewed 8/1/2021; https://www.ncta.com/broadband-facts

[8] SCTE Standard 238 2017; https://www.scte.org/standards-development/library/standards-catalog/scte-238-2017

[9] NCTA Broadband Stats and Facts; Viewed 8/1/2021; https://www.ncta.com/broadband-facts

[10] iGR, 2020, A Strand of Hope: How strand-mounted small cells can address the demand for 4G and 5G mobile data, iGR Media Center, Viewed 3/15/2021 https://igr-inc.com/media-center/white-papers/strand-mounted-small-cells/strand-mounted-small-cells.asp

[11] *"Why Cable's Broadband Network Is Handling the Pandemic and Ready for the Future", NCTA; 4/29/2020; https://www.ncta.com/whats-new/why-cables-broadband-network-is-handling-the-pandemic-and-ready-for-the-future*

[12] L Hardesty, 2019, Altice's 19,000 small cells in Long Island don't help Sprint's network much, say analysts, Fierce Wireless, Viewed 3/1/2021, https://www.fiercewireless.com/wireless/altice-s-19-000-small-cells-long-island-don-t-help-sprint-s-network-much-say-analysts

[13] "Cable's wireless biz 'ready for its star turn' – analyst"; LightReading; J Baumgartner; 6/17/21; https://www.lightreading.com/cable-tech/cables-wireless-biz-ready-for-its-star-turn---analyst-/d/d-id/770301

[14] DOCSIS® Network vs. Fiber Backhaul for Outdoor Small Cells: How Larger Footprint of DOCSIS Networks Lowers TCO in the Outdoor Use Case, https://www.cablelabs.com/docsis-vs-fiber-backhaul-outdoor-small-cells

[15] CableLabs Low Latency DOCSIS® Technology Launches 10G Broadband into a New Era of Rapid Communication; https://

www.cablelabs.com/cablelabs-low-latency-docsis-technology-launches-10g-broadband-into-a-new-era-of-rapid-communication

[16] "Internet Facts and Stats; NCTA; viewed 7/20/2021; <https://www.ncta.com/broadband-facts>

[17] *"Why Cable's Broadband Network Is Handling the Pandemic and Ready for the Future", NCTA; 4/29/2020; < https://www.ncta.com/whats-new/why-cables-broadband-network-is-handling-the-pandemic-and-ready-for-the-future >*

[18] Ibid

[19] Ibid

[20] Image taken from Airspan website; 8/1/2021; https://www.airspan.com/cbrs/

[21] https://www.scte.org/standards-development/library/standards-catalog/ansiscte-124-2011/

[22] https://www.scte.org/standards-development/library/standards-catalog/ansiscte-91-2015/

[23] https://www.ul.com/

[24] https://www.csagroup.org/

[25] SCF 2021 Market Forecast document 050.10.5; http://smallcellforum.org/scf-market-forecast; viewed 7/28/21

[26] Ibid

Software Reliability Engineering

Scaling the Cloud with Automation

A Technical Paper prepared for SCTE by

Brian Gray
Sr. Manager, Software Engineering
Comcast Cable
1800 Arch St., Philadelphia, PA 19103
267.634.5540
Brian_gray@comcast.com

Sriram Ramakrishnan, Principal Architect, Comcast Cable

Fei Wan, Sr. Principal Architect, Comcast Cable

1. Introduction

Operations teams have long functioned under a primary mandate of assuring customers the most solid and uninterrupted experience possible. The recent advent of software reliability engineering (SRE) introduced engineers to the formalized automation of toil, leaving more time for creative problem solving of service interruptions. However, twin issues remain: how to automate a virtualized cloud environment in practice, and how to measure and prioritize repeated tasks to be automated for the greatest impact. In this paper, we offer a case study of the evolution of a complex cloud infrastructure from a state of manual deployment, scaling, failover, and upgrading, to one of push-button control and automated self-management. Driving this evolution is a pair of mathematical tools developed in Comcast's Core Application Platforms (CAP) group that use the financial concepts of "net present value"/NPV and "internal rate of return"/IRR to organize and value automation opportunities simply and objectively.

2. A Look Inside the Elements Deployment Architecture (A Starting Point)

Elements is the internal name of an asynchronous, fast, flexible push notification service which supports both individual and fan-out notifications across such diverse network destinations as outbound web hooks and mobile push notifications. The service is built as a publisher/subscriber (pub-sub) service, in which the senders of the messages, also known as 'Publishers/Producers' are decoupled from the receivers of the messages, called the 'Subscribers/Consumers'.

High Level Architecture

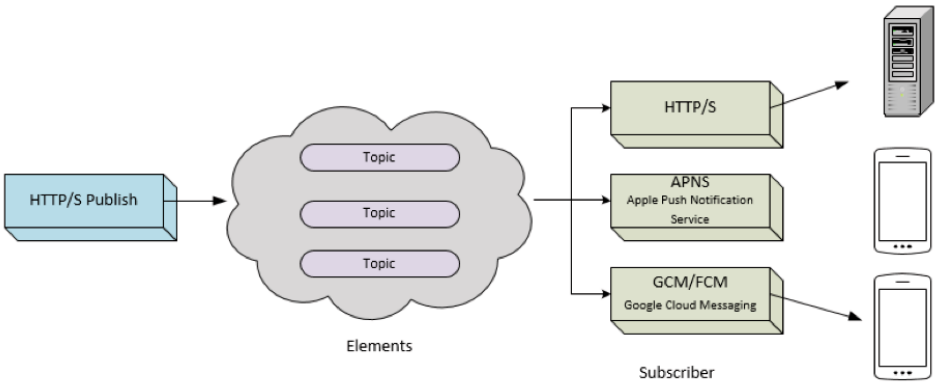


Figure 1 – Elements Architecture Layers

Complicating the equation is the complex requirement to operate across many layers of Comcast's internal network – or the "plant" – to reach destinations such as an X1 set-top box with notifications. Because we need deep access to underlying plant details, we lose out on the option of hosting this particular system in a public cloud like AWS with convenient, high-level facilities to assist our operations. Such a thing would technically be possible in a split-architecture design, using public cloud for pub-sub management, configuration, and global logic, and DirectConnect-ing into a private space that

hosts intelligent routing gateways. That solution, however, introduces many more points of failure, some of which actually lose the ability to report back that a failure happened. To compensate, you would then need **more** complexity to implement a run-time approach on a distributed trace algorithm.

2.1. Challenges

The infrastructure for the solution we maintain therefore runs entirely in our legacy private cloud environment and has many components, including Mesosphere DC/OS for container orchestration, Consul for service discovery, Couchbase for database persistence, and HAProxy for load balancing. Besides these, the system integrates with external Comcast systems for authentication, authorization, secrets management, and observability. As you can see given the complexity of this system, we see immense leverage from following the best practices for SRE. Unfortunately, when we started, we didn't have any such best practices. We were more operations-focused, which is reactive and time intensive, and not following many SRE guidelines.

To give an example, let's suppose we find out that HAProxy CPU usage is ridiculously high and that it's affecting the performance of our application. The extant tools we have available are Ansible – a simple IT automation facility – configured via "playbooks", and the HAProxy configuration schema. One option, then, is to check the config and verify if the value for nbproc is set > 1, where nbproc is an HAProxy configuration key that allows the proxy to spin up multiple processes. To change this config, since it's an emergency, we would create a one-time playbook with this HAProxy config and push the changes to our proxy servers. Once the change has been pushed, we would do a manual rolling restart of our proxy servers.

Obviously, there are multiple things wrong here. This change leads to inconsistent configuration, having no testing or automation. Also, this one-time run causes changes not to be source-controlled in GitHub, team-members might not be aware of this change, and there is no source-of-truth for the configuration. Beyond all these issues, time is critical: time we did spend on manual changes or figuring out who made the change. For all teams, but especially small teams like ours, time is the most valuable resource, and we really need to perform efficiently.

Another instance is a routine scaling of the database processing instances in our Couchbase infrastructure. If we wanted to increase the size of our cluster, we used Terraform (TF) to create the virtual machines (VMs) from the legacy cloud console and performed the configuration for adding this new Couchbase node to the cluster. Obviously, many things are wrong here, as well: No automation, lots of opportunities for failures, and no consistent way to scale the infrastructure.

In summary, these are the challenges we faced, which ultimately affected us as a team, and applications we maintain:

- Lack of automation
- Lack of Testing
- Inconsistent Configuration
- One-Off runs with Ansible
- No continuous integration & continuous deployment (CI/CD)

3. The Elements End State

3.1. Improved Configuration Management

The Elements end state continues to leverage Ansible as the configuration management tool. It's a useful tool that allows the declarative specification of your Infrastructure as Code (IaC), and having Ansible enables us to make changes to the configuration in a more consistent fashion. But we still run into the issue of one-time runs. Teams were making changes to the infrastructure by making one-time runs through Ansible. There are couple of issues with this approach:

- Changes are not shared with the other team members, and others might not know what has changed in the infrastructure.
- There is no single source of truth.

We really wanted to avoid this scenario and the only way to achieve that was to ensure teams follow the guidelines in making Ansible changes:

- All Ansible playbook changes needs to be in source control.
- Changes to playbooks must be done through Pull Requests (PRs) and can be merged only once review is completed by peers.

The obvious downside is changes might take time to be deployed, and reviews might take longer. We tried to reduce review times by pairing once a week (preferably after scrum) and approving the PR's. Remember, this process takes time and discipline among team members. But, in the long term, this helped us to have a predictable and consistent infrastructure that's easier to maintain.

3.2. Infrastructure Testing

Consider our scenario where we need to scale our Couchbase cluster. We don't have any consistent way to test our infrastructure, and we need one to test our Terraform, Packer and Ansible playbooks. Terratest is a Golang library that provides patterns and functions for testing the infrastructure built through Terraform. We baked Couchbase and all other standard tools into an image, which we created using Packer. Terratest can test the build of this image as well. Once the image has been built, Terraform will use those image IDs to build the Couchbase infrastructure. Ansible can be used to configure additional changes. Inspec is used to Test Ansible playbooks. Inspec is an open-source auditing and automated testing framework to describe and test for regulatory concerns, recommendations, or requirements. See Figure 1 – Elements Architecture Layers

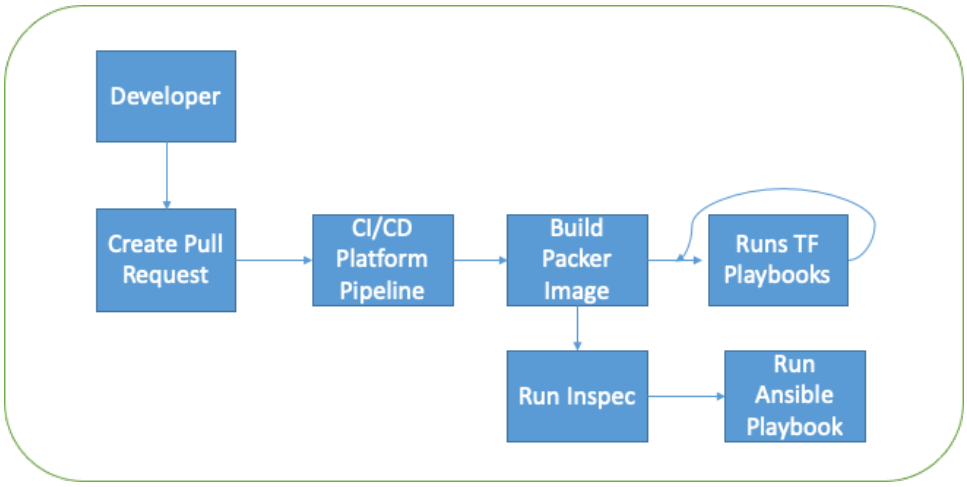


Figure 2 – PR -> Pipeline -> Playbook Process

Developing Terraform tests using Terratest requires Golang skills. This is not an insurmountable barrier, but some learning is needed. Inspec is more human-readable and might be easier for teams to learn.

Software infrastructure testing takes process, organization, and discipline, but will pay-off in long-term.

Just having testing is not enough, however; we needed an automated way to trigger these tests. That's where we decided to follow CI/CD guidelines for our infrastructure playbooks.

3.3. Continuous Integration and Deployment

This is where we really dig into the application of software development best practices to infrastructure code. These follow a couple of simple guidelines:

- All infrastructure code, including Terraform, Packer and Ansible playbooks, is in our source-control GitHub Enterprise repository.
- Any changes to the playbooks go through a PR process and, we have branch protection enabled to avoid direct merging into main. All PRs require at least 1 approving review.

It's good practice to have custom, application-specific templates for creating issues and PRs. For every PR that gets created, we have our CI pipeline that validates the syntactical correctness of the playbook, validates the format (using Terraform fmt for Terraform playbooks and Ansible lint for Ansible playbooks), and runs the test. We run automated tests in an environment that's totally separate from production for Terraform playbooks. We create all the resources (servers, load balancers, machine images) using namespaces with a unique name, to ensure that we don't accidentally overwrite any "production" resources in that environment and accidentally clash with other tests running in parallel. We ensure the tests always clean up after themselves, so we don't leave a series of zombie resources lying around untracked. Most times, we can anticipate the time taken for our tests and will stay within the default limits. But on occasion we might have to increase those timeouts.

Once PR builds are successful and review is complete, PRs can be merged into the main branch. In our case of scaling the Couchbase cluster, it's as simple as modifying the Terraform playbook to increase the number of instances. All changes get propagated or deployed to a lower environment automatically as they merge to main. But we don't do automatic deployment to production environments. We follow a

release-based, gated model to ensure we validate the infrastructure job was successful before we proceed to production.

4. Lessons Learned from the End State and the Path Forward

Configuration management, automated testing, and continuous integration are the key building blocks of a SRE approach to software reliability. The operations tasks required to ensure optimal uptime, performance, scalability, redundancy, and observability can overwhelm any team, but also can often be reduced to logical procedures: the very things where programming shines as a solution!

By leveraging a team's abilities as programmers, they can analyze a process (for example adding nodes to a database cluster), reduce that process to a task in the form of code, and have a pipeline trigger that task either in response to identified and alerted criteria, or else on demand, when a human merges a PR that changes the configuration. Thus do we address the primary culprit costing the team time and resources: toil.

5. SRE, and the Plight Thereof

As an SRE team, we spend a great deal of our capacity on automation of toil (in a rather circular definition, toil can be thought of as work that can be automated). You could say it's this very thing that most distinguishes an SRE team from a DevOps team. We see an opportunity, where we're doing work by hand, over and over, and look for ways to write code to do that work for us. Then we iterate to find the next most insistent itch to scratch.

This kind of ad-hoc, subjective process works well enough and can be great for morale. I mean, who wouldn't want to throw themself into automating away their personal bugaboo? So, let's assume we allow engineers to retain that flexibility. What else can we do to optimize our selection and evaluation of automation opportunities? Turns out we can get some help with tools invented for use in evaluating financial investments. Not only do these tools assist us in decision-making, but they can also act as indicators to report the real leading and lagging value of the SRE team to its parent organization.

6. The Financial Side

You know that thing you're an expert in? And then they make a movie about that thing, and you complain about all the ways they get it wrong? That's what this section will read like to financial experts. The situation here is simply that an engineering team with a need to formalize and quantify the priority of the automation work described above found a useful tool for doing exactly that, so – disclaimer provided – let's just dive in without checking the depth…

6.1. What Is Net Present Value?

Put simply, Net Present Value (NPV) is the amount a potential investment is worth today, including all outgoing costs, all incoming returns, and discounting each number to account for the time value of money.

What is the time value of money then? It's a formalization of the concept that money now is worth more than the same amount of money later. Sure, $100 is literally worth more now than $100 will be in ten years, because of inflation. But in addition to that, $100 in 2021 dollars is worth more now than $100 in 2021 dollars will be in 2031, in part because the ten years in the middle give you more options for what to do with the money.

To calculate the NPV of a proposal, you first need to decide what rate to assign to this time value. This is called the discount rate. When a person wins the lottery and is given the option of taking $X now or $Y/year for 30 years, the "now" value is calculated by applying the discount rate – for Powerball, 4% – to each payment's amounts and summing them to yield an NPV.

How is the discount rate figured? The simplest way is to determine the best low-risk investment available as an alternative to whatever else you may be thinking of doing with the money. For regular people, maybe this is a high-yield money market or bond fund. For a business that tracks metrics, it could be the return on marginal advertising dollars or scaling up workforce to accommodate more service delivery. Something you have control over and could throw money at. For the purposes of this discussion, we're conflating discount rate and "cost of capital" here. For a full treatment of the difference between them, review Cost of Capital vs. Discount Rate: What's the Difference?, but very briefly and inexactly, we can imagine that cost of capital is the real-world stuff above (advertising ROI, etc.), and discount rate is what it's called when it assumes the form of a number used in a mathematical formula.

Anyway, each time money changes hands, the present value of that transaction is calculated as:

$$\frac{R_t}{(1+i)^t}$$

Figure 3 – Present Value of a Single Event

…in which R_t is the cash flow (positive or negative) at the time of the event, i is the assigned discount rate, and t is the amount of time in the future the event takes place. To get the NPV of an entire series of events including the initial expenditure, we simply sum up all discrete events:

$$\mathrm{NPV}(i, N) = \sum_{t=0}^{N} \frac{R_t}{(1+i)^t}$$

Figure 4 – Net Present Value of a Series of Events

This result is a scalar value, in dollars, representing the magnitude of value that the project will net in excess of the discount rate over its lifetime. In other words, how much better or worse this investment is than going the risk-free route.

In the case of the lottery? Look into long-term equity or debt investments and see if you can find anything that beats 4%. It is for this reason that winners are advised to take the lump sum and invest initially in an index fund while figuring out better options. Historically, the average of the market as a whole has generated ~10% returns over the long term, minus inflation. The lump sum is far greater than the true NPV of the long-term payments because the discount rate applied to the winnings is so low.

6.2. What Is Internal Rate of Return?

Internal rate of return (IRR) is a fascinating inversion of NPV, which manages to factor out the discount rate. Thus, you're not concerning yourself with inflation, the risk-free investment alternatives, or cost of capital, only the factors internal to the investment. It does this by taking the NPV formula, setting the answer to 0, and solving for the discount rate:

$$\mathrm{NPV} = \sum_{n=0}^{N} \frac{C_n}{(1+r)^n} = 0$$

Figure 5 – Internal Rate of Return Formula

The mechanics of how to solve this are not at all easy or straightforward, so we'll gloss over them in favor of just using built-in Excel formulae. For a treatment of numerical methods to solve for IRR, start here.

The result is the discount rate at which this investment would yield no NPV. How is this useful? By itself, it is not. That is to say, in order to make use of the discount rate in an NPV calculation, it needs to have a real-world investment to compare against or you're just guessing. On the other hand, when you compare the IRR to other potential investments, it makes for a wonderful way to rank them by pure return percentage internal to the investment itself. So, to start, you can say, "Who cares what my cost of capital is? Investment #5 shows the highest IRR of all my options."

When combined with cost of capital though, you can first prune out those non-starters whose IRRs do not exceed it.

6.3. Which One Is Better?

The attempt to answer this question is effectively an attempt to decrease the number of tools you have at your disposal. It can be comforting to have and know your hammer and treat everything like a nail, but you become more effective when you learn about screwdrivers as well. As for IRR and NPV, they both have their place because they both describe very different viewpoints on your investment options:

- IRR gives you the rate of return but says nothing about the scale of the investment.
- NPV gives you the total value of the investment but gives no perspective on how quickly you recoup your initial layout.

An easy example is a pair of investments, one of which gives you an IRR of 25% over 2 years, returning $50,000 NPV. Another sustains a 20% IRR over 10 years and yields $200,000 NPV. The first option is a higher rate of return and gives control of your original investment back sooner, allowing reinvestment in year 3. The other locks in a $200,000 NPV investment and allows you to be hands off for 10 years. Are you sure you're going to have a >20% IRR option 2 years from now? What would a likely 10-year portfolio starting with the 25% investment look like, and would it compare favorably or unfavorably with a static 20%?

So, it pays to calculate both and use your own judgement to decide which investments to fund, depending on your risk acceptance, your demand for flexibility, and your list of known options.

7. Factoring out Dollars

Now that we've covered the financials, let's throw away the dollar signs. These calculations are great mathematical tools to leverage, but we're not here to talk about money. We're here to talk about time, so what happens when we change all our units to units of time?

I don't know if you noticed, but the IRR calculation placed $0 on the left, canceling out the R_t (expressed in dollars) on the right. The resulting percentage rate is unitless, indicating that we could use anything as the original units: potatoes, rainbows… or time. If you invest 40 hours of work automating a task, so that now there are 4 hours of work per month you don't have to do anymore, the formula works exactly as if you were investing money.

IRR doesn't care about units. Time works identically to money and gives you useful, sortable results.

Bam, done. We can go home now.

7.1. What about NPV?

Now this value does have units, namely dollars. And because it does, we must re-assess the validity of one of our concepts: the time value of money. If we're to switch our units to time, what does it mean to refer to the "time value of time"? Conceptually, we need to resolve the premise that time now is more valuable than time later. Is it?

This could be argued both ways. Let's say you have 2 weeks of vacation, and further that these weeks can rollover year over year until whenever you wish to use them. Maybe there is more value in using them now; after all, you could be dead in a year. But then Universal keeps opening exciting new theme parks. What if in 2 years they add a Jurassic Park Cruise where the room stewards dress as raptors and you have a port of call on Isla Sorna, and you could have gone had you not used the 2 weeks on Orlando? Decisions, decisions.

This 100% sounds like the kind of thing I'd have written a thesis on in school, and they'd have been all, "this is interesting, but it's not economics until it's quantified", but I'm not going to do that here. All that we need to recognize here is that you're going to have to make your own decision and assign your own value to the NPV discount rate (and it's ok for that value to be negative). Just take the time to think through your situation and come up with a value that makes sense for the priorities of your organization.

Ask your executives, and you'll likely hear that they'd rather realize a return on time sooner than later, because executives tend to keep an eye on mitigating risk. A project that completes sooner decreases risk because it represents a shippable, (possibly) revenue generating asset. If it succeeds, those revenues kick in sooner, leading to them literally being worth more due to the time value of money concepts discussed above. If it fails, the company takes the loss but has stopped pouring money into development sooner. Either way, there is a strategic net benefit to software leaders to apply a positive time value of time in calculating the NPV of potential automation efforts.

Once you've decided on a reasonable discount rate, NPV can work just fine using time as its unit of measurement. Exercise care assuming this extends to rainbows.

8. The Evaluation Process

The process is quite straightforward, given a step #0: create a spreadsheet that calculates IRR and NPV based on initial investment, with slots to input expected hours returned at various dates in the next 3

years. Why 3 years? That's arbitrary. You can extend it to 5 or 8 or whatever, but in CAP's technical sphere, rare is the infrastructure script that can be expected to survive more than 3 years of re-architecture. Once you've replaced your EC2 cluster with Fargate, or started managing Kubernetes with GKE, your carefully written scripts to resize, rehydrate, or redeploy are of no use. So, given you have a spreadsheet, the process is:

1. Do a task the first time, and remember you did it
2. Notice that a task is toil, by virtue of having to do it a second time
3. Record how long it takes to do it by hand
4. Mentally break down the process into automatable chunks
5. Estimate how many dev-hours would go into coding the automation, along with how often and when this toil would have to be done in the future
6. Enter #5 into the spreadsheet (as a negative value), and #3 whenever you expect to have toil replaced by automation in the future (accounting for the fact that until the automation is complete, you'll still have to do it by hand)

That's all you need to document the opportunity, which you will do in the "Data" tab as illustrated below in Figure 7. Then whenever you get capacity to do some automating, pull up the spreadsheet, sort by IRR, NPV, and Breakeven time, and compare the top few of each to find a combination you like. Maybe you only have a couple weeks until a larger project hits, so you take a smaller automation task with a high IRR. Maybe you've got a lull and a lot of time to do something big, so you shoot for high NPV.

Let's look at a sample spreadsheet. Here's what tab 1 – "Dashboard" – looks like:

Opportunity Name	IRR	NPV (hrs)	Breakeven	Idx
OpenStack Usage Accounting	-49.40%	(37.57)	Never	1
xCloud Usage Accounting	56.24%	69.35	Jan 2022	2
Big Good Project	35.71%	590.14	May 2022	3
Small Borderline Project	11.23%	0.75	Mar 2023	4
Medium Good Project	23.38%	58.42	Mar 2023	5
Big Bad Project	1.39%	(426.32)	Never	6
Doesn't Even Make Back Investment	-19.40%	(197.74)	Never	7

Figure 6 – Example IRR/NPV Dashboard

These values are calculated based on data entered in a second tab (named "Data") along the guidelines listed above. The first thing to note in the top left corner of tab 2 is the discount rate. We set this to 10% as a first guess as to how much our "time value of time" is, well… valued, and over time and experience we have had no reason to change it. It turned out our instincts about the time value of money for our team were good enough to be useful to our prioritization efforts. Notice also how the larger a project is, the greater the number of empty cells at the start to account for elapsed time spent in implementation:

Discount Rate: 10.00%	Development	Hours of Toil Avoided																			
Opportunity Name	Hours (Month 0)	Year 1												Year 2							
		1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8
OpenStack Usage Accounting	-100		6	6	6	6	6	6	6	6	6	6	6								
xCloud Usage Accounting	-100				6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
Big Good Project	-1600			75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75
Small Borderline Project	-40						8						8						8		
Medium Good Project	-240												120								
Big Bad Project	-2800													120	120	120	120	120	120	120	120
Doesn't Even Make Back Investment	-480				10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

Figure 7 – Example IRR/NPV Data Input Tab

Something to think about here is the difference between "Big Good Project" and "Big Bad Project". The returns (120h/mo vs 75h/mo) basically scale with the investment (2800h vs 1600h), and so if one is good, it stands to reason so is the other. The difference here that sets them apart is calendar time to complete. Big Good Project completes in 2 months, whereas Big Bad Project – at less than twice the number of required hours – takes a year. In real-world terms, what we're looking at here is the difference between a team of specialists, and a team that has cross-trained. Big Good Project is done by a full team made up of engineers all qualified to work this particular effort, and so can break it up and work in parallel. Big Bad Project is executed by a team in which only one or two engineers is qualified for this work, so it takes longer. But in each case, the applicability of the technology is still valid for only 3 years from inception, so the time to develop eats into the return.

To convert the raw data entered in tab 2 into the easily readable and sortable dashboard in tab 1, we need some Excel magic. We mentioned above that Excel offers built-in IRR and NPV formulae, which means we can skip the numerical methods for calculation and do this is a way that is accessible to the average engineer. Take a look at the formula for IRR in the first row of the dashboard:

```
=IF(Data!C4 <> "", XIRR(Data!C4:AM4, Months!$A$1:$A$37), "")
```

A sharp eye will notice that there is a third tab we have not yet mentioned. This 3rd tab, named "Months", is a simple, static list by row of dates representing the first day of each of the next 36 months. This allows us to drive our calculations by date. The XIRR function supplied by Excel takes the data from the "Data" tab, applies the dates from the "Months" tab, and calculates all the discounted values, sums them up, and solves for 0. The final piece to the formula is that "IF" at the beginning that simply allows us to show an empty cell if the data is not found, rather than a confusing internal error message.

The NPV cell is extremely similar in its formula:

```
=IF(Data!C4 <> "", XNPV(Data!$B$1, IF(ISNUMBER(Data!C4:AM4),Data!C4:AM4, 0),
Months!$A$1:$A$37), "")
```

Within the same basic structure, we now use the XNPV function, which requires the discount rate (from Data!$B$1), our effort data, and the months, plus the requisite embedded IFs to account for any possible missing information.

Getting back to the dashboard as a holistic entity, the first two rows (sorted by "Index" order, the order entered in the Data tab) represent a real case documented by the process above. We had to audit our usage of cloud resources and would have had to every month for the foreseeable future. We recorded time during the task, and it came out to 6 hours. A rough estimate of the time needed to automate the task came to 2½ weeks of work for one person.

Another bit of information that must be considered is that cloud teams sometimes migrate from one private cloud platform to another. The old system gets named after the legacy software on which it is based. The new system uses an internally branded codename, in our case "xCloud", and with a phased "go live" schedule that deprecates the old platform as hardware is shifted over.

Now at first glance that seems like a big investment: 2½ weeks to code something that only saves 6 hours a month. Which is exactly why we do this mathematically. This example, when applied to xCloud usage happens to provide the largest return on the sheet, at an IRR of 56.24%! By anyone's standards that would be a great candidate for a project to undertake. On the other hand, if we use the same effort to automate

legacy cloud usage, we see that it's a non-starter due to the limited lifespan of the solution. Because we will only (generously) get a year of use out of the automation attributable to migrating off of the platform, we never make back the time investment.

At other places in the sheet, you can see examples of good projects, projects right on the border with NPV near 0, and a couple that are negative, but that highlight the difference between IRR and NPV. "Big Bad Project" has a positive IRR – though less than the discount rate – but a hugely negative NPV. We'd technically end up saving time on it, but at the cost of missing out on a whole lot of better options along the way. "Doesn't Even Make Back Investment" is a loser from the start with a negative IRR, though overall it loses far less than "Big Bad Project".

9. Example

Taking as an example the case of xCloud Usage Accounting, we can walk through this process:

1. The team's VP has requested an accounting of how much we are spending in operational expenses for the cloud resources we use in the Elements project. We need to break this down into categories for compute, storage, network, etc. The work is manual but assisted by an internal web interface.
2. For the second consecutive month, the same request arrives to collect usage data. The work is done manually again but flagged for automation.
3. We record that the task required 6 hours of work to complete manually.
4. The internal tool has an application programming interface (API) that provides programmatic access to the categories we need to create this monthly report. We determine that we can write a script that queries each category we need information on, merge for each component by category, and assemble the data into a simple report of use at an executive level.
5. The work is estimated to require 100 developer-hours, which we can schedule over the course of 3 months based on our capacity as a team.
6. We open the spreadsheet to the "Data" tab. -100 is entered in the column for "Development Hours" to account for the time to automate. 3 months are skipped to account for elapsed calendar time spent developing, and 6 is entered in each month thereafter to realize returns on our investment from month 4 to the end of year 3.

Flipping back to the "Dashboard" tab, the information has been instantly calculated to yield an IRR of 56.24% and a NPV of 69.35 hours.

10. Conclusion

The basic concept of using data to support business decision making is hardly groundbreaking. We use various graphs and equations in the preparation of business cases, project performance reports, and managerial accounting as just a few examples. The higher levels of corporations are filled with Masters of Business Administration (MBAs) and other people comfortable speaking in finance as a second language, but often this perspective does not filter down to those closer to everyday decision making.

This paper presented a complicated, manually-operated system of interconnected dependencies and walked through a breakdown of the automation it took to render it into a self-managing entity. We also discussed one method of leveraging established financial models, at the individual task level, to yield a simple, plug-and-play tool for prioritizing toil automation.

It could be argued that the most valuable part of all of this is establishing a pattern of discipline to collect and organize the data (steps #1-5 of "The Evaluation Process"). Go ahead and edit your user story template to add pre-filled acceptance criteria with entries for:

1. Elapsed time spent on the task
2. The automation effort estimate

…and you're most of the way there.

Abbreviations

API	Application Programming Interface
CAP	Core Application Platforms
CI/CD	Continuous Integration / Continuous Deployment
CPU	Central Processing Unit
DC/OS	Distributed Cloud Operating System
IaC	Infrastructure as Code
IRR	Internal Rate of Return
MBA	Master of Business Administration
NPV	Net Present Value
PR	Pull Request
Pub-sub	publish/subscribe
SRE	Service Reliability Engineering
TF	Terraform
VM	Virtual Machine

Bibliography & References

Cost of Capital vs. Discount Rate: What's the Difference?, Christina Majaski; https://www.investopedia.com/ask/answers/052715/what-difference-between-cost-capital-and-discount-rate.asp

Internal Rate of Return: Numerical Solution, Wikipedia; https://en.wikipedia.org/wiki/Internal_rate_of_return#Numerical_solution

Solving The Mysteries of the Distributed Access Architecture

A Technical Paper prepared for SCTE by

Matthew Stehman
Comcast
1800 Arch St, Philadelphia, PA
Matthew_Stehman@comcast.com

Ramya Narayanaswamy
Comcast
1800 Arch St, Philadelphia, PA
Ramya_Narayanaswamy@cable.comcast.com

Jude Ferreira
Comcast
1800 Arch St, Philadelphia, PA
Jude_Ferreira@comcast.com

Robert Gaydos
Comcast
1800 Arch St, Philadelphia, PA
Robert_Gaydos@comcast.com

As the benefits of a distributed access architecture (DAA) continue to be seen from real world applications, Comcast is continuing to convert its analog cable modem termination systems (CMTSs) to virtualized CMTSs (vCMTS). The new technology of DAA not only allows for increased performance of the network with the switch to all-digital components, but also increased visibility with more sophisticated real-time telemetry. The DAA framework emits high fidelity telemetry data from many components, from the primary headend to the customer premises equipment (CPE). The scale of our DAA footprint is growing rapidly, and it is no longer feasible for humans to monitor all of the raw telemetry data and identify patterns of interest and issues. This paper introduces a computational framework and analysis methodology for automated monitoring and alerting for events of interest.

With analog CMTSs, telemetry data is acquired via polling MIBs from the CMTS OS, typically hourly and even down to five-minute intervals in some cases. Our vCMTS implementation has a dedicated telemetry core that constantly emits and writes all telemetry data at 15 second intervals to a time series database, so the real-time data can easily be acquired and analyzed by the DAA team. Each vCMTS captures thousands of telemetry streams, comprising over 1 billion samples per day from a single physical server, which houses many vCMTS cores. With this volume of data, we needed a telemetry analysis tool that could make sense of the data in its current form and continue scale up with DAA in the coming years. Comcast is currently developing a tool for this purpose, internally named “Sherlock.”

The name isn’t entirely coincidental. As the title of this paper implies, the very act of distributing an access architecture tends to uncover many infrastructural mysteries that could benefit from a sleuth. Some relate to the huge amount of data that flows in every 15 seconds from the thousands of broadband-foundational components within our physical infrastructure. The upstream signal path in particular is a trove of noise-related anomalies, as one example referenced within this paper illustrates. It represents an excellent network segment to expose to machine learning (ML) – which thrives on large amounts of data.

While the DAA telemetry covers most of the active components in the network, there are external tools and data that can significantly enhance the capabilities of Sherlock. As such, Sherlock interfaces with other systems such as: customer contact, automated support tickets, existing performance metric tools, etc. The core component of Sherlock is its ability to interface with a wide variety of data sources and create a single, time-aligned view of the entire system for analysis.

Once the single, time-aligned view of DAA is created, event identification and alerting can be implemented. Initially, logic-based event tagging is implemented based on common thresholds for events like partial service, plant-based noise as well as system statuses such as DAA cores offline, remote PHY device (RPD) reboots, and CPE connectivity. Once these events are determined, analytics can be performed to evaluate the frequency and severity of events. Using the event statistics, rankings are created to support the DAA team in prioritizing issues to address as well as keep track of persistent issues. The analysis and rankings are performed at different levels of aggregation: RPD, physical server, site and even division and national.

Future research utilizing the core functionality of the tool includes advanced ML techniques to find patterns/events outside of the standard events identified from traditional logic-based checks. This paper introduces several active research areas in ML in the DAA space. An overview of the architecture is introduced, and an example is discussed.

The development of this tool has had a large impact on the successful deployment of DAA in Comcast's footprint. A sample of example findings from Sherlock is discussed as well.

1. Overview of DAA Telemetry

Remote PHY (R-PHY) has taken hold as the technology of choice for deployment of DAA solutions. The concept of a distributed architecture decreases the amount of equipment that traditionally sits in a cable headend and then connects via hybrid fiber/coax (HFC) to neighborhoods and eventually to customer homes. DAA moves the PHY, or physical RF layer, closer to the user by deploying RPD-equipped R-PHY nodes that sit on the access edge of the network. DAA allows for higher speeds to the end-user because it uses digital fiber optics in place of legacy analog optics. Digital fiber links improve signal quality and support higher modulation orders. DAA also offers operational savings related to the cost of headend equipment, power and more, as small hub sites or curbside equipment act as the PHY layer of the network.

vCMTS technology enables us to shift to a DAA by disaggregating the CMTS. It also allows us to move to IP-based connectivity and converge voice and data services with video and other legacy services, with an added benefit of no longer needing to manage and maintain traditional, bulky CMTS gear.

The transition from legacy hardware to a distributed server-based architecture that can run external software applications allows Comcast extreme visibility into the DAA platform. The scalability and openness of DAA means the platform can now support applications such as real time telemetry streaming that would have been too demanding to run on legacy CMTSs. The DAA system is rich in telemetry, where individual components within the network transmit data as frequently as every 15 seconds to indicate the health of system/network. Mining DAA telemetry data to identify and detect issues in the network that could potentially lead to bad customer service is not only challenging but also involves combing through a lot of existing tools and data sets to build a smart access network.

1.1. Problem Statement

As we scale our DAA deployment to thousands of digital nodes and hundreds of vCMTSs and sites in the next few years, we anticipate operational challenges. Among them, monitoring and going through all telemetry data points to determine system health, and correlating the impact of one component in the architecture to the other key components, while identifying impairments proactively, before customers are affected.

Sherlock is a tool designed to address those challenges by looking at all the relevant metrics, creating a time-aligned view at the most granular level, scoring the health of the system, and identifying root cause of issues. It also proactively identifies anomalous patterns that lead to poor system performance. The goal of Sherlock is to analyze and identify patterns of impairments and rank them based on several criteria, that are outlined in Section 4.

1.2. DAA Topology and Telemetry

As mentioned previously, the DAA architecture offers rich telemetry, from the headend all the way to the CPE. A simplified representation of the telemetry coverage across the DAA topology is show in Figure 1. A variety of telemetry metrics are reported across the topology, ranging from system statuses to traffic and even HW/SW versions of the RPDs.

In our topology, the primary headend houses the core servers and provides statuses of the principal and auxiliary cores for global system health. The primary headend comprises multiple PPODs, or physical

point of deployments. A PPOD is a cluster of servers running the necessary services to operate the connected RPDs. The PPOD the connects to a set of DAAS (distributed access architecture switches) that transfer data to and from the RPDs. The primary core or GCPP, for Generic Control Protocol Principal, provides containerized services for automating deployments, managing applications, the initial authentication of the RPDs, and configuring RPD features and video services. The principal core does not provide any services (video or data).

The GCPP core performs the following three primary functions:

- Initial authentication of the RPD.
- Initial configuration of the RPD, including the list of cores to which it connects and the resources that those other cores will configure.
- Configuration of the multicast sources that the RPD uses to populate QAM video (broadcast and narrowcast) channels. The GCPP allows integrating videos on a standardized, single video platform.

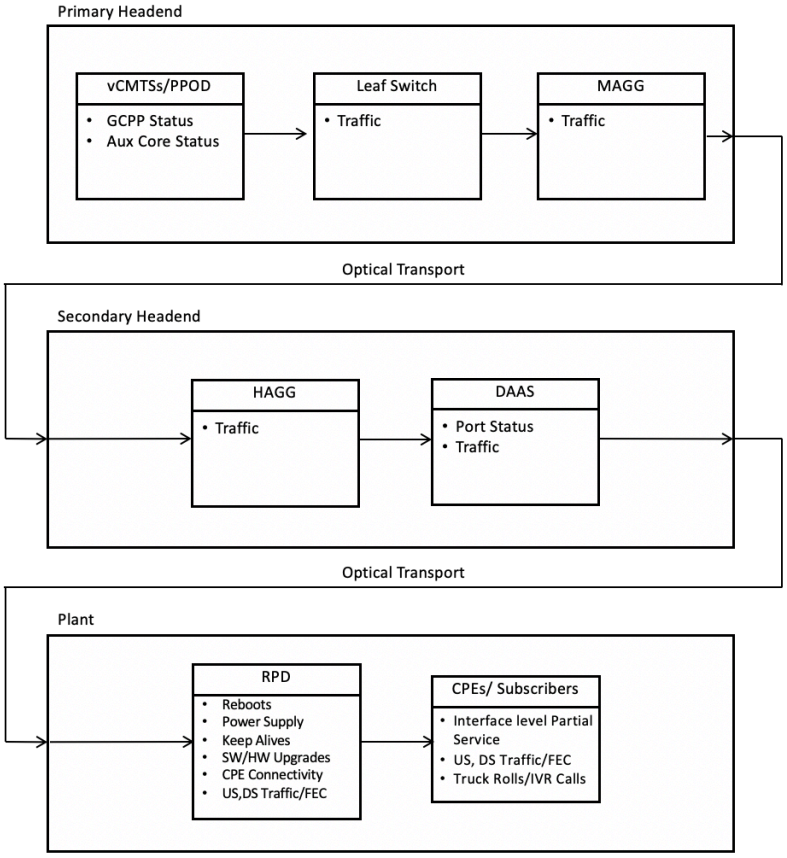


Figure 1 - DAA Topology Metrics Overview

The auxiliary core or GCP (generic control plane) is the second of the two main control planes within DAA architecture: The GCP, which sets up a control plane tunnel over a generic transport protocol such as TCP or UDP. GCP is used to program the R-PHY system upstream and downstream parameters from

the CMTS. It is also used to control the R-PHY system, thus if the AUX core is offline, no data can flow to/from the RPDs.

The secondary headend houses a variety of switches that prepare and breakout the signals prior to sending/receiving data from the RPDs. The main switch type of interest is the DAAS (distributed access architecture switch), which aggregates 10 Gbps connections to remote nodes. The DAAS switches report connection statuses which show the heath of the connection from the fiber port on the switch to the RPD.

In the plant region of the DAA topology, RPD metrics include RPD meta information, interface traffic and even CPE-level information. Having telemetry in different regions of the network topology allows for easier identification of where in the topology an issue may have originated. Table 1 includes more detail of the main metrics that are collected from the DAA network. These metrics were chosen through discussions with DAA experts as well as iterative exploratory data analyses during the initial development phases. These metrics, while not exhaustive, cover the key aspects of system/network health as well as customer experience. The list of metrics is continuing to grow as Sherlock is used in the development and deployment of DAA.

Table 1 - List of Metrics

Metric Type	Metric	Description
Platform Status	GCPP Status	GCPP Status captures the state of the Global Control Primary Plane and indicates if it is operational, offline or initializing
	Aux/GCP Core Status	The Aux Core Provides HSD services and the status indicates if it is online, offline or being configured
Network Status	Keep Alives	Keep Alives track the status of the TCP network connection between the principal cores and the switch interfaces
Hardware	RPD Reboots	Details about RPD reboots such as a reason for reboot, type and recovery time
	DAAS Port Status	DAAS port status captures the status of the DAAS switches located in the secondary headend
Traffic	Device, RPD, Routers-US and DS Traffic	Upstream and Downstream traffic is collected from various components from PPOD to CPE
Device Status	CPE Registration	Registration status captures the CPE devices attempts to pair with the vCMTS that must happened every 30s as dictated by the DOCSIS specifications
	US/DS Bonding Status	Bonding Status for each US/DS interface per CPE device is captured
FEC	Corrected and Uncorrected Codewords	CCW and UCCW are collected at device and interface levels. These can be an indicator of impairments within the plant/faulty modems that need to be proactively identified and addressed
Customer Contact	Truck Rolls, IVR Calls	Truck Rolls and IVR Calls capture customer contacts and are key metrics used within Comcast to measure operational efficiency

Originally, we were focused on finding metrics that correlated with trouble calls/truck rolls, since customer-facing technician appointments were thought to be a good indicator of system health issues.

However, it was found that trouble calls were highly sporadic, because of inherent human aspects, e.g., different tolerances to service interruptions and when calls happen relative to an issue. After exploring the customer-facing aspects, it was determined that Sherlock should focus on the engineering side of the DAA data and let the data drive the insights. Customer contact and technician visits are still evaluated, since those are valid indicators of issues, albeit not as straightforward to identify as data that comes directly from the system data itself.

1.3. Need for Data Aggregation

DAA telemetry has several dimensions, such as frequency of data, type of data (event-based and telemetry-based), the level at which telemetry is captured (device, RPD, PPOD, etc.), and traffic direction (downstream and upstream). This makes it challenging for a tool to thoroughly mine, to provide views at different levels in the network that help identify the health of the system, or identify areas of problem spots in the network.

In addition to the multi-dimensionality, data is transmitted and stored at different locations, which creates the need for a central data repository. As well, standardization of the data elements across different time intervals is needed, so as to have the data accessible for analysis and modeling while minimizing data transfer and storage costs.

To solve the problems stated in Section 2.1 as efficiently as possible, identification of the common components across DAA and outside data logs would enable the aggregation of information in either signal direction and still get the desired visibility of network health and customer experience. Considering the current DAA architecture, aggregating telemetry data points at the RPD level would enable us to focus on a specific RPD and its associated cable modems, or aggregate it to PPOD/site/vendor level. Aggregating at a device level, by contrast, would create millions of rows of data per metric, which would be computationally intensive and would not provide a view that would help operations or the DAA engineering team in understanding network/platform health. Aggregating data at a PPOD level would mask the issues encountered at an RPD level, given the mix of device types, software versions running on the RPD or vendor type.

2. Implementing Sherlock: a Big Data Analysis Architecture for DAA

As discussed, our DAA data is generated from many different components and are stored in a variety of different specialty database systems. The individual systems are customized specifically for the applications. While having compartmentalized data storage solutions for each type of data is simpler from a development and maintenance standpoint, it can make analysis tasks that require several data sources quite cumbersome.

To allow for efficient analysis of all relevant DAA data with minimal manual operations to join, clean and analyze the data, Sherlock was built using a big data analysis architecture. Sherlock has the ability to interface with a variety of existing cloud and on-premise data storage solutions (APIs, SQL databases, Prometheus, AWS), and combine all the relevant data for efficient analysis.

Building a centralized framework to combine all the different data sources, however, is only half the battle. This task is even more challenging considering the growing scale of the DAA data streams. Consider: A typical RPD has thousands of metrics that are stored at 15 second intervals, and each PPOD can link hundreds of RPDs, so, in total, a single PPOD will generate billions of data points a day. Given that we are continuously deploying new vCMTSs, Sherlock needs to be able to scale with the growing DAA footprint and require minimal maintenance. It is easy to see why this operation must be automated,

since it is not feasible to expect a team to manually process data at this scale and frequency. This section discusses the requirements, implementation, and core features of Sherlock.

2.1. Requirements

Sherlock is meant to be used by the DAA deployment and operations teams to actively monitor and address any DAA deployment and operational issues. Thus, the tool must meet the following requirements:

- Full footprint coverage
- Scheduled analysis reports
- Ad hoc analysis abilities
- Fast computation

Those requirements ultimately allow for ML to discover hidden trends and patterns in the data. An overview of the requirements is presented below.

Since the vCMTS deployments are occurring nationwide, Sherlock must be able to analyze data at different levels of aggregation, from a single RPD to a headend and all the way up to the national level. The varying levels of analysis allow experts to not only understand issues with a single RPD but understand if that same issue exists elsewhere and to what extent.

Sherlock should be able to perform automated analyses and generate reports on a schedule so the deployment team can consistently monitor performance. The scheduled runs can be weekly or even daily if needed. The automated runs should produce a concise and consistent output to enable efficient tracking of performance metrics.

Even though scheduled analysis runs are great for consistent summaries of deployment statuses over a known time window, there will inevitably be ad hoc analysis tasks that require a specialized analysis and output. Therefore, Sherlock should also have a manual interface to easily interact with the core data structures and perform a specialized analysis if needed.

Finally, Sherlock must be computationally efficient when performing operations. There is no specific metric for this requirement, but the general motivation is that the computational framework should support the frequency of the scheduled runs in the above requirement. In addition, ad hoc analyses should be able to be completed in a reasonable time frame. That is, if it takes 10 hours to compile the data and prepare an analysis, the tool would not be useful. To allow for efficient interactions and analyses, computational operations should take only a few minutes in general, such that the analyst can stay engaged while working with the data.

2.2. Implementation

Given the requirements listed in Section 3.1, significant effort went into developing Sherlock's implementation, such that all requirements would be met, while still allowing future scaling as DAA continues to grow. The initial stage of development was to become familiar with the data sources, and in this stage, it became very apparent that an advanced solution would be required.

The first implementation attempt was to load the DAA telemetry data with basic Python packages directly via API requests to the DAA time series database. While this was quickest way to start accessing the data and start developing a plan for how the data should be compiled, cleaned, represented, and analyzed, it was not performant enough to meet the design requirements. Using this approach along with standard

Python parallel processing packages, it was taking hours to load a week's worth of data from just a handful of RPDs (representing a very small fraction of Comcast's RPD footprint). It became evident that a more computationally-efficient and robust solution would be required. At this stage, the first implementation was deliberately done at very small scale for exploratory analysis (for determining useful telemetry metrics and experimenting with different processing/visualization methodologies).

Once it was clear what data was important and how it was going to be analyzed, the team designed a production system to meet all the requirements. The diagram of the implementation is shown in Figure 2. The chosen implementation solution utilizes Amazon EC2 (Elastic Compute Cloud) computing resources, which can scale to the meet the needs of a specific task. The source code is written in Apache Spark, which is an open source distributed processing framework specifically for big data. Databricks is used as the resource management system that manages Spark sessions and coordinates the EC2 instances to complete the computation tasks. The raw data is processed and stored in Delta Lakes[1] that are optimized for efficient reading and writing of big data on the distributed Spark framework.

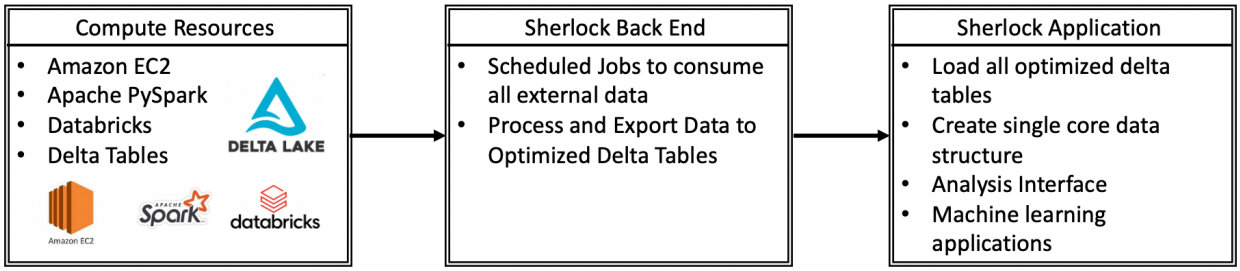


Figure 2 - Sherlock Implementation Diagram

Once the computing and storage frameworks were stood up for use, the data engineering team built the Sherlock backend. It consists of a pipeline of scheduled jobs that consistently read in the raw DAA data from the variety of sources previously discussed, to process them into efficient formats. The resulting data structures are saved to Delta Lakes for efficient access from the main Sherlock application. With the backend responsible for acquiring the data and pre-processing it, the main Sherlock application can then just reference these extremely efficient tables at analysis run time. The main application then uses this centralized data structure to perform a variety of analyses, which are discussed in later sections, as well as to provide the base data set for ML applications. An in-depth discussion about the main features/modules of the Sherlock application can be found in Section 3.3.

Once the end-to-end architecture was developed, the performance against two of the main requirements (full footprint coverage and fast computation) were evaluated. An analysis pipeline, including data loading, processing, event detection and plotting, was run for the entire DAA footprint (representing thousands of RPDs) for a one-week duration. This pipeline finished in less than one hour – the same amount of time it took to merely load 20 RPDs worth with standard Python processes. Similar analyses on subsets of the footprint take less than 20 minutes, and in the future the aim would be to build specialized pseudo-real-time operations that can be performed much quicker (such as real-time event detection with ML). These performance metrics surpass the design requirements discussed earlier. Scalability with the DAA framework also doesn't seem to be a concern at this point, either, due to the scalable and distributed computing framework provided by Databricks and Spark.

[1] Delta Lake "…is an open-source project that enables building what is called a Lakehouse architecture on top of existing storage systems such as S3, ADLS, GCS, and HDFS." For more information, see https://delta.io/

2.3. Features

Sherlock has five modules that handle different aspects of the analysis functionality: core, visualization, event classification, ranking and ML. The breakdown of module responsibilities is shown in Figure 3.

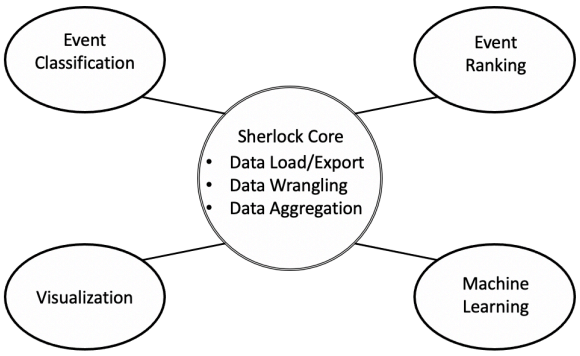


Figure 3 - Diagram of Sherlock Modules

The core module handles most of the data integration, such as loading the data from all the supported sources and converting to a time aligned view for each RPD. The core data structure in Sherlock is a Spark DataFrame with rows for each five-minute timestamp, for every RPD of interest, and columns for all the available metrics. Having the lowest level of aggregation at the RPD level was chosen since this allows for easy aggregation up in higher levels like sites, divisions, vendors, etc., while still being able to meaningfully aggregate CPE-level metrics (for example, upstream transmit and receive power). This core data structure is the foundation for the other modules.

Once the single time aligned view is created, the event classification module identifies events of interest. This module has a fully configurable pipeline that runs through a variety of event detection algorithms and combines the results into a summary table with the event type, start/end times and any other useful metadata about the event. The specifics of the event classification pipeline and logic are discussed in detail in Section 4.1. At this stage, any events identified for individual RPDs are available for use by other modules to support in-depth analyses.

The ranking module uses the events identified from the event classification module to rank the RPDs/sites, from worst to best, and prioritize any issues on the network. The ranking is performed at the RPD level and can thus be aggregated up to other levels as desired. Weights are assigned to each event based on several factors, and the final RPD ranking is the weighted sum over the events for each RPD. The ranking algorithm is discussed in further detail in Section 4.2.

The RPD/site-level rankings are then used to filter down the raw data to regions where there are a lot of interesting characteristics requiring investigation. The visualization module is then used to create charts that highlight the specific areas where the issues occurred. Currently, Sherlock generates plot files on request, given that the project is still in development at this writing (summer 2021). However, once the concept views are finalized, a dashboarding solution with all the views will be stood up to allow for easier access to plots. The two main plot types are RPD timeline and site timeline. An example RPD timeline plot is shown in Figure 6.

The RPD timeline views combine a wide variety of data. Although these plots are currently stand-alone files, they are fully interactive HTML plots. Even though Sherlock is in the development stage and production dashboards are not implemented yet, users of the output significantly benefit from the ability

to zoom/pan on events of interest. Having this ability allows for more productive interactions with the visualizations and serves as a proof of concept (PoC) for production visualization, to gather feedback for the production views. Figure 6a contains the AUX and GCPP core statuses. Figure 6b shows the battery level for the backup power supplies feeding the RPD. Figure 6c displays the CPE status information including counts of devices online/offline, and in different partial service states across the RPD. Figure 6d shows US/DS traffic in the form of total octets transferred, as well as US forward error correction (FEC) percentages (including unerrored codewords [UECWs], corrected codewords [CCWs] and uncorrected codewords [UCCWs]) across all interfaces on the RPD. Figure 6e contains all event-based information, including customer calls, automated alerts, technician repair tickets, and RPD SW/HW updates. Figure 6f contains the DAAS switch status between the RPD and vCMTS core.

Sherlock is the first tool in Comcast to make all this data readily available and digestible in a concise visualization. It is easy to see how powerful this timeline view is, as it allows clear visibility into events across the entire architecture, from vCMTS statuses in the headend all the way to customer experience and contact.

While the RPD view is very useful for deep-diving into specific events affecting areas of the DAA network, it doesn't easily allow for accessing the scale of the event. For example, power outages would likely affect multiple RPDs at time, whereas other issues, like noise ingress, are likely to be very localized. To help visualize and assess the scale of the events across the footprint for a given time window, Sherlock produces event heatmap plots, as shown in Figure 4. This view aggregates the individual RPD levels to the site level. The time dimension is hourly and the heatmap shows the count of RPDs that experienced a specific event in each hour. This type of view allows for a very quick review and determination of how widespread specific issues are across a given aggregation level.

The final module in Sherlock is the ML module, which can utilize the results of the other modules. The ML portion of Sherlock is talked about in detail in Section 6. The main goal of this module is to use all the core data generated through Sherlock's operations as training data for ML algorithms. Given the obvious richness in the DAA data set, there is a clear benefit to applying novel ML applications to mine the dataset and uncover complex patterns. The immediate use cases are:

1. Finding similar events via clustering.
2. Using pattern recognition to discover complex patterns and relationships across the vast dimensions of the data set.
3. Prediction of future issues based on current data.

These applications, if successful, have the potential to completely transform the DAA space. More discussion on the ML aspect of Sherlock is presented in Section 6.

3. DAA Event Classification and Rankings

As part of building a reliable and robust access network to deliver fast speeds to customers, we need to ensure our plant, network and platform health are constantly monitored to proactively detect and mitigate issues and reduce impact to customers.

There are cable industry standards and specifications which are widely used within Comcast to characterize the health of the HFC plant, CMTS and the connection to cable modems. Some of the more common metrics that are tracked are signal-to-noise ratio (SNR), modulation error ratio (MER), transmit power, receive power, and FEC, to and from CMTS. Each of these metrics has acceptable ranges. When telemetry data point goes above or below those ranges, the variance and the duration could indicate

different issues within the network. Comcast tooling currently classifies and captures those anomalies and alerts are sent out. Sherlock leverages those existing alerts to measure DAA plant health.

As stated previously, it becomes an operational challenge to mine through all the data using Grafana or existing dashboards, to pinpoint where and when things go wrong. Thus, Sherlock implements a scalable event detection pipeline to automatically detect events of interest. Sherlock can then leverage the detection of events to rank sites, PPODs and RPDs, based on the occurrence of said events, and help prioritize teams to address issues.

3.1. Event Classification

Sherlock makes use of the centralized core data structure with the telemetry metrics discussed in Section 2.2 to implement an automated event detection pipeline at several levels of aggregation. The lowest level of aggregation is currently the RPD level, while events can also be detected at PPOD level and above. The individual event criteria are specified as objects in the pipeline, and then the pipeline executes each object on the raw data to detect events.

Currently, the events are logical/threshold-based, because that is a great starting point to easily identify any known types of events. The types of events include offline status, anomalous trends in each metric, RPD reboots, SW/HW upgrades and even customer contact. While a single event only looks at specific metrics, the pipeline groups multiple events into a single event to infer when a more complex event is happening. This is especially useful in cases of known maintenance, such has RPD SW/HW upgrades, since these events will undoubtably cause outages and anomalous traffic patterns. This ability allows us to connect any maintenance/upgrade events to corresponding outages so they can be scored differently than unplanned outages.

The event pipeline is typically run on a weekly basis, which allows for analytics and tracking to be performed to document the type and frequency of events across the footprint. The event findings are summarized and automatically distributed to the DAA team to evaluate.

A brief description of the currently implemented events for the RPD level are presented in Table 2. As we continue to add new telemetry metrics and develop the existing list of events, this area will be constantly fine-tuned.

Once these events are identified in the RPD data, they can be aggregated up to higher levels to determine whether events are local, or more widespread. As previously mentioned, Figure 4 displays a visualization known as an event heatmap, which shows the number of RPDs in each site exhibiting a given event at a given time throughout a week. In this example, most events are a spread across a few RPDs, however, there are pockets of “no event found” flags that occur daily at the same time. These specific events were determined to be nominal nightly CPE reboots where the DAA system is healthy, but most CPE devices are offline performing scheduled updates. These types of dense views provide an extremely useful view of the network at a glance and easily display any major issues on the network that would require further investigation.

Table 2 - Description of event classifications

Event Name	Event Description
auxOfflineEvent	Aux core is offline
DAASOfflineEvent	DAAS port is offline
datagapEvent	Telemetry drop outs
gcppOfflineEvent	GCPP core is offline
keepAlivesEvent	RPD keep alive counter is non zero
noEventFound	No other event is flagged
rpdBatteryDrainEvent	RPD is on backup battery power
rpdHealthyEvent	RPD is online with expected CPE connectivity
rpdIopEvent	Automated ticket assigned to RPD
rpdIvrEvent	Customer support call
rpdMacUpdateEvent	RPD hardware was changed
rpdRebootEvent	RPD rebooted
rpdSWUpdateEvent	RPD software was changed
rpdTcEvent	Technician dispatched
usFecEvent	US FEC UCCW exceeded threshold
usOctetsEvent	US RPD traffic out of family
usPartialServiceEvent	Number of CPEs in US partial service above threshold
usPerfEvent	OpTek US System Alert

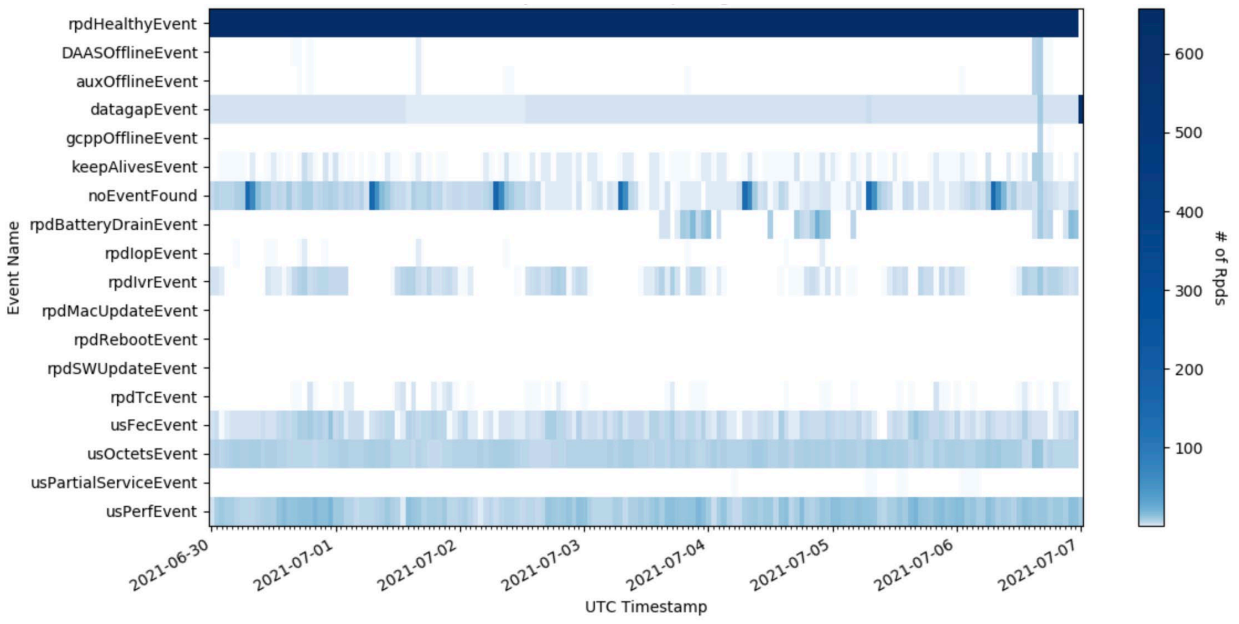


Figure 4 - Event Heatmap for a Single Site Over a Week

3.2. Ranking Methodology

Based on the events mentioned in the previous section, Sherlock ranks sites, PPODs or RPDs with events that are impacting customers for any given duration the most, and assigns them the highest rank. Ranking is aimed to identify sites, PPODs or RPDs with widespread events or events that occur frequently, causing service disruption to customers.

Individual events are assigned weights based on severity, duration, and customer impact. The final ranking is thus the weighted sum of the event coefficients to determine the most impaired sites/PPODs/RPDs. There are some nuances that go into ranking, where events that occur during a nightly maintenance window are ranked lower than events during non-maintenance windows for the same duration and frequency. Additionally, most of the events that occur immediately after RPD HW/SW changes are assigned a lower rank or aren't counted, as the entire system resets to clear current configurations and make the assigned changes. While customer contact events are not directly used in the weighted sum, because of the reasons discussed in Section 2.2, they can be used to break ties when two items have the same score.

Figure 5 shows the workflow that takes the telemetry data through event classification and ultimately to the final rankings. Once the rankings are complete, they are sent to the DAA team for evaluation. Section 4.3 discusses how the rankings are used.

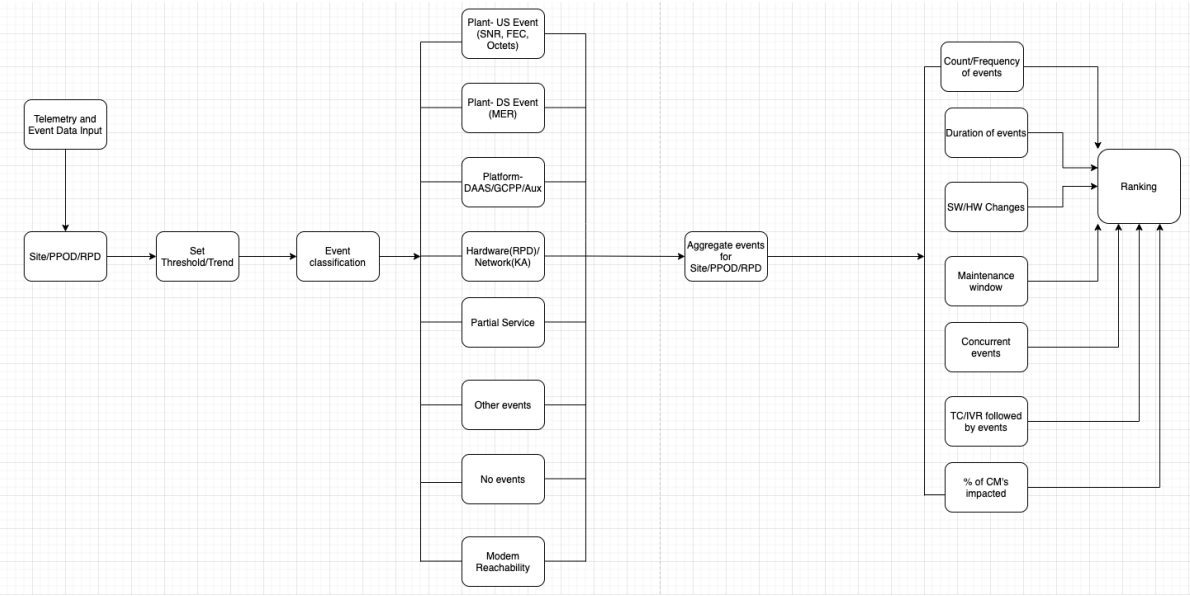


Figure 5 - Sherlock Workflow for Event Classification and Ranking

3.3. Ranking Usage

Once Sherlock generates weekly rankings for the entire footprint, or ad hoc rankings based on business needs, reports and views are stored within AWS. Reports capture highest ranked sites and RPDs based on events listed in Section 4.1, and list all the different events that were captured for that time period. This report serves as a starting point for engineering and operations teams to pinpoint any potential issues within the vCMTS architecture.

Using the ranking report, the heat map views at the site and PPOD are used to quickly see if the issue is widespread or isolated, and if there are specific events that occurred more frequently than the other events. The next layer analyzes the time series view at the RPD level and compares all the metrics that feed into Sherlock to ascertain end-to-end system health. RPD-level views also provide an events summary, with event duration and type, with the functionality to zoom into those events. Each event is assigned a unique ID, which helps in further analysis. This way we can quickly identify events at various levels in the network, and interdependencies between events as well as impact on customers. Ranking and views generated by Sherlock also help in flagging changes made to the system, such as hardware or software upgrades that caused a specific set of events and customer impairments.

Sherlock reports provide week-over-week trending, which helps in highlighting chronic vs. transient issues. Since homes-passed-per-RPD are relatively small in DAA, compared to analog nodes, some of our existing tools can prioritize the number of customers affected by events that would otherwise under-rank DAA issues. As such, Sherlock reports are specifically designed to better understand the DAA network and highlight problem spots in digital nodes, regardless of the number of homes passed.

Sherlock has an integration point with our internal messaging service where reports and visualizations are posted. Plans are underway to migrated to a web-based UI to provide enhanced analysis functionalities to DAA teams.

4. Practical Use Cases and Example Findings

In this section we illustrate the power of Sherlock with examples of specific events and views that highlight the health of the system.

4.1. Noise/Ingress

Table 3 and Figure 6 indicate an upstream noise event which is seen in the form of elevated UCCW. During the upstream noise event, plant and hardware elements (Figure 6a, b, c, e and f) appear to be functioning normally, whereas Figure 6d shows a drop in CCW and a corresponding increase in UCCW. In a DOCSIS plant, transient noise is a normal upstream event and the impact to customer service is minimal. Several tools already exist to sends alert on such events. Table 3 shows a sample of events determined by Sherlock during a portion of this noise ingress example. Most notably, Sherlock identifies usFecEvents as well as usPerfEvents (OpTek) events, indicating that our existing upstream performance monitoring tools are catching these events as well. The exact details of the OpTek events could be overlayed to determine more specific information, such as number of interfaces affected, etc.

Table 3 - Sample Event List for a Single RPD During US Noise Event

Event Type	Event Start Time (UTC)	Event End Time (UTC)
usFecEvent	6/25/21 9:35	6/25/21 9:35
usPerfEvent	6/25/21 16:35	6/25/21 22:35
usPerfEvent	6/25/21 22:50	6/26/21 2:15
usFecEvent	6/25/21 23:50	6/25/21 23:50
usFecEvent	6/26/21 1:05	6/26/21 2:00
usPerfEvent	6/26/21 2:35	6/26/21 3:10
usPerfEvent	6/26/21 4:40	6/27/21 0:35

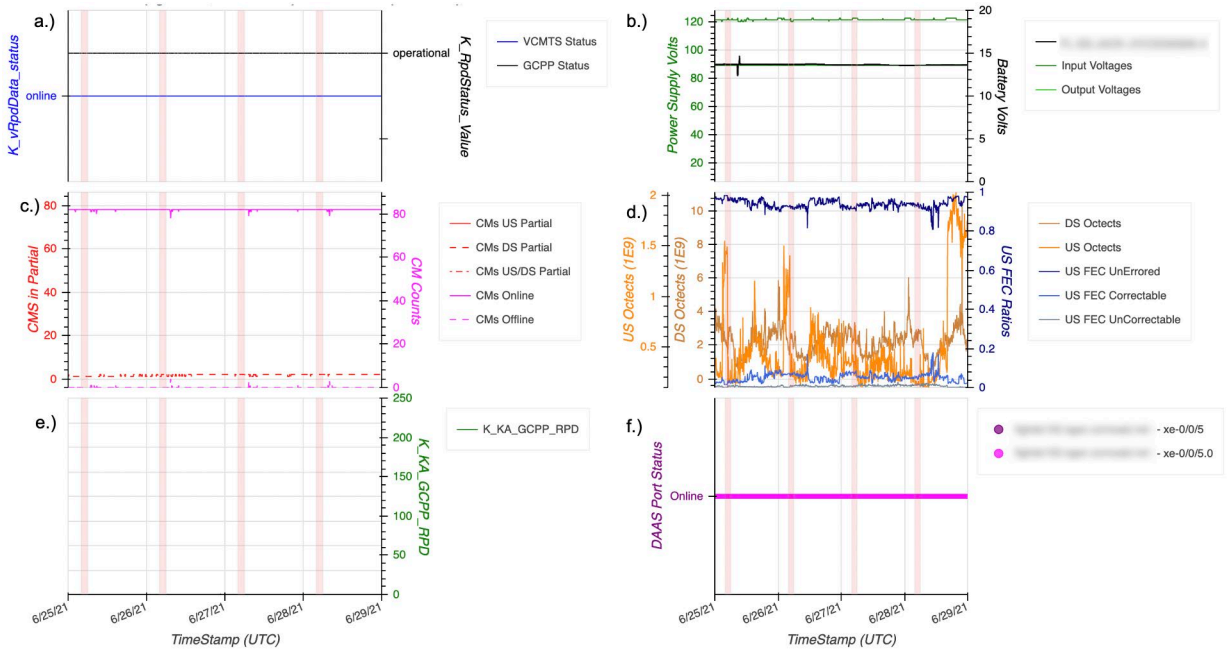


Figure 6 - Time Series View of RPD Noise/Ingress Event: a.) vCMTS/GCPP Statuses, b.) RPD Power Supply, c.) CPE Counts, d.) Traffic and US FEC e.) System Events (Customer Calls, Automated Alerts, RPD SW/HW Changes and f.) RPD Switch Status

4.2. SW/HW Upgrades

Table 4 and Figure 7 show a planned software update. This event is marked in Figure 7e with the initiation of the event as the RPD reboot/reset. Then corresponding dynamics across the other telemetry metrics are shown in Figure 7a, c, d and f. When the RPD software is updated, the AUX core goes offline, the CMs go offline (with some partial service along the way), traffic dips below nominal levels and the DAAS port also goes offline. Having the context of a software upgrade is key, since under other circumstances these types of system responses would be not ideal. However, software upgrades are performed during maintenance windows to minimize customer impact.

A sample of the event list is given in Table 4 showing a subset of the events identified in this time frame. It is worth noting that since the RPD reports its SW version at any given time, software upgrades like these can be determined directly from the data without relying on a ticketing system or external dependencies.

Table 4 - Sample Event List for a Single RPD During SW Update

Event Type	Event Start Time (UTC)	Event End Time (UTC)
auxOfflineEvent	6/17/21 14:50	6/17/21 14:50
usOctetsEvent	6/17/21 14:50	6/17/21 14:50
rpdRebootEvent	6/17/21 14:50	6/17/21 15:45
rpdIopEvent	6/17/21 14:50	6/17/21 14:50
keepAlivesEvent	6/17/21 14:50	6/17/21 14:50
DAASOfflineEvent	6/17/21 14:50	6/17/21 14:50
datagapEvent	6/17/21 14:50	6/17/21 14:50
rpdSWUpdateEvent	6/17/21 14:55	6/17/21 14:55

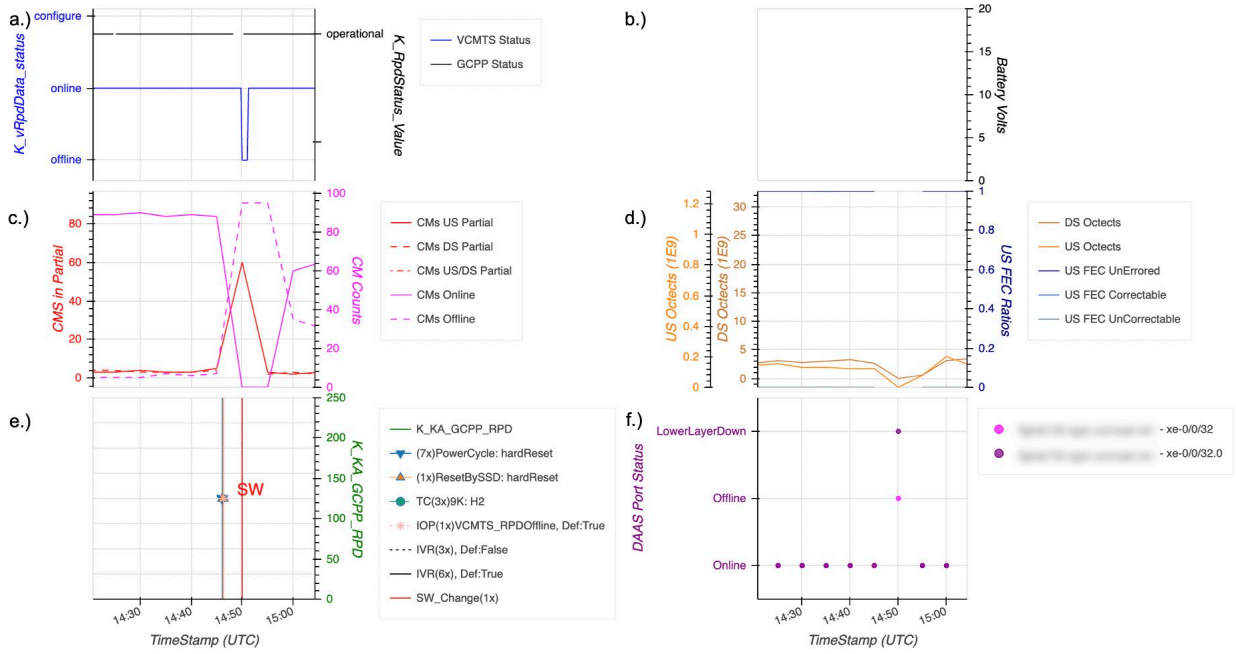


Figure 7 - Time Series View of RPD HW/SW Event: a.) vCMTS/GCPP Statuses, b.) RPD Power Supply, c.) CPE Counts, d.) Traffic and US FEC e.) System Events (Customer Calls, Automated Alerts, RPD SW/HW Changes and f.) RPD Switch Status

5. Machine Learning Applications

As discussed earlier, the data aggregations and processing done by Sherlock represent an ideal setup for ML applications. The DAA data spans many dimensions, and ML/data mining techniques should be used to extract as much useful information as possible to optimize deployments and, ultimately, customer experiences. While the Sherlock ML module is nascent, three immediate use cases are currently being explored.

Clustering: The first application of ML is clustering. Clustering attempts to find smaller groups of similar samples in the raw data. This is especially useful when trying to determine if certain populations of data are more impacted than others, as well as understanding if there are subgroups of data inside major groups. An example of how this could be applied in the DAA space would be taking an RPD event classified via the methods in Section 4.1 and attempting to see if there are sub-populations of RPDs that experienced a given event for different reasons to help triage the event. Specifically, if partial service events are identified, clustering would be able to determine if some RPDs have CPEs in partial service mode because of noise ingress, platform-related issues, configuration issues, scheduled maintenance or even CPE-specific issues. This information could then be used to address the root cause of the individualized partial service issues. A simplified example of this is shown in Section 6.1.

Pattern Recognition: The next application of ML is advanced pattern recognition. As discussed earlier and as per industry standards, events on the network are typically identified via logic-based threshold exceedances, where an event is identified when a certain metric exceeds a predetermined value. While this is useful is many cases, it is limited when it comes to multi-dimensional events with complex relations, because completing a comprehensive detection algorithm with nested if/then logic becomes very cumbersome and hard to maintain. This is where ML shines: If example patterns in the data can be labeled by experts, models could be trained to find the important relationships across many different metrics to identify more complicated patterns than traditional logic-based approaches. An example use case of this in DAA could be identifying RPD backup battery degradation by looking at current and voltage drain during power outages. This application is in development.

Prediction: The third initial application of ML in DAA is the prediction of future issues given real-time data. At Comcast, customer experience is paramount and the ability to forecast and address issues before customers are aware of them is groundbreaking. Given the expansive coverage and real time nature of DAA telemetry, it is possible to use ML methods to find leading indicators of customer impacting events that are classified by the methods discussed earlier. An example of this for DAA could be forecasting when core server load will be too high, to the point of potentially shutting down.
This can be proactively addressed to obviate an outage. This application is also in development.

5.1. Clustering Example

This section presents a real-world use case for clustering DAA data. The example used here is trying to identify clusters of issues that cause partial service events at the RPD level. Using the Sherlock event classification module, partial service events were identified across all RPDs, where an event is classified as 25% or more of CPEs are in US partial service mode for at least 15 consecutive minutes.

Since the Sherlock data is very high dimensional and time-based, the first action is to perform a dimensionality reduction, to help the model focus on important features of the data. Several methods are possible here: traditional feature extraction/engineering, principal component analysis and even auto-encoding neural networks. This compresses the data into a smaller feature domain for the model to learn patterns. The architecture for clustering is shown in Figure 8.

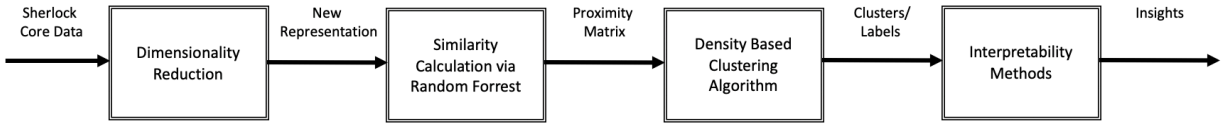


Figure 8 - Clustering Architecture

Clustering algorithms require a distance/proximity matrix that contains the distance between every pair of points in the sample data. This is often an area that requires a significant amount of tuning, because the choice of distance metric used has a huge impact on results. In most cases, Euclidean distance is used. However, in cases where the data set contains mixed continuous and categorical data, Euclidean distance hardly makes sense and typically custom distance metrics are derived for the specific problem at hand. Developing custom distance metrics can be extremely time consuming. For this reason, Sherlock is using a relatively novel distance calculation that relies on a type of ML method called "random forests," which can handle continuous and categorical data simultaneously with minimal pre-processing. The random forest is run on the data with training mode off, which essentially splits the data based on inherent similarity (typically entropy or Gini index.) The result can be turned into a proximity matrix as the number of times pairs of samples ended up in the same leaf node across all the trees in the forest.

The proximity matrix is then passed to a clustering algorithm to attempt to find clusters. Since the underlying structure of the data is not known a priori, representation-based clustering algorithms such as K Means are likely not a good fit. Instead, density-based methods such as density-based spatial clustering of applications with noise (DBSCAN) are utilized, since they make no assumptions about the shape/structure of the clusters. DBSCAN is also a good choice, since it does not require the desired number of clusters to be specified and instead attempts to identify the ideal number of clusters as well as any outliers. Once the clusters are identified, interpretability techniques should be employed to identify what clusters represent in the real world.

Once the cluster labels for each sample are determined, the data can then be passed to an interpretable ML classification algorithm. Essentially, the raw data and the corresponding cluster labels are used to train a classification model. In this case, the model is a single decision tree, to determine the path a sample takes to its classification target. Once the model is trained, it can be investigated to understand if the cluster labels have any real-world meaning.

In the example of US partial service clustering, the event pipeline discussed in Section 4.1 was run on the full footprint of DAA for three weeks, during which 897 partial service events were identified. The features for the clustering algorithm are a wide data table with Boolean flags for other events that occurred in proximity to the USPartialServiceEvents. In this example, the only other events considered were usFecEvents and rpdRebootEvents (to simplify the analysis); all events would be considered in a full analysis. The clustering algorithm was able to identify four clusters. Those four clusters were then passed as labels in addition to data as training samples to a decision tree. The resulting decision tree is shown in Figure 9.

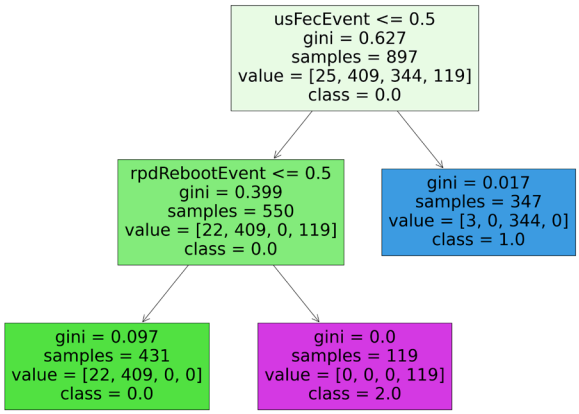


Figure 9 - Interpreting Cluster Results via a Decision Tree

From Figure 9, one can see the learned set of rules the tree uses to assign the cluster labels to a given US partial service sample. A tabular summary of the tree paths for each class is given in Table 5. From here it is easy to see that the decision tree can learn the cluster meanings relatively well. It is worth mentioning that the cluster label -1 is provided from the DBSCAN algorithm as outlier points that don't fit in to particular clusters, thus the decision tree has no path to correctly label those samples.

Table 5 - Decision Tree Paths for Clusters

Cluster Label	Number of Samples	Decision Tree Path	Decision Tree Label Accuracy
0	409	(usFecEvent <= 0.5) and (rpdRebootEvent <= 0.5)	95%
1	344	(usFecEvent > 0.5)	99%
2	119	(usFecEvent <= 0.5) and (rpdRebootEvent > 0.5)	100%
-1	25	N/A	N/A

The results from the above clustering example can be used to identify why RPDs experience widespread partial service events and lead to further mitigation-related enhancements to try in the future. In this case, partial service seems to be driven by usFecEvents and rpdRebootEvents. Further work could be done to understand the samples that had no usFecEvent and no rpdRebootEvent. While the results from this analysis are not too surprising, the architecture is a springboard for correlating different types of events and trying to identify where to dig deeper in understanding non-trivial issues. It is easy to see how this type of analysis could be expanding to more complex issues like understanding sporadic auxOfflineEvents, provided the correct data was fed to the ML architecture.

6. Conclusion

When Comcast began deploying DAA, the need for an automated big data analysis framework was immediately apparent. The DAA framework enables extremely rich telemetry with high frequency sampling rates, making two things true: 1) manual data analysis was infeasible, and 2) exposing the large amounts of data that is perfect for ML-based analysis. Our solution, internally called Sherlock, combines high fidelity data from a variety of sources across our physical infrastructure into a single centralized data structure that can be easily accessed for an assortment of analyses. Creating a centralized data structure with relevant DAA data proved to be instrumental in providing actionable insights from Sherlock analyses.

Sherlock was implemented using state-of-the-art technology that will allow for future scaling as we continue to grow our DAA footprint. Sherlock's core data structure allows for a multitude of analysis implications including event detection, event ranking, visualization, as well as ML advancements. These features allow us to identify and prioritize system issues at a glance, whereas such analyses were previously much more involved and required many manual operations. These analyses are currently being used by our internal teams to monitor DAA deployments and overall system stability.

While Sherlock is a relatively new tool, it is already starting to expand with applications to enhance the power of the insights provided to the DAA teams. As part of this work, we are exploring ML applications to find important trends in the complex DAA data. The immediate ML applications include clustering, pattern recognition and future event prediction. We are also working to integrate Sherlock into our expanding network topology graph, which will also open up new possibilities for advanced insights on the DAA network. The goal was and is to build a platform that can identify issues and recommend preventive maintenance before customers are impacted. Sherlock has proven to be extremely powerful in

its present state and is expected to become even more useful as the next generations of features are developed.

Abbreviations

API	application programming interface
AUX	auxiliary
AWS	Amazon Web Services
CCW	corrected codewords
CM	cable modem
CMTS	cable modem termination system
CPE	customer premise equipment
DAA	distributed access architecture
DAAS	distributed access architecture switch
DBSCAN	density-based spatial clustering of applications with noise
DOCSIS	Data-Over-Cable Service Specifications
DS	downstream
EC2	[Amazon] Elastic Compute Cloud
FEC	forward error correction
GCP	generic control plane
GCPP	Generic Control Protocol Principal
HAGG	headend aggregation switch
HFC	hybrid fiber/coax
HTML	hypertext markup language
HW	hardware
ID	1) identification; 2) identifier
IP	Internet Protocol
MER	modulation error ratio
MIB	management information base
ML	machine learning
OS	operating system
PHY	physical layer
PoC	proof of concept
PPOD	physical point of deployment
QAM	quadrature amplitude modulation
RPD	remote PHY device
R-PHY	remote PHY
SCTE	Society of Cable Telecommunications Engineers
SNR	signal-to-noise ratio
SQL	structured query language
SW	software
TCP	Transmission Control Protocol
UCCW	uncorrected codewords
UDP	User Datagram Protocol
UECW	unerrored codewords
UI	user interface
US	upstream
vCMTS	virtualized cable modem termination system

Bibliography & References

https://www.cisco.com/c/en/us/td/docs/cable/remote-phy-devices/rpdsw51/b_rphy_system_startup_config_5_x/gcpp_support_for_remote_phy.pdf

https://www.cisco.com/c/en/us/td/docs/cable/remote-phy-devices/configuration/guide/b_rphy_management_8_x/rpd_reset_8x.pdf

http://mibs.cablelabs.com/MIBs/DOCSIS/

https://www.nctatechnicalpapers.com/Paper/2018/2018-node-provisioning-and-management-in-daa

Strategies for Continuous Integration and Continuous Deployment at Scale at the Network Edge

…a.k.a. The Pursuit of the Zero-Downtime Headend…

A Technical Paper prepared for SCTE by

Quincy Iheme
Engineering Manager
Comcast TPX
Philadelphia
267.260.2923
Quincy_Iheme@Comcast.com

1. Introduction

The next several years will go down as the time when an astounding amount of work happened at the edges of the network, in part to get to the coveted "zero downtime headend." A perpetually-up headend matters because it sets the stage for edge compute services. Much of that work at the edge of the network is benefitting from the role of software, and in particular, the branch of software engineering that is Continuous Integration / Continuous Deployment (CI/CD).

This paper discusses how operators can take advantage of the new flexibility that comes with the deployment of software at the network edge, more quickly and securely. It will provide an overview of CI/CD, with a specific focus on how it is being applied at the edges of the network, in virtual Cable Modem Termination System (vCMTS) deployments and related efforts.

Software engineering is unquestionably infiltrating many aspects of the physical infrastructure that historically were hardware-only. As a result, the establishment of and adherence to a common CI/CD platform can set the stage for the anticipated increase in software deployment at the network's edges, especially for services or activities that require very low latency and/or wider throughput.

This paper will describe how Comcast arrived at a common CI/CD platform, then put it to work at the network edge to maintain network uptime and increase upstream throughput – such as blue/green and automated deployments, as well as Distributed Access Architecture (DAA) components like vCMTS and edge switching, to enable continuous configuration and deployment at scale. It will also discuss tools and automation, and how to configure and push software to the edge to enable customers to experience faster broadband speeds.

2. A Brief History of the Edge of the Network

The so-called "edge of the network" is both a moving target and a destination that carries different actual locations, depending on who's describing it. It's a moving target largely because of 60+ decades of technological advancement – what was "the edge" changed with each new chapter in capacity expansion, from microwave to coaxial tree-and-branch to modern HFC (Hybrid Fiber-Coax) topologies. Generally speaking, as capacity increases, the "edge of the network" moves closer and closer to consumers.

Network engineers tend to define "the edge" as being somewhere near where optical signals hand off to RF, or, near the output of optical nodes. Other disciplines define the "edge" as being at the output of the set-top box or broadband gateway. From a competitive perspective, there are still only three physical wires – three edges – that reach directly into U.S. households: The power line, the phone line, and the "cable" line. The power utility industry, despite commendable effort, has yet to successfully provide broadband connectivity. The telephone industry is showing, through its heavy investments in wireless 5G, what it knows to be true about the throughput potential of twisted-pair copper wires. Our network edge really is the only one that can consistently deliver high-bandwidth services, over a wire (or not), to IP/connected devices.

For the purposes of this paper, the "edge of the network" is located at the headend, where headend is defined as where we are actively transforming traditional CMTS devices, with integrated QAM modulator, into virtual CMTSs, where the modulation function is relocated into Remote PHY Devices, or RPDs.

The network edge is a high-stakes place, coveted by all in the broadband ecosystem, from cloud purveyors to networked gamers to any of a dozen adjacent industries that wish to be as close as possible to end consumers. As engineers, it is our responsibility to make sure our network edges are as healthy, fortified and ready as they can be, for whatever is coming. A big part of being healthy, fortified and ready is applying the principles of Continuous Integration and Continuous Deployment.

3. Why Continuous Integration/Continuous Deployment Matters

Continuous Deployment (CD) is a software engineering methodology that takes any code change that is deemed stable and (normally through parallel methods of Continuous Integration, or CI) and makes it available in production. This usually includes automated building, testing, and deployment, without human intervention.

Our CI/CD progression began as it often does within hardware-centric industries that began long before software engineering principles emerged: Organically, within certain parts of the company responsible for software delivery. In the beginning, many of the commercially available and/or open-source CI/CD systems were used throughout the organization (see Figure 1 as an example of the vast CI/CD landscape).

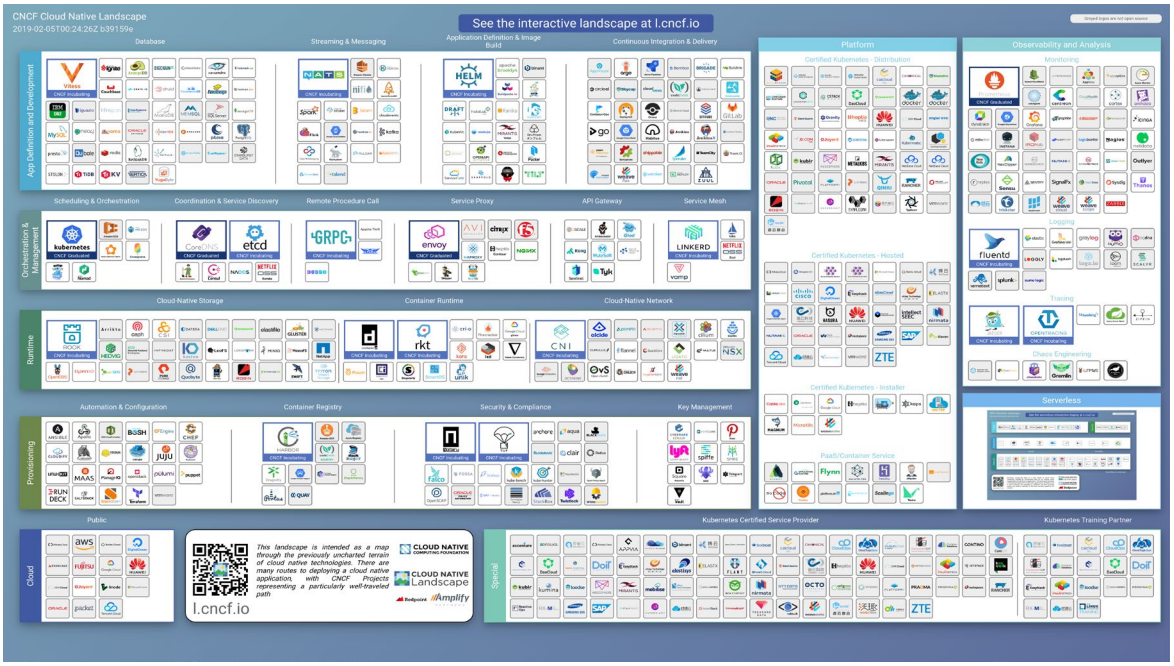


Figure 1 - CNCF Cloud Native Interactive Landscape

The practice of automating software delivery was one that was developed and documented outside the company, by the software industry at large. As more internal software teams began individually doing CI/CD themselves – including maintaining the cost of their own infrastructure to do so. Jenkins [1] (another common choice for implementing CI/CD within the engineering lifecycle) was one of the most popular options, as was Concourse, and a few others.

As more teams adopted CI/CD and related tooling, a need arose for a reasonably standardized pipeline tool. The main driver: Eliminating the technical debt associated with maintaining multiple infrastructures necessary to host different CI/CD tools, while creating a unified language that could be leveraged by all (or most.) By adopting a configuration-based CI/CD platform, component re-use can be applied to share many of the patterns and implementations that are very similar from one team to the next team. By establishing a shared platform and community around one tool, teams spend less time "reinventing wheels," among them code onboarding, because they can leverage already-public configurations.

As the reach of CI/CD platforms widened, a need arose to take stock of everything that was being done, in software, within Comcast's Technology, Product and Experience (TPX) organization. (Answer: A lot!) The CI/CD survey and its results happened within the construct of an internal TPX Architecture Guild, which consists of a group of engineers and technologists who were brought together to help guide technological decisions for the broader TPX organization.

The Architecture Guild uncovered a large amount of duplicate effort: Most teams were running their own instances of CI/CD software. The guild also identified varying degrees of implementation inconsistency: Teams were either as far along as deploying code into production every day (continuously deploying/integrating), or as far back as manually dropping compiled artifacts onto a web server.

Therefore, the guild recommended to Comcast TPX leadership that there be a dedicated team for maintaining CI/CD infrastructure, and that a single platform be chosen as the solution of choice. The recommendation was made with the idea that if you can get to a common platform and community, it would be easier to learn and onboard new teams, because they would only need to learn about a new CI/CD tool once. Questions like "has anyone done _____ with this platform?" could be answered by the more advanced users, further reducing the maintenance burden. The Architecture Guild was the ideal spot for a discussion as it was somewhat removed from the already formed opinions of specific teams. Asking team members to relinquish favored software tools, in favor of a different tool, is never easy; for more on this, see "The Tooling Abyss" by Joann Schuman, in this year's Fall Technical Forum papers.

Our then-Chief Software Architect and Senior Fellow, Jon Moore, was the main facilitator of the CI/CD unification discussions. As he describes it, in an interview conducted specifically to inform this paper, finding consensus on a unified CI/CD tool involved reconciling a couple of key points, including concerns around items like security and team separation. "The idea was that if we could get to a common CI/CD platform, with a community, it would be easier to learn and to on-board new teams, hand over ownership, because it wouldn't involve learning a new CI/CD tool," Moore explained [2].

4. The Architecture Decision Record (ADR)

With a few options then being proposed, an Architecture Decision Record (ADR) was formed to begin to track updates, offer reviews and other general comments. An ADR is a document that

captures decisions, including context (why was the decision made?) and consequences (what will adopting this decision impact?).

Since the ADR was in source control, interested/involved parties offered comments via pull requests and subsequent reviews. After a short period of proposals and reviews, a couple of options remained. Technically, any of the final options would have worked; as such, the goal was to pick the CI/CD tool that would cause the least amount of unhealthy friction, especially for those who would have to move over from another platform.

Instead of an outright vote, the guild proposed the use of a confidence poll, where each CI/CD software platform was ranked on a scale of 1-5, with a 5 representing, essentially, "the best," and 1 signaling "this would be a mistake." After rating the options on that scale, the scores closest to a 5 were interpreted as "acceptable to a majority of people." Ultimately, and instead of a straight-up vote for the "right" platform, Moore made the decision, based on the results of the CI/CD review process and confidence poll. As a result, the company now uses an open-source CI/CD tool called "Concourse" [3] to build CI/CD into various engineering lifecycles.

5. CI/CD at the Network Edge with vCMTS

With a tool of choice in place, teams could begin to implement an automation-driven software lifecycle to further increase the speed at which we could roll out changes. Such an effort was leveraged by several teams; in particular, and as it relates to this paper, a team that manage the Virtualized Cable Modem Termination System (vCMTS). Of course, there is a good deal of manual work to be completed beforehand. The racks that support the vCMTS need to be built. All the servers and switches that are part of the rack are shipped to headends. In the headend, wires are connected, to the RUs, power, and the GPS antenna. At that point, all the manual work needed for the connection is complete.

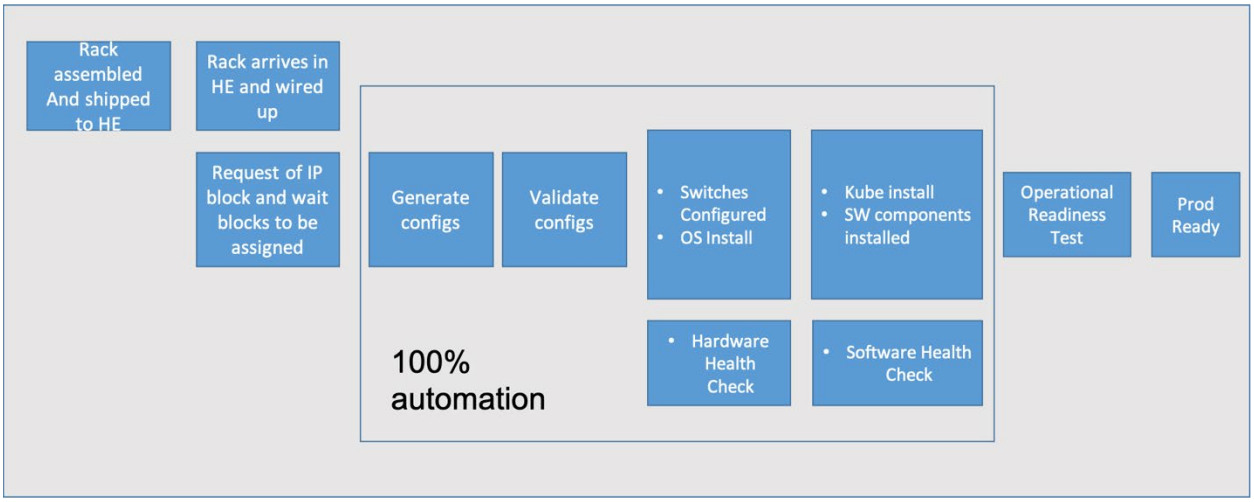


Figure 2 - Automated vCMTS Cluster Build Process

At this point, automation of the switch configuration is possible. Previously, if provisioning was needed, a ticket would need to be filed and a technician dispatched, to make the change onsite. An IP block would have to be requested and used to perform the requested provision. With automation, the IP block and configuration is known beforehand so the requested provision can be generated automatically. All of that now happens in Concourse, using a template to input the IP addresses, and to generate the configuration. Every cluster includes at least 2 switches and around 42 servers; up to 24 clusters can serve a headend. Looking at each pipeline (a set of tasks that run pre-defined operations), there is one pipeline per cluster. Each can have up to 192 vCMTSs. In that sense, each service group is a tiny vCMTS.

Every time there is a need to trigger a pipeline, a set of questions is asked: What am I deploying, and is it service-affecting or not? What kind of tests do I need to perform? The config, generated by Concourse, is then pushed to all the devices (servers and switches, numbering in the tens of thousands.) Afterwards, the OS (Operating System) is installed, and then software is installed. After automated hardware and software health checks are performed, the clusters are deemed production ready.

One thing to remember is all these different processes are resources. Concourse is a resource, running on servers. It needs a good amount of hardware to perform the operations being requested of it. One of the consequences of a shared CI/CD ecosystem is the sharing of resources. More resources could be allocated to speed up deployments to the vCMTS, but that would risk less availability of resources for other teams to use. This project in particular limits itself to 30 processes at a time.

For every change that needs to be made, a plan needs to be built to map out its journey to production. That mapping is fronted by a couple of questions, starting with reach: Changes are not rolled out to every single cluster at once. The focus is usually on new clusters that do not have new software yet. After clusters with older software are prioritized, the rest of the software updates are planned out. Based on the deployment type, and the service-affecting level, the deployment plan is calculated.

Once the deployment plan is approved, it is automatically rolled out during a maintenance window. First, a “pre” snapshot is taken of the current state of the systems in production. Health checks are performed, to make sure current systems are working as expected. Every change is made inside of a maintenance window. During that period, alarms are silenced because of the need to take systems down. After alarms are paused, the change is deployed to production. After the deployment, multiple post-deployment checks are run, with the aid of machine learning (ML). If there is a failure, there are 3 retries to access the systems and to ensure that the any change-related activity has settled down. Any failure that happens is logged, to let us know exactly what happened.

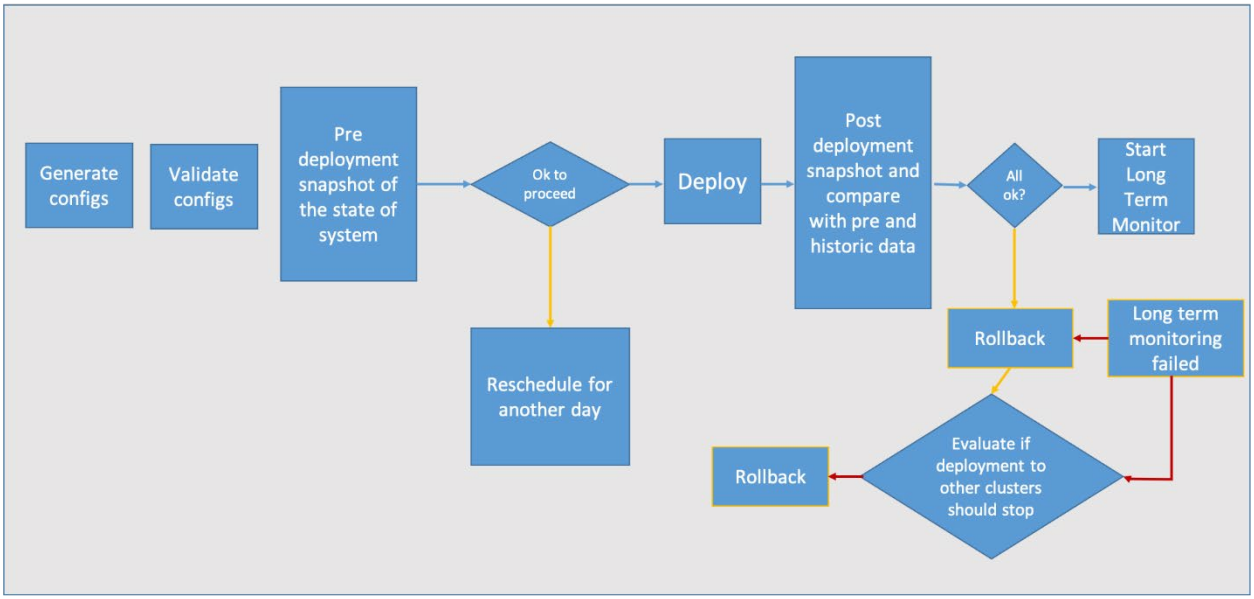


Figure 3 - Software and Network Change Deployment Pipeline

The process of connecting the RPDs ("Remote PHY Devices" / QAM modulators) can then begin. On a daily basis, and during the transition from integrated (which is to say, "traditional") CMTSs to their virtual counterparts, around 20 to 30 new clusters are being built in various places. Given common errors that can occur during config generation, or when scripts are run in a different order, or when switches malfunction, this process would traditionally take about three to four days. With the process of automating vCMTS deployments through Concourse, changes like these take anywhere from 2-4 hours. The variability of time depends on the platform, how many other jobs are running, platform health, etc. The beauty of automation is that there is a single source of truth from which configurations are read.

In the case that a rollback is needed for an upgrade that was deployed, for instance, checks exist to ensure that the rollback instances are healthy and ready to receive traffic. If there needs to be a change to the vCMTS core software, there's a check to see if it supports all CPEs (Customer Premise Equipment / set-tops and gateways). With a CI/CD methodology in place, you would expect any rollback to be done as quickly as (if not faster than) the roll out to production. In this case, rolling back is a progressive process. There is sampling of all kinds of data (with the aid of machine learning) that gives visibility around what specific failures there were and whether a rollback is needed in a certain area.

A zero-downtime headend necessarily means removing changes that can cause stoppages, and removing elements that can become points of failure. The iCMTS, with integrated RF modulators, was an example: From the analog node, you connected all the RFs. With vCMTS, the RF portions are removed from the CMTS, and put into RPDs. The physical component – the PHY, is in the RPDs; the nodes, in fiber-deep configurations, could be on a pole, in a pedestal, or underground. Upgrading them should be non "hitful" – which is to say, not service impacting.

At the headend, where the CMTS servers run, the intent is to complete upgrades with minimal scheduled downtime, on the order of 3 minutes. With a traditional iCMTS, the downtime tied to a scheduled upgrade is typically around 2 hours; so far, with the vCMTS, scheduled downtime related to upgrades is 15 minutes.

Our ultimate goal is to reduce headend downtime to close to or exactly zero. If there a catastrophic failure occurs with a small "blast radius," can take 15 minutes to redeploy software to a vCMTS, rather than take the risk of a lengthy time to diagnose and fix.

We are currently (summer 2021) also working on an in-service upgrade functionality. The intent is to get to 15 minutes or less of downtime for 2 components: The RPD, and the vCMTS, to simultaneously address the service-affecting route.

6. Conclusion

The journey into adopting a culture of Continuous Innovation and Continuous Deployment has been one that mirrors the actual process of continuously iterating and deploying: Problems are defined by processes or issues in progress. Updates and improvements are added frequently, without sacrificing stability. Changes are rolled out while ensuring reliability and adequate steps to rollback. Rolling out CI/CD (along with a unified platform and growing community) was done without jeopardizing the customer experience. What would've once been considered a phenomenon is now common practice in software engineering. Those lessons are now able to be transferred into the hardware side of engineering, to help drive improvements at a wider scale and without the need of constant manual intervention.

I would like to take the time to thank those who lent their time and expertise into the completion of this paper and presentation. Frankyln Athias, Sherita Ceasar, Leslie Ellis, Jon Moore, Max Knee and Bhanu Krishnamurthy.

Abbreviations

ADR	Architecture Decision Record
CI/CD	Continuous Integration/Continuous Deployment
CPE	Customer Premise Equipment
DAA	Distributed Access Architecture
HFC	Hybrid Fiber Coax
iCMTS	Integrated Cable Modem Termination System
PHY	Physical Circuit
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency
RPD	Remote PHY Device
vCMTS	Virtualized Cable Modem Termination System

Bibliography & References

[1] https://landscape.cncf.io/

[2] Informational interview, Jon Moore, Senior Fellow, Comcast

[3] https://concourse-ci.org/

Additional references:

[4] *Solving The Mysteries of the Distributed Access Architecture*; Matthew Stehman, Ramya Narayanaswamy, Jude Ferreira, Robert Gaydos, 2021

[5] *The Tooling Abyss*; Joann Shumard, 2021

[6] *Humanoids Optional: Deploying vCMTS at Scale with Automation*; Bhanu Krishnamurthy, Gregory Medders, 2021

[7] *Architecture with 800 of My Closest Friends;* The Evolution of Comcast's Architecture Guild; Jon Moore, 2019: https://www.infoq.com/articles/architecture-guild-800-friends/

Successful Wi-Fi 6 Deployment to Customer Homes

A Service Provider's Guide to Intelligently Controlling and Optimizing the Wi-Fi 6 Home Network

A Technical Paper prepared for SCTE by

Bill McFarland
Chief Technology Officer
Plume Design, Inc.
290 California Ace. #200
Palo Alto, CA 94306
1-844-69-75863
bill@plume.com

1. Introduction

Hailed as a new era for Wi-Fi connectivity, 802.11ax—known better as Wi-Fi 6—has created much excitement and anticipation. Capable of more than doubling the speeds of its predecessor, 802.11ac (or Wi-Fi 5), Wi-Fi 6 promises to transform the home network.

As well as boosting throughputs, this new technology improves overall system capacity, security, and even power consumption and battery life. And the timing couldn't be better. Not only are consumers relying on home Wi-Fi more than ever, but the smart home is also evolving to include more connected devices, immersive experiences, and bandwidth-hungry applications.

The Wi-Fi Alliance anticipates nearly 2 billion Wi-Fi 6 devices to be shipped in 2021 to consumers and organizations.1 Wi-Fi 6 products entering the market will spark renewed consumer interest in Wi-Fi upgrades, including deployments of Wi-Fi 6 home networks. It's an exciting opportunity for Communications Service Providers (CSPs) to elevate their offerings to subscribers.

There's no doubt that the new capabilities of Wi-Fi 6 will improve the user experience in all connected spaces. But maximizing its full potential means optimizing Wi-Fi networks in ways that are both more critical and more complex.

Wi-Fi 6 connectivity in any space is only as good as the system controlling it. To deliver on the promise of this new technology for their customers, CSPs must understand the limitations of Wi-Fi 6 home deployments and the requirements for intelligently managing the networks.

2. The State Of The Wi-Fi 6 Industry And The Market

Wi-Fi has become as ubiquitous as the internet itself. Considering that Wi-Fi 6 brings the first major improvement to 2.4 GHz band in a decade, the buzz surrounding it is understandable.

And a lot has changed in that decade. The home network is much more complicated and customers are increasingly relying on that Wi-Fi connectivity for work, school, and entertainment:

- The average US household served by Plume has 14.5 connected devices.
- 28% of consumers use smart home devices such as cameras and thermostats. (Deloitte)
- The number of global shipments of smart devices is expected to grow 71% between 2019 and 2023 (from 814.8 million to 1.4 billion). (eMarketer)

The excitement about Wi-Fi 6 can be best summarized in the words of Federal Communication Commission (FCC) Chairman Ajit Pai, who said, "The American consumer's wireless experience is about to be transformed for the better."

The Wi-Fi 6 market is just emerging. But the technology's potential to transform the home network is tremendous, especially as the smart home and work-from-home (WFH) trends gain momentum. Research firm Strategy Analytics sees the wireless home becoming "one of the leading technology trends of the early 21st century."

Here's just a glimpse of what the industry forecasts:

- 17 billion home devices will be in use by 2030 globally, with Wi-Fi 6 accounting for a third of device sales by 2023. (ABI Research)

- Wi-Fi 6 will become the predominant Wi-Fi standard both for consumers and enterprises by 2023. (Cisco)
- On the public Wi-Fi front, Wi-Fi 6 hotspots will grow 13-fold between 2020 and 2023, comprising 11% of public Wi-Fi hotspots by 2023. (Wi-Fi Alliance)
- 2 billion Wi-Fi 6 devices will be shipped in 2021 to consumers, enterprises, and public agencies. (Data aggregated from the Wi-Fi Alliance Product Finder, last accessed February 2, 2021.)

2.1. Market Developments

Rapid growth of the Wi-Fi 6 market began when the Wi-Fi Alliance launched its Wi-Fi CERTIFIED 6™ certification program, based on the IEEE 802.11ax standards, in September 2019. Even before that, some industry players—including networking vendors—began building the next-generation Wi-Fi infrastructure.

By February 2021, the Wi-Fi Alliance had certified nearly 1,400 different Wi-Fi 6 products. More than a third of those were in the router category, including home gateways. While the smart home category included only a few certified devices initially, some smart TVs, tablets, and even a couple of gaming systems were already available. It's only a matter of time before the consumer market explodes.

Momentum was also built in 2020 for Wi-Fi 6E (the E stands for Extended). Wi-Fi 6E is simply Wi-Fi 6 operating in the 6 GHz frequency band. The 6 GHz band has recently been allocated for Wi-Fi in the US, EU, and other geos. While this band promises more spectrum, better availability of 160 MHz wide channels, and lower interference levels, it requires more sophisticated management.

Notable Wi-Fi 6E events include:

- Broadcom announced a portfolio of residential and enterprise access point (AP) solutions for 6 GHz WLAN in January 2020, followed by a February announcement of a Wi-Fi 6E chipset.
- In April 2020, the United States became the first country to open the 6 GHz spectrum, and some have called the decision by the FCC historic and monumental.
- In December 2020, FCC authorized the first Wi-Fi 6E device, a Broadcom low-power indoor transmitter. In his statement, FCC Chairman Ajit Pai said this was "an exciting glimpse of America's Wi-Fi future."
- In early January 2021, the Wi-Fi Alliance introduced Wi-Fi 6E certifications, which will use enhanced, WPA3 security. The Alliance has certified more than a dozen products, including routers, from vendors such as Intel, Qualcomm, and Samsung, as of February.

2.1.1. Wi-Fi 6 Vs. 6E Devices

Wi-Fi 6E devices will be backward compatible with Wi-Fi 6 and previous Wi-Fi standards. But to take advantage of those new 6 GHz channels in Wi-Fi 6E, you'll need to be using devices that support it. In other words, you'll only be using Wi-Fi 6E once you pair a 6E-enabled client device (like a laptop or smartphone) and a 6E-enabled access point. Wi-Fi 6 devices paired with a 6E-enabled router will still be using the typical 5 GHz or 2.4 GHz channels.

3. Key Wi-Fi 6 Features And Benefits

In the home environment, the network is becoming more congested as a larger number of devices connect to Wi-Fi and customers run more bandwidth-heavy applications such as 4K video streaming, virtual reality (VR) gaming, and video conferencing. Wi-Fi 6 brings key changes to several areas, including:

- 2x higher throughput in low-congested environments
- Improved power efficiency
- 2x more devices that can be supported efficiently (to 8)

The new features are designed to better handle different types of traffic simultaneously from multiple users, as well as improve coverage for overlapping networks and dense environments. In practice, however, many of the new capabilities have limitations and drawbacks. Mitigating the issues will require advanced tools that manage and optimize the home networks.

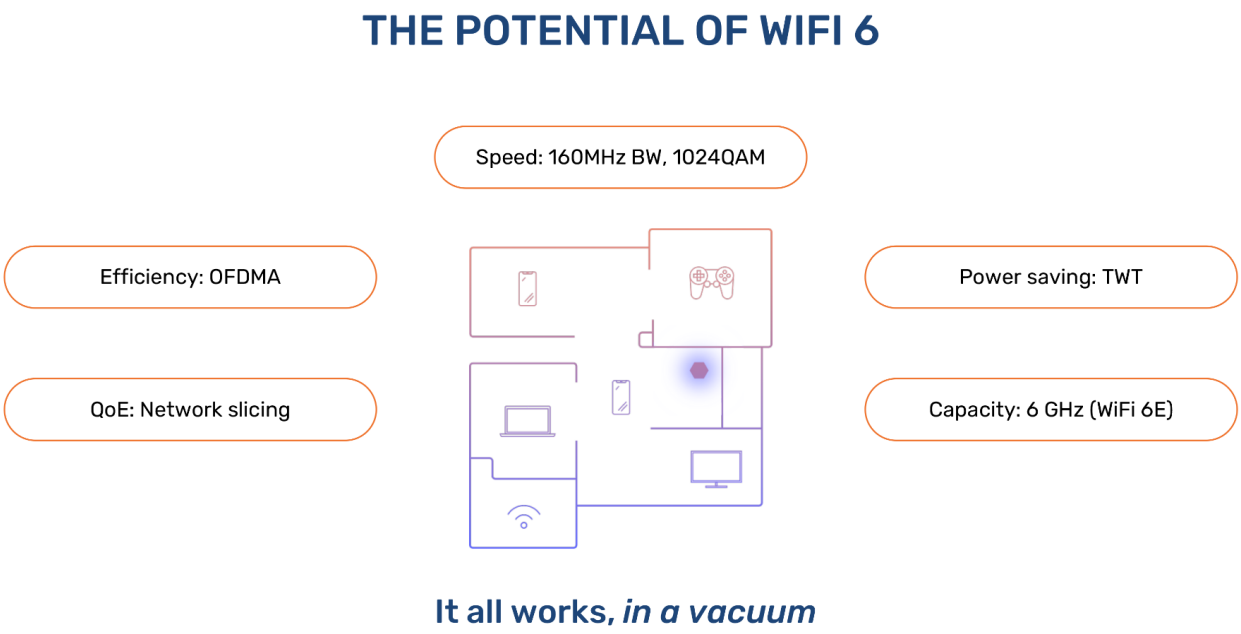


Figure 1 - The Main Advantages of Wi-Fi 6

3.1. An Overview Of New Wi-Fi 6 Features

3.1.1. 160 MHz Channel Bandwidth

The wider the channel used, the higher the data rate. While the Wi-Fi 5 technology officially supported 160 MHz capabilities, few devices actually offered channel bandwidths greater than 80 MHz. Although this is technically not a new feature, the expectation is for most Wi-Fi 6 devices coming to the market to support 160 MHz channel bandwidths.

3.1.2. 1024-QAM

Wi-Fi 6 goes from 256 quadrature amplitude modulation (256-QAM) to 1024-QAM, increasing the physical layer data rate by 25% at short range. Essentially, this modulation brings a more dense packing of bits into the signal, thereby boosting the network speed by increasing the data rate for the channel.

3.1.3. OFDMA

Uplink and downlink orthogonal frequency-division multiple access, or OFDMA, is a marquee feature of Wi-Fi 6, which allows for a single transmission to communicate with a large number of devices. It greatly improves efficiency and capacity by subdividing the channel into smaller frequency allocations (resource units) that are transmitted from one AP in parallel.

3.1.4. Resource Unit Reservations

The resource units, or smaller frequency slices, that OFDMA operates with can get allocated to particular clients. This allocation changes dynamically over time, which is what allows a single, wide, efficient transmission to serve multiple clients at once. Reserving resource units for particular clients or data flows provides guaranteed Quality of Service (QoS) to those clients or flows because a fixed amount of bandwidth within the network is allocated specifically for that traffic.

3.1.5. Target Wake Time

First developed for 802.11 ah (900 MHz Wi-Fi, branded "Wi-Fi HaLow" by the Wi-Fi Alliance), target wake time (TWT) is expected to be deployed widely with Wi-Fi 6. This mechanism improves battery life for devices that are transmitting only occasionally or at a low-duty cycle. The AP creates a schedule with specific times for each client to be awake and reserves that time so no other devices can transmit during that window, giving the waking devices clear airwaves to quickly communicate and return to sleep.

3.1.6. UL-MU-MIMO

The Uplink Multi-User Multiple Input Multiple Output (UL-MU-MIMO) is the companion to Downlink (DL) MU-MIMO that was standardized and implemented in Wi-Fi 5. UL-MU-MIMO allows multiple devices to be transmitting at the same time to the same AP, improving efficiency and uplink capacity. With OFDMA, the different devices transmitting at the same time use different parts of the frequency spectrum. In the MU-MIMO case, the devices involved rely on multiple antennas to separate the traffic by spatial means, so the AP can independently receive signals coming from clients in different directions.

3.1.7. BSS Color

Basic Service Set (BSS) color added with Wi-Fi 6 allows more efficient airtime usage between overlapping networks that are on the same frequency channel. This technique marks (or "color-codes") shared frequencies at the very beginning of the packet to indicate which network the packet belongs to and enables devices to make a very quick assessment if they are safe to transmit or must defer to traffic coming from a network with a different color.

3.1.8. 6 GHz Frequency

Wi-Fi 6E extends Wi-Fi 6 capabilities to the 6 GHz spectrum, previously only available to licensed users. This is considered a big step, as the need has become more urgent to prevent congestion on existing frequencies. The industry has been advocating for some time for the allocation of more bandwidth for

unlicensed use. The new spectrum makes available up to seven superwide 160 MHz channels that can support high-bandwidth applications such as unified communications and industrial IoT. For home networks, use cases that can benefit from the low-latency, higher-speed Wi-Fi include augmented reality and virtual reality.

4. Wi-Fi 6 Network Optimization And Controls

As highlighted earlier, the new or enhanced Wi-Fi 6 capabilities are not without issues. To ensure these features work to their full potential and truly benefit customers, CSPs need to implement sophisticated controls and optimize the network. These controls are especially critical in today's congested home Wi-Fi environments and in dense areas where issues such as interference arise.

4.1. 160 MHz Channel Bandwidth

Doubling the channel width from 80 MHz to 160 MHz should double throughput. So, in theory, customers would see twice-as-fast downloads, double resolution, and so forth, with the double data rate. The problem is that, in most countries, only two independent 160 MHz channels are available. In a dense environment, such as an apartment complex or even suburban homes on small lots, two different customers are likely to use an overlapping channel, leading to significant interference.

This issue can be mitigated with proper network configuration, which requires three factors:

- Intelligent sensing and prediction of the interference on 80 MHz subchannels of the 160 MHz transmissions.
- Intelligent channel allocation and bandwidth selection.
- Consideration of the complete interference picture in an environment, optimizing channel allocation across entire apartment complexes or neighboring homes.

Simple, locally managed networks can't achieve these factors. CSPs will need to handle them centrally through the cloud. For example, cloud-based interference analysis factors, loads, and device types. When considering what channel bandwidth to assign to each AP, the platform knows the history and current set of clients and loads that are present in the network at each access point. The system can then apply 160 MHz bandwidth channels where they are most needed and avoid them where they are not. Central cloud management also allows the optimal selection of channel bandwidths and frequency channels across an entire apartment building. Frequency channels can be “tiled" across the apartments, minimizing the conflicts between neighboring apartments. In cases where re-use of the same channel is unavoidable, the cloud-based optimization system can select apartments that are most able to share frequency channels based on their historical activity.

4.2. OFDMA

This signature Wi-Fi 6 feature stemmed from observing the poor efficiency of the small packets typically emanating from Internet of Things (IoT) and other low-data-rate devices. Each 802.11 packet has significant overhead, and to avoid collision with other transmitters, each transmitter must listen first to the medium for some time. Even if each transmission sends a modest amount of data, the overhead consumes a lot of airtime.

OFDMA allows transmission to multiple users simultaneously in one packet, eliminating overhead and wasted time. This capability only improves efficiency when a significant number of IoT devices are on

the same AP, each device sending or receiving a modest amount of data. But, modern Wi-Fi networks in smart homes will not always end up with a significant number of IoT devices on the same AP.

Here's why: Home networks have migrated to multi-AP topologies using mesh networks, repeaters, or multiple gateways. If the IoT devices in the home were to simply connect to the nearest AP, each AP would have few devices appropriate for grouping into OFDMA transmissions. On the other hand, forcing devices to connect to an AP that is too distant will force the data rates of these connections to drop, compromising efficiency in a different way.

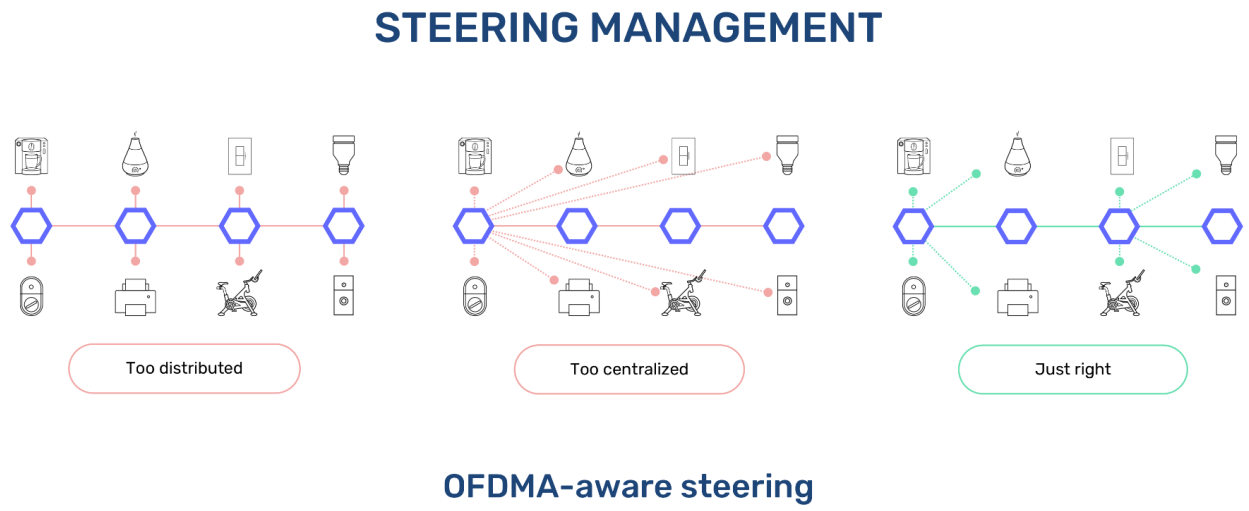


Figure 2 - OFDMA-Aware Steering

Consequently, OFDMA operation requires OFDMA-aware client steering. To make complex decisions, a centralized, intelligent network controller would need:

- Knowledge of which APs and clients are Wi-Fi 6 capable.
- Historical observation and forward prediction of the data needs of each device in the home.
- The ability to make intelligent, optimized choices about which of the multiple APs each device in the home should connect to considering the capabilities, traffic load, signal strength, and data rate each device can achieve.
- The ability to steer and hold devices on the correct AP.

For all the APs and clients to work efficiently, the best arrangement must be created through rigorous optimization. This should include the ability to steer clients to APs and hold them there, even if those aren't the closest APs. The steering mechanism must be specific to each type of device because different devices behave differently to various steering mechanisms.

4.3. Target Wake Time

Oftentimes, the home network has several APs operating on the same frequency channel. This creates potential overlap when multiple APs try to schedule the TWT for different devices. To avoid conflict, a central scheduler can ensure that TWT periods at the co-channel APs do not overlap.

The scheduler must know:

- The transmission requirements of all the TWT clients
- The shared channels among APs
- The clients connecting to each AP
- The signal strength between the APs and clients in the home

With this data, the system can create an optimized arrangement of TWTs at each AP, covering all clients across the home network.

In a multi-dwelling unit (MDU), the optimizer could go a step further, looking across overlapping networks that it controls in the MDU and plan TWT assignments and groups across the entire building. The system would factor the cloud-based controller's knowledge of which apartments interfere with one another, and the client capabilities and load requirements in each apartment.

4.4. Network Slicing

A QoS mechanism, network slicing is another way of allocating frequencies and time among different clients. This is essentially a time-division, multiple access (TDMA) method (where transmissions occur at reserved times) but has the added overlay of a divided frequency spectrum. In a home with multiple APs or in an MDU, network slicing can lead to high collision rates when both networks schedule the same time periods. The result is poor QoS, defeating the point of the time/frequency reservations. Centralized, intelligent controls can mitigate this issue by coordinating among the multiple APs.

4.5. 6 GHz Frequency Band

To protect point-to-point microwave systems that operate on the same 6 GHz spectrum, Wi-Fi 6E devices must use one of two modes:

- A low power (18dBm/63mW) that uses the low-power spectrum rules: This mode can't transmit as far or at as high a data rate because the signal is not very strong.
- Under the command of an AP that has an automated frequency control (AFC) system (30dBm/1 Watt allowed): In the US, for example, the AFC systems must check the data in an FCC database to ensure there are no microwave systems in the vicinity of the AP. There are currently around 100,000 microwave links in the US used by mobile carriers, industrial and business entities, and public safety agencies. By avoiding a frequency channel used by nearby microwave systems, the AP and its clients can operate at the high-power level, enjoying full range and data rates.

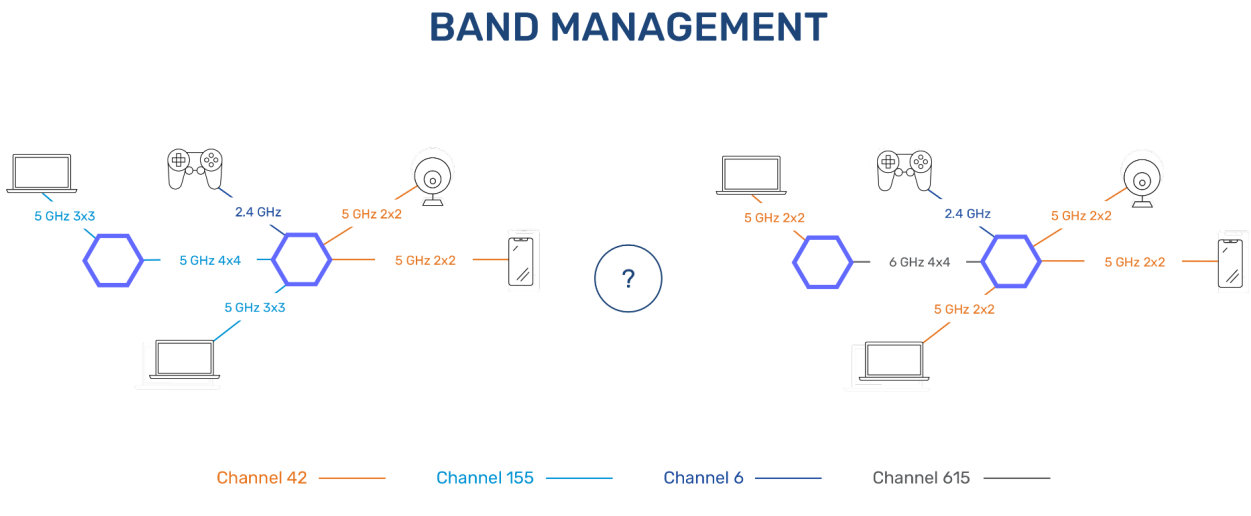


Figure 3 - Planning Band Management

These operating modes will require different types of controls:

- For the low-power transmission, more complicated, multi-AP configurations will be required for an optimization system to select the appropriate network topologies, frequency assignments, and client-steering options.
- For the AFC systems, the AP will need to communicate with a smart controller that can look up the FCC database, factor in the geodata, calculate interference levels, then deliver instructions back to the AP.

For either mode, the control system needs to take into account the client types, loads, and capabilities to decide how to allocate the network's radio resources. Depending on the capabilities of the clients in the network, it is not always optimal to put one of the AP's radios in the 6 GHz band. For example, using a 6 GHz channel for the backhaul connection may help the backhaul, but it may take away the high-performance radio in the AP from the 5 GHz band. High-performance clients that do not have a 6 GHz capability may therefore connect at lower speeds, actually degrading the experience in the home.

5. Intelligent Management Imperative For CSPs

Wi-Fi 6 brings long-anticipated improvements to a technology that's now 20 years old. The new features will solve many challenges created by the ever-expanding, increasingly more congested home network. It's undisputed that this technology takes the home network to a new level. However, Wi-Fi 6 brings additional complexities, such that simple controls can no longer satisfy the requirements of this evolved ecosystem.

To maintain the Quality of Experience (QoE) for their customers, CSPs will need to adopt an intelligent management approach. With centralized controls, ideally based in the cloud, CSPs can ensure that they're achieving the performance potential of Wi-Fi 6. More sophisticated controls, coordination, and

optimization will become especially critical as smart home adoption picks up the pace, putting further pressure on connectivity and performance.

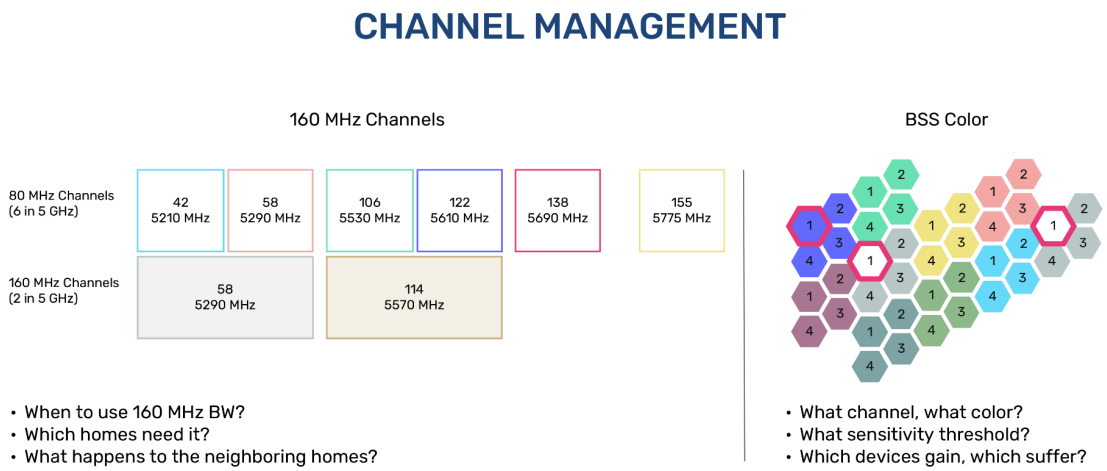


Figure 4 - Planning Channel Management

CSPs should harness the power of the cloud and artificial intelligence (AI) to make sure their subscriber's home networks are ready to support the exciting capabilities of Wi-Fi 6. The cloud offers practically infinite memory and computing power, supporting AI-driven innovation.

6. Measured Results

The previous sections outline the need for centralized cloud management in a Wi-Fi system, particularly when Wi-Fi 6 capabilities are being exploited. This section presents measured results of what can be achieved when centralized management is applied to real-world networks.

6.1. Selection Of 80 Vs. 160 MHz Channels

As described earlier, Wi-Fi 6 devices include the option to operate in 160 MHz channels, but this may not result in optimum performance. The optimization system measures the interference in homes and chooses the optimum channel bandwidth to maximize the throughput in that home. The chart below shows the percentage of homes that the system configures into 160, 80, 40, or 20 MHz. This study was conducted across a sampling of households in the United States.

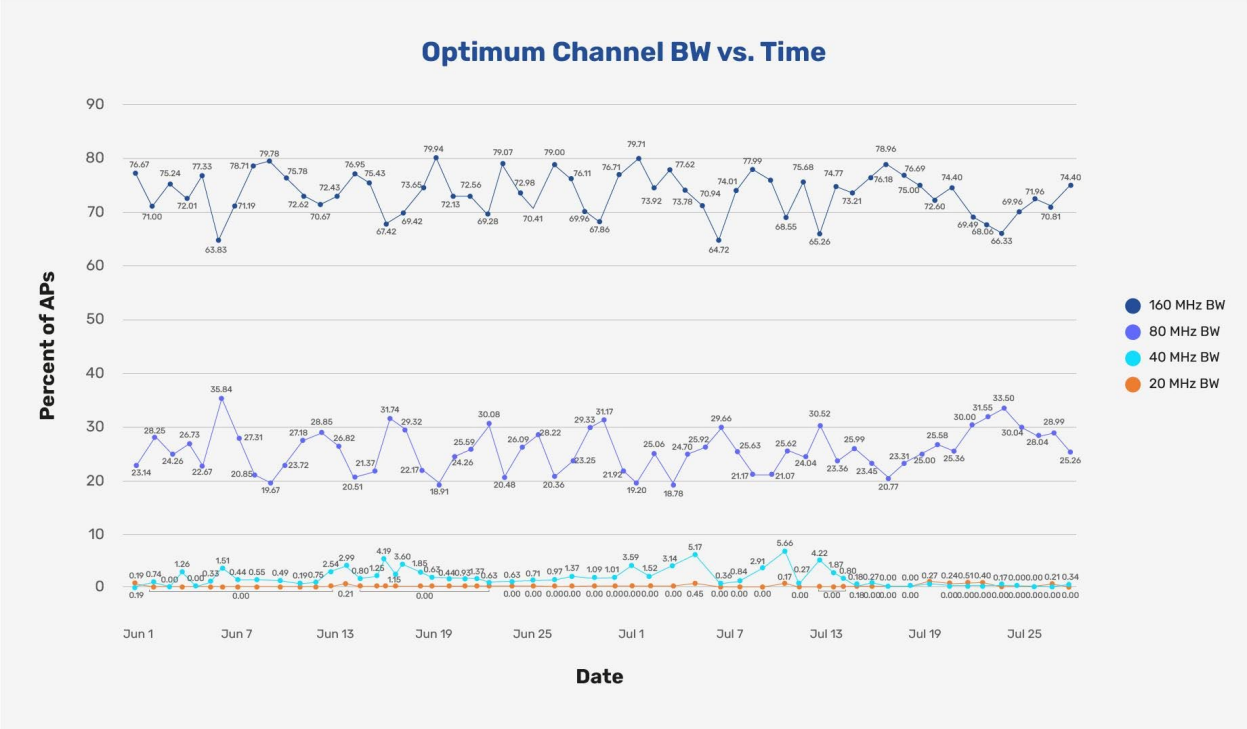


Figure 5 - Channel Bandwidth: SuperPod_AX to SuperPod_AX links

Several interesting observations can be made. First, roughly 25% of homes will achieve better throughput in 80MHz mode than in 160 MHz mode. Second, the exact distribution varies on a day-by-day basis due to variance in network usage and therefore neighbor interference. Finally, there are a small number of locations that achieve the best throughput in 40 MHz mode.

Clearly, a system that sets all networks into 160 MHz mode, or one which makes a one-time selection of 160 vs. 80 MHz will result in sub-optimal performance in a significant percentage of homes.

6.2. DFS Impact On 160MHz

Much of the 5 GHz frequency band is shared with radar systems. Wi-Fi networks must search for the presence of radar signals, and vacate the channel if they are observed. These radar events cause significant disruption to users as the entire network switches to different frequency channels to avoid the signal that looks like radar. Often these signals are not coming from a real radar system, but are generated by other types of devices operating in the 5 GHz band, or unusual signals generated by the collision of Wi-Fi packets from multiple networks. Therefore, these radar effects can be seen anywhere in the country, but are more common where radars are present, such as near airports.

Theoretically, doubling the channel bandwidth roughly doubles the odds that a radar event will happen within the channel, requiring the system to suspend operation and switch to a different channel. However, we have studied these radar events extensively and find that they are not at all evenly distributed by location, or by frequency channel at a given location. The table below shows the measured radar event rates reported by APs in a series of locations, some with multiple APs, some with a single AP. The rates are expressed as the number of radar events per hour of operation (0.04 = one radar event per day on average). The color-coding in the diagram highlights the location, AP, and frequency channel triplets that have high (orange), medium (yellow), and low (gray radar event frequency. Note that in this diagram each

numbered channel is 20 MHz wide. Therefore an 80 MHz channel will span 4 of these channels, while a 160 MHz would span 8.

Long Term Radar Rate Per DFS Channel Per AP (Aggregate Learning from July-2021)

locationid	nodeid	52	56	60	64	100	104	108	112
58bbb63798c6eff642b1d1a2	EM7F60018C	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
5fbd3b8b66b7453aa7b48808	EM7F300045		0.10	0.10		0.00	0.00	0.01	0.00
5f933d571f6ecf4457c695f0	EM7F600049	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	EM7F600065	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	EM7F600083	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
	EM7F600097	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5f9f8d1a4a576e734d231590	EM83A00143				0.49			0.04	0.04
57f3193f3ce0d96ed09d0217	EM7F60014C	0.00	0.00	0.00		0.06	0.06	0.06	0.21
5f920576bfddfc59ee6dd846	EM7F60007B			0.00	0.00			0.00	
	EM7F600024	0.01		0.00	0.00	0.00		0.00	0.00
	EM7F600054					0.00	0.00	0.00	0.00
	EM7F600064	0.00	0.00	0.00	0.00	0.03	0.03	0.03	0.10
57f489f83ce0d96ed09d6570	EM7F6000C1		0.00	0.00	0.00	0.00	0.00	0.00	
	EM7F300006	0.00	0.00	0.00	0.00				
	EM7F300040	0.03	0.03	0.05	0.00		0.00	0.00	
	EM7F600119		0.00	0.00		0.00			
60480b07328add3166edda7d	EM7F30001D	0.00	0.01	0.01				0.00	0.00
	EM7F30004D		0.00	0.00	0.00	0.00	0.00	0.00	0.00
	EM7F300001	0.02	0.07	0.07	0.06				
60391ea536d7902c9a44a93c	EM7F6000C6	0.01	0.01	0.01	0.01	0.00	0.00	0.02	0.02
	EM7F6000D8	0.00	0.02	0.03	0.01	0.02	0.04	0.03	0.04
	EM7F60012B	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00

SUM (Long Term Radar Rate)

0.0000 0.10000

Figure 6 - Long Term Radar Rate For DFS Backoff

Looking at the chart, it is easy to see that some locations can use 160MHz channels without concern, while others would be hit by radars at a high rate, frustrating users with frequent channel changes. However, many of the homes that could not support a 160 MHz channel can support an 80 MHz channel if it is properly positioned.

To minimize service interruptions, the system should monitor and record the radar rates that occur at each individual AP, and use that to intelligently avoid channel usage that would result in frequent radar events.

6.3. Clustering

Interference can occur in any type of environment, degrading or eliminating the benefits of Wi-Fi 6. However, the most likely location for high interference is in apartment complexes, or MDUs. Interference in these types of environments can be minimized using optimizations that consider MDUs in a holistic way, optimizing across the entire MDU. Because it would be impossible to jointly optimize the entire planet, the first step is to identify the MDUs, or sets of locations that present strong interference to each other.

Unfortunately, there is no simple database for knowing which locations will interfere with other locations. While conceptually this would be aligned along MDUs, actual interference patterns are complicated by any number of factors. The necessary grouping is accomplished through a clustering step. Each AP reports the neighbors that it can see, including the neighbors that it can see on different frequency channels using off-channel scanning. This database is then used to group the locations into clusters that should be optimized together because the locations have a high degree of interaction.

Community network machine learning algorithms can be used to accomplish this task. The diagram below shows a conceptual picture of how this works. Tightly intertwined locations are grouped together (e.g. locations in the same apartment complex), while only mildly connected groups (e.g. neighboring apartment complexes) are segmented.

Figure 7 - Tightly Intertwined Locations Are Grouped Together (Left) While Only Mildly Connected Groups Are Segmented

The following table summarizes the results when clustering algorithms were applied to two large service providers' customer bases. Keep in mind that the system clusters only the locations that have an AP that is managed by the cloud system. Because the cloud system cannot control neighbors that are not part of the system, there is no reason to cluster them together.

Table 1 - Cluster Identification For Community Network Machine Learning Based On Two Large Service Providers

Parameter	Service Provider A	Service Provider B
Number Of Plume-Managed Locations	12,527,567	5,017,865
Total Number Of Clusters	2,012,417	928,133
Size Of Largest Cluster	1,173	412
No. Of Clusters With >100	8,367	7,049
No. Of Clusters With 50-100 Locations	21,633	85,036
No. Of Clusters With 10-50 Locations	195,747	74,748
No. Of Clusters With 1-10 Locations	1,786,670	760,277

With the clusters identified, a joint optimization of all APs within the cluster can be performed, optimally assigning frequency channels and channel bandwidths across each of the clusters.

6.4. Interference Improvement

Studies of the improvement that can be achieved when optimization is applied across complete clusters as opposed to optimizing at each location individually were performed. These studies were done in a selection of individual MDUs, which were identified by our service provider partners.

The following diagram shows a histogram of the number of locations in the set of MDUs vs. the "pain" metric coming from interference. The pain metric is a function of the interference seen in a certain home times the load that devices in that home place on their own network. The pain metric does a better job of

estimating the true customer QoE than just the interference. For example, a home that is completely idle (no traffic flowing), is not impacted by a high level of interference from a neighbor. On the other hand, a home that has a very high load, consuming nearly 100% of the airtime, can be degraded with only minor interference from a neighbor. At a pain index of ~0.5 consumers experience occasional interruptions in real-time services such as voice and video.

In the diagram, the "Before" distribution represents the histogram of the homes in the MDUs prior to joint clustered optimization, and the "After" distribution represents the histogram of the same set of homes after joint clustered optimization.

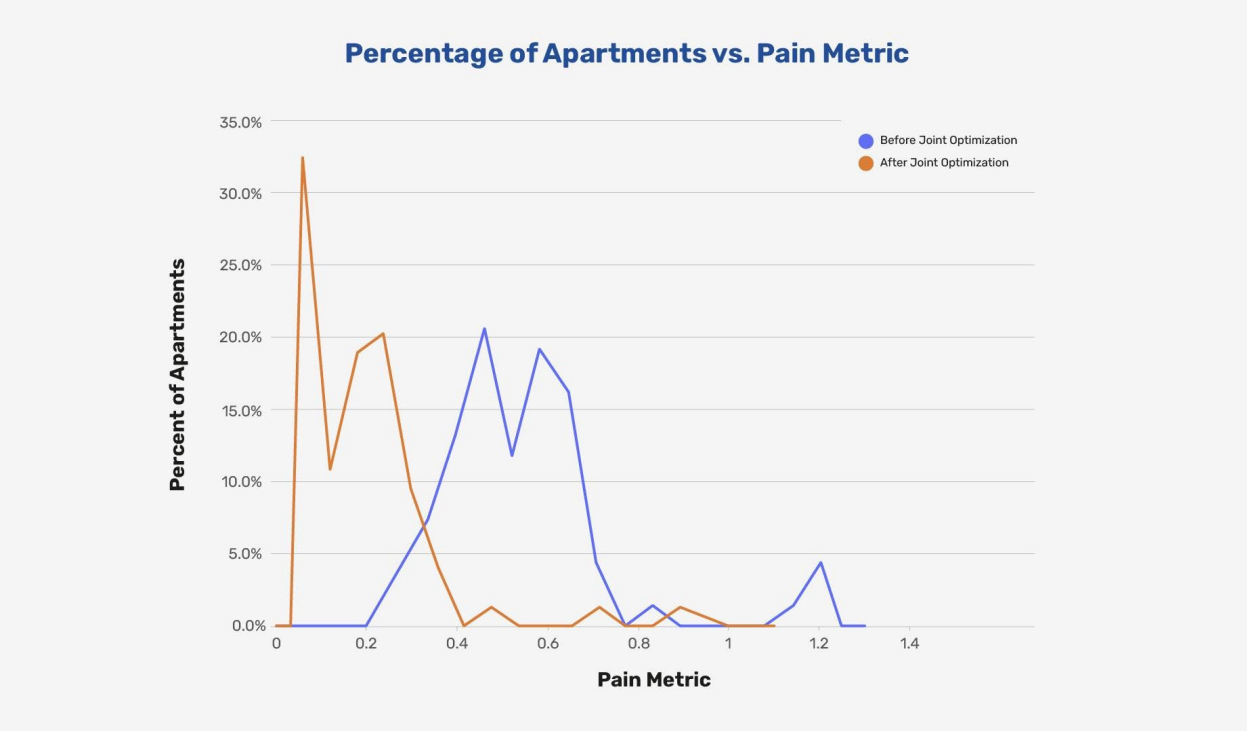


Figure 8 - No. Of Locations In The Set Of MDUs Vs. The Interference "Pain" Metric

The histogram shows that the number of homes at high levels of pain (exceeding the threshold at which real-time services are impacted) can be significantly reduced by the clustered optimization process.

7. Conclusion

As with most new technologies, it will take a few years for Wi-Fi 6 to become ubiquitous. But the possibilities are thrilling once the technology is fully developed and implemented. The greatly enhanced speed and performance of Wi-Fi 6 will open new doors for emerging technologies such as IoT and AR/VR. It's not only a new era for consumers but also an opportunity for CSPs to take advantage of the technology to enhance and expand services.

Wi-Fi 6 is, indeed, powerful. But it doesn't eliminate the need to optimize the network—on the contrary, the complexity of this technology creates an even greater demand for intelligent management. CSPs should take advantage of the advanced management solutions that future-proof their customer deployments while meeting customer demand and maintaining QoE.

Abbreviations

802.11ac	current generation of Wi-Fi (Wi-Fi 5)
802.11ax	next generation of Wi-Fi (Wi-Fi 6)
AFC	automated frequency control
AI	artificial intelligence
AP	access point
AR	augmented reality
BSS	basic service set
CSPs	communications service providers
dBm	decibel-milliwatts
DL	downlink
FCC	Federal Communications Commission
GHz	gigahertz
IEEE 208.X	Institute of Electrical & Electronics Engineers Local Area Network Standards
IoT	Internet of Things
MHz	megahertz
mW	milliwatt
OFDMA	orthogonal frequency-division multiple access
QAM	quadrature amplitude modulation
QoE	Quality of Experience
QoS	Quality of Service
SCTE	Society of Cable Telecommunications Engineers
TDMA	time-division, multiple access
TWT	target wake time
UL	uplink
MDU	multi-dwelling unit
MU-MIMO	multi-user multiple input multiple output
VR	virtual reality
Wi-Fi HaLow	wireless networking standard IEEE 802.11ah
WFH	work from home
WLAN	Wireless Local Area Network

Bibliography & References

Wi-Fi Alliance Wi-Fi Predictions For 2021; Wi-Fi Alliance
https://www.wi-fi.org/news-events/newsroom/wi-fi-alliance-wi-fi-predictions-for-2021

Cyber-Intrusion Protection For The Smart Home, Susmita Nayak; Plume
https://blog.plume.com/cyber-intrusion-protection-for-the-smart-home

2021 Connectivity And Mobile Trends Survey; Deloitte
https://www2.deloitte.com/us/en/insights/industry/telecommunications/connectivity-mobile-trends-survey.html?id=us:2el:3dp:wsjspon:awa:WSJCIO:2020:WSJFY20

Smart Homes 2020: The End of Interruptive Marketing As We Know It, Victoria Petrock; eMarketer.com
https://www.emarketer.com/content/smart-homes-2020

Chairman Pai Statement On FCC Authorization Of First 6 GHz WI-FI Device; Federal Communications Commission
https://www.fcc.gov/document/chairman-pai-fcc-authorization-first-6-ghz-wi-fi-device

Smart Home Will Drive Third Wave In Wireless Home Evolution; Strategy Analytics
https://news.strategyanalytics.com/press-releases/press-release-details/2019/Smart-Home-Will-Drive-Third-Wave-in-Wireless-Home-Evolution-Strategy-Analytics/default.aspx

Work/School from Home Fuels 223 Million SOHO Consumer Wi-Fi CPE Shipments in 2020; ABI Research
https://www.abiresearch.com/press/workschool-home-fuels-223-million-soho-consumer-wi-fi-cpe-shipments-2020/

Cisco Annual Internet Report (2018–2023) White Paper; Cisco
https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

What Is The Status Of Global Wi-Fi 6E Efforts?, Catherine Sbeglia; RCR Wireless News
https://www.rcrwireless.com/20210120/network-infrastructure/wi-fi/what-is-the-status-of-global-wi-fi-6e-efforts

Wi-Fi Alliance Delivers Wi-Fi 6E Certification Program; Wi-Fi Alliance
https://www.wi-fi.org/news-events/newsroom/wi-fi-alliance-delivers-wi-fi-6e-certification-program

What You Should Know About Wi-Fi 6 And The 6-GHz Band; Test & Measurement Tips
https://www.testandmeasurementtips.com/what-you-should-know-about-wi-fi-6-and-te-6-ghz-band/

Synchronous Ethernet (SyncE) Usage for DAA and Mobile X-haul over DOCSIS

A Technical Paper prepared for SCTE by

Chris Zettinger
Principal Systems Engineer
CommScope
2400 Ogden Ave., Suite 180, Lisle, IL 60532
+1 630 281 3272
chris.zettinger@commscope.com

Yair Neugeboren
Sr. Manager, Architecture
NVIDIA
8 Habarzel St., Tel Aviv, Israel
+972 54 2205051
yneugeboren@nvidia.com

1. Introduction

Synchronous Ethernet (SyncE) is used mainly in telecom networks to provide physical layer clock delivery for frequency synchronization (a.k.a syntonization). The cable industry has not yet adopted SyncE as a common supported functionality as IEEE1588 is more commonly used for time synchronization (mainly for DAA use cases).

Adopting SyncE in conjunction or in addition to IEEE1588 may add advantages both from a performance perspective, and from a simplicity and cost perspectives as well.

This paper will describe different use cases for both DAA RPHY/RMACPHY and mobile backhaul (LTE and 5G) where SyncE can add a true value. Analysis of performance compared to traditional IEEE1588 usage will be shown and the benefits in simplicity will be discussed.

2. The Need for Synchronization

Traditional network synchronization has been based on the accurate distribution of frequency. Wireless networks have evolved to require the distribution of accurate time and phase information. In order to deliver this information, network operators have to distribute a reference timing signal of suitable quality to the network elements processing the application.

There are two basic approaches for distributing synchronization information. The first is to follow a distributed primary reference time clock (PRTC) approach, implementing a global navigation satellite system (GNSS) receiver in the end application, and the second is based on a master-slave hierarchical strategy.

Master-slave synchronization uses a hierarchy of clocks in which each level of the hierarchy is synchronized with reference to a higher level, the highest level being the Primary Reference Clock (PRC). Clock reference signals are distributed between levels of the hierarchy via a distribution network which may use the facilities of the network. The hierarchical strategy can be used for physical layer frequency distribution as well as higher layer packet-based frequency, phase, and time distribution.

Packet-based methods for distribution of frequency, phase, and time have been developed using precision timing protocol (PTP) and network time protocol (NTP). PTP is typically used instead of NTP in applications with stricter phase and time synchronization requirements. The improved accuracy for PTP is a result of its hardware-based timestamping.

2.1. Synchronous Ethernet (SyncE) Overview

The ITU has defined mechanisms to use the Ethernet physical layer to distribute frequency information across a network that are similar to the physical layer methods used with synchronous digital hierarchy (SDH)-based network synchronization. ITU-T G.8261 provides the network limits for transferring timing across a packet network. G.8262 defines the timing characteristics of an Ethernet Equipment Clock, while G.8262.1. defines improvements for the Enhanced Ethernet Equipment Clock.

Synchronous Ethernet (SyncE) uses the edges in the Ethernet data signal to define the timing content of the signal and distribute a physical layer clock across a packet network. Each system recovers and forwards the network timing through the distribution path. A reference timing signal traceable to a PRC provides a reference timing signal to an external timing port (e.g., BITS, GPS) on the first Ethernet switch in the path. The system clock function of the switch synchronizes its Ethernet transmit bit stream to the

reference timing signal. Each subsequent Ethernet switch in the path recovers the timing from the incoming Ethernet bit stream on a designated port, and synchronizes its Ethernet transmit bit stream to the recovered timing signal via its system clock function. Like all physical layer frequency synchronization techniques, all network elements between network segments need to be capable of recovering and passing the frequency downstream.

A system clock function that supports SyncE timing contains an Ethernet Equipment Clock (EEC) or a physical layer clock. The EEC provides filtering of noise on the external reference or recovered timing signals. A holdover function is also provided to allow continued operation in the event of a failure in the distribution path.

ITU-T G.8264 adds support for an Ethernet Synchronization Message Channel (ESMC) based on the IEEE 802.3, Organization Specific Slow Protocol (OSSP). This message is used to communicate the clock quality levels (via a Synchronization Status Message, SSM) of the PRC to all EECs in the network. The EEC may use this received information to determine the appropriate actions to take, such as failover to a different SyncE port or enter holdover mode, when network failures occur. As an example, if an EEC loses its connection to the PRC and transitions to a holdover state, it will replace the Quality Level (QL) of the PRC in its outgoing SSMs with the QL of its internal oscillator.

The higher information rate and lower noise of SyncE enables a higher clock bandwidth than for a packet-based clock. The bandwidth of the clock PLL in an EEC as defined in G.8262 is in the range of 1 - 10 Hz. The clock bandwidth for an enhanced EEC as defined in G.8262.1 is in the range of 1 - 3 Hz. The higher bandwidth generally allows for a faster convergence.

The filtering associated with an EEC defined in G.8262 is sufficient for the cable network use cases identified and do not require the tighter filtering associated with enhanced EEC in G.8262.1.

2.2. Precision Timing Protocol (PTP) Overview

IEEE 1588, Precision Timing Protocol (PTP) is a standard for enabling precise synchronization of real-time clocks for devices in communications networks with system-wide synchronization accuracy in the sub-microsecond to micro-second range.

PTP relies on the transmission of dedicated packets that form the significant instants of a packet timing signal. The timing of these significant instants is precisely measured relative to a master time source, encoded in the form of a time stamp, and distributed to a packet slave clock.

PTP timing synchronization starts with a Grandmaster Clock which typically derives its time from a Primary Reference Clock (PRC) such as a GPS Receiver. Timing synchronization then propagates across the network through the exchange of PTP messages between each Master and Slave, allowing each Slave node to synchronize to the PTP timing reference provided by its Master.

There are three 1588 profiles defined by the ITU-T. G.8265.1 supports frequency synchronization, while G.8275.1 and G.8275.2 both support both frequency and phase synchronization.

G.8275.1 profile is defined with full timing support from the network, so all network elements are PTP aware and participating in the protocol. The message rate for this profile is fixed at 16 messages per second. Only two-way communication is supported in this profile. The current version of the specification requires the use of a physical layer clock, like SyncE. A version of the profile that does not include SyncE is considered for further study. R-DTI mentions this as an alternate PTP network profile based on full

timing support in the network. G.8273.2 defines the performance parameters for boundary clocks supporting this profile. The PTP clock in a T-BC defined in G.8273.2 has a bandwidth of 0.05 - 0.1 Hz.

G.8275.2 is defined with only partial timing support from the network, so not all network elements need to be aware and participating in the PTP protocol. PTP messages that transit a non-participating network element are subject to the additional delay variation associated with queuing of traffic before transmission. This can degrade performance. The profile provides flexibility in PTP message rates, from 1 message per second to 128 messages per second. The higher message rates allow support for applications that require higher accuracy and lower rates can be used for less demanding applications. Slower message rates also reduce the load on the processor on the master clock. Only two-way communication is supported in this profile. Because [R-DTI] adopted G.8275.2 as the profile to be used for R-PHY, it may be used for the RMD as well, assuming only frequency information is used. G.8273.4 defines the performance parameters for boundary clocks supporting this profile.

A PTP clock hierarchy consists of multiple devices that are synchronized within the same Time Domain. PTP timing synchronization starts with a Grandmaster Clock (GMC) which derives its time from a Primary Reference Clock (PRC) such as a GPS Receiver. Timing synchronization then propagates across the network through the exchange of PTP messages between each Master and Slave, allowing each Slave node to synchronize to the PTP timing reference provided by its Master.

PTP defines classifications for clock functionality, including Ordinary Clocks and Boundary Clocks. An Ordinary Clock (OC) is a PTP clock with a single PTP port. An OC may be either a Master Clock or a Slave Clock. An Ordinary Master clock cannot be slaved to another PTP clock. A Boundary Clock (BC) refers to a PTP clock with multiple PTP ports. A BC can synchronize to a Master clock via its Slave PTP port and serve as a Master via its Master PTP port to one or more Slave clocks.

2.3. Hybrid Mode

A node that uses SyncE with an EEC in combination with a timing protocol such as PTP or NTP with a Packet Equipment Clock (PEC) may be referred to as a hybrid EEC/PEC clock. The source of SyncE within a network should be generated from same source as the PTP domain (e.g., GPS or PRTC) to avoid frequency drift.

When SyncE, with its higher level of frequency accuracy and lower level of noise, is used with a timing protocol such as IEEE 1588, a higher level of timing accuracy is achievable across a packet network. Since SyncE is a physical layer clock, it is not subject to packet delay variation as the signal transits the network.

G.8273.2 provides the functional requirements for a telecom boundary clock or telecom time slave clock when used with full timing support from the network. The following figure from G.8273.2 provides an illustration of the clock function in a hybrid EEC/PEC clock.

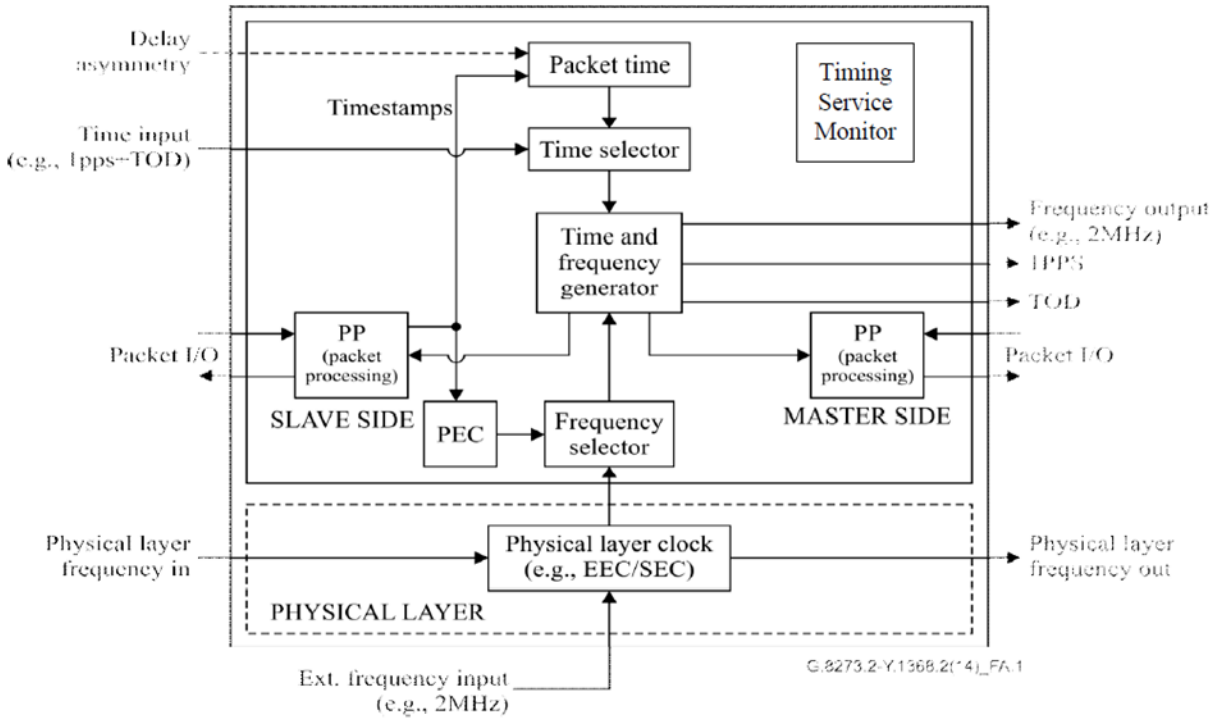


Figure 1 – Telecom Boundary Clock Model [G.8273.2]

The time information carried in the timestamps is used to establish the local time scales. The frequency information carried in the timestamps is used in the PEC to generate the local frequency.

The frequency selector block may select either the frequency information recovered from the timestamps, or the frequency recovered from a physical layer clock (e.g., Synchronous Ethernet).

3. Cable Network Timing

Timing and synchronization requirements for cable networks come from areas including existing DOCSIS specification requirements, Modular Headend Architecture v2 system requirements, and support of precision timing services, like Mobile Backhaul and other Mobile X-haul use cases.

In the MHAv2 architecture, the CMTS Core and the R-PHY are two entities located in separate locations. The DS PHY and US PHY are located in the R-PHY Device, and the DOCSIS MAC is located at the CMTS Core. Frequency and phase synchronization are required between the CMTS Core and the R-PHY Device so that they have a common knowledge of the DOCSIS time and to allow correct burst reception and to align the MAC scheduler with the PHY timestamping at the R-PHY. DOCSIS services require both frequency and phase synchronization while other services only require frequency synchronization, including Video (Sync mode), OOB (55-2 Sync mode), NDF/NDR, or Leakage Detection Signal Generation.

Certain MAC-NE (RMD) services do not require any external timing synchronization. These services can include DOCSIS services, Video (Async mode), or OOB (55-1 and/or 55-2 Async mode). Other services require frequency synchronization of the MAC-NE (RMD to an external frequency source. These services can include Video (Sync mode), OOB (55-2 Sync mode), NDF/NDR, or Leakage Detection Signal Generation.

The Mobile Backhaul application connects the mobile switching core (i.e., evolved packet core, EPC, for LTE or next-generation core, NGC, for 5G) and the radio access network (RAN) node. Long-Term Evolution (LTE) Frequency-Division Duplex (FDD) requires frequency synchronization between neighboring cells. Long-Term Evolution (LTE) Time-Division Duplex (TDD) requires additional phase synchronization between neighboring cells.

3.1. RPD Use Case

Remote PHY Device (RPD) is one type of a Distributed Access Architecture (DAA) that separates the Integrated CCAP into a CCAP Core that remains in the operator's headend and an RPD that resides in a remote fiber node or remote shelf. The RPD provides all PHY-related circuitry needed for HFC high-speed data and video services, including DOCSIS packet timestamping. The CCAP Core contains the DOCSIS MAC and the upper layer DOCSIS protocols, including all signaling functions, downstream and upstream bandwidth scheduling, and DOCSIS framing. It also contains all the video processing functions that an EQAM provides.

The CCAP Core and RPD synchronize their DOCSIS clocks in both frequency and phase so that they have a common view of DOCSIS time. This common view is enabled by Remote DOCSIS Timing Interface (R-DTI), which requires support for IEEE 1588v2 precision timing protocol (PTP). R-DTI further requires G.8275.2 with G.8275.1 and SyncE identified as being optionally supported.

The RPD also provides timing synchronization to subtended cable modems. Frequency synchronization between the RPD and the cable modems is achieved through the DOCSIS symbol rate, and DOCSIS time stamping [via Sync messages in DOCSIS 3.0 or OFDM preamble in DOCSIS 3.1] used to provide relative clock phase synchronization.

[R-DTI] identifies the timing architecture where the Core and the RPD are slaved to an external timing master.

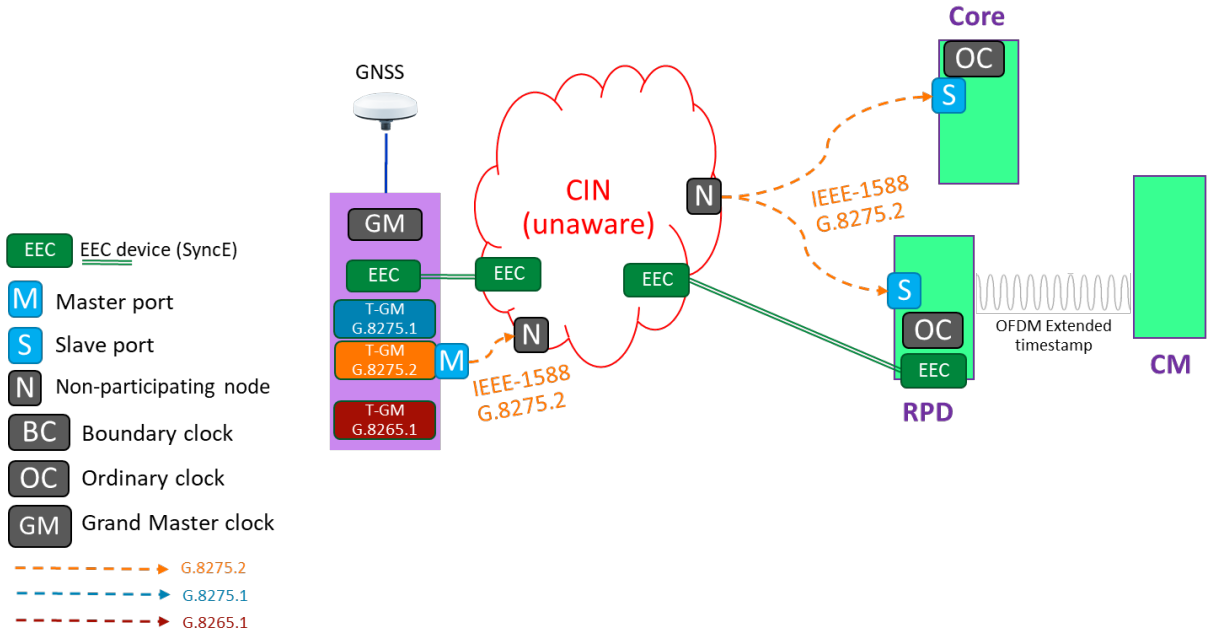


Figure 2 – RPD Use Case

External timing support is needed by the R-PHY application for the following features:

- DOCSIS services (phase and frequency sync)
- Video sync mode (frequency sync)
- Signal leakage detection (frequency sync)
- 55-2 sync mode (frequency sync)
- NDF/NDR (frequency sync)

Frequency synchronization for video sync mode avoids the need to support MPEG PCR restamping.

A summary of timing requirements from [R-DTI] for an RPD in R-PHY mode are listed below.

- The RPD MUST support 1588 OC slave
- The RPD MUST meet time/phase synchronization accuracy of ± 1 ms in reference to the 1588 GM.
- The RPD MUST comply with the T-TSC-P requirements of [G.8275.2].
- The RPD SHOULD syntonize the frequency of the local NDF/NDR clock with the assistance of the frequency recovered from Ethernet.
- The RPD MUST meet the frequency accuracy requirements (≤ ±530 ppb) specified by [R-OOB] when supporting signal leakage detection.
- The RPD SHOULD support Synchronous Ethernet for Precision Timing Services.
  Note: Precision Timing Services is a term used in [R-DTI] to describe Mobile BackHaul for wireless applications. Synchronous Ethernet is not required and is not even classified as a "truly optional" feature in [R-DTI] for other services, including DOCSIS services.

A summary of timing requirements in [D-RFI] are listed below.

- The 10.24 MHz Master Clock MUST have:
  - Frequency accuracy ≤ ±5 ppm
  - Maximum drift ≤ 10-8 per second (≤ ±.01 ppm)
  - An edge jitter of <=10 ns peak-to-peak (±5 ns)
  - DOCSIS timestamp jitter < 500 ns p-p.

The system clock in an RPD with SyncE support will typically consist of two components, a physical layer clock or EEC that uses SyncE as its timing reference and a packet equipment clock or PEC that uses PTP as its timing reference. The two timing references are normally traceable to a common frequency source since in DOCSIS the frequency and phase are coupled together. The output of the EEC portion may be used to assist and accelerate the PEC portion in achieving lock to the PTP timing reference. A node that uses SyncE with an EEC in combination with a timing protocol such as PTP with a PEC may be referred to as a hybrid EEC/PEC clock.

There are three timing related performance issues that exist in typical RPD implementations that can be improved by the use of a hybrid EEC/PEC clock system. The first is that CIN networks may introduce large PDV that can affect the frequency and phase servo algorithm convergence and accuracy of the RPD. The second is that typical RPD holdover performance is limited while there isn't a holdover specification defined for an RPD in R-DTI. The third is that services can take a while to be restored during initialization of an RPD since they rely on timing lock notification from the RPD which can last a non-negligible amount of time.

Another optional benefit of using SyncE is the ability to use NTP as the time of day (TOD) source instead of PTP for the RPD. Using NTP will eliminate the need to install a pricier PTP GM, since the frequency is already locked via SyncE. This is a subject for more detailed further study.

3.1.1. Benefit #1: Time accuracy can be improved

R-DTI defines the time/phase synchronization accuracy of an RPD to be within ±1 ms when referenced to the 1588 GM for DOCSIS Timing. R-DTI also requires support for the PTP network profile defined in G.8275.2. This profile is defined with partial timing support from the network. Since not all network elements need to be PTP aware, packet delay variation will be imposed on the PTP messages in both directions by each non-participating network element in the timing distribution chain. Long chains of non-participating elements accumulate the packet delay variation and degrade the time accuracy. The more stable and accurate frequency associated with the SyncE timing distribution and the physical layer clock assisting the packet equipment clock can result in lower output frequency and phase noise on the output signals from network element and enable improved time accuracy over using PTP alone.

Assuming SyncE and PTP are traceable to the same PRC, the servo algorithm for PTP could be simplified to lock to phase while frequency is already locked.

3.1.2. Benefit #2: Improved holdover performance

A TCXO is normally be used in RPD applications to meet the timing requirements defined in R-DTI without the increased power and cost associated with other technologies, like OCXOs. When the packet timing reference from the 1588 GM is interrupted, the RPD's clock will transition into holdover and the RPD's clock output will drift according to the characteristics of the TCXO and the quality of the timing reference frequency estimate prior to the holdover event. While R-DTI does not have an explicit holdover specification, an implied specification can be derived from the time/phase synchronization accuracy of ±1 ms and the most restrictive frequency accuracy requirements in [R-DTI] which is ±530 ppb. Currently, the holdover-out-of-specification notification is a vendor-specific implementation which relies on estimates of worst case frequency drift scenarios. The phase drift of a clock x(S) can be calculated using the following equation.

$$|\Delta x(S)| \leq \{(a_1 + a_2)S + 0.5bS^2 + c\}[ns]$$

Where:

a_1 represents an initial frequency offset during the entry into holdover

a_2 accounts for temperature variations after the clock went into holdover

b corresponds to the frequency drift caused by aging

c accounts for any additional phase shift during the entry into holdover

The following table provides some representative values for the phase drift coefficients for the equation above for a clock using TCXO technology and a clock using OCXO technology.

Table 1 – Representative Phase Drift Coefficient Values for a Clock

	Units	TCXO-based	OCXO-based
a_1	ns/s	10	10

	Units	TCXO-based	OCXO-based
a_2	ns/s	200	20
b	ns/s^2	5×10^{-5}	2×10^{-5}
c	ns	10	10

A TCXO meeting the holdover specifications in G.8262 or G.8262.1 will remain within this ±1 ms level for roughly an hour and 20 minutes. Replacing the TCXO with an OCXO extends the time the RPD would remain within the ±1 ms accuracy to roughly 1 day (25 hours), but it does come with additional cost and power associated with the OCXO. By using SyncE and a physical layer clock to enhance the performance of the RPD's packet equipment clock and providing a stable frequency reference that is traceable to the 1588 GM, the time output from the RPD will continue at the correct rate and maintain the correct time for an almost indefinite amount of time. This period of time is primarily limited by failures in the SyncE timing distribution network that interrupt the traceability to the PRC. These values are shown in the following diagram.

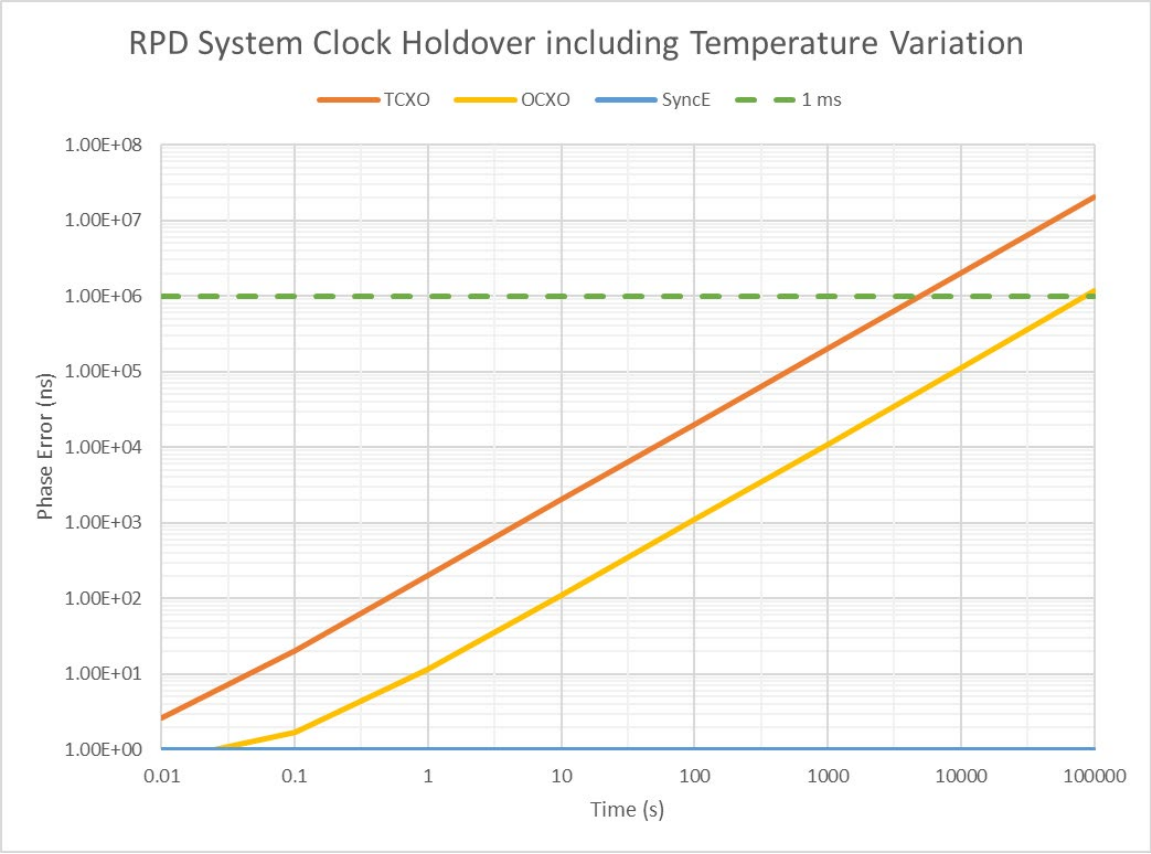


Figure 3 – RPD System Clock Holdover with No Initial Frequency Offset

Since R-DTI contains a frequency accuracy requirement of ±530 ppb, an RPD which is in frequency locked state might go into holdover with an additional initial frequency offset of ±530 ppb. This will cause the RPD to cross the ±1 ms threshold much faster. RPDs with either a TCXO or an OCXO will remain within this threshold for roughly 25-35 minutes. By using SyncE and a physical layer clock, the time output from the RPD will continue at the correct rate and maintain the correct time for an almost indefinite amount of time. These values are shown in the following diagram.

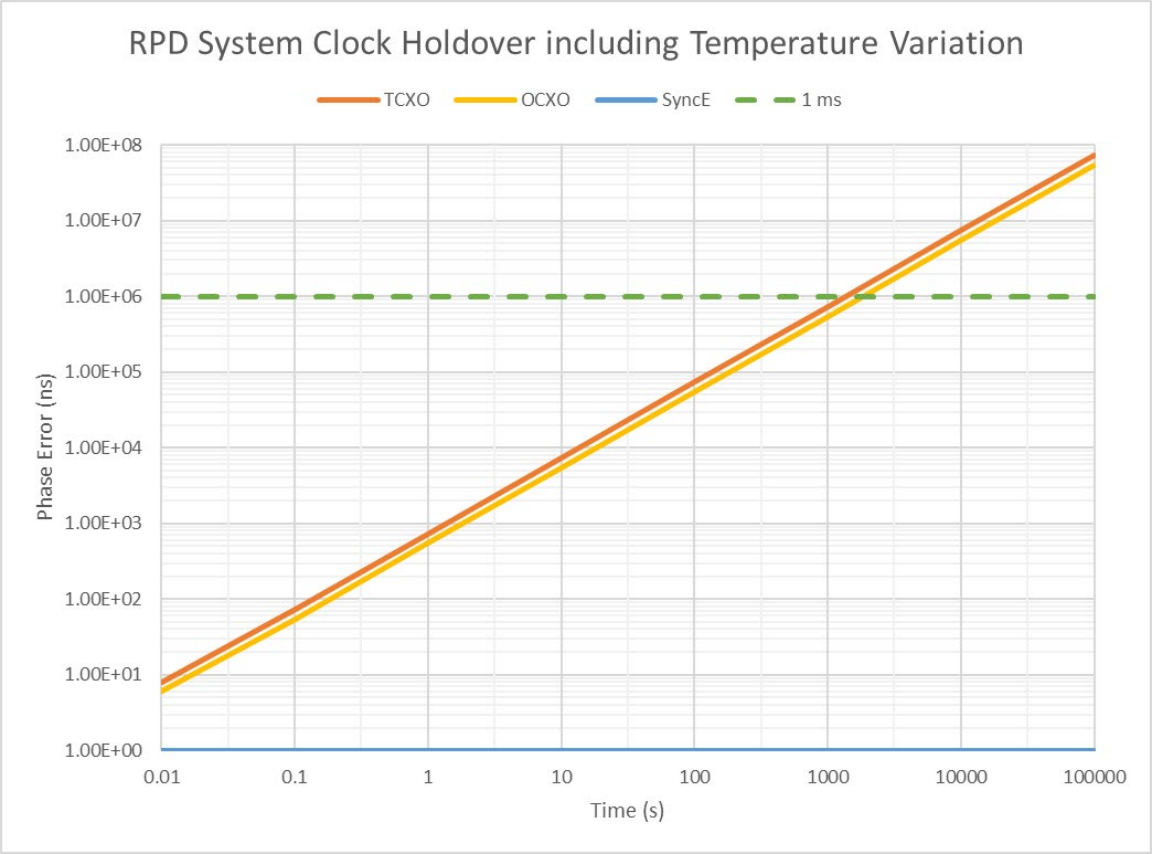


Figure 4 – RPD System Clock Holdover with 530 ppb Frequency Offset

3.1.3. Benefit #3: Reduced time to lock and recover after reset

SyncE as a timing reference to the physical layer clock that assists the packet equipment clock in the RPD's system clock can shorten the time for certain services to be available following an RPD reset.

As required in [D-RFI], an RPD receives its timing reference via PTP. The packet equipment clock is the source for all downstream timing. When exiting reset, the RPD must wait until the packet equipment clock achieves frequency lock and phase lock before it can begin to provide services, even services that only require frequency synchronization, like video and OOB services. Frequency lock can be achieved in a shorter amount of time than the time required to achieve phase and frequency lock. Phase and frequency lock with the ±1 ms specified in R-DTI can take multiple minutes.

Techniques can speed the acquisition of phase and frequency lock for the RPD's packet equipment clock after a reset. A soft reset with warm start relies on frequency information about the operating conditions of the RPD's packet equipment clock prior to the reset. Warm start is most effective if the reset time is brief and network conditions are stable during the reset time. If the conditions are not suitable, then a hard reset will be required. A hard reset is similar to a power-up reset in that it does not rely on information about the packet equipment clock's prior operating and can take longer to achieve lock. [R-PHY] provides an overview of soft reset and warm start functionality.

Reducing the time-to-lock for an RPD system clock using SyncE assistance is partially due to the difference in PLL bandwidths between the two portions of the system clock. The bandwidth of the SyncE physical layer clock PLL is in the range of 1 - 10 Hz for the non-enhanced case defined in G.8262 and in the range of 1 - 3 Hz for the enhanced case defined in G.8262.1. In contrast, the PTP clock in a T-BC defined in G.8273.2 has a bandwidth of 0.05 - 0.1 Hz. While the frequency locking process is non-linear, the difference in bandwidths between the SyncE physical layer clock and PTP clock should result in the SyncE clock obtaining frequency lock at least ten times faster than the PTP clock. Additionally, since the frequency lock is dependent on the SyncE timing reference, a larger amount of PDV can be tolerated on the PTP timing reference without significantly increasing the lock time. Frequency-based services using an RPD clock with SyncE assistance should be available in less than one tenth of the time for the same services to be available for an RPD using only a packet timing clock.

The following table shows some representative time frames for achieving frequency lock after exiting reset. The simplifying assumption is that the time to obtain phase lock after frequency lock is achieved is constant for each of the three cases.

Table 2 – Frequency Lock

	Hard Reset	Soft Reset / Warm Start	Hard Reset with SyncE Assist
Frequency lock achieved	minutes, depends on network PDV	within a minute	within seconds

Because of the faster frequency lock that results from the physical layer clock assistance, the RPD can begin to deliver video services while phase lock is still being acquired by the packet equipment clock portion of the RPD system clock. DOCSIS services may be delayed until phase lock is achieved by the packet equipment clock portion.

3.2. RMD Use Case

The Remote MACPHY Device (RMD) is a Distributed Access Architecture (DAA) device that contains all PHY-related circuitry needed for HFC high-speed data and video services as well as the DOCSIS MAC and the upper layer protocol support. This includes all DOCSIS signaling functions, downstream and upstream bandwidth scheduling, and DOCSIS framing. It also provides the digital interface to an Auxiliary Core for MPEG video and OOB signals as well as the analog interface for transmission over RF or linear optics.

Since the DOCSIS MAC and the PHY are collocated within the device, the RMD application does not have the complexity associated with frequency or phase synchronization between two devices for DOCSIS operation that the RPD requires. Non-DOCSIS functions, including some video and OOB functions, require the RMD to be frequency synchronized to the Auxiliary Cores that provide those functions. The Auxiliary Cores could be free running with an internal clock that meets the frequency accuracy specification for the specific application (e.g. ±30 ppm for video) or be synchronized to the PRC using technologies including SyncE, PTP, NTP, GNSS, BITS, etc. This allows the timing synchronization requirements of the RMD to be less strict than the requirements for the RPD.

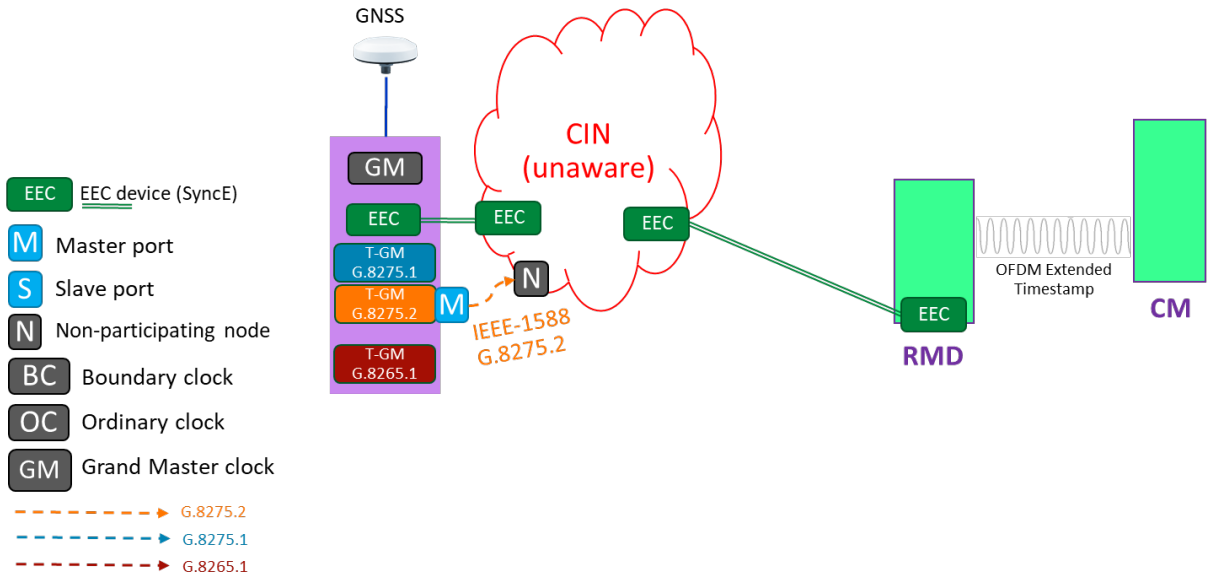


Figure 5 – RMD Use Case

Without a requirement for phase synchronization in the RMD application, there are four options for timing the RMD: free-run, frequency synchronized using the PTP profile defined in G.8265.1, frequency synchronized using SyncE, and frequency and phase synchronized using the PTP profile defined in G.8275.2. Using G.8275.2 is consistent with the requirements for an RPD in [R-DTI].

External timing support is needed by the RMD application for the following features:

- Video sync mode (frequency sync)
- Signal leakage detection (frequency sync)
- 55-2 sync mode (frequency sync)
- NDF/NDR (frequency sync)

For RMD applications that require external timing, frequency synchronization may be provided via SyncE, or by PTP. The timing reference signal in the two PTP options will have a higher amount of noise as a result of the packet delay variation on the PTP messages in the timing distribution network. The higher rate of significant instants in the SyncE timing reference and the lower amount of noise in the SyncE timing distribution network allow for a wider bandwidth of the physical layer clock in the RMD system clock function. This enables a faster lock time and faster restoral of service after a reset for a SyncE-based or SyncE-assisted RMD system clock over a PTP-only timing approach. In contrast to the RPD application, the faster restoral of service when using SyncE for frequency synchronization applies to both DOCSIS and non-DOCSIS services.

An RMD typically provides multiple Ethernet interfaces. The interfaces provide redundant SyncE timing references from the PRC to the RMD and allow the physical layer clock function to monitor the quality of the signal on the interfaces and select the better quality one, if appropriate.

The less complicated filtering and monitoring functions associated with SyncE timing references for the physical layer clock allows functions to be implemented in hardware and reduces the number of functions that need to be implemented in the CPU of the RMD. This reduces the load for timing functions on the

RMD's CPU and allows additional non-timing functions to be implemented in the same class of RMD's CPU.

3.3. MBH Use Case

A mobile backhaul (MBH) network connects the mobile switching core (i.e., evolved packet core, EPC, for LTE or the 5G core) and the radio access network (RAN) node. The SYNC specification [SYNC] describes the architecture and requirements to enable cable operators to use DOCSIS technology to carry precision frequency and phase synchronization signals over the hybrid fiber-coax (HFC) plant. This allows cable operators to take advantage of their rich infrastructural assets to provide backhaul services comparable to fiber for their own mobile traffic.

Although many wireless networks previously required only frequency synchronization, including LTE Frequency-Division Duplex (FDD), other networks require time and phase synchronization. Long-Term Evolution (LTE) Time-Division Duplex (TDD) is an example of a technology that requires frequency and phase synchronization.

3.3.1. Frequency Synchronization

From [SYNC], networks that require only frequency synchronization have a target performance for frequency accuracy of ±16 ppb [G.8261.1]. This is based on the requirement of a ±50 ppb radio frequency accuracy.

3.3.2. Time Error budget

[SYNC] provides a comprehensive view of the end-to-end performance budget of 1.5 µs for maximum time error for networks that require frequency synchronization and phase synchronization. One section covers the DOCSIS portion of the network, and specifically the portion of the budget allocated to CMTS and DAA equipment. The performance target for constant time error of the CMTS, RPD, or RMD is ±200 ns for class A equipment and ±100 ns for class B equipment. The jitter value in the CMTS, RPD, RMD component is been reduced to 5-10 ns (1/204.8 MHz plus phase noise jitter) as a result of using OFDM-based DOCSIS systems.

DOCSIS Time Protocol (DTP) allows IEEE 1588 protocol information to be passed over the DOCSIS network with high frequency and phase accuracy by eliminating the jitter resulting from network buffering in the DOCSIS network. DTP takes advantage of the fact that DOCSIS is a synchronous system and distributes a physical layer clock to distribute frequency information to the cable modems. It also uses the DOCSIS 3.1 timestamp to distribute time information. A signaling path determines the downstream timing offset, which is used as a correction factor for PTP.

3.3.3. Mobile Backhaul Use Cases

Four synchronization use cases for mobile backhaul are identified in [SYNC].

- Physical Layer Timing Support for Frequency Synchronization
- Full Timing Support for Phase Synchronization
- Partial Timing Support for Phase Synchronization
- Partial Timing Support for Frequency Synchronization

The current version of [SYNC] identifies requirements associated with Physical Layer Timing Support for Frequency Synchronization and Full Timing Support for Phase Synchronization. The other two use cases will be addressed in a future version of the document.

3.3.4. Physical Layer Timing Support for Frequency Synchronization

For Physical Layer Timing Support for Frequency Synchronization, the RPD or RMD is synchronized to the PRC using SyncE, and the cable modem provides the frequency reference to the end application using SyncE.

The RPD uses PTP to align its frequency and phase for R-PHY operation and generates the downstream DOCSIS frequency and timestamp traceable to the PTP clock domain. For the downstream clock to be SyncE-traceable, the input SyncE and PTP should be generated from the same source. The RPD operates in Hybrid mode so that SyncE may be used to assist the PTP clock for functions like holdover and fast-lock. The frequency source of SyncE should share a common frequency source with the PTP domain (e.g., GPS or PRTC) since DOCSIS couples the frequency and phase together and to avoid frequency drift between the two domains. Cores usually only support G.8275.2 as defined in [R-PHY]. Supporting different profiles for RPD and Core will require different GMs.

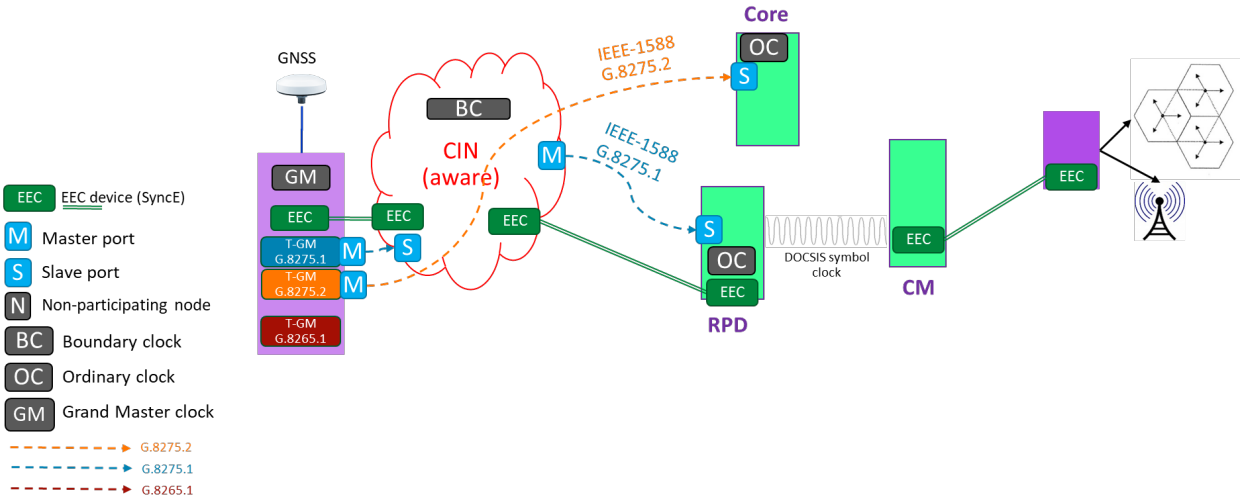


Figure 6 – RPD Physical Layer Timing Support for Frequency Synchronization Use Case

The RMD's system clock is locked in frequency via SyncE. Without a requirement for phase synchronization in this application, the RMD may use an arbitrary time of day or use NTP to establish the system's time of day. The lower amount of noise on the SyncE timing reference allow for a wider bandwidth of the physical layer clock in the RMD system clock function. This enables a faster lock time, faster restoral of service after a reset and more accurate holdover for a SyncE-based RMD system clock over a PTP-only timing approach.

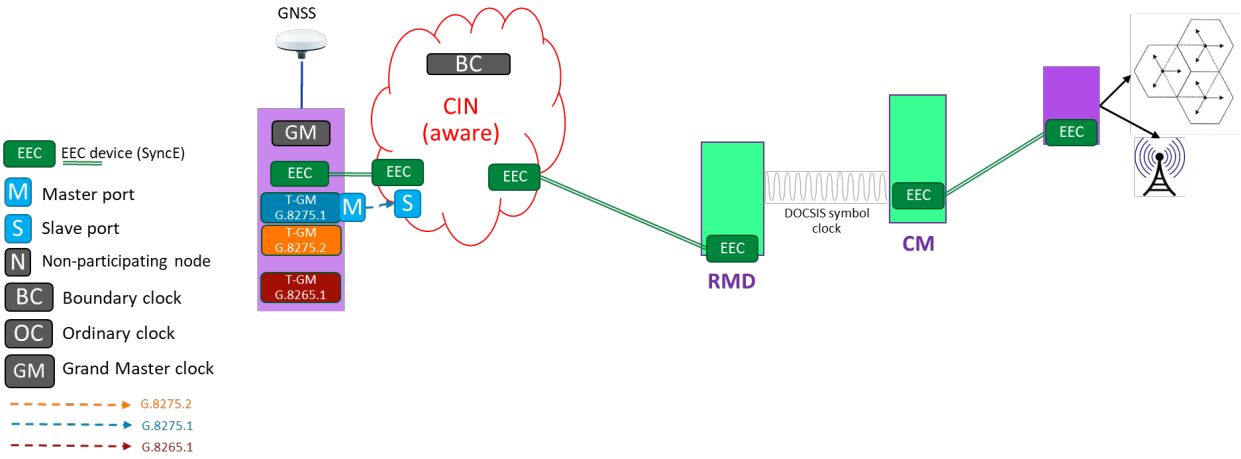


Figure 7 – RMD Physical Layer Timing Support for Frequency Synchronization Use Case

3.3.5. Full Timing Support for Phase Synchronization

Full Timing Support for Phase Synchronization in [SYNC] requires the DOCSIS interworking function to support the network profile defined in G.8275.1 with each switch in the network having PTP awareness (such as Ordinary Clock or Boundary Clock). It also requires the use of a physical layer clock, like SyncE, at each element in the network.

R-PHY has timing requirements specified in [R-DTI] that are different from the timing requirements for MBH support, i.e., the timing accuracy requirements (for frequency and phase) and the timing delivery requirements (profile, 1588/SyncE, etc.). The challenge is to have the R-PHY Device (RPD) clock support the requirements for two timing applications when the set of requirements for one application is not a superset of the other.

The following table from [SYNC] contrasts the requirements for an RPD operating in a MBH application and an R-PHY application.

Table 3 – R-PHY Requirements vs. MBH "Full Timing Support for Phase"

Item	R-PHY	MBH	R-PHY requirements when supporting MBH precision timing services
Lock threshold	1 ms	50 ns	See Section C.2.1 of [SYNC]
Phase lock time	Few minutes	Not defined (should be short)	See Section C.2.2 of [SYNC]
Profile	[G.8275.2]	[G.8275.1]	See Section C.2.3 of [SYNC]
SyncE	Recommended	Required	See Section C.2.4 of [SYNC]
Holdover requirements and duration	Not specified	[G.8273.2]	See Section C.2.5 of [SYNC]
Phase steps	Not allowed when locked	Not specified	See Section C.2.6 of [SYNC]

Item	R-PHY	MBH	R-PHY requirements when supporting MBH precision timing services
Frequency change rate	10 ppb/s	Not specified	See Section C.2.6 of [SYNC]
Precision timing frequency and phase budget	Table based on old DTP section	Specified in Section 8 of [SYNC]	See Section C.2.7 of [SYNC]
"soft reset" support	The RPD holds in holdover for <1 min during a soft reset in order to have a quick reset using a warm start. The RPD goes operational even before it re-locks to the GM.	Convergence after soft reset is for further study	See Section C.2.8 of [SYNC]
BC functionality	Allowable	Required on every Ethernet hop	See Section C.2.9 of [SYNC]

The RPD uses PTP to align its frequency and phase and generates the downstream DOCSIS frequency and timestamp traceable to the PTP clock domain. For the downstream clock to be SyncE-traceable, the input SyncE and PTP should be generated from the same source. The RPD operates in Hybrid clock mode so that SyncE may be used to assist the PTP clock for functions like holdover and fast-lock. The frequency source of SyncE should share a common frequency source with the PTP domain (e.g., GPS or PRTC) to avoid frequency drift between the two domains.

DTP functions are distributed between the RPD and the Core. Cores usually only support G.8275.2 as defined in [R-PHY]. Supporting different profiles for RPD and Core will require different GMs. PTP announce message delivery is from the RPD to Core via UEPI to forward to the CMs. Core needs to encapsulate the PTP announce messages on the UEPI PW from the RPD and forward to all CMs on relevant DSIDs.

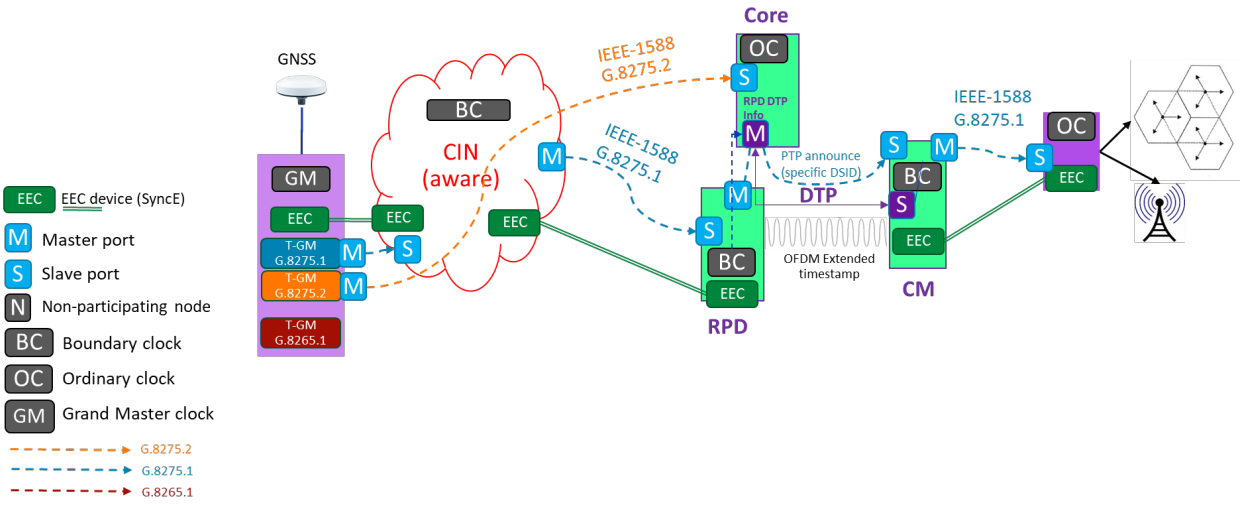


Figure 8 – RPD Full Timing Support for Phase Synchronization Use Case

The RMD's system clock is locked in frequency and phase via PTP using G.8275.1 with SyncE assistance. The lower amount of noise on the SyncE timing reference allow for a wider bandwidth of the physical layer clock in the RMD system clock function. This enables a faster lock time, faster restoral of service after a reset and more accurate holdover for a SyncE-based RMD system clock over a PTP-only timing approach. The lower noise on the system clock as a result of the SyncE assistance allows for lower noise and better accuracy of the timestamps.

All DTP functions are co-located in the RMD so there isn't the complexity associated with coordinating and synchronizing between two devices.

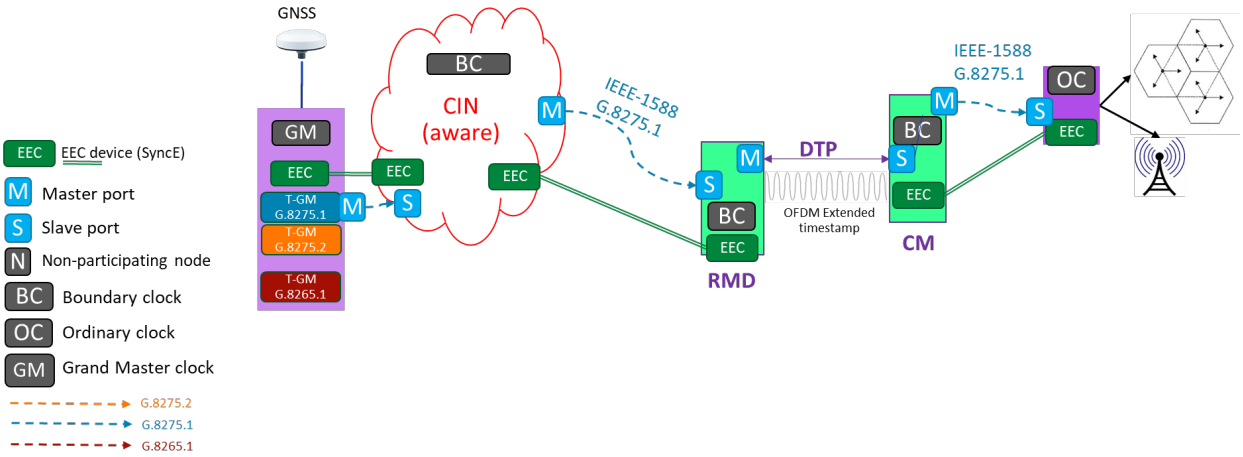


Figure 9 – RMD Full Timing Support for Phase Synchronization Use Case

4. Conclusion

Synchronous Ethernet (SyncE) is used mainly in telecom networks to provide physical layer clock delivery for frequency synchronization (aka syntonization). The cable industry has not yet adopted SyncE as a common supported functionality while IEEE1588 is more commonly used for time synchronization (mainly for DAA use cases).

The paper investigated the impact of providing SyncE support in DAA R-PHY and R-MACPHY applications as well as two use cases for mobile backhaul (LTE and 5G) applications. The RMD implementations provided additional benefits over the RPD implementations for each use case due to the integrated DOCSIS MAC and PHY functions in a single system and simplified timing interface. SyncE provided multiple benefits over the traditional PTP-only based implementations in these applications including:

- Improved timestamp accuracy
- Improved holdover performance
- Reduced time to lock and reduced time to recover services after reset

These improvements are realized without additional complexity. As a result, it is clear that SyncE is a valuable functional addition to the cable network.

Since converged interconnect networks (CINs) are generally new networks, requiring SyncE support for equipment in the CIN is a reasonable requirement. Even in case where a portion of the network does not support SyncE, there are reasonably priced boundary clocks that can be plugged into the CIN equipment to provide SyncE support.

Abbreviations

BC	Boundary Clock
BITS	Building Integrated Timing Supply
CCAP	Converged Cable Access Platform
CIN	Converged Interconnect Network
CMTS	Cable Modem Termination System
DAA	Distributed Access Architecture
D-RFI	Downstream Radio Frequency Interface
DTP	DOCSIS Time Protocol
DS	Downstream
EEC	Ethernet Equipment Clock
E-EEC	Enhanced Ethernet Equipment Clock
FDD	Frequency Division Duplex
GM	Grandmaster
GMC	Grandmaster Clock
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HFC	Hybrid Fiber Coax
LTE	Long Term Evolution
MBH	Mobile Backhaul
MHA	Modular Headend Architecture
MHz	Megahertz
MPEG	Moving Picture Experts Group
NTP	Network Time Protocol
OC	Ordinary Clock
OCXO	Oven Controlled Crystal Oscillator
OOB	Out of Band
PEC	Packet Equipment Clock
PRC	Primary Reference Clock
PRTC	Primary Reference Time Clock
PTP	Precision Time Protocol
RAN	Radio Access Network
R-DTI	Remote DOCSIS Timing Interface
RMD	Remote MACPHY Device
RPD	Remote PHY Device
SyncE	Synchronous Ethernet
TDD	Time Division Duplex
TDD	Time Division Duplex
TOD	Time of Day
T-TSC-P	Telecom Time Slave Clock for Partial timing support
US	Upstream

X-haul	Fronthaul, Midhaul, and Backhaul applications

Bibliography & References

[DRFI] *Downstream RF Interface Specification, CM-SP-DRFI-I16-170111, January 11, 2017, Cable Television Laboratories, Inc.*

[FMA] *Flexible MAC Architecture (FMA) System Specification, CM-SP-FMA-SYS-I02-210526, May 26, 2021, Cable Television Laboratories, Inc.*

[G.8261] *ITU-T Recommendation G.8261, Timing and synchronization aspects in packet networks, August 2019.*

[G.8262] *ITU-T Recommendation G.8262, Timing characteristics of synchronous Ethernet equipment slave clock, November 2018.*

[G.8262.1] *ITU-T Recommendation G.8262.1, Timing characteristics of enhanced synchronous Ethernet equipment slave clock, January 2019.*

[G.8265.1] *ITU-T Recommendation G.8262.1, Precision time protocol telecom profile for frequency synchronization, April 2016.*

[G.8273.2] *ITU-T Recommendation G.8273.2, Timing characteristics of telecom boundary clocks and telecom time slave clocks, August 2019.*

[G.8275.1] *ITU-T Recommendation G.8275.1, Precision time protocol telecom profile for phase/time synchronization with full timing support from the network, June 2016.*

[G.8275.2] *ITU-T Recommendation G.8275.2, Precision time protocol telecom profile for phase/time synchronization with partial timing support from the network, June 2016.*

[IEEE 1588] *IEEE 1588, IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems, July 2008.*

[MULPIv3.1] *MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.1-I21-201020, October 20, 2020, Cable Television Laboratories, Inc.*

[R-DTI] *Remote DOCSIS Timing Interface Specification, CM-SP-R-DTI-I08-200323, March 23, 2020, Cable Television Laboratories, Inc.*

[RESULTS] *Experiment Results for Supporting LTE-FDD, LTE TDD, and 5G Timing Synchronization Over DOCSIS CAA and DAA, SCTE 2019*

[R-PHY] *Remote PHY Specification, CM-SP-R-PHY-I15-201207, December 7, 2020, Cable Television Laboratories, Inc.*

[SYNC] *Synchronization Techniques for DOCSIS® Technology Specification, CM-SP-SYNC-I02-210407, April 7, 2021, Cable Television Laboratories, Inc.*

Terahertz Spectrum

Challenges, Potential And Applications

A Technical Paper prepared for SCTE by

Lakhbir Singh
Principal Engineer
Charter Communications
6463 S Fiddlers Green Circle
+1 7036751414
Lakhbir.singh@charter.com

1. Introduction

The way we create, share and consume information has led to a tremendous increase in wireless data traffic. The demand for wireless data continues to grow exponentially and that puts extreme requirements on the future communications systems. Each generation of mobile technology, from the first to the fifth (5G), has been designed to meet the new needs of end users and network operators. While the number of mobile connected devices is increasing rapidly, from 8.8 billion in 2018 to a predicted 13.1 billion by 2023, at the same, as anticipated by Edholm's law of bandwidth, the achievable data-rate is scaling new heights, from 20 Gigabits per second (Gbps) of peak data rate in 5G to expectedly 1 Terabit per second (Tbps) within the next few years. Figure-1 shows the roadmap for the wireless data rates up to year 2035 and is a good indicator of the past and anticipated growth in data rates.

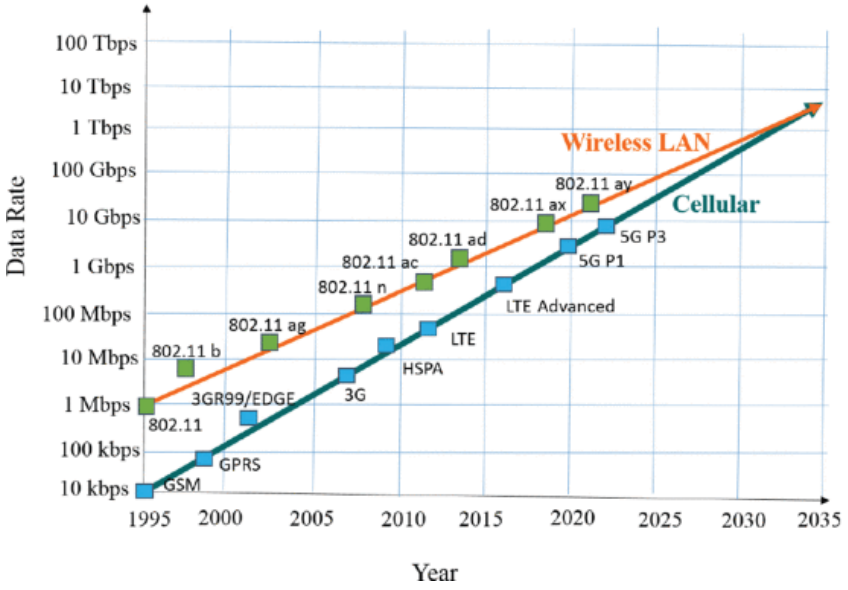


Figure 1 – Wireless Roadmap Outlook

Societies continue to become more and more intelligent, data-centric, data-dependent, and automated. Autonomous systems are hitting our roads, oceans, and air space. Millions of sensors will be embedded into cities, homes, and production environments, and new systems operated by artificial intelligence residing in local cloud and fog environments will enable a plethora of new applications. To glue this all together, communication networks will provide the nervous system for these new smart system paradigms. The demands, however, will be daunting. Networks will need to transfer much greater amounts of data at much higher speeds and lower latencies. Furthering a trend already started in 4G and 5G, sixth generation (6G) connections will move beyond personalized communication toward the full realization of the Internet of Things (IoT) paradigm, connecting not just people, but also computing resources, vehicles, devices, wearables, sensors, and even robotic agents. Beyond improving existing communication and signal processing solutions, there is a need to explore untapped frequency bands for communications. In order to enable the transmission of very high data rates (10s of Gbps), two approaches are possible. The first one is to improve spectral efficiencies beyond 10 bit/s/Hz for the systems operating at moderate bandwidths, which will be very challenging. The second

approach requires ultra-high bandwidths beyond 20 GHz while operating at moderate spectral efficiencies. But, such a large amount of spectrum is only available in the Terahertz (THz) frequency range, i.e., beyond 300 GHz.

THz band communications is being projected as a key enabling technology for next-generation wireless systems that promises to integrate a wide range of data-demanding and delay-sensitive applications. As more and more data move to the edge, the last meter solutions are getting more and more attention. The higher bandwidth and contention free data communication promised by terahertz spectrum is expected to be an important part of ITU (International Telecommunication Union) Network-2030 vision. As the wireless world moves toward 6G, radio layer Terabit per second (Tbit/s) communications and the supporting access and backhaul network infrastructures are expected to become a predominant technology trend. However, with the current approach, certain severe limitations affect the capability of future wireless communications systems to meet the combined requirements of high data rate, near-zero latency, and high spectral and energy efficiency. In this context, utilizing THz frequency bands for wireless transmissions, as an extension to optical fiber, is a promising enabler to bridge this gap and provide ubiquitous high-speed Internet access beyond 5G.

Since most of the traffic is generated indoors, the future wireless systems are also expected to rely on a significant number of indoor solutions to massively offload cellular networks. To enable this, the focus has been millimeter wave (mmWave) frequencies such as 28, 39 and 60 GHz. However, the use of mmWave solutions still leads to certain limitations as the shared throughputs will only approach a few Gbit/s. As an intermediate step between millimeter-wave and optical wireless communication systems, the THz band offers unique trade-offs between channel bandwidth and propagation properties that make it a key player in future wireless communication systems, including 6G. The THz band has sufficient resources not only to satisfy the 5G requirements of 10 Gbit/s peak data rate but to enable a number of new tempting high data rate applications.

However, the THz band brings a few novel challenges in terms of propagation (THz wall) and device availability (THz gap). To achieve wireless throughput of 100 Gbps and beyond, THz spectrum seems like a logical solution due to the availability of large bandwidths. The global community have been putting a lot of effort in developing THz band-based solutions and has shown some promising early development, but still have a long way to go before we see commercial solutions available.

As the THz band continue to gain noticeable attention within the global community. Seamless data transfer, unlimited bandwidth, microsecond latency, and ultra-fast download are all features of the THz technology that is anticipated to revolutionize the telecommunications landscape and change the way people communicate and access information. Motivated by the potential of THz technologies to shape the future of wireless communications, this paper seeks to identify the unique characteristics of THz waves, the critical technology gaps in system design, recent developments and some potential use cases/applications.

2. Terahertz (THz) Spectrum Overview

The Terahertz spectrum region of electromagnetic radiation lies between the microwave and infrared portions of the spectrum. Basically, THz range extends from the highest frequency radio waves to the lowest frequency infrared light. Theoretically, THz band should start at 1000 GHz, but for all practical purposes spectrum beyond 100 GHz is considered a part of THz band as shown in figure-2.

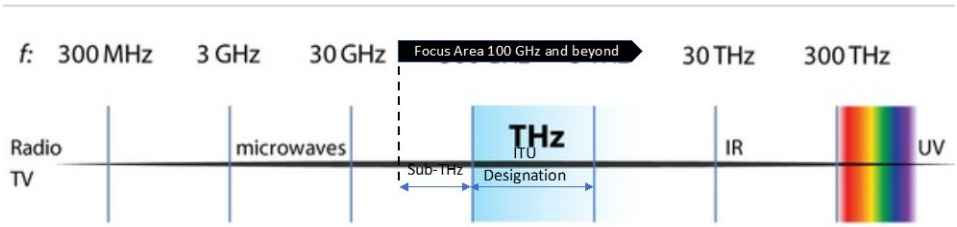


Figure 2 – Terahertz Spectrum

THz waves, also known as T-waves, T-rays, sub-millimeter waves or micrometer waves, are the electromagnetic waves in the frequency band of about 0.1 to 10 THz. ITU (International Telecommunications Union) has designated 300 GHz to 3 THz as the THz band. The spectrum between 100 GHz to 300 GHz is known as sub-THz band.

3. Terahertz Wall - THz Waves Propagation and Other Characteristics

The THz link suffers from several path loss mechanisms, the free space path-loss (FSPL) due to signal spreading, the effective antenna aperture, and the molecular absorption. The latter is a distinguishing feature of the millimeter and sub-millimeter bands. The main difference between the mmWave and the THz band is the progressively increasing molecular absorption loss.

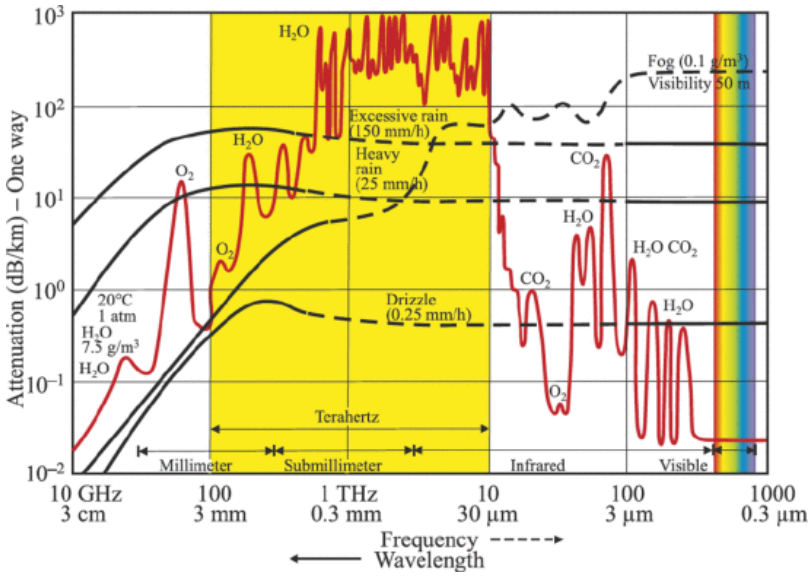


Figure 3 – Terahertz Propagation

At terahertz frequencies the propagation losses are very high, meaning the signal die down very fast. The impact of atmospheric attenuation for THz links is shown in figure-3. Although propagation losses are very high, a closer look indicates that not all frequencies suffer the same amount of propagation loss and the losses are not linear as we move towards high frequency region. Despite the existence of absorption peaks centered at specific frequencies, the availability of transmission windows allows for viable communication in the THz frequency band. These windows of opportunity (low absorption windows) for transmission at 94, 140 and 220 GHz are shown in figure 4 below.

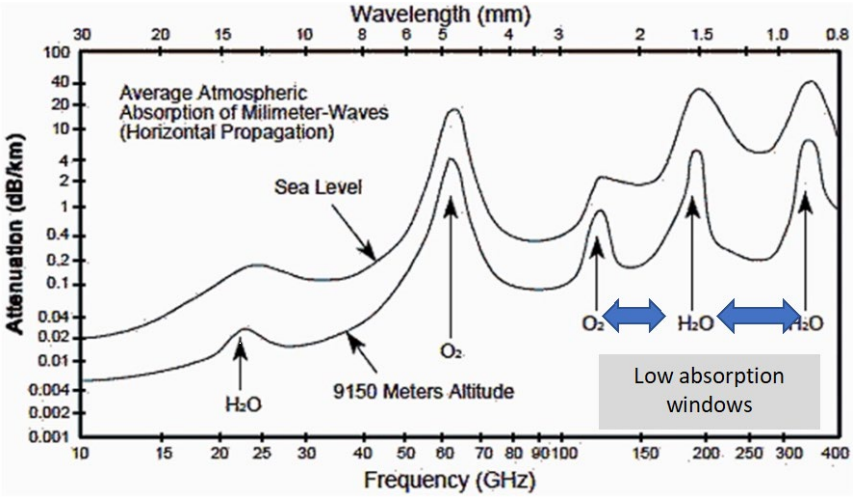


Figure 4 – Terahertz Propagation: Low Absorption Windows

The propagation distance of THz waves is significantly impacted by very high atmospheric loss. Based on the propagation studies, table-1 below shows the anticipated loss at a certain frequency using a specific bandwidth. Since water and oxygen are seen as the strongest absorber of the THz waves. For designing a THz link, it is very important to estimate the weather impact on high-capacity data links. As the THz band channel is considered highly frequency selective, the transmission distance is limited by attenuation and the appropriate carrier frequency should be determined according to the application.

Table 1 – Terahertz Ranges For Fixed Links

Frequency Range (GHz)	Contiguous Bandwidth (GHz)	Loss (dB/km)
275-320	45	<10
335-360	25	<10
275-370	95	<100
380-445	65	<100
455-525	70	<100
625-725	100	<100
780-910	130	<100

At terahertz frequencies the propagation takes the form of narrow a beam (similar to laser) and not a broadcast, which makes them more suitable for the point-to-point links. As shown in the

figure-5, THz waves quickly achieves a point where further power increase does not make any improvement in the link distance. This phenomenon is also known as terahertz wall. Figure-5 shows a comparison of different scenarios for horizontal transmission at sea level during good weather, bad weather, a range of distances (from 1 meter up to 6 kilometers), and specific frequencies between 35 gigahertz and 3 terahertz, to determine how much the signal strength degrades as conditions vary. The higher losses with rain and fog can be clearly seen as represented by dotted lines.

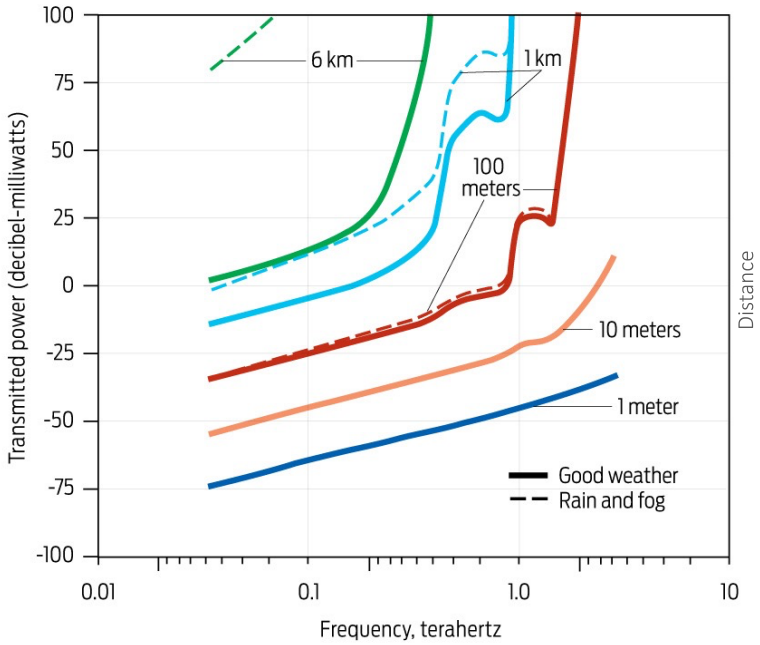


Figure 5 – Terahertz Wall: Frequency, Power versus Distance

For short-range operation, for example, signals traveling 10 meters or less—the effects of the atmosphere and bad weather don't really come into play. Any effort to send anything farther than that will quickly hit the terahertz wall. No matter how much the signal is boosted, essentially nothing gets through. A 1 watt signal with a frequency of 1 THz, for instance, will dwindle to nothing after traveling just 1 km, it only retains about 10-30 percent of its original strength. Even if transmit signal power is increased to the ridiculously high levels say a petawatt, it would get attenuated to a mere femtowatts by the time it reached its destination. Currently there are no terahertz sources capable of producing anything approaching a petawatt; the closest is a free-electron laser, which has an output in the low tens of megawatts and isn't exactly a field-deployable device. Basically, the extreme attenuation or propagation loss rules out using the terahertz region for long range terrestrial communications.

Some of the other unique properties of the terahertz waves are being harmless to the tissue and ability to pass through opaque materials. These properties make them promising candidate for use in noninvasive imaging in industry, medicine, and security. Unlike X-rays, terahertz waves are too low in energy (0.004 electron volt) to knock electrons off atoms, which could damage living tissue. Also, because of their shorter wavelength, they can produce pictures that are far sharper than those made with microwaves, the current safe imaging alternative. Terahertz waves also occupy a unique window of the electromagnetic spectrum where a large

number of molecules emit and absorb radiation. The signals produced when a molecule jumps among rotational modes form a unique and highly distinctive chemical fingerprint. Terahertz waves also passes thru most dielectrics, opaque materials for example textile, paper, plastics, cardboard etc. Since water absorption is high, terahertz waves can be used to monitor water content in food and chemicals. How these properties can be used for imaging and spectroscopy applications are briefly covered in the section 5.2 and 5.3 of this document.

4. Terahertz Gap – THz Technology Development Challenges

Terahertz region sits between the microwave (electronics) and far infrared optical (photonics) regimes. To make something as simple as a terahertz transceiver, a device that's capable of both emitting and detecting terahertz radiation, a few adjustments to a conventional electronic or photonic transceiver should have been possible. But neither technology is really suitable for the task. For example, trying to adjust a radio transceiver so that it operates at terahertz frequencies, fundamental physical limitations are quickly encountered. The rapidly flipping voltage of a high-frequency alternating current can have strong electromagnetic effects on the transistors in a transceiver, giving rise to parasitic resistances and capacitances that take power away from the devices. More fundamentally, there is a speed limit on how fast electrons can move. Beyond a few tenths of a terahertz, the frequency of oscillations becomes so high that an electron can't cross a transistor channel from the source to the drain before the voltage flips its polarity, forcing the electron to reverse direction and causing transistor power to drop precipitously. As shown in figure-6 below, the current devices exhibit very poor power-frequency scaling in THz range.

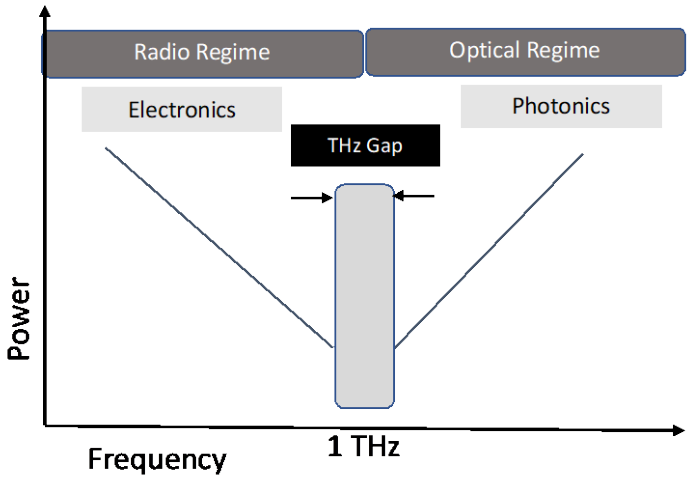


Figure 6 – Terahertz Gap: Power versus Frequency

This poor power versus frequency response exhibited by the current technologies have resulted in the device unavailability. This lack of devices in the terahertz spectrum as shown in figure-7 below is known as THz gap - a region in electromagnetic spectrum that does not have devices available for generation and detection. THz gap results from the lack of compact, energy-efficient device technologies. Most of the currently available devices are bulky, expensive and energy inefficient with power efficiencies in a single digit. At higher frequencies due to small antenna, device sizes create a lot of fabrication challenges, for example difficulty in ensuring

noise and inter-component interference suppression. In fact, integrated electronics are becoming large in comparison to the size of the corresponding antennas. Also, the high power and current density requires adequate cooling to prevent the device from disintegration.

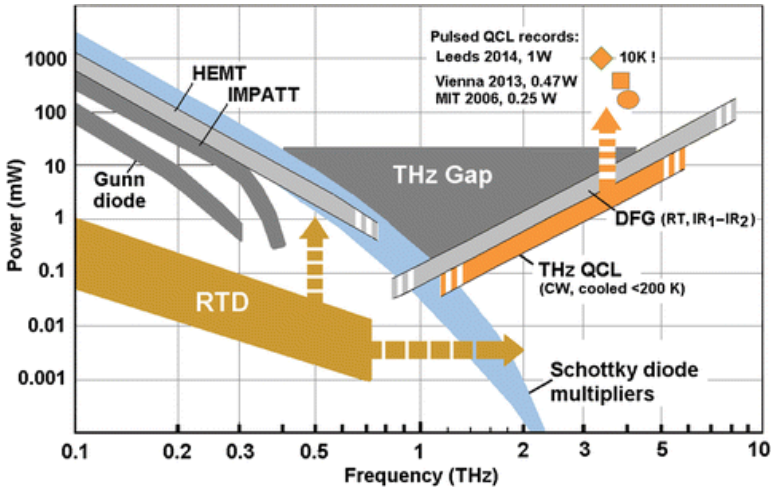


Figure 7 – Terahertz Gap: Current Device Technologies

These limitations are driving the invention of new materials. Progresses in electronic, photonic, and innovative plasmonic device technologies is continuously bridging the THz gap. Some of the promising technologies to achieve good output power and efficiency are Gallium Arsenide (GaAs), Gallium Nitride (GaN), Indium Phosphide (InP), complementary metal oxide semiconductor (CMOS), Silicon Germanium (SiGe), and fully depleted silicon on insulator (FD-SOI) CMOS. Figure-8 provides a high-level view of capabilities being offered by current technologies in terms of supported frequency and power output with still a long road ahead.

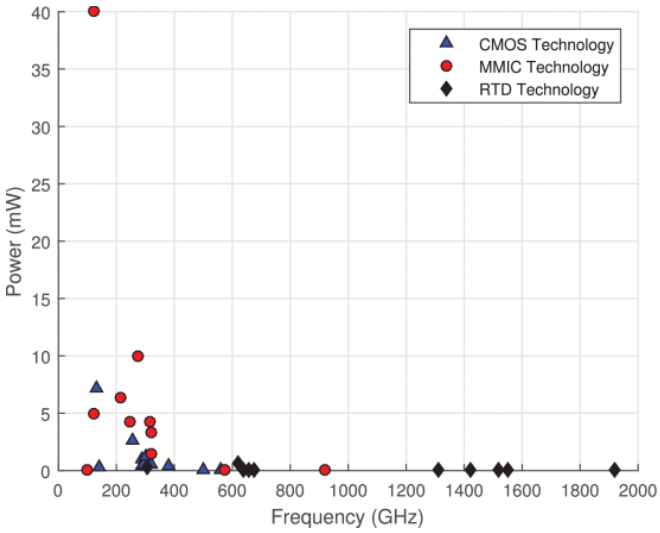


Figure 8 – Solid State Electronics: Frequency of versus power

Although today's technology has scarce high-power THz transmitters, it is expected that signal generators with high power can be created in the near future considering ongoing studies on graphene and InGaAs (Indium Gallium Arsenide) mHEMT (Metamorphic High Electron Mobility Transistor) monolithic microwave integrated circuits (MMICs). In the electronics approach, frequency multiplying and mixing chains enabled by transistor and diode technologies, resonant tunneling diodes, and traveling wave tubes are some of the major players. In a photonics approach, the beating of two lasers or frequency-difference generation, photoconductive antennas and quantum cascade lasers are examples of THz technologies aimed at closing the gap coming from the other end of the spectrum.

In addition to the device availability some of the other major areas that require a lot of work include radio frequency frontend and antenna design, propagation and channel modeling, waveforms, signals, and coding, beamforming and (ultra-massive) MIMO as well as resource management and medium access control schemes. All of these challenges demand for higher processing powers thus driving the need for new advanced processing devices.

THz spectrum offer the potential for breakthrough high data throughput applications. However, there are four key considerations for optimizing the performance of sub-terahertz systems operating over wide or extreme bandwidths: optimizing signal to noise ratio (SNR), minimizing phase noise, addressing linear and nonlinear impairments and making a waveform selection. In order to optimize SNR to achieve the best Error Vector Magnitude (EVM) performance, maximizing signal power (S) can achieve the highest SNR, but reducing the signal power becomes necessary to avoid harming any components along the signal chain. The noise contributions for SNR can be problematic for wideband applications since the noise power is integrated over wide signal bandwidths. For example, 10 GHz bandwidth signal will have 1,000 times more integrated noise than a 10 MHz signal. Basically, due to power level and high bandwidth contributing to poor SNR, basically SNR is effectively "squeezed" both on the high end and the low end. This is a key consideration in moving to extreme bandwidth test systems, and the SNR can often translate into what residual EVM is achievable. Also, up conversion from an intermediate frequency (IF) to sub-terahertz frequencies involves frequency translation with a local oscillator (LO) signal source(s) and frequency converter(s). A frequency multiplier will increase the phase noise by 20 * Log (N), where N is the multiplication factor. Furthermore, the multiplier can introduce additive phase noise that will further degrade the multiplied LO phase noise, dependent on the quality of the multiplier used. Low residual EVM test system performance at sub-terahertz frequencies requires high-quality, low-phase-noise LO signal sources.

To make the most out of the very large bandwidth provided by the THz band channel, new communication and signal processing techniques are needed. These include, among others, time, frequency and phase synchronization techniques, when transmitting at Tbps and in the presence of phase noise, real time channel estimation and equalization of tens to hundreds of GHz-wide bandwidths; and modulation strategies that can make the most out of the distance-dependent bandwidth of the THz channel. All these techniques need to be realized while keeping in mind that the sampling frequency of the fastest digital to analog and analog to digital converters is far below the required Nyquist frequency, which motivates the exploration of heavily parallelized systems as well as analog signal processing solutions. As opposed to traditional wireless networking paradigms, where the channel bandwidth is the scarcest resource that needs to be efficiently shared, the vast spectral resources in the THz band flip the script and change the focus to effectively everything else. Innovation is required in the higher layers of the

protocol stack, including dynamic power and bandwidth multi-hop relaying strategies that efficiently leverage the THz channel and real-time buffer-less routing protocols able to overcome the traditional delays associated with store-and-forward policies at the network layer; or innovative end-to-end transport layer solutions, which can accommodate frequent lack of connectivity between highly directional nodes.

Though the list of challenges in designing ultra-high frequency large bandwidth systems is very large, but the research community is continuously working on bridging this gap with new materials, processes and techniques.

5. Terahertz (THz) - Use Cases/Applications

Based on their properties discussed in the section 3, the applications of the THz waves can be broadly divided into three areas:

- Wireless Data Communications - THz can support high bandwidth (GHz), high data rates (Tbps) in the short range
- Imaging – THz waves are non-ionizing and harmless to tissues, can pass through opaque surfaces
- Spectroscopy - THz wave rotational and vibrational frequency matches those of molecules resulting in the creation of unique spectral fingerprint

Since the focus of this paper is wireless data communication, therefore communication related use cases will be discussed in more details, while imaging and spectroscopy will be covered at a very high level.

5.1. THz based Wireless communication applications

This section outlines some of the applications/use cases that will require ultra-broadband connectivity to work effectively. Due to high atmospheric losses the long-range communications (a few kilometers) with THz is not possible, but still there is a huge potential in using THz waves in the short range. At a high level the communication uses cases can be divided into three categories: Nanoscale (nano to centimeter), Microscale (centimeter to meter) and Macroscale (a few meters).

5.1.1. Terahertz Nanoscale Applications

Very small size of THz transceivers and antenna is driving a lot of application at a nanoscale. Over a very short distance referred to as nanoscale, the THz waves can be used to support intra-chip, inter-chip or intra-device communications. Chip-to-chip communications means wireless links inside computers or any other electronic devices. It is of high relevance because wired connectors and micro strip lines on printed circuit boards potentially become a bottleneck of upcoming bus systems and inter chip connections.

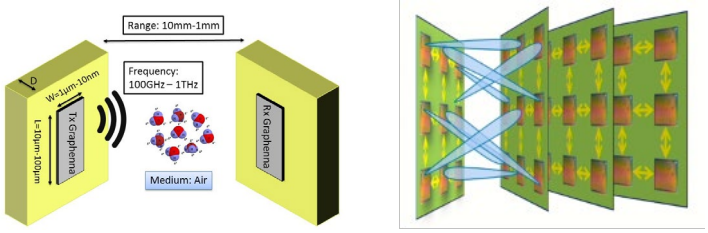


Figure 9 – Terahertz: On chip or chip to chip communications

Since transmission ranges are a few cm and the fully static environment allows for a fixed alignment during design process with automatic beam steering as an option for sequential multipoint operation between more than just two chips. Energy and chip real estate are two scarce resources in devices (e.g., smartphones, tablets, laptops). With the new advances in the coding and modulation used in the physical layer (PHY), chip to chip communications can be used to reduce the area and energy overheads of a given interconnect. Thus, terahertz waves can be used to implement simple and energy efficient on chip or inter-chip communications. The THz wireless communication is quickly evolving into a key enabler of applications involving operations inside computers and devices for a typical range of a few centimeters which includes chip-to-chip, board-to-board and device-to-device communications.

Additionally, nano-antennas enable wireless interconnection amongst nano-sensors deployed inside and over the human body resulting in many bio-nano-sensing applications.

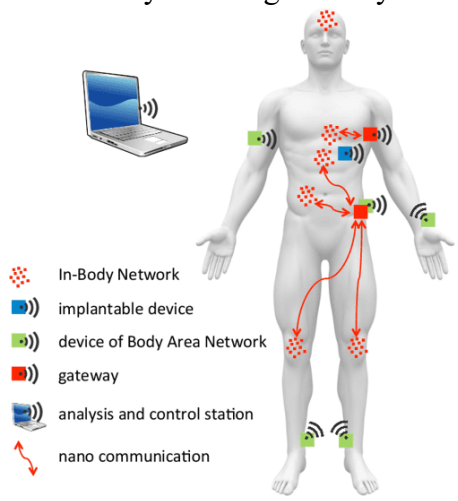


Figure 10 – Terahertz: Body area network (BAN)

As shown in figure-10 above, THz sensors can be planted inside and outside body to gather and transmit data to the wireless gateway connected to the diagnostic application.

5.1.2. Terahertz Microscale Applications

Wireless local area network (WLAN) and wireless personal area network (WPAN) form the basis of microscale applications which include high-definition television (HDTV) in home distribution, wireless displays, seamless transfer of files, and THz access points in the areas with device congestion. The THz band provides small cell communication for mobile cellular networks, where ultra-high data rate can be provided to mobile users within transmission range up to 20 m. As shown in figure-11 below, wireless personal area networks can be used for ultra-high-speed ad-hoc connections between devices over short distances, for example between a printer and a laptop. The typical THz deployment will be indoors and the alignment of high gain antennas should be ideally done by automatic beam steering, but rough manual alignment may be thinkable in case of less directive antennas.

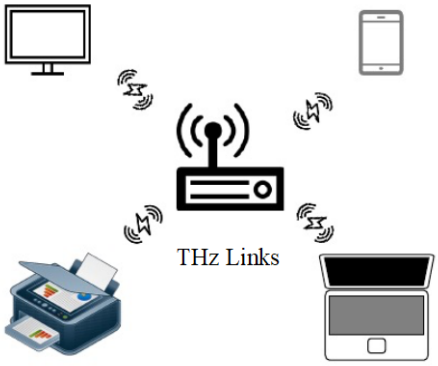


Figure 11 – Terahertz: Wireless Personal area network (WPAN)

THz frequencies can be used to provide transmission solutions in ad-hoc networks and for nomadic users by facilitating connection to access points for example at railway, stations, public buildings, shopping malls, etc. Microscale wireless communication at the THz band can also be used to transmit uncompressed high definition (HD) videos for education, entertainment, telemedicine, as well as security purposes. Kiosk downloading is another example of microscale application at THz frequencies, which offers ultra-high downloads of digital information to users' handheld devices. For instance, advertisement posters in metros, trains or streets can be the front interface for downloading pre-fixed contents such as newly released movie trailers, CDs, books, and magazines.

Wireless data centers are considered another promising application at the microscale. THz links in data centers can be deployed as a parallel technology to cabling, whereas the THz links can be used for the rack to rack or inter-cluster connectivity as shown in figure-12 below.

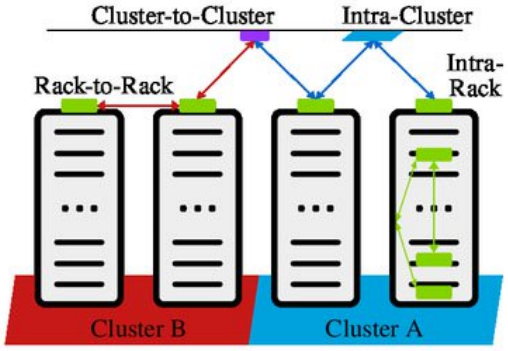


Figure 12 – Terahertz: Data Center Networks (DCN)

Such deployment in data centers results in an enhanced performance experience along with immense savings in cabling and time to deploy. For the optimization of capacity within data centers, a frequent and dynamic reconfiguration of the architecture is required, which is extremely difficult to achieve with fully wired systems. This introduction of ultra-high data rate wireless hops shown in figure-12 above, are considered as an attractive alternative. For wireless hops in a data center, racks are equipped with highly directive steerable antennas on top of the

racks. Connections between different racks are realized by aligning the beams of the corresponding antennas.

5.1.3. Terahertz Macroscale Applications

On a macroscale, THz wireless communication facilitates potential outdoor applications which range from few meters up to kilometers. For example, wireless backhauling/fronthauling is one of the envisioned applications for the standard 100 Gbps transmission solutions. THz links can be used for backhaul for base stations of macro cells especially in the areas where optical fiber is not available. For fronthaul THz point-to-point links can be used to connect the radio controller of a base station and the remote radio head (radio unit). An example of fronthaul application is shown in figure-13 below. The increasing number of mobile and fixed users in both the private, industrial and service sectors will eventually require hundreds of Gbps to be carried either to or between cell towers (backhaul) or between cell towers and remote radio heads (fronthaul).

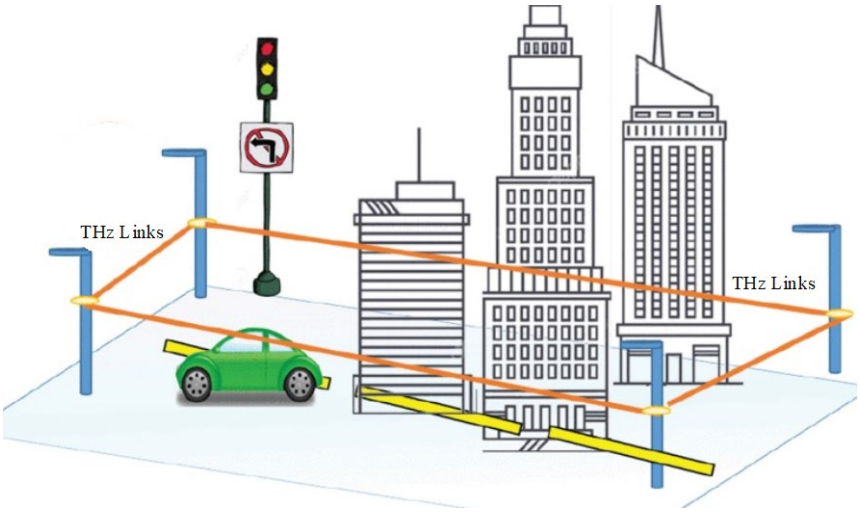


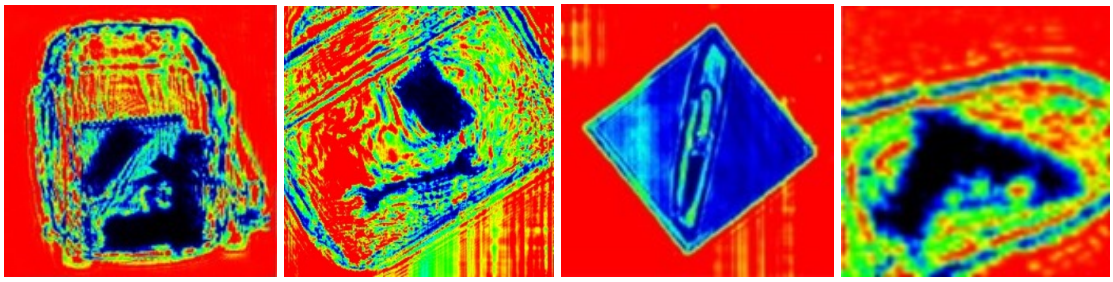
Figure 13 – Terahertz based fronthaul

Future advancements which include massive deployment of small cells, implementation of cooperative multipoint transmission and Cloud Radio Access Networks (C-RAN) may increase the required data rates for either fronthauling or backhauling or both. Therefore, fixed links with ultra-high data rates can be utilized for the wireless extension of backbone networks, for the aggregation of multiple backhaul links with lower data rates as well as for the wireless backhaul of future cellular base stations with very high overall throughputs.

Even though the path loss increases with frequency, the antenna gain is related to the square of the operating frequency. Therefore, high gain antennas can cope with the extreme path loss in the THz band. As known, antennas with high operating frequency have relatively narrow beams compared to lower frequencies; hence, THz antennas can be utilized in point-to-point links to backhaul small cells. Moreover, since the antenna size decreases with frequency, it is possible to place a massive number of antennas on small surfaces. For example, it is possible to cover a 1 mm2 area with four antennas at 300 GHz. Beyond massive multiple-input multiple-output (mMIMO), namely ultra-massive (UM-MIMO), can be employed to increase the communication distance as well as the data rate enhancement.

5.2. Terahertz Imaging

Terahertz waves are nonionizing, meaning its photons are not energetic enough to knock electrons off atoms and molecules in human tissue, which could trigger harmful chemical reactions. THz waves can also penetrate a variety of non-conducting, amorphous, and dielectric materials, such as glass, plastic, and wood. These see through and non-ionizing properties makes them a good candidate for imaging application. Due to shorter wavelength compared to microwave radiation, THz waves create higher resolution images as well. Absorption and reflection of terahertz waves are highly material-dependent, and helps in providing the contrast in the images. Some examples of terahertz imaging with see-through capability are shown in figure-14 below. The images demonstrate THz imaging ability to reveal hidden objects.



Hidden Objects (Notebook/gun in a bag, wrench in a bag, knife in a book, gun in a box)

Figure 14 – Terahertz see-through imaging

The applications of THz imaging are endless and have been the key focus of research community. A lot of commercial products are available for industrial inspection, quality control, security scanning, medical imaging etc., for example, to screen for skin cancer and tooth decay. In the future, imaging capability in a reflection mode could be incorporated in smartphones for seeing through packages to detect what is inside without opening them, walls to locate wires and pipes, or decorative pieces to detect air voids and cracks another application for 3D THz imaging is to use it for providing accurate position determination and object detection capability. Terahertz radiation also offers a safe, non-ionizing solution to the problem of screening travelers passing through airports and crossing national borders.

5.3. Terahertz Spectroscopy

Terahertz waves also occupy a unique window of the electromagnetic spectrum where a large number of molecules emit and absorb radiation. The signals produced when a molecule jumps among rotational modes form a unique and highly distinctive chemical fingerprint. These unique fingerprints at terahertz radiation can be used for chemical detectors. As an example, in figure-15 below, unique fingerprints for a drug and explosive are shown.

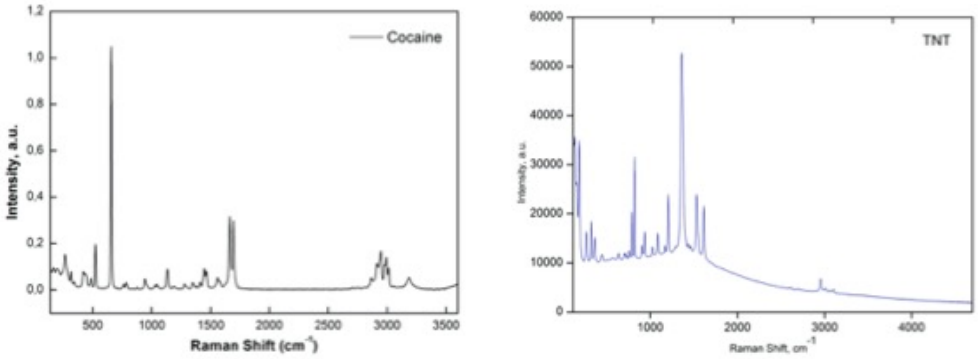


Figure 15 – Terahertz unique chemical fingerprints

Unique chemical fingerprints at THz waves makes them a good candidate for material and gas sensing, for example, environmental monitoring includes detection of pollutants and contaminants to meet the requirements of health and safety, monitoring of atmospheric components for weather and climate observation. Analyzing gases exhaled in human breath is a highly promising application, with the potential for early detection of disease biomarkers. Such detection would be fast and non-invasive, facilitating early diagnosis and state-of-health monitoring. Potential detectable diseases include lung cancer, diabetes, some neurological disorders, and smoking and alcohol consumption. In biochemistry, THz applications have been developed to explore the bioactivity of chemical compounds and identify protein structures.

For security and public safety applications, many explosives (for example, C-4, HMX, RDX and TNT) and illegal drugs (for example, methamphetamine and heroin) have characteristic features in the THz range. In an example (source RIKEN, Japan) below use of THz spectroscopy to inspect a package containing different chemicals has been demonstrated.

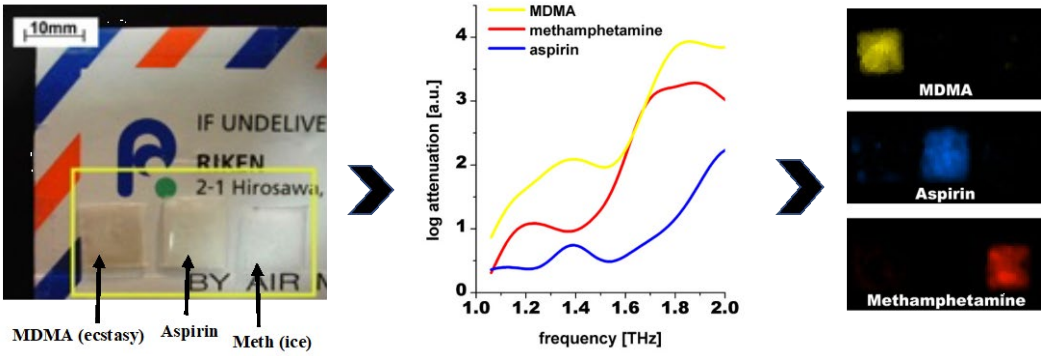


Figure 16 – THz spectroscopy for substance detection

As shown in figure-16, three different substances produce a unique fingerprint which can be used to detect the substance without opening the package. The combination of terahertz imaging and spectroscopy has resulted in a lot of commercially available products for non-destructive testing and new applications are growing every day.

6. Terahertz (THz) Development status

Seamless data transfer, unlimited bandwidth, microsecond latency, and ultra-fast download are all features of the THz technology that is anticipated to revolutionize the telecommunications landscape. These features have resulted in the terahertz band gaining noticeable attention in the global research community. The enthusiasm of the research community is reflected in the increased number of publications issued in the recent years on IEEE, SPIE etc. The potential associated with THz technology has attracted the broader research community. The combined efforts of active research groups are resulting in new designs, materials and fabrication methods that demonstrate endless opportunities for THz development. In table-2 below are some examples of various groups conducting THz research. The list below is a testimony of the focused research in this area being conducted in the laboratories across the globe.

Table 2 – Research Groups Working on Terahertz

Research Group/Lab	Location	R&D Activities
Mittleman Lab at Brown University	USA	THz PHY layer, spectroscopy, THz probes
Broadband Wireless Networking Lab at Georgia Tech	USA	THz PHY and MAC layer, THz Nano-communications, THz devices
NaNoNetworking Center in Catalunya	Spain	THz Nano-communications
Ultra-broadband Nano-Communications Lab at Univ. of Buffalo	USA	THz PHY and MAC layer, THz Nano-communications, THz devices
THz Electronics Lab at UCLA	USA	THz devices, reconfigurable meta-films, spectroscopy
MIT Terahertz Integrated Electronics Group	USA	Sensing, Metrology, security and communication.
Fraunhofer Institute for Applied Solid State Physics	Germany	THz PHY and MAC layer, THz electronics
Terahertz Communication Lab	Germany	Channel investigation and THz reflectors
Core Technology Lab Group NTT Corp	Japan	Terahertz IC and modularization technology
Texas Instrument Kilby Lab	USA	Ultra-Low Power sub-THz CMOS systems
Tonouchi Lab at Osaka Univ.	Japan	THz Nanoscience, THz bio-science, THz bio-sensing and industrial applications
THz Electronics Systems Lab at Korea Univ.	Korea	THz PHY and MAC layer, THz electronics
Nano-communications Center at Tampere Univ. of Technology	Finland	THz PHY layer, THz Nano-communications

Various regulatory and funding agencies have been supporting THz projects and opening up new horizons in communications and devices for deployment beyond 5G technology. The regulatory framework for spectrum above 100 GHz is beginning to take shape, as discussed in the subsequent sections.

6.1. International Telecommunications Union (ITU)

In November World Radiocommunication Conference 2019 (WRC-19) which meets every four years to discuss international radio rules (RRs) and other topics, made 190 GHz (252 to 356 GHz) of spectrum available for THz communication with no specific EESS (Earth Exploration Satellite Services) protection requirements. Before WRC 2019, there were no allocations above 275 GHz. Allocations were made for fixed and mobile communications for in 275-296 GHz, 306-313 GHz, 318-333 GHz and 356-450 GHz bands. Resolution was also made for a study plan for ITU-R for increased sharing above 71 GHz. Bands of interest include 100 - 102 GHz, 116 - 122.25 GHz, 148.5 - 151.5 GHz, 174.8 - 191.8 GHz, 226 - 231.5 GHz and 235 - 238 GHz. This came as a very encouraging news for the worldwide research community focused on THz research.

6.2. Federal Communications Commission (FCC)

The United States, through the Federal Communications Commission's (FCC) Spectrum Horizons First Report and Order, is taking a leadership role in the future of 5G and beyond. As a result of this initiative, the FCC is making more spectrum available. On March 21, 2019, the FCC adopted new rules to encourage development of new communication technologies and expedite the deployment of new services above 95 GHz. As a part FCC 19-19, the FCC opened 95 GHz- 3 THz band for unlicensed and experimental communications under "Spectrum Horizons" program. The FCC made the entire band available for experimental authorizations. In addition, the FCC made a total of 21.2 GHz spectrum available for unlicensed use. As shown in figure-17 below the unlicensed spectrum would span several band segments (116-123 GHz, 174.8-182 GHz, 185- 190 GHz, and 244-246 GHz).

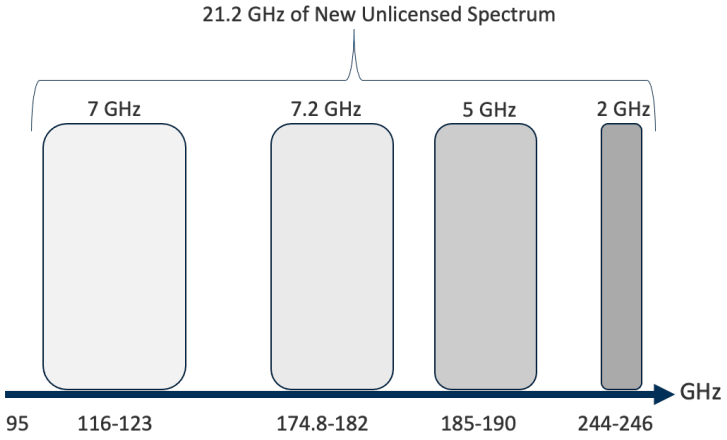


Figure 17 – FCC Spectrum Horizons Initiative: Unlicensed Spectrum

21.2 GHz of new unlicensed spectrum between 95 GHz and 3 THz can be used by devices that do not interfere with existing governmental and scientific operations and that operate within a maximum threshold of average EIRP (Effective Isotropic Radiated Power) of 40 dBm, as stated in the FCC order. This initiative gives researchers, innovators, and entrepreneurs the flexibility to conduct experiments blocks between 95 GHz and 3 THz through experimental licenses lasting up to 10 years. Researchers may also more easily market and demo equipment during the trial period. This will encourage the development of communication technologies and services above

95 GHz, such as data-intensive high-bandwidth applications; wireless cognition; and imaging, positioning, and sensing operations.

6.3. IEEE (Institute of Electrical and Electronics Engineering)

In 2008, the IEEE 802.15 established the 802.15.3d Task Group (Terahertz Interest Group) as a milestone towards investigating the operation in the so called “no man’s land” and specifically for frequency bands up to 3 THz. The new group was tasked to identify THz band communications as a feasible wireless technology for extremely high access rates of up to 100 Gbit/s. The group has worked closely with the International Telecommunication Union (ITU) and the International Radio Amateur Union (IARU) regarding the description of the frequency bands higher than 275 GHz.

The ratification of the IEEE 802.15.3d amendment to 802.15.3, was a significant step towards standardization consumer wireless communications in the sub-THz frequency band. IEEE 802.15.3d offers switched point-to-point connectivity with data rates of 100 Gb/s and higher at distances ranging from tens of centimeters up to a few hundred meters. IEEE approval of channelization standard is considered a major milestone towards THz wireless communication.

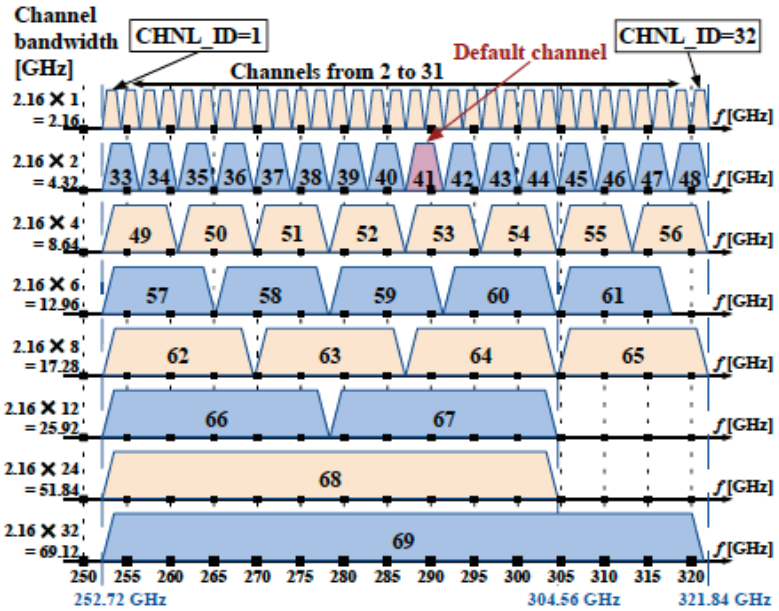


Figure 18 – Bandwidth and channels for IEEE 802.15.3d standard

The major highlights of IEEE 802.15.3d standard are:

- 8 different channel bandwidths (as multiples of 2.16 GHz)
- 2 PHY-modes (THz-SC PHY, THz-OOK-PHY) with 7 modulation schemes: BPSK, QPSK, 8-PSK, 8-APSK, 16-QAM, 64 QAM, OOK
- 3 channel coding schemes: 14/15-rate LDPC (1440,1344), 11/15-rate LDPC (1440,1056), 11/15-rate RS(240,224)-code.

As shown in figure-18 above, IEEE 802.15.3d may operate over the sub-THz frequencies between 252.72 GHz and 321.84 GHz. In total, there are 69 overlapping channels. There are 8 supported channel bandwidths ranging from 2.16 GHz (same as in mmWave IEEE 802.11ad 2012) up to massive 69 GHz channel bandwidth. Depending on the use case and the hardware capabilities, either the entire frequency range can be allocated for a single 69.12-GHz-wide channel (CHNL_ID = 69) or can be shared between several smaller channels. The bandwidths of the channels are integer multiples of 2.16 GHz. The channel number 41 with the bandwidth of 4.32 GHz (2x2.16 GHz) is the default channel.

Following the decision at World Radio Conference 2019 (WRC-2019), all the bands depicted in Figure-18 are available for THz communications globally if specific conditions to protect radio astronomy and earth exploration satellite service (EESS) are met. These conditions do not explicitly specify any transmit power limits and are applicable, in practice, primarily to the narrow areas surrounding the ground radio astronomy stations.

6.4. European Union ICT (Information and Communication Technologies)

There is a sharp increase in recent years in the research funding by European Union, body like Horizon 2020 (H2020). This funding ranges from enhancing research on the device technology to communication aspects including channel, physical, MAC and network layer characterization. The THz band related projects currently funded by H2020 are summarized in the table-3 below. Some of the projects are based on establishing the feasibility of communication windows within Terahertz Gap, for example Ultrawave, Dream, EPIC, and WORTECS.

Table 3 – EU Horizon 2020 Terahertz Related Projects

Project Name	Band	Target Speed	Focus
iBrow	60 GHz-1 THz	10 Gbps	Innovative ultra-broadband ubiquitous wireless communications through THz. Low cost and simple wireless transceiver architecture. Integrated semiconductors emitters and detectors with seamless fiber-wireless links
TERRANOVA	300 GHz	100 Gbps	Tbps wireless connectivity by THz innovative technologies to deliver optical network quality of experience in systems beyond 5G
TERAPOD	300 GHz	100 Gbps	Terahertz based ultra-high bandwidth wireless access networks for data centers
ULTRAWAVE	300 GHz	100 Gbps	High capacity backhaul links to enable 5G cell densification by exploiting bands beyond 100 GHz.
EPIC	-	1 Tbps	Enabling practical wireless Tbps communication with next generation of coding. Develop new FEC (forward error correction) codes for Tbps rates.
DREAM	D-band	100 Gbps	D-band reconfigurable meshed radio solution at 100 Gbps by exploiting radio spectrum bands like 130-174.8 GHz with beam steering functionality to reach optical systems speed.
WORTECS	90 GHz	10 Gbps	Wireless optical radio terabit communication system. Use 90 GHz for proof of concept with Gbps rates.

Project Name	Band	Target Speed	Focus
ThoR	300 GHz	100 Gbps	Front and backhaul solutions at 300 GHz for data rates beyond 5G.
TERAPAN	300 GHz	100 Gbps	Adaptive wireless point-to-point terahertz communication systems for indoor environments for distances of up to 10 m at data rates of up to 100 Gbps
ROOTHz	THz	-	Development of solid state nanodevices at room temperature of THz communications

The project which involves the research on advancing the communication methodologies while considering the MAC layer issues and challenges are TERAPOD, TERANOVA, and ThoR. Their aims and objectives include the device, channel and antenna characterization. They are also focusing on designing a simple and efficient MAC protocol for point to point and multipoint scenarios. However, each project is looking at a specific scenario. For example, the TERAPOD project is aiming to design a communication methodology for a data center environment, which involves potentially the channel, antenna and Physical layer considerations for an indoor environment only. The TERRANOVA on the other hand focuses on the backhaul point to point scenario for outdoor long range environment including small cells. Each scenario requires a different modelling approach for antenna, channel and propagation model and therefore requires different strategies to access the channel to establish communications. Similarly, ThoR is also looking at high speed link upto 100 Gbps over 300 GHz band for backhaul with partial involvement of point to point scenario for MAC layer channel access. These projects have helped creation of new THz devices, communication techniques which will lay new foundations for the rapid development of practical systems.

7. Conclusion

The paper has presented the current status in research, spectrum regulations and standardization activities in the field of THz communications. Though there are challenges developing terahertz products i.e., closing the THz gap by discovering new devices and the THz wall, the inherent high propagation loss caused by atmospheric absorption. Still, the field of THz communications has drastically changed in the last ten years and, what used to be a relatively niche discipline, is now identified as one of the key enablers of the future ultra-broadband networks, 6G and beyond. Major progress has occurred worldwide in the last decade both in terms of materials and devices as well as communication techniques. With these advances, wireless Tbit/s communications and the supporting backhaul network infrastructure are expected to become the main technology trend within the next ten years and beyond. To facilitate this, an unprecedented amount of 160 GHz of the spectrum was recently opened for THz communications after WRC 2019. Moving forward, besides addressing the technical open challenges, the adoption of the THz spectrum for communications needs to be accompanied by supporting spectrum policies that ensure an efficient, fair and secure use of the spectrum by both active (communications) and passive (sensing) users. Based on the recent progress, the commercial introduction of THz communications can be expected in the not-too-distant future and will help coping with the ongoing tremendously increasing demand for wireless data rates.

Abbreviations

5G	5th Generation
BAN	body area network
bps	bits per second
BPSK	binary phase shift keying
C-4	composition-4
CD	compact disc
CMOS	complementary metal oxide semiconductor
C-RAN	Cloud Radio Access Networks
DCN	Data Center Network
EVM	Error Vector Magnitude
FCC	Federal Communications Commission
FD-SOI	fully depleted silicon on insulator
FEC	forward error correction
FSPL	free space path loss
GaAs	Gallium Arsenide
GaN	Gallium Nitride
Gbps	giga bits per second
GHz	gigahertz
HDTV	high-definition television
H2O	chemical formula for water
HMX	high melting explosive
ICT	Information and Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
ITU	International Telecommunications Union
InP	Indium Phosphide
LO	local oscillator
MAC	medium access control
mHEMT	metamorphic high electron mobility transistor
MMIC	monolithic microwave integrated circuit
mMIMO	massive multiple-input multiple-output
O2	chemical formula for oxygen
OOK	on off keying
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RDX	royal demolition explosive
RTD	resonant tunneling diode
SCTE	Society of Cable Telecommunications Engineers
SiGe	Silicon Germanium
SNR	signal to noise ratio
SPIE	Society of Photographic Instrumentation Engineers
THz	terahertz
TNT	trinitrotoluene
UM-MIMO	ultra-massive multiple-input multiple-output
WLAN	wireless local area network
WPAN	wireless personal area network
Tbps	terabit per second

Bibliography & References

V. Petrov, J. Kokkoniemi, D. Moltchanov, J. Lehtomaki, Y. Koucheryavy and M. Juntti, "Last Meter Indoor Terahertz Wireless Access: Performance Insights and Implementation Roadmap," in IEEE Communications Magazine, vol. 56, no. 6, pp. 158-165, June 2018, doi: 10.1109/MCOM.2018.1600300.

Michael C. Wanke and Mark Lee, "Transceivers to Conquer the Terahertz Frontier New ICs harness the untamed terahertz band" IEEE Spectrum, 31 Aug 2011

J. M. Jornet, "Terahertz Communications: From Nanomaterials to Ultrabroadband Networks," 2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz), 2020, pp. 1-2, doi: 10.1109/IRMMW-THz46771.2020.9370577.

Simon Rommel, Thiago R. Raddo, Ulf Johannsen, Chigo Okonkwo, and Idelfonso Tafur Monroy, "Beyond 5G - wireless data center connectivity," Proc. SPIE 10945, Broadband Access Communication Technologies XIII, 109450M (Presented at SPIE OPTO: February 06, 2019; Published: 1 February 2019); https://doi.org/10.1117/12.2506269.

H. Elayan, O. Amin, B. Shihada, R. M. Shubair and M. -S. Alouini, "Terahertz Band: The Last Piece of RF Spectrum Puzzle for Communication Systems," in IEEE Open Journal of the Communications Society, vol. 1, pp. 1-32, 2020, doi: 10.1109/OJCOMS.2019.2953633.

Carter M. Armstrong "The Truth about Terahertz"IEEE Spectrum August 17, 2012

V. Petrov, T. Kurner and I. Hosako, "IEEE 802.15.3d: First Standardization Efforts for Sub-Terahertz Band Communications toward 6G," in IEEE Communications Magazine, vol. 58, no. 11, pp. 28-33, November 2020, doi: 10.1109/MCOM.001.2000273.

Dressler, F. and S. Fischer. "Connecting in-body nano communication with body area networks: Challenges and opportunities of the Internet of Nano Things." Nano Commun. Networks *6 (2015): 29-38.*

The Dennis Botman Story

A Tale of Next-Level Chatops

A Technical Paper prepared for SCTE by

Michael Winslow
Senior Director, Software Development and Engineering
Comcast Corporation
michael_winslow@comcast.com

Ryan Emerle
Senior Principal Engineer
Comcast Corporation
ryan_emerle@comcast.com

Mia Kuang
Software Engineer
Comcast Corporation
mia_kuang@comcast.com

1. Introduction

What do you do when you have an amazing team and you want to make them even better? The 1995 Chicago Bulls added Dennis Rodman. Well, our team at Comcast created Dennis BOT-man: An automated Chatops bot that helped us re-imagine how we support our internal engineering teams.

1.1. What is Chatops?

ChatOps is a collaboration model that connects people, tools, process, and automation into a transparent workflow (see Figure 1). With our Chatbot, Dennis Botman, we've taken an iterative approach to add functionality over time. As we will get into later in the paper, our primary objective is to connect people with knowledge resources like frequently asked questions FAQs and how-to documentation.

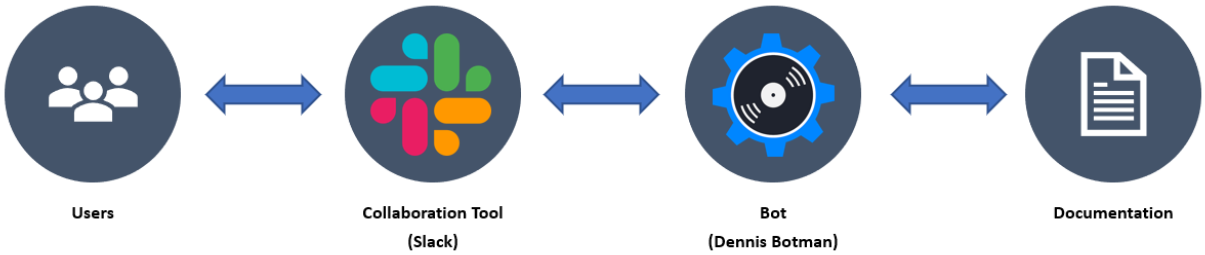


Figure 1 - Typical Chatops flow for Dennis Botman

1.2. Why do we need Chatops?

To understand why we needed this level of automation, it helps to know the scale of the application we support.

Comcast's open-source domain name system (DNS) management tool, VinylDNS, allows millions of DNS entries to be controlled by hundreds of individual groups throughout the company (Cleary, 2020). This decentralized governance model is a marvel on its own. One software developer, Stephanie Hingtgen, described her team's usage of VinylDNS in the following way:

> *"We leverage the VinylDNS system to instantly provision A and AAAA records for all new VIPs in our zone, as well as giving the user the ability to provision CNAME records. This was crucial for rolling out IPv6 VIPs in RDEI (Rapidly Deployed Elastic Infrastructure) so that users would have an easy way of remembering their VIP name."*

To ensure the reliability of this critical application, our group adopted a DevOps[1] model where our software engineers rotated to support the product. The primary means by which to engage our support team was by asking questions in our Slack support channel.

Before long, our support channel was filled with questions from our users; many of which were already answered in our support documentation. Valuable time that our engineers could have been spending developing new software and features were instead spent repeatedly answering the same basic questions. Something had to be done.

[1] DevOps is the combination of cultural philosophies, practices, and tools that increases an organization's ability to deliver applications and services at high velocity

In this paper, we'll tell the story of how we coded and delivered Dennis Botman to be a true member of the support team. We'll explain how Dennis Botman was able to answer over 35% of the questions being asked in our support channel. We will show how adding powerful tooling within the bot allowed engineers all over the company to solve DNS issues instantly. And perhaps most fascinating, we will use metrics to demonstrate how our support bot empowered our users by arming them with a self-service mindset.

After reading this paper, if your team would like to take advantage of the features we developed in Dennis Botman, feel free to visit our open-source project on Github: https://github.com/vinyldns/vinyldns-bot

2. Problem Statement

2.1. Small Team Supporting a Large Organization

The world of DNS is massive. Everything that is on the internet, from servers to automated teller machines (ATM) machines to your smartphone, are all addressable by internet protocol (IP) addresses which are often mapped to DNS records.

When tasked with managing millions of these records, which service hundreds of teams throughout Comcast, we set out to build a service to enable users to manage their own records. Born out of the necessity to reduce time-to-market, VinylDNS is the user-facing service we created that enables self-service for DNS.

What happens when thousands of users across hundreds of organizations suddenly have access to DNS records and the inherent complexity therein? They're going to have questions and they're going to need answers.

Suddenly, what was once a utopian dream of a self-sufficient, self-service, tool becomes a conduit to an influx of support requests. This deluge of questions simply shifted the burden and complexity of DNS management from the DNS engineering team to the VinylDNS service team.

Quickly we were scrambling to extricate ourselves from the critical path of thousands of users. To do so, we needed to find ways to offload our support personnel while still getting answers to the users and getting DNS records provisioned.

2.2. Customers Do Not Read Documentation

One of the quickest and easiest ways to get answers for users was to thoroughly document the service. With VinylDNS, the service and its corresponding APIs and tooling were all thoroughly documented before even being released to users.

Documentation of every aspect of the system was written. Reams of documentation detailed every nuance of the service and its corresponding user interface. Months of writing and revising lead to a thorough reference of the system and its behavior.

The problem? Users unfortunately do not tend to read documentation. This isn't because users are lazy, or unwilling; they simply have more important problems to solve. As authors of a user-facing service, we cannot assume that we are the center of our users' universe. Our service is likely a small part of the

overall problem being solved, and as such we cannot expect thousands of users to take several hours to pore over reams of documentation.

2.3. Providing Real-Time Support

In order to overcome this tendency to ignore the documentation, we needed to provide a means by which users could be directed to the subset of documentation that addresses their questions.

A logical next step was to provide real-time support. This entailed creating a channel in Slack where users could come and ask questions. The channel would be staffed by VinylDNS engineers who would help users get to their answers.

Upon establishing real-time support, the support personnel were inundated with questions. Among these were a set of questions that kept appearing. As such, the natural next step was to distill some of the documentation down into FAQs. The FAQs provided more direct answers and saved the users and support personnel time sifting through documentation.

While creating FAQs helped to reduce the influx of support questions, a pattern started to emerge. Users would come to the support channel, ask a question, and then a support person would find the relevant FAQ, copy the link, and paste it into the support channel as a response.

During this period, there were three to four engineers spending upwards of thirty percent of their time answering questions, searching the documentation on behalf of users, and pasting FAQ links into the support channel.

2.4. Providing After-Hours Support

Our team monitored our Slack support channel for one week and discovered that 13% of the questions from our customers happened after hours. There is no guarantee that a chatbot would solve their problems, but our customers would have more options when our engineers are not available. A simple reminder of our on-call hours is better than no response at all.

2.5. Is Automation the Answer?

It soon became clear that the repetitive nature of the real-time support that was being offered was untenable. We had engineers whose time is better spent improving the product than acting as glorified search engines.

When faced with repetitive tasks, there is a tendency toward automation. As engineers, we evaluated the landscape and began to consider how we might automate the repetitive task of searching documentation, copying links, and pasting them into the support channel.

The cost of automation is not free, however, so we needed to estimate the return-on-investment (ROI) for the automation.

The general formula for ROI *(Fernando, 2021)* is:

$$ROI = \frac{Net\ Value\ of\ Investment}{Cost\ of\ Investment} \times 100\%$$

In the case of automation, our value and cost parameters are in units of time. As such, we can restructure the equation as follows:

$$ROI = \frac{Time\ Saved - Time\ Spent}{Time\ Spent} \times 100\%$$

This simple formula can be a bit deceiving, as Time Spent needs to include the initial time to build the automation plus all future time spent maintaining it. Similarly, Time Saved is also cumulative. Further, determining, or predicting, Time Saved can be challenging. In this case, we're focused on the estimated time saved by the predicted reduction in time spent on support duties. Time saved by reduction in human error, training personnel and other factors are not being considered.

For the purposes of calculating a value, we can limit the ROI projections and calculation to a discrete time period. For the following example, we limit the calculation to one year.

Assume we have three engineers spending 30% of their time supporting users. Given a 40-hour work week, we can calculate the support cost, in time, for a single week:

$$Support\ Cost_{week} = 3\ engineers \times (40\ hours\ \times 30\%)$$
$$Support\ Cost_{week} = 3\ engineers \times\ 12\ hours$$
$$Support\ Cost_{week} = 36\ hours$$

Let's assume that we can implement some automation in four weeks at 40 hours per week.

$$Automation\ Cost_{total} = 4\ weeks\ \times 40\ hours$$
$$Automation\ Cost_{total} = 160\ hours$$

Let's further assume that our analysis indicates that the automation is projected to save 25% of our support personnel's total time.

$$Automation\ Savings_{week} = Support\ Cost_{week} \times 25\%$$
$$Automation\ Savings_{week} = 36\ hours \times 25\%$$
$$Automation\ Savings_{week} = 9\ hours$$

If our cost of automation is a total of 160 hours, then we can calculate the ROI for a year:

$$ROI_{year} = \frac{(Automation\ Savings_{week} \times 52\ weeks) - Automation\ Cost_{total}}{Automation\ Cost_{total}} \times 100\%$$

$$ROI_{year} = \frac{(9\ hours\ \times 52\ weeks) - 160\ hours}{160\ hours} \times 100\%$$

$$ROI_{year} = \frac{468\ hours - 160\ hours}{160\ hours} \times 100\%$$

$$ROI_{year} = \frac{308\ hours}{160\ hours} \times 100\%$$

$$ROI_{year} \approx 192\%$$

The resulting calculation shows that if we invest 160 hours in automation to recover 25% of the time spent performing support activities, we'll see a projected return of 192%. For every hour spent automating, 192% of that time is returned as time saved over a one-year period.

3. Getting Started Getting Started

3.1. To Build or to Buy

With an estimated ROI of 192%, we were confident that we were going to take steps toward automation. But before deciding to invest our engineering time to creating a custom chatbot for our internal customer base, we did some initial research on the costs of purchasing a solution. We found that adopting a chatbot service can possibly get quite costly.

According to mobilemonkey.com (MobileMonkey, 2021), once you factor in the cost of the platform along with the salary related costs for setup, development, and maintenance, you could be paying over $60k in the first year alone.

Table 1 - Estimated cost for a chatbot paid solution

	In-House Chatbot Costs	Agency Chatbot Fees
Chatbot Software Platform	$50-$500/month	$50-$500/month
Chatbot Setup and Development	Salaries (5-100 hours of work)	$500-$2,500
Ongoing chatbot support and maintenance	Salaries (0-10 hours of work per week)	$50-$5000/month

We decided it would be better to develop our own bot using opensource solutions and allowing our engineers to solve our specific problems. We chose to start with the Github's Hubot Framework (Metz, 2015) which is written in CoffeeScript and node.js. Hubot handles basic chat communication, which saved us a lot of time. But we required much more intelligence built into our bot if it was going to be useful. We would need to extend the functionality of Hubot, but by how much?

3.2. The "Big Ideas" Phase

During our initial brainstorming sessions, one thing that the team insisted was that the bot should be conversational rather than a glorified command line interface (CLI). This aspect did cause some team members to really think BIG:

- "We need to learn natural language processing / understanding (NLP/NLU)."
- "We cannot do this without artificial intelligence."
- "Machine learning is essential."

After several meetings that appeared to get us no closer to a starting point, the team began to lose interest. It was starting to feel like this task was simply too large an undertaking. As discussed in section 2.5 (Is Automation the Answer?), there just did not seem to be enough value to justify this level of effort.

Our problem was that we were trying to design the perfect system right from the planning phase. The sheer size and complexity of the solution we were proposing was putting us in a state of analysis paralysis. We would either need to narrow our scope or scrap the idea all-together.

3.3. Perfect is the Enemy of Good

Anyone who has ever read Stephen R. Covey's best-selling book **7** Habits of Highly Effective People (Covey, 1989) will tell you that to "begin with the end in mind" is critical. But be careful not to misconstrue this message with the idea that everything must be known and understood before beginning work on a software project. There are so many benefits to be had by taking an iterative approach.

In a 2019 article published by Forbes on the topic of Machine Learning projects, they state that "87% of projects do not get past the experiment phase and so never make it into production" (Dans, 2019). This is because software development teams far too often attempt to "boil the ocean" rather than identifying bite-sized chunks of functionality to deliver fast. It is better to have a product mindset and get users to begin interacting with your software in order to give valuable feedback.

If we wanted to be a part of the 13% of projects that actually see production, we were going to have to focus on the core problem we are trying to solve. We decided to clearly state our purpose and simplify our initial feature set. Once we were able to agree on a small set of essential requirements, Dennis Botman was ready to come to life!

4. Dennis Botman is Born / How it works

4.1. Initial Feature Set

We decided on the following technical requirements:

- The bot will be available in 0..n Slack support channels
- The bot can accept free text "interactions" as well as exact-match "commands"
- The bot can respond with answers to FAQs
- The bot can respond with links to support documentation and real-time dashboards
- Users can have direct message (DM) conversations with the bot
- The bot MUST log all interactions with the bot (to improve support)

A high-level design of our implementation is shown in Figure 2.

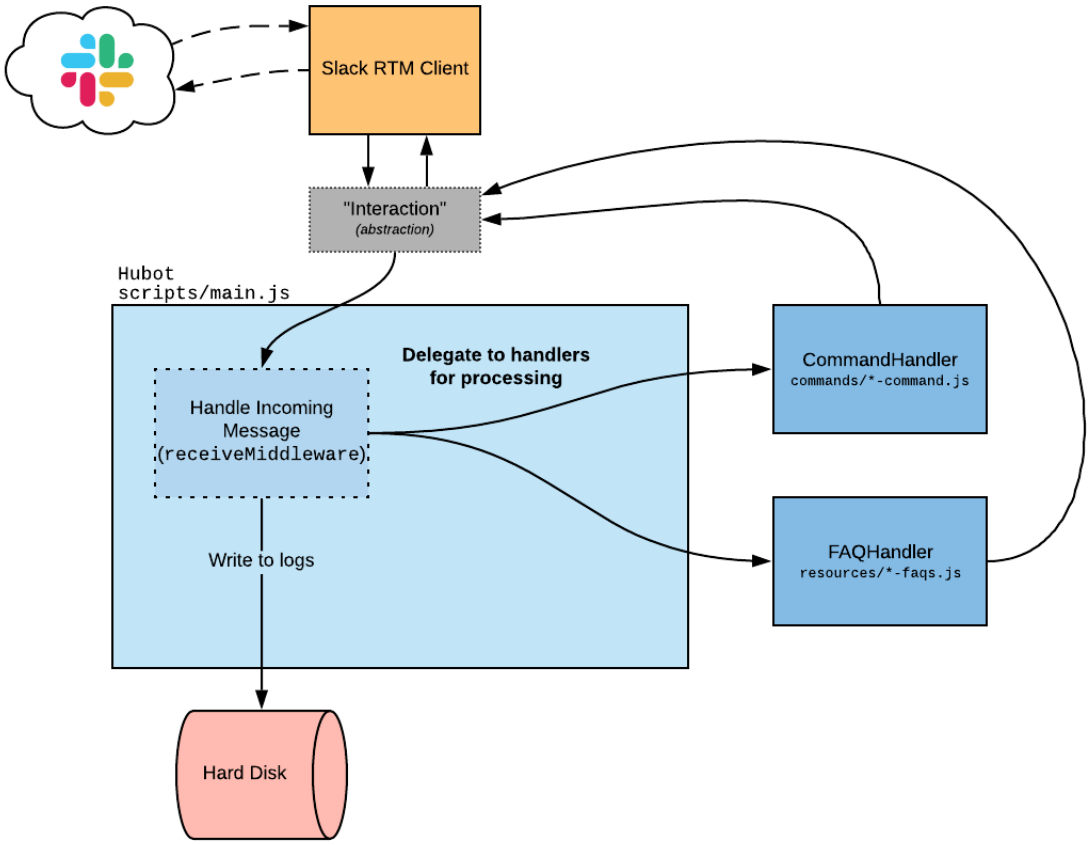


Figure 2 - Vinyldns-bot high level design

4.2. Basic Interaction (FAQs)

Figure 3 shows a common example of a user interacting with Dennis Botman.

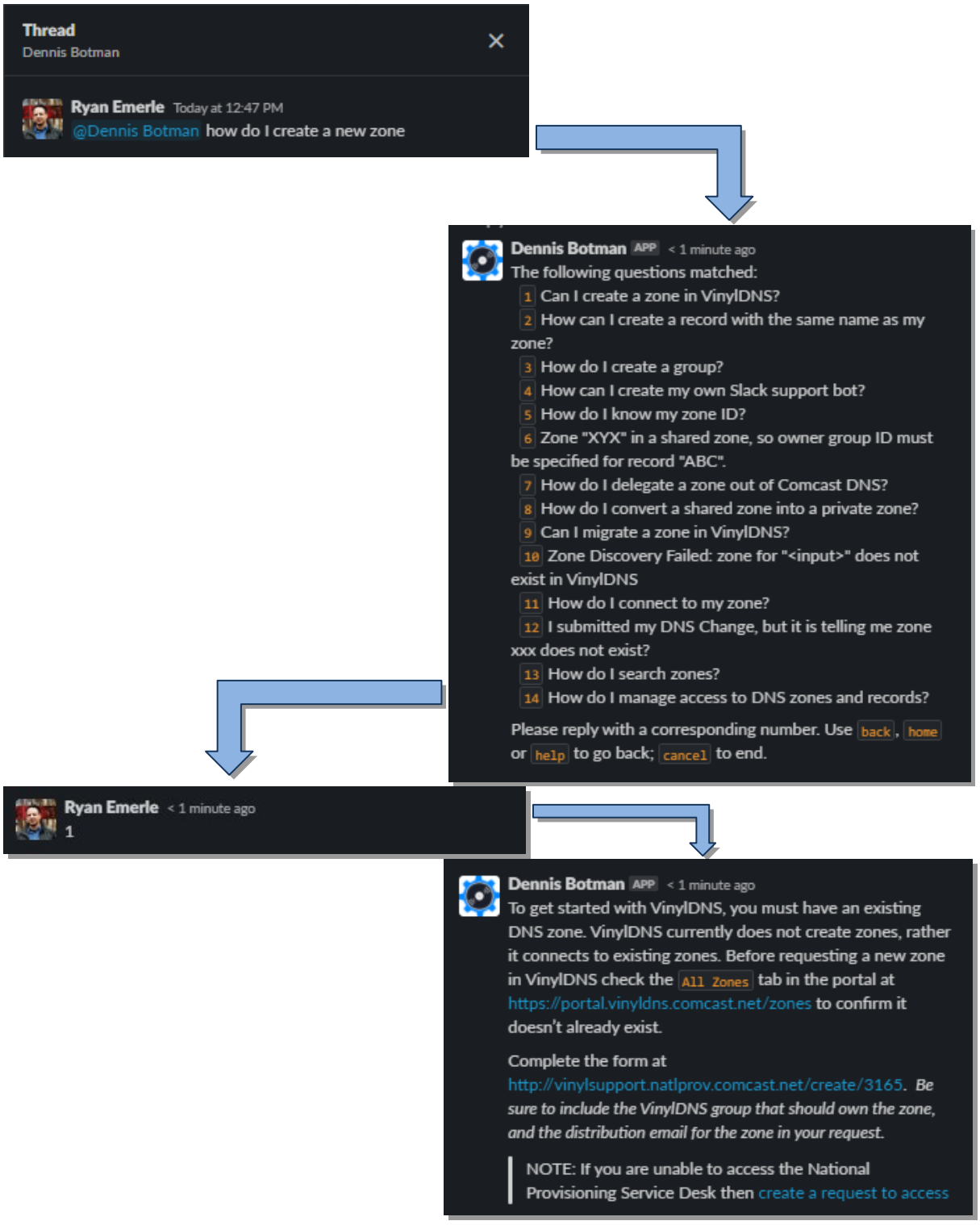


Figure 3 - Basic Dennis Botman conversation

4.3. Pre-Processing Text/Questions From Users

Users can interact with Dennis Botman by asking any question via direct messaging or in a Slack channel. Dennis Botman can process a user's question by searching through a set of FAQs files and return a subset of matched questions.

While the Hubot Framework handles the directing of traffic between the user and the bot, we still had to program the logic in determining how the incoming questions were processed. To do this, we utilized some common techniques around "Text Cleaning" and "Pre-Processing."

The bot pre-processes a user's input and extracts meaningful words. Then it performs the search for matching words from all the FAQs in the repository. Any FAQ that has one word or more matched is a candidate. Each FAQ is presented as a question form itself, with hidden extra terms that can be used to match with user input, and exclusion terms to be removed from the search. Depending on the number of matched questions, the number of matched meaningful words in each question, and the configured limit for number of results, the search will return a list of FAQs in a specific ranked order. If there is only a single match, the response to that FAQ will be provided directly. If there are multiple possible matches, the user will be prompted to choose from them using an enumerated response.

Using the example question from above, let's step through how Dennis Botman interprets the input:

Starting String:

"How do I create a new zone?"

Step 1: Convert text to lowercase

Converting all the incoming request text to lowercase allows us to do exact string comparisons against our library of FAQs given that we also convert the library to lowercase strings.

- result: *"how do i create a new zone?"*

Step 2: Remove unwanted characters

Special characters and punctuation can cause our results to be inaccurate. It helps to scrub the incoming text of any characters that may be troublesome.

- result: *"how do i create a new zone"*

Step 3: Remove stop words

Stop words are words which are filtered out before or after processing of natural language data (text). Words like "a", "do", and "is" are considered "stop words." We maintain our own list of stop words to increase our chances of getting a high-confidence result to our text search.

- result: *"create new zone"*

Step 4: Tokenization

Tokenization (in lexical analysis) is the process of demarcating and possibly classifying sections of a string of input characters. In other words, we separate each word in our string so that they may be used individually. This will help create the word match count needed for the final step.

- result: [*"create", "new", "zone"*]

Step 5: Ranking

By counting the number of matches that we have of our tokenized words within our library of FAQs, we can rank the confidence of our results. You can clearly see why the most relevant results appear highest in our list.

- "Can I **create** a **zone** in vinyldns?" – (2 matches)
- "How can I **create** a record with the same name as my **zone**?" – (2 matches)
- "How do I **create** a group?" – (1 match)

5. Increasing the Usage of the Bot

5.1. Roadshows and Reminders

In order to realize the benefit of our investment of time into creating a bot to automate support of the VinylDNS service, we need our users to engage with the bot. Not only must users engage with the bot, but they must also be able to get the answers they're looking for, so that they do not require human intervention.

The first step in increasing engagement is a logical one – tell the users that the bot exists. After initial development of the bot, the VinylDNS team found opportunities to speak at internal conferences in order to raise awareness across organizations. Through technical conferences and demonstrations, we were able to promote the existence of the bot and its many benefits.

While we evangelized the bot across multiple organizations, we also needed to bring attention to the bot where support was happening – in our Slack channel. To do so, we tackled the low-hanging fruit first. We simply added details about the bot in the channel topic. The channel topic is the place where users can find information about the channel and the various resources offered.

While include the bot in the channel topic was essential, it did not draw immediate attention to the bot. So, in order to supplement this, we added a daily "message of the day." This is a message that is posted to the support channel every morning. The message is posted by the bot itself and it details the existence of the bot, how to use it, and other details about the support channel.

With the addition of the "message of the day" we saw a marked increase in bot interaction. However, on occasion, the message of the day would be scrolled out of view by support discussions. As such, we also added a notification for all users who first join the support channel. Upon joining, new users receive a message detailing the existence of the bot and how to interact with it to get answers.

5.2. One-on-One Conversations with the Bot

After making a concerted effort to promote the existence of the bot, we needed to find ways to encourage users to interact with it.

First, we wanted to reduce the anxiety of users experimenting with a bot in a Slack channel with over 1,000 users. As such, one of the first features we added was the ability to DM the bot. This allowed users to interact with the bot in a private channel, independent of the main support channel. All the functionality remains the same; the only difference is that the experimentation is not broadcast to the wider community.

5.3. Using the Bot as a Proxy

As Albert Einstein once said, "[E]xample isn't another way to teach, it is the only way to teach." With that in mind, we wanted to encourage usage of the bot by interacting with it ourselves. To do this, instead of copy-and-pasting links to FAQs on our documentation website, we issued commands to the bot on behalf of the user.

To further extend this concept, we added a special command: "for <user>". This allows support personnel to issue commands on behalf of another user but establish a dialog between the bot and that user (see Figure 4).

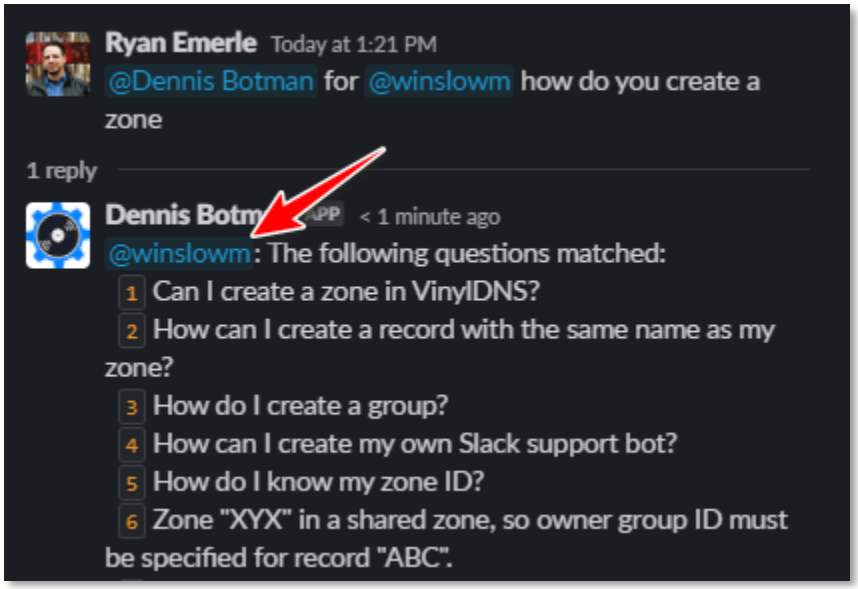


Figure 4 - Creating a dialog between Dennis Botman and a third party

5.4. Letting the Bot be the Bad Guy (Unsolicited Responses)

This feature was the result of collaboration with engineers in other groups, namely the groups that manage the DNS servers. One issue that they identified is that users would often create a support ticket for some DNS change, then immediately go to the real-time support channel and ask about the status.

In order to reduce the need for human intervention when a ticket was already created, we introduced the concept of "unsolicited messages". Unsolicited messages are messages sent by the bot that are not the result of a user interaction with the bot. As an example, a user might come to the support channel and say, "I created a ticket XYZ-123, can someone take a look?" The bot uses pattern matching to detect the presence of a ticket number, and automatically responds to the requestor with a message letting them

know that the ticket will be addressed in the order in which it was received and no further action is required on their part.

5.5. Keeping Our Engineers Involved

Once user engagement began to increase, we found additional opportunities to add functionality. Instead of simply providing answers to questions, the bot could also retrieve data from external sources in order to answers questions.

Keeping engineers involved in the automation of support activities was the key to making the bot more useful and, as a result, better equipped to resolve user questions without human intervention.

An example is the addition of a “lookup” command. This command queries the DNS backend servers to return record resolution details back to the user (see Figure 5). This is in lieu of asking the user to use cryptic command line utilities to query DNS resolvers. In the case of “lookup,” the user could simply message the bot “lookup <FQDN>” (where FQDN is a fully qualified domain name; e.g., “lookup www.comcast.com”). The result is a human-readable DNS lookup.

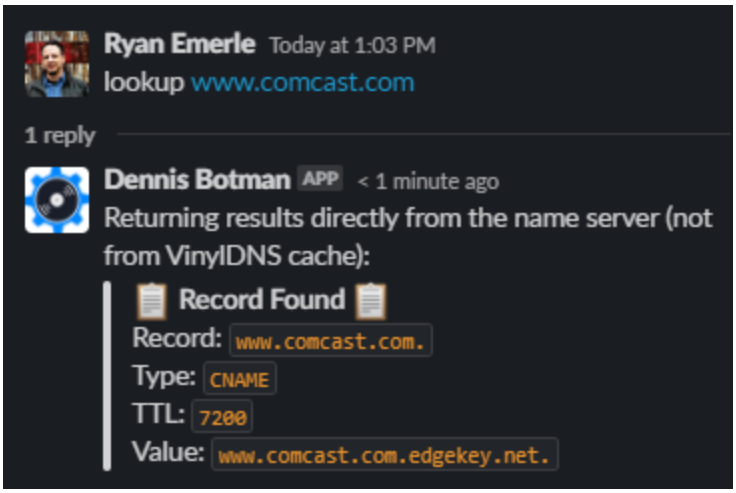


Figure 5 - An example of the “lookup” command

Similar types of commands were added that increased the utility of the bot and reduced the prerequisites required of users trying to interact the DNS services.

Engagement with a wider community proved to be invaluable. Through collaboration with engineers across multiple organizations we were able to significantly increase the ability for users to get to answers more quickly.

Seeing the power of collaboration in action, and given that VinylDNS is an open-source project, we open-sourced the VinylDNS bot code as well. Through greater exposure we hope to evolve the project to meet the needs of larger and more diverse audiences.

6. Observing the Lasting Effects

6.1. Overview of the Data

Determining the impact of our conversational chatbot would require us to log every interaction. This enables us to compare the number of times our internal customers were able to get the answers they needed without requiring human involvement. In addition to being able to put these results into buckets of human versus bot support, we were also able to determine exactly how our customer prefer to interact with the bot (see Figure 6).

The following results were over a 30-day period:

- 711 total interactions
- 347 interactions with the bot
- 260 of those were direct messages
- **7** of those were unsolicited

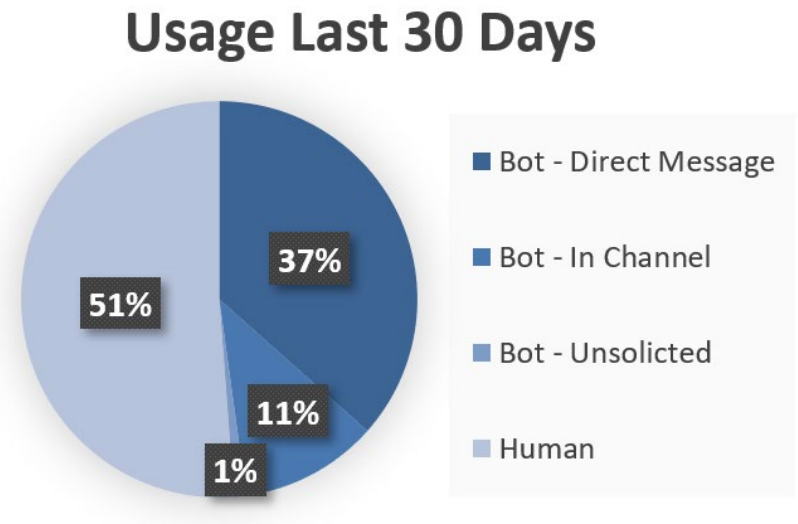


Figure 6 - VinylDNS Support Channel Interactions

According to our findings, total bot interaction was at 49% as opposed to 51% interactions with humans. This number alone does not tell us if our users were able to get to the answers they were looking for with each bot interaction. So we did a little more research.

By reviewing the users and time stamps associated with the 347 interactions with the bot, we investigated how many times those users subsequently returned to the support channel to get answers from a human. We found that 271 (78%) of those interactions required no human follow-up.

6.2. Re-calculating ROI based on real numbers

Automation Cost: Our initial time spent to develop Dennis Botman was approximately three weeks of one full-time engineer and one part-time engineer. It should be noted that we do spend small amounts of time on a continual basis to maintain the bot, but our initial time invested was less than a month.

1.5 engineers x 40 hours x 3 weeks = 180 hours of initial Automation Cost

Automation Saving: Using the statistics we gathered from the logs, we can replace our assumed time savings (25%) with the actual time savings we've observed from actual interactions with Dennis Botman.

271 bot handled requests (no human) / 711 total interactions = 38% Automation Saving

If our Support Cost remains constant at 36 hours and we multiply that by 38% instead of 25%, we get the following:

$$Automation\ Savings_{week} = 36\ hours \times 38\%$$
$$Automation\ Savings_{week} = 14\ hours$$

If our cost of automation is a total of 180 hours, then we can calculate the ROI for a year:

$$ROI_{year} = \frac{(Automation\ Savings_{week} \times 52\ weeks) - Automation\ Cost_{total}}{Automation\ Cost_{total}} \times 100\%$$

$$ROI_{year} = \frac{(14\ hours\ \times 52\ weeks) - 180\ hours}{180\ hours} \times 100\%$$

$$ROI_{year} = \frac{728\ hours - 180\ hours}{180\ hours} \times 100\%$$

$$ROI_{year} = \frac{548\ hours}{180\ hours} \times 100\%$$

$$ROI_{year} \approx 304\%$$

7. Conclusion

Software engineers solve some of the most complex problems for our customers every day to improve their product experience. Far too often, we do not use those same skills to solve our own problems to make our engineering experience more enjoyable and productive. We simply accept that tasks like on-call support, status reporting, deployments, and testing are things that have to be done manually in addition to our development tasks.

With Dennis Botman, we did more than simply play around with a Chatbot as a side-project. We, in essence, created another teammate. A teammate that could make all of our jobs easier, but only after we spent time training it, as well as making our customers comfortable with getting help from a bot. In the end, we were able to free up more time to tackle larger, more interesting problems.

Tasks to improve engineering efficiency can have a huge impact on how you grow and operate your engineering teams. Setting aside time for your engineers to solve their own problems should always be looked upon as an investment of time or resources that carries many beneficial outcomes.

Abbreviations

API	application programming interface
ATM	automated teller machine
CLI	command line interface
DNS	domain name system
DM	direct message
FAQ	frequently asked question
FQDN	fully qualified domain name
IP	internet protocol
RDEI	rapidly deployed elastic infrastructure
ROI	return on investment
NLP/NLU	natural language processing / understanding
NLU	natural language understanding
VIP	virtual ip (internet protocol)

Bibliography & References

Cleary, P. (2020, August 28). *Why Comcast open sourced its DNS management tool*. Retrieved from opensource.com: https://opensource.com/article/20/9/open-source-dns

Covey, S. R. (1989, August 15). *The 7 Habits of Highly Effective People*. Retrieved from franklincovey.com: https://www.franklincovey.com/the-7-habits/

Dans, E. (2019, July 21). *Stop Experimenting With Machine Learning And Start Actually Using It*. Retrieved from Forbes.com: https://www.forbes.com/sites/enriquedans/2019/07/21/stop-experimenting-with-machine-learning-and-start-actually-usingit/

Fernando, J. (2021, April 8). *Financial Ratios > Return on Investment (ROI)*. Retrieved from Investopedia: https://www.investopedia.com/terms/r/returnoninvestment.asp

Metz, C. (2015, October 23). *The Most Important Startup's Hardest Worker Isn't a Person*. Retrieved from WIRED: https://www.wired.com/2015/10/the-most-important-startups-hardest-worker-isnt-a-person/

MobileMonkey. (2021, August). *Chatbot Pricing: Everything You Need to Know About Chatbot Prices for SMBs*. Retrieved from https://mobilemonkey.com/: https://mobilemonkey.com/blog/chatbot-pricing/

The Evolution Towards Autonomous Networks

A Comprehensive Overview of Frameworks and Applications of AIOps

A Technical Paper prepared for SCTE by

Claudio Righetti
Chief Scientist
Telecom Argentina S.A.
crighetti@teco.com.ar

Mariela Fiorenzo
Tech Scientist
Telecom Argentina S.A.
mafiorenzo@teco.com.ar

Emilia Gibellini
Tech Scientist
Telecom Argentina S.A.
egibellini@teco.com.ar

Martin Juiz
Tech Scientist
Telecom Argentina S.A.
mjuiz@teco.com.ar

Abstract

Artificial Intelligence (AI) and cloud computing are two factors (among others) that allow Communication Service Providers (CSPs) to become Digital Service Providers (DSPs). CSPs, through AI, have the possibility of transforming their networks, and the operation of services. The natural evolution of applying AI to our networks ("knowledge plane"), is what has been called Autonomous Networks. Networks that can be self-configured, self-healing, self-evolving and self-optimizing. In this paper we will present the journey that we have started in the data-driven operation of our networks and services (AIOps) and the use cases with which we seek to reach autonomous networks. Like also the discussion of some frameworks, the challenges and how long this evolution will take.

Content

1. Introduction

AI Operations (AIOps) is one of the most talked-about acronyms in the IT world these days. In the last year, the term AIOps has been introduced strongly in the Telecommunications industry and has become a special topic in the main conferences [1] [2].

The term AIOps was originally coined by Gartner in 2016 and have pushed the concept into the marketplace. “AIOps combines big data and machine learning to automate IT operations processes, including event correlation, anomaly detection and causality determination” [3].

This definition applies to AIOps in enterprise IT environments: CSP typically have bigger and more complex IT environments with multiple sets of IT challenges. For a long time at the SCTE Cable-Tech Expo and other conferences, data analytics and AI / ML applications have been presented that today we would call AIOps.

A first definition of AIOps in the telecommunications industry is the use of AI for the operation of networks and services. In some way the automation of many of the processes of the operation, being the long-term objective to reach what is called autonomous networks.

Certainly, there has been a lot of activity in trying to use AIOps to improve cable network operation and customer experience. On the other hand, future networks (full virtualization and containerization) and services will generate a volume of data with exponential growth, and this poses a great challenge.

In 2018, after the merger of Cablevision S.A. (MSO) and Grupo Telecom (Telco, MNO), Telecom Argentina began a process of digital transformation, which has accelerated since the COVID 19 pandemic. We are going from being an MSO and MNO, that is, a Communication Service Provider (CSP), to a Digital Service Provider (DSP). Telecom is going to a Multiservice Convergent Network and Multi devices approach, where the Client / User can consume their own and third-party services (platforms) from any device and connected to any of our access networks.

The DSP is not merely a dumb pipe offering shared access to a common utility; it is an online, real-time business that deals with countless transactions every day, managing high volumes of data traffic and

multiple devices per user, and often multiple users per account. The mobile and fixed landscape has changed dramatically and CSP's are fine-tuning their businesses, and their network infrastructure, to cater for the digital needs of the data-hungry customer.

In our networks and services, AIOps has the potential to change, the way we operate, and to become the foundation of the transformation that leads to the fourth industrial revolution. But this requires hard work, a long-term commitment, and a deep cultural change. All Operations Support Systems (OSS) in our current and future networks generate a huge amount of data.

The final aim of all these efforts is to be able to offer our services in an adequate way for our next generations of clients. They are nowadays putting the requirements in the market and driving the evolution of technologies. Our clients do not buy technologies, they buy services. Operating, managing, and provisioning future services with automation processes becomes essential. If we want a complete automation, we will need AIOps.

This technical paper proposes a comprehensive overview of frameworks and applications of AIOps, and we will present the journey that we have started in the AI-driven operation of our networks and services (AIOps) and the use cases with which we seek to reach autonomous networks. This technical paper is organized as follows. Following this section, we introduce AIOps. In section 2, we expose AIOps Service Management Framework. In section 3, we present the state of the art in the evolution towards autonomous networks, the concept of the knowledge plane and our reference architecture. Section 4 we present the AIOps use cases in Telecom Argentina. The last section, which is section 5, outlines the key challenges in our digital transformation journey that we started in 2018.

1.1. AIOps Overview

While AIOps was developed to give scalability to the management of IT systems, given their increasing complexity, the associated platforms and the AIOps framework, we can extend its applications to Telecommunications operations. AIOps has 3 main parts: *observe*, *engage,* and *act*. (Figure 1).

AIOps brings together three different IT disciplines - Service Management, Performance Management, and Automation - to achieve your continuous improvement and information goals. AIOps is the recognition that in our new accelerated and hyper-scaled IT environments, there must be a new approach that takes advantage of advances in big data and machine learning to overcome legacy tools and human limitations.

Implementing an AIOps strategy goes beyond obtaining better analysis of existing data. To create the foundation of a machine learning system that provides continuous information, you need:

- Open access to data including multiple consumable sources of streaming and historical IT data
- Machine learning and algorithms that learn behavior patterns from data and generate automated information
- Automation to act on analytical information and collaborate with the ITSM customer service center

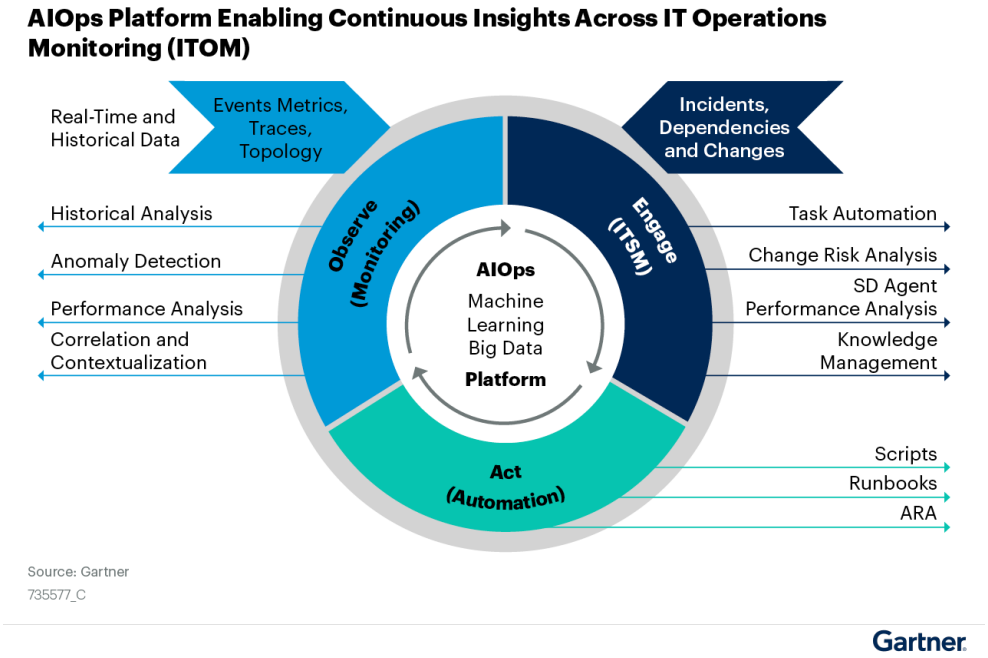


Figure 1 - AIOps for ITOM by Gartner

Various AIOps platforms have been developed in the IT world. “Nowhere is this more evident than in the world of DevOps, a data-rich, back-office practice that presents a perfect sandbox for exploring the power of artificial intelligence. The teams in charge of operations now have a burgeoning collection of labor-saving and efficiency-boosting tools and platforms on offer under the acronym AIOps, all of which promise to apply the best artificial intelligence algorithms to the work of maintaining IT infrastructure” [4].

AIOps platforms enhance decision making across I&O personas by contextualizing large volumes of operational data. I&O leaders should use AIOps platforms to improve analysis and insights across the application life cycle, in addition to augmenting IT service management and automation.

Enterprises are increasing their use of AIOps across various aspects of IT operations management (ITOM) and maturing their use cases across DevOps and SRE practices.

Enable continuous insights across ITOM by supporting these three aspects of AIOps platforms: observe, engage and act [5].

In Figure 2, we present where the AIOps platforms are in the Hype Cycle for Monitoring, Observability and Cloud Operations.

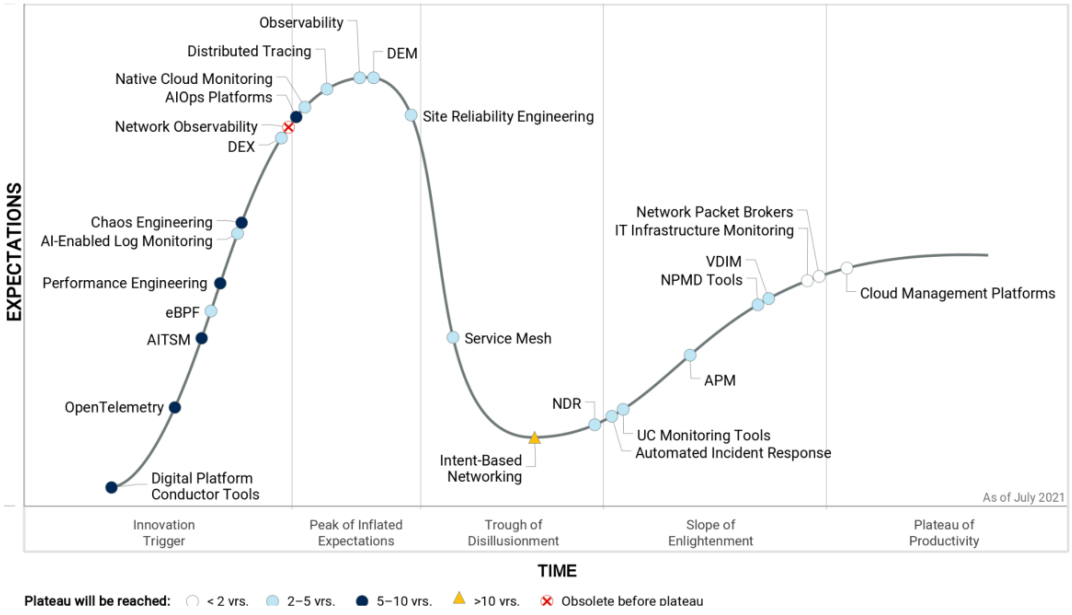


Figure 2 - Hype Cycle for Monitoring, Observability and Cloud Operations, 2021
Source: Gartner

1.2. AIOps and Cloud Native

We are building cloud native applications as a collection of smaller, self-contained microservices has helped organizations become more agile and deliver new features at higher velocity. In the Figure 3 we present the evolution in the IT world from a traditional architecture to one based on containers and the role of AIOps.

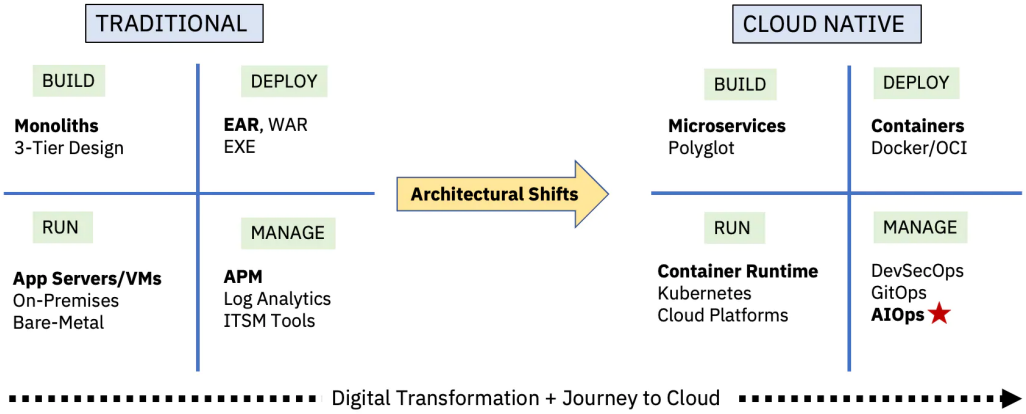


Figure 3 – Evolving IT landscape
Source: IBM[1]

[1] https://www.ibm.com/cloud/blog/aiops-a-path-to-reliability-at-cloud-scale

1.3. Future Networks and AIOps

According to the Ericsson Mobility Report [6] in 2026, 3.5 billion 5G subscriptions are forecast. The Massive IoT technologies NB-IoT and Cat-M1, in 2026, these technologies are forecast to make up 46 percent of all cellular IoT connections. Two-thirds of the world's data didn't exist five years ago, and the datasphere will grow more than 5X by 2025, according to IDC.

At Telecom Argentina our vision of 10 G and 5G is that it is not just an access technology, it is a *digital ecosystem* to create services tailored to our clients, end to end.

"5G promises to be revolutionary if CSPs, vendors, application developers and companies find how to co-create, manage and participate in a digital ecosystem" [7].

The 5G ecosystem is conceived with the technologies and architectures used by the "cloud natives[2]", its functions, services and attributes are controlled by software applications and can be exposed as an "API" or as a service adjusting to the measure of each need. Allowing to expose network capabilities and combine them with AI.

Figure 4 shows the journey towards a native cloud architecture and the level of automation, which was presented in "CableLabs Envision Vendor Forum 2021 Mobile and Convergence".

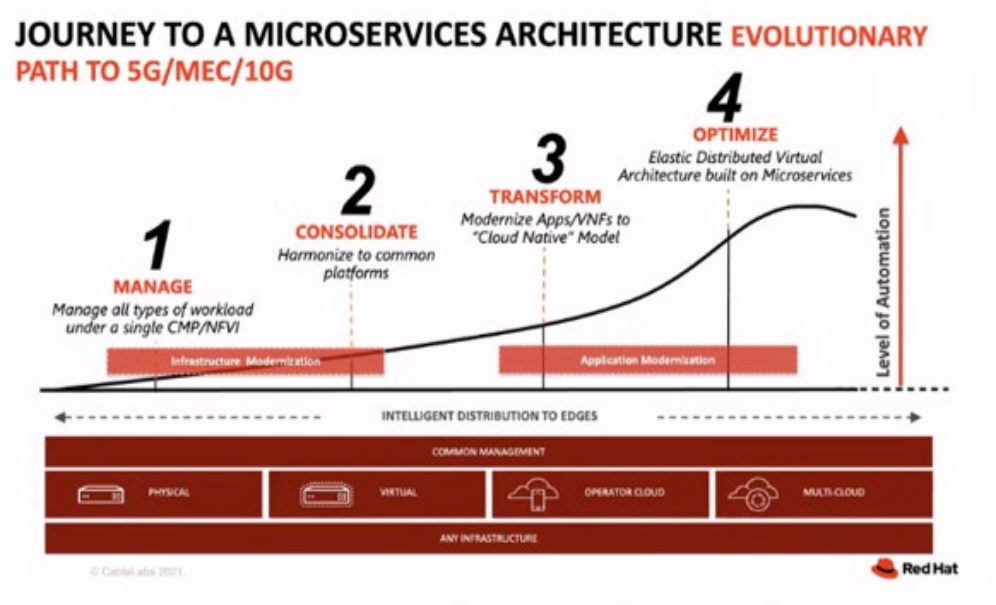


Figure 4 - Journey to a Microservices Architecture
Source: CableLabs[3]

[2] When we deploy network functions in the cloud, we must distinguish between cloud ready, that is, network functions on virtual machines (Virtual network functions, VNFs) with deployment automation and cloud native where network functions are based on containers (cloud-native network functions, CNFs).
[3] https://community.cablelabs.com/wiki/display/COMMUNITY/Envision+2021 (members only)

The computing capacities necessary to generate the services are also distributed and can be located where it "makes the most sense" according to the need.

5G was born AIOps, since from its conception the so-called *plane of knowledge* has been included. That is, a layer within the architecture oriented to the operation and orchestration of networks and services.

1.3.1. Future Networks and AI

Although AI / ML is part of the architecture of future networks, it is important to highlight the combination of 5G, 10G and AI for the generation of new services. Future networks will provide the scalable bandwidth and remote computing power needed to collect and process the growing volumes of data that will drive the proliferation of artificial intelligence, distributing intelligence everywhere. Figure 5 shows the economic value generated by 5G and AI.

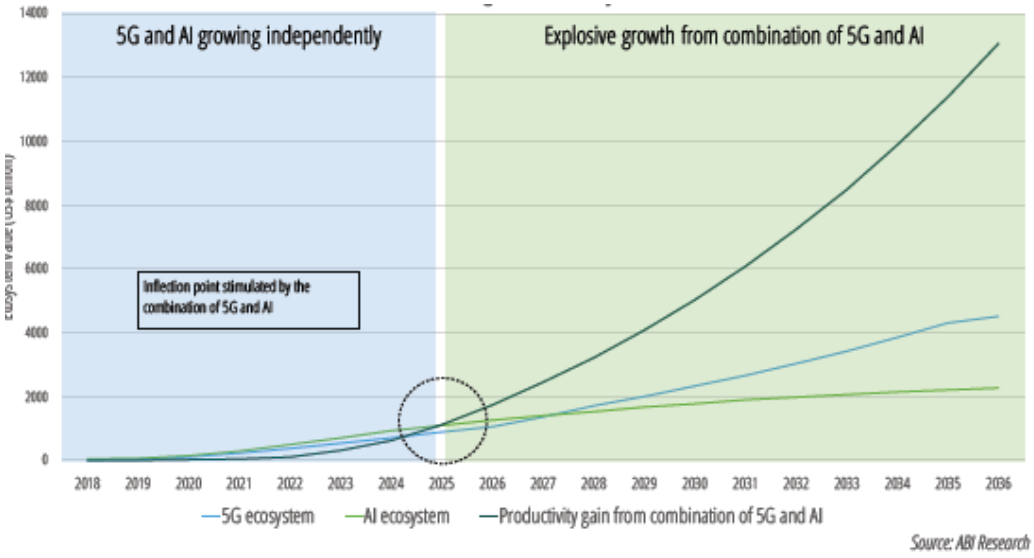


Figure 5 - Economic value generated by 5G and AI

Source: Intel[4]

2. AIOps Service Management Framework – TM Forum

The transformation affects all aspects of a CSP business, including operations that have undergone constant change since a few decades. CSP's role has shifted from offline records related to telephony services to processes that design, install, provision, activate, maintain, and manage inventory for each network-based service. AIOps, envisions a high level of AI-assisted or driven automation in IT and network operations [8].

In addition to the challenges faced by any large enterprise, CSPs include specific business and operations systems and processes, as well as new software-defined and controlled networks. Each of these environments is in flux, moving toward cloud-native architectures that can run on public, private, and hybrid (a mix of public and private) clouds. Automation of operations is the scope of AIOps for CSP,

[4] https://www.intel.la/content/www/xl/es/wireless-network/5g-ai-foundations-business-society-abi-report.html

although the teams working on AI and automation are largely separate today. Additionally, data tends to be disparate, processes can be semi-documented and inconsistently automated, and organizations tend to work in silos.

It is important to understand how different is working with AI from traditional operations. It is based on intention rather than procedure and employs non-deterministic logic and non-predictable outputs from specific inputs.

2.1. AIOps motivation

AIOps envisions a high level of AI-driven or assisted automation in IT and network operations. It's a radical leap, but critical if CSPs want to achieve their automation goals. How they get from here to there is the most important question that AIOps helps to answer.

While current technology (without AI) can automate nearly 75% of manual effort, AI is the best choice for processing extremely large data sets that combine heterogeneous data from multiple sources and where multi-domain, or cross-functional correlation is required.

In the context of our work, we adopt AIOps definition as in Figure 6, where Operations include:

1. Business processes enabled and driven by AI.
2. Service Management frameworks, processes, and tools that have been properly reengineered and adapted to support the operations management of AI-based applications and their components in Production. We call this layer *AIOps Service Management*.
3. AI-based applications that are actively running delivering business and operational services. In AIOps, we assume that key systems and their components in Production are deeply infused with AI capabilities forming a blend of AI and traditional software. In figure 1 we indicate these systems as AI-based BSS, AI-based OSS, AI-based Data Analytics, AI-based ERP, third-party AI platforms, other AI applications.
4. Infrastructure components where automation is driven by AI.

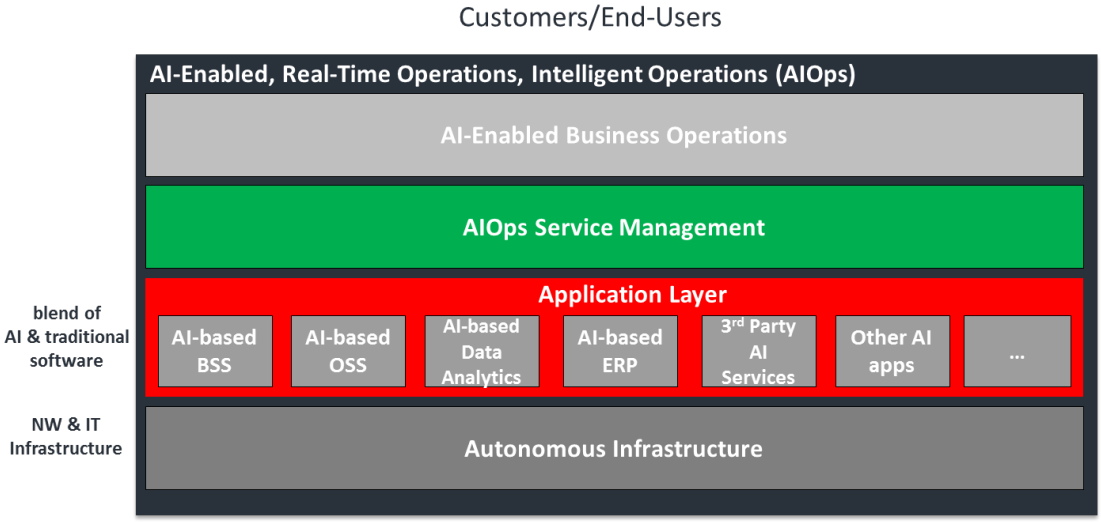


Figure 6 - AIOps and AIOps Service Management
Source: TM Forum

AI can also be a better fit than traditional software to automate responses to both repetitive analytical requirements and ever-changing network configuration and customer experience requirements.

Figure 7 – Goals for AIOps
Source: TM Forum 2020, ACG Research

CSPs should take a few steps on the automation path to increase the degree of hands-on service they can deliver, regardless of the achievement of large-scale autonomous networks - many more milestones remain on that path.

2.2. AIOps Service Management processes

Within the life cycle of traditional software there are two stages: Implementation and Production.

- *Implementation*: it is the set of all the processes and activities that carry a given software and its corresponding components from the Development stage to the Production stage of the life cycle, making that software and its capabilities available to end users.

- *Production:* is the set of processes and activities that operate, monitor, support, maintain all applications and all their components in live production environments, and ensure that the capabilities and services provided by those applications are available and appropriately consumable by end users in accordance with SLA. Production is usually the stage where apps spend most of their life. We design and develop applications in days, weeks, or months. We then deploy them to Production in minutes or during an overnight rollout deployment. There, in production, applications can last for years and even decades.

In AIOps this segmentation is indistinguishable. The dynamic nature of AI software and its ability to learn and evolve autonomously while running in Production create a continuum between the Implementation and Production stages. AI components will permanently move from the deployment state to the live state and vice versa, challenging the distinguishable boundaries that currently exist between the deployment and production stages.

2.3. Real-world use cases

Some use cases based on the real-world business needs of CSPs have been developed. They are focused on customer experience, service quality, business performance, and efficiency.

These include:

- Prevention and prediction of poor customer experience
- Prediction of churn and proactive customer retention
- Accurate service level monitoring
- Proactive root cause identification, communication, and resolution in 5G networks
- Customer complaint prevention
- Preventive Maintenance
- Smart Operations and Maintenance (O&M) for Broadband Services in the Home
- Closed-loop service assurance

2.4. AIOps and Automation

The information provided by machine learning and analytics can drive automation to save time and reduce costs. An AIOps solution can provide functionality for high-value use cases, such as automated event remediation, closed-loop compliance processes, and event-driven automation. Smart ticketing can be particularly valuable as it automatically generates service tickets based on automatic anomaly detection and then optimally routes the tickets to the expert who can best fix the problem or fix the problem automatically. As a result, operators can manage a growing number of assets without increasing labor costs, freeing up staff for more valuable activities, and delivering better quality to customers.

2.5. How to start

AIOps continues an evolution in operations that began some decades ago, when teams focused primarily on offline record keeping, increasingly automated processes. This journey started with observability, moved to actionability, and now culminates in closed-loop automation. The solutions used to implement AIOps fall along similar lines.

- Analyze complex data sets to identify patterns in monitoring, capacity and automation data in hybrid on-premises and multi-cloud environments.
- Provide information that guides immediate actions to reduce costs, solve problems more quickly, and improve quality.
- Closed-loop automation: artificial intelligence and machine learning enable problems to be predicted, found, and solved without human intervention, often before they affect service quality, and to improve quality in a modern environment of software defined networks (SDN).

3. Autonomous Networks

Given the growing virtualization and cloudification (Telco-Cloud)[5] of our networks make possible the evolution towards autonomous networks. At Telecom we adopt the AIOps framework, understanding that it is the from automation path to autonomous networks (AN). "Autonomous networks, put simply, enable our engineers and technicians to go off and chase really interesting problems, leaving the less interesting problems to be operated by software systems" [9].

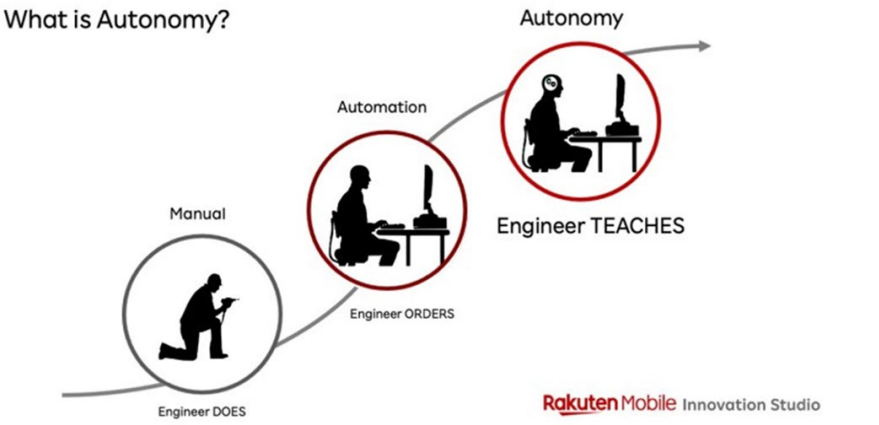


Figure 8 – What is Autonomy?
Source: Rakuten Mobile

We understand AIOps as the way to automate autonomous networks. Although today there are no autonomous networks (AN). Rakuten is the first fully virtualized network in the world, they have the privilege of being able to begin the journey of creating a truly autonomous network [10].

There are several organizations and working groups proposing frameworks, standards and how to apply the AI to automation in the networks. There are three main thrusts of AN: ETSI, ITU, and the TM Forum, which are driven by operators and vendor (Telecom Argentina is participating or is in consultation with them).

- TM Forum - Autonomous Networks Project [11].
- ITU Focus Group on Autonomous Networks (FG-AN) [12].
- ETSI Experiential Networked Intelligence Industry Specification Group (ENI ISG) [13].

3.1. Definitions

We have taken some definitions from those working groups and we started to build our own AIOps and AN framework.

- **Data Analytics**: monitoring data to look for patterns and anomalies (without applying intelligence) and applying those patterns towards effective decision making.

[5]Much of our core of the mobile network is virtualized, however the APIs are proprietary to the vendor.

- **Artificial Intelligence**: the development of computer systems capable of performing tasks that normally require human intelligence; this includes visual perception, speech recognition, decision-making, and translation between languages.

- **Machine learning**: a type of AI that gives machines the ability to learn automatically and improve from experience without being explicitly programmed.

- **Deep learning**: takes machine learning further by processing information in layers, where the result or output from one layer becomes the input for the next.

- **Automation:** within MSOs and MNOs, this means automation of processes that were previously carried out by people; AI is an enabling technology that may (or may not) help with the process of automation. Within MSOs and MNOs, this means automation of processes that were previously carried out by people like configuration, management, operation and testing of physical and virtual devices within the network. With growing costs and the daily emergence of bandwidth-hungry applications, networks cannot be managed manually. Increased levels of network automation help to reduce complexity and are essential for businesses to keep up in the digital world. AI is an enabling technology that may (or may not) help with the process of automation. What it culminates is a network that is highly predictable and highly available improving the business outcomes.

- **Cognitive computing:** like AI, cognitive computing is based on the ability of machines to sense, reason, act and adapt based on learned experience, but whereas AI acts on its analysis to complete a task, cognitive computing provides the information to help a person decide. Like AI, cognitive computing is based on the ability of machines to sense, reason, act and adapt based on learned experience. Cognitive computing refers to computing that focuses on reasoning and understanding at a higher level and in a manner that is analogous to human cognition, rationale, and judgement. Applications of cognitive computing include speech recognition, sentiment analysis, face detection, risk assessment and fraud detection. The difference between AI and cognitive computing lies in the way they approach the purpose of simplifying tasks. AI is used to augment human thinking and solve complex problems. Cognitive computing mimics human behavior and reasoning to solve complex problems like the way humans solve problems.

- **Autonomous networks:** are those that possess the ability to monitor, operate, recover, heal, protect, optimize, and reconfigure themselves; these are commonly known as the self-properties. The impact of autonomy on the network will be in all areas including planning, security, audit, inventory, optimization, orchestration, and quality of experience. At the same time, autonomy raises questions about accountability for non-human decision that affect customers" [14].

3.2. Intent-Based Networking

"Intent" is the keyword in this technology, which describes a network's business objective or an outcome.

Intent-based networking (IBN) is an emerging technology concept that aims to apply a deeper level of intelligence and intended state insights to networking. Ideally, these insights replace the manual processes of configuring networks and reacting to network issues. The goal is networking that uses machine learning and cognitive computing to enable more automation and less time spent on manual configuration and management. With intent-based networking, network administrators define the intent and the network's software finds how to achieve that goal using AI and ML by

performing routine tasks, setting policies, responding to system events, and verifying that actions have been done.

These systems not only automate time-consuming tasks and provide real-time visibility into a network's activity to validate a given intent, but they also predict potential deviations to that intent, and prescribe the action required to ensure it. This greater intelligence makes the network faster and more agile and reduces errors [15].

3.3. Levels

Six levels of autonomous networks were defined by the members of the TM Forum. As shown in Figure 9.

Autonomous Levels	L0: Manual operation & maintenance	L1: Assisted operation & maintenance	L2: Partial Autonomous Networks	L3: Conditional Autonomous Networks	L4: High Autonomous Networks	L5: Full Autonomous Networks
AN services (Zero X)	N/A	Individual AN case	Individual AN case	Select AN cases	Select AN services	Any AN services
Execution	P	P/S	S	S	S	S
Awareness	P	P	P/S	S	S	S
Analysis/ Decision	P	P	P	P/S	S	S
Intent/ Experience	P	P	P	P	P/S	S

Personnel (manual) Systems (autonomous)

Figure 9 – Six Levels of Autonomous Driving Network

Source:TM Forum

Level 0 - manual management: The system delivers assisted monitoring capabilities, which means all dynamic tasks must be executed manually.

Level 1 - assisted management: The system executes a certain repetitive sub-task based on pre-configured to increase execution efficiency

Level 2 - partial autonomous networks: enables closed-loop O&M for certain units based on AI model under certain external environments, lowering the bar for personnel experience and skills.

Level 3 - conditional autonomous network: Building on L2 capabilities, the system with awareness can sense real-time environmental changes, and in certain network domains, optimize and adjust itself to the external environment to enable intent based, closed-loop management.

Level 4 - high autonomous network: Building on L3 capabilities, the system enables, in a more complicated cross-domain environment, analyze and make decision based on predictive or active closed-loop management of service and customer experience driven networks.

Level 5 - full autonomous network: This level is the goal for telecom network evolution. The system possesses closed-loop automation capabilities across multiple services, multiple domains (including partner's domains), and the entire lifecycle, achieving autonomous networks.

3.4. The Knowledge Plane

The research community has considered in the past the application of AI techniques to control and operate networks. In 2003 David Clark et. al propose the knowledge plane (KP) as a *pervasive system within the network that builds and maintains high level models of what the network is supposed to do, to provide services and advice to other elements of the network. The knowledge plane is novel in its reliance on the tools of AI and cognitive systems* [16].

The knowledge plane (Figure 10) paradigm proposes the evolution to a cognitive network, where the devices learn, decide, and act to achieve end-to-end goals. This emerging paradigm is clarifying a set of new cognitive-based protocols and algorithms that optimize network's performance.

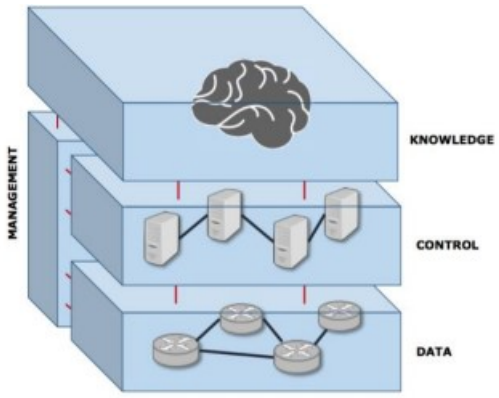


Figure 10 - The four planes in network architecture.
Source: TM Forum.

Several different KP based approaches have been proposed [17]. But it is not until the development of NFV and SDN that such proposal once again takes hold in communities such as the TM Forum, IETF, 3GPP, ETSI etc and the Industry.

In [18] progress is made in the definition of a new paradigm based on this plane. This is knowledge-defined network (KDN) operates by means of a control loop to provide automation, recommendation, optimization, validation, and estimation. The KDN paradigm is also taken by the TM Forum, ETSI and ONAP as a proposal to specify future architectures.

An example of the plane of knowledge in an SDN is presented in Figure 11 [19].

The introduction of AI and automation using AI, in short, seeks to ensure greater performance and efficiency of networks.

The paradigm shift that will be brought about by the introduction of these new technologies includes a substantial shift from a focus on network operations to a focus on the user experience.

We have conceptualized these AIOps tools that interact with the different types of networks as Knowledge Plane, a "place" where the massive amount of data obtained from the network is processed with the different

AI tools, either in real time or in a post-processing, and that, based on results, produces modifications in the network itself. This is called closed-loop automation[6].

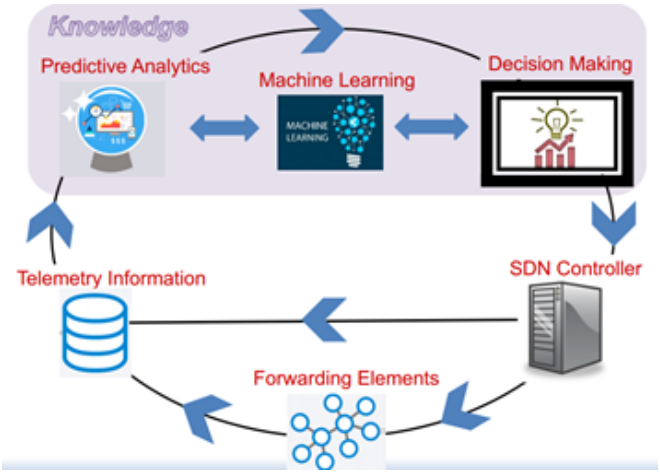


Figure 11 – Knowledge Plane in a SDN

3.5. The Knowledge Plane at Telecom

More than a decade has passed since the emergence of a paradigm of autonomous computing in the world of telecommunications. Back then, there was a gap between that paradigm and the capacities of the networks. However, a path has been taken in recent years with the adoption of cloud computing, NFV and SDN.

These technological advances have made available a more agile infrastructure, computing capacity and storage as resources more abundant than ever. Motivated by this evolution, together with the ever-growing need to improve the management and administration of networks and services, we present this first approach to the plane of KP. We adopt the knowledge plane paradigm for our Telco Cloud project.

We define a reference architecture (Figure 12) with the purpose of automating the services from end to end, with a holistic view of the network and services. It is based in the different WGs and other initiatives of other CSPs and vendors.

At the beginning, and to have a common language, a framework was defined in which the preponderance of the knowledge plane in the different layers of the network, both virtual and physical, is highlighted.

[6] Note: There are many different names for closed loop, cognitive loop, Monitor Analyze-Plan-Execute over a shared Knowledge (MAPE-K), observe–orient–decide–act (OODA), etc.

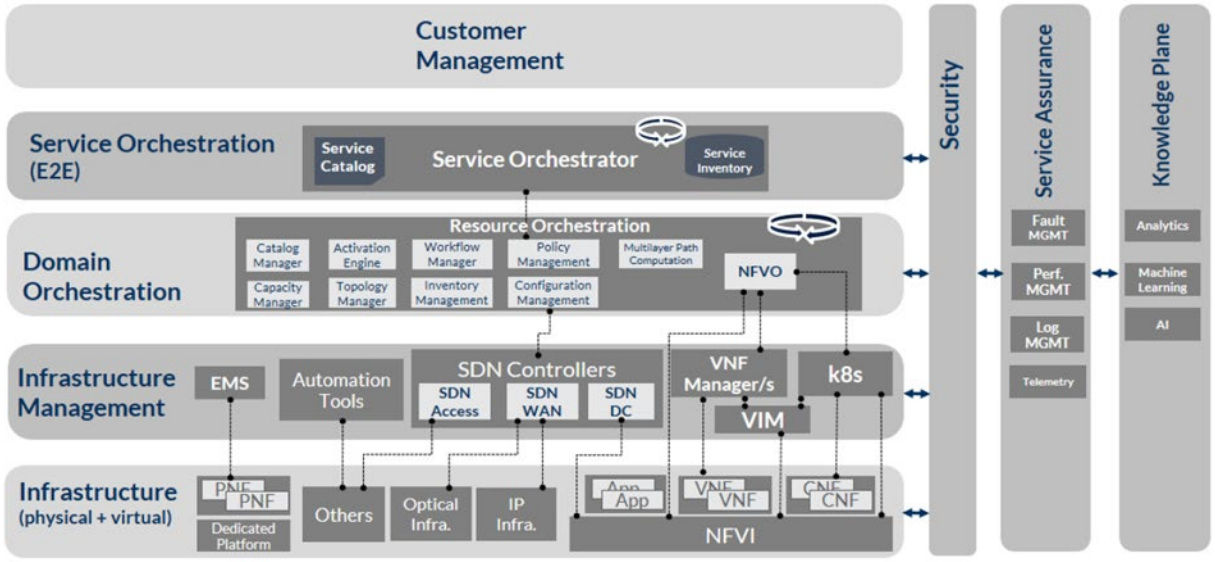


Figure 12 - Knowledge Plane Integration

There is no doubt that in the course of time this general vision will not only be modified but also specified. A result that only experience can provide.

3.6. An outline of Knowledge Plane architecture

Based on the different standards and our own experience in recent years, we can propose a conceptual architecture where two modalities of KP and Close Loops are distinguished.

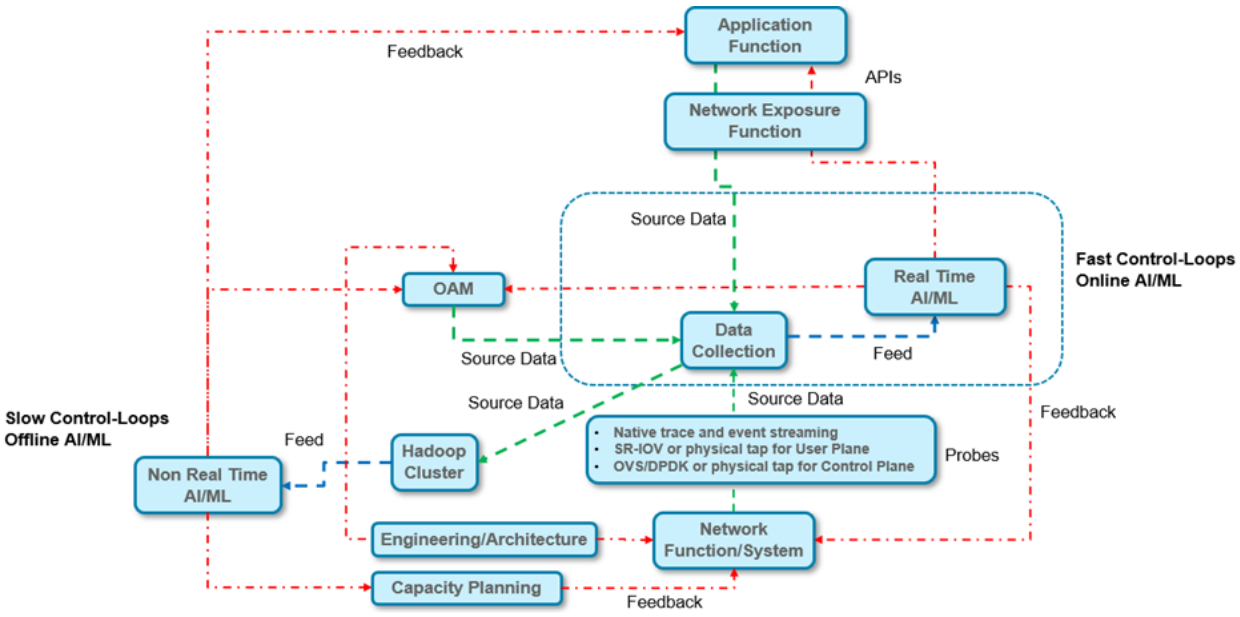


Figure 13 - Knowledge Plane architecture

Offline AI/ML (or Slow Control-Loops): based on a large amount of historical data, stored in a data lake, statistical models or other algorithms are applied (for example, time series) and the results obtained by data scientists lead to implement changes in the network, in OAM and in the long-term architecture.

Online AI/ML (or Fast Control-Loops): This instance is very dependent on:

- type of network to which it is applied
- if it is Access or Core

The information obtained from the network through probes, applications and OAM is stored in a Data Collection in real time. Then they could be combined in new tables to feed AI/ML algorithms (note that we are not putting training because it could be that the algorithm is a type of Clustering).

The Data Collection and Real Time AI/ML could be embedded in a network element or could be external to it.

As the algorithm works in real time, it would be necessary to prepare it, to test it with real data, making a pre-processing in laboratory before taking it to production, with a portable solution (for example Kubernete/KubeFlow) in a process of continuous improvement of the software. The result of the algorithm feeds back directly to the network elements, OAM, and applications. Obviously, the supervision of a human expert should be constantly validating the automated work of the algorithms ("He Teaches").

3.7. AI Techniques in Network Automation

Many techniques that can be used for decision making with each providing different cognitive capabilities. Many offer inference capabilities and only a few offer perceptive capabilities, so it is important to identify the problem requirements to choose the right technology. Figure 14 summarizes the comparative capabilities of the technologies [20].

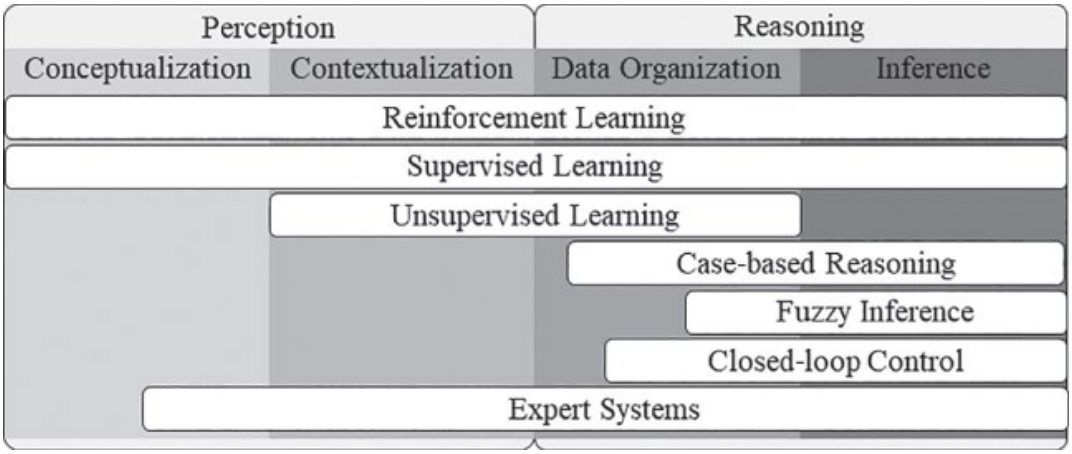


Figure 14 - Capabilities of the different automation techniques.

4. AIOps at Telecom Argentina

At Telecom, there are a few initiatives that we have considered within the AIOps framework. We believe that it is important that they are seen under the focus of Augmented Intelligence. Nowadays, as a STEM team one of our missions is to lead AIOps in our current and future networks. Our recipe within STEM team is diversity, work in cells, agile mindset, and self-learning. In the following sections we present two of the initiatives we have been working on.

4.1. Ada

At Telecom Argentina, we have been working on a tool for network dimensioning for some time now [21]. In 2017 we developed STEM-ML for Cablevisión, a tool that assisted the decision-making process for the dimensioning of the HFC network. Later, we decided to incorporate data from the radio-access network (RAN) and relaunched the project under the name Ada[7]. The goal is to have a tool that combines machine learning and SME input to assess decisions on CAPEX and OPEX investments.

Historical data is collected for every HFC node, as well as for every cell on the mobile network. The idea is that for decisions in the short term (up to a year from the present), ML-based forecasts are provided, and the role of experts is to set priorities and act based on them. For long-term decisions (5-10 years periods), on the other hand, ML-based forecasts would require a longitude of history that hasn't yet occurred. In our experience, however, it is possible to provide reasonable approximations if forecasts are supported not only by historical data but also by experts' knowledge. Hence, we are currently working on finding ways for the experts to pass input to ML algorithms.

In the past, planning engineers used to research on potential scenarios and make a series of calculations to approximate what they thought was going to happen in the long-term, based on their knowledge of the average client and use cases. Working together, we built profiles for the operating sectors, so they can refine on what would be expected from a variety of use cases. To better understand this, let's see that for example, the case of a group of households where the service is mostly used to check social media, and the income is low will not be the same as in another group of households where there are heavy streamers, and the income is higher. We are providing the planning teams with better information to start with.

As this paper is being written, we are working on how to enable engineers to pass information to a model about what they expect the CAGR would be at different kinds of operating sectors, in the next 5-10 years. This will involve a combination of simulation and forecasting techniques. Also, it will require a certain level of automation that would not be achievable outside a ML framework.

4.2. Customer claim prediction

Within the domain of AIOps one of the most popular use cases is customer claim or tickets prediction [22] [23]. The main idea is to use information derived from different elements of the network, such as: CMTS, cablemodems, electronic devices, etc. to estimate the probability of a customer to generate a complain.

Information is generally collected through OSS systems and ingested in a machine learning (ML) pipeline where preprocessing, analysis and ML model testing is performed. We are currently developing this kind of solution to ultimately increase customer satisfaction.

[7] The Ada name of the AIOps tool for planning the mobile network and HFC, is in honor of who is considered the first programmer in history, Ada Lovelace (1815-1852)

Using hourly collected information from over 1.2 million DOCSIS 3.0 cablemodems we are trying to anticipate customer complains two days in advance. To handle this amount of data we have partnered with Google[8] to develop and deploy this project using Google Cloud Platform (GCP) services. Although we are in preliminary stages (MVM1), many steps have been given towards the final implementation.

Potentially relevant variables (e.g., signal to noise ratio, consumed bytes, average Rx, t3 and t4 time outs, etc.) have been identified and our Service Assurance team have been able to efficiently transfer these data, collected by our OSS systems, to GCP.

Daily, cablemodem and customer claims are transferred to a relational database mounted on GCP. After the data is successfully ingested, an automated data cleansing process is run using Big Query[9] to guarantee a proper information quality level. In addition, a Data Studio[10] dashboard has been created to double check this already curated information. We have found very useful to quickly visualize quality KPIs, such as: MAC Addresses count, percentage of missing values, customer claim evolution, etc.

Even though the data cleansing stage is run over the whole data set, we are using a different strategy to develop and test different ML models. Two random samples containing 10,000 and 50,000 cablemodems are being taken to accelerate the processes of feature engineering, hyper parameter tunning, model evaluation and model selection.

We are exploring different approaches to maximize precision and recall metrics[11].

At time this document is being written, we have tried a series of XGBoost models with different sets of hyperparameters[12] and feature engineering scenarios. We are also working on a different approach, reframing the whole task as a survival analysis problem. Furthermore, we are considering using recurrent neural networks to capitalize the sequential structure of the cablemodems metrics data.

Regarding technical challenges, two of the most difficult tasks have been dealing with an extremely unbalanced dataset and label noise. Although from a business perspective having a low proportion of customer claims is a good indicator, some ML models can struggle to learn from these kind of data sets. In our case, after filtering the target population (DOCSIS 3.0 residential customers), the positive class dropped to 0.1%, which added more complexity to the problem. To reduce the impact of this issue on model performance we applied hyperparameter tunning[13] and a technique called SMOTE [24] combined with an under sampling of the majority class. Both approaches led to improvements in model performance.
On the other hand, records are not always properly labeled. Some customers, for example, don´t make a complain, even when their cablemodem is not working adequately. This leads to label overlapping. Some cases that should have been labeled as positive get labeled as negative, hence, remaining in the same region in the feature space of the positive class with the opposite ground truth label. To handle this problem, we have used a Python package for confident learning called Cleanlab [25]. The main idea of this approach is to run a model and remove cases that have been confidently classified within a class when their ground truth label corresponds to another class.

[8] Telecom Argentina adopted a multi-cloud strategy On-Premises, Amazon, Azure and Google.
[9] As stated from Google in their web page, Big Query is a: “Serverless, highly scalable, and cost-effective multicloud data warehouse designed for business agility”
[10] Google tool for data analysis and visualization.
[11] We found these two metrics to be the best way to measure model performance based on the characteristics of the data set (e.g., highly unbalanced) and business needs (maximize precision to avoid as many false positives as possible)
[12] We have performed random search, and we are currently changing to a Bayesian optimization approach.
[13] We searched for the best value of “scale_pos_weight”, the hyperparameter used by XGboost to handle unbalanced classes.

We have been able to overcome many challenges and, although we are still in initial phases of this project, we hope to deploy a state-of-the-art ML system to give one more step towards AIOps goal: fully automated networks.

Finally, with the aim of showing a high-level view of our project, the Figure 15 below describe the data flow and the desired outcomes.

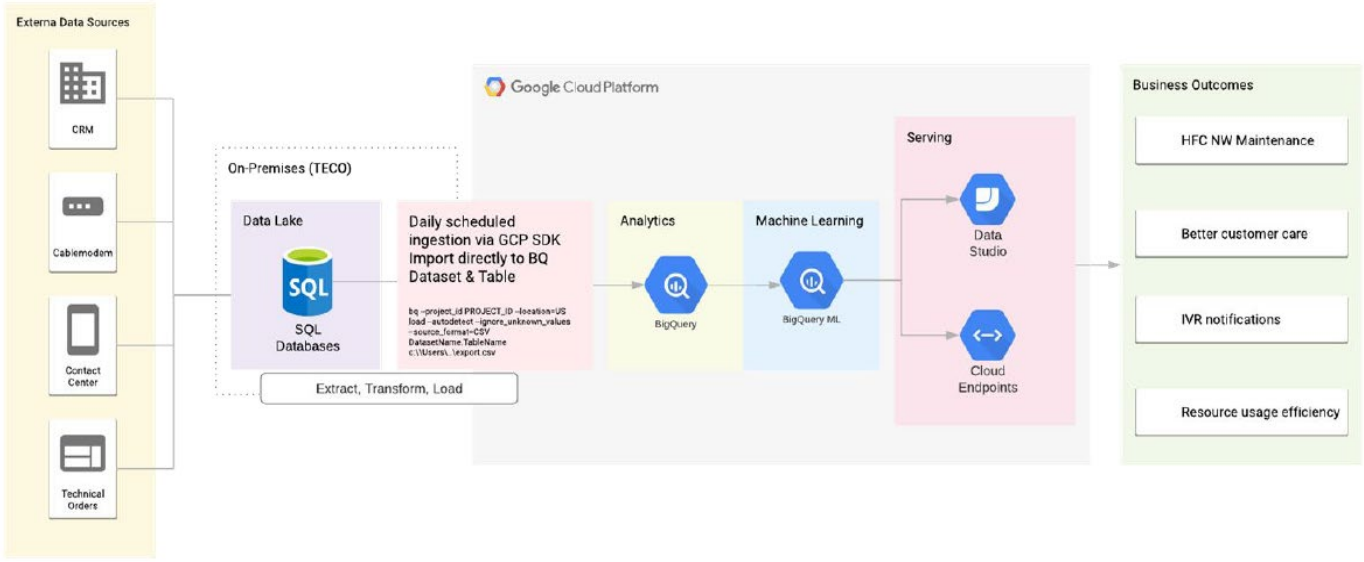


Figure 15 - Data Flow and outcomes

We are currently including DOCSIS 3.1 cablemodems in the first model (MVM 2) and starting field tests working with Service Quality, Field Service and Customer Experience teams.

On the other hand, we are beginning to define the MVM 3 related to a closed loop system to impact the QoE of some of our services. That is, with AIOps frameworks, starting our journey from automation to autonomous networks.

5. Conclusions

Throughout this technical paper we have seen the need to automate the operations of our current and future networks using artificial intelligence, presenting the AIOps framework that we have adopted in Telecom Argentina. This adoption was accelerated during the beginning of the pandemic that has produced an enormous change in our organization. However, we are still on the path of digital transformation and AIOps plays a fundamental role.

Within this transformation, however, there is something in which we cannot apply AIOps, it is in interpersonal skills that will not be able to be replaced. We must develop the empathy, collaboration, and autonomy of our work teams. On the other hand, although the path of AIOps is from automation to autonomous networks, we must not forget that we must see it from the perspective of Augmented Intelligence (AgI).

"While Artificial Intelligence is creating machines to work and react like humans, Augmented Intelligence is using those same machines with a different approach to improve human capabilities. In fact, AgI involves people and machines working together, leveraging their own strengths to achieve greater business value. In other words, the primary goal of AgI is to empower humans to work better and smarter" [26].

Acknowledgement

The authors want to thank the sponsorship of Miguel Fernández (CTO at Telecom), Gustavo Ramos, Gastón De Arriba and Mauricio Ferreira in relation to the work carried out and future work on the AIOps framework, including this technical paper. And from Service Quality Management to Leonardo Di Pilato, Jorge Pittana and Fabio Maggiore, from Capacity Planning to Alejandro Ambrosio, Natalia Clivio and Alex Ferraro and from Mobile Networks Planning to Oscar Rodríguez and Germán Montes de Oca and from Qualcomm to Alex Florea, for the teamwork that we have been doing in recent years.

Abbreviations

10G	10th generation
3GPP	3rd Generation Partnership Project
4G	4th generation
5G	5th generation
AgI	augmented intelligence
AI	artificial intelligence
AIOps	artificial intelligence operations
AN	autonomous networks
API	application programming interface
BSS	business support system
CMTS	cable modem termination system
CNF	cloud-native network function
CSP	communication service provider
DOCSIS	data over cable service interface specification
DSP	digital service provider
ENI ISG	experimental networked intelligence industry specification group
ERP	enterprise resource planning
ETSI	European Telecommunications Standards Institute
FG-AN	focus group on autonomous networks
GCP	Google Cloud Platform
HFC	hybrid fiber-coaxial
I&O	infrastructure and operations
IBN	intent-based networking
IETF	Internet Engineering Task Force
IoT	internet of things
IT	information technology
ITOM	IT operation management
ITSM	IT service management
ITU	International Telecommunication Union
KDN	knowledge-defined network

KP	knowledge plane
KPI	key performance indicator
MAC	media access control
MAPE-K	monitors analyze-plan-execute over a shared knowledge
ML	machine learning
MNO	mobile network operator
MSO	multiple system operator
MVM	minimum viable model
NB-IoT	narrow band IoT
O&M	operations and maintenance
OAM	operations, administration, and maintenance
OODA	observe–orient–decide–act
OSS	operations support systems
QoE	quality of experience
RAN	radio-access network
SCTE	Society of Cable Telecommunications Engineers
SDN	software defined networks
SLA	service level agreement
SMOTE	synthetic minority over-sampling technique
SRE	site reliability engineering
STEM	science, technology, engineering, and mathematics
VNF	virtual network function
WG	working group

Bibliography & References

[1] AIOPS 2020 International Workshop on Artificial Intelligence for IT Operations Available: https://aiopsworkshop.github.io/.

[2] AIOps Expo Florida 2021 Available: https://www.aiopsexpo.com/.

[3] Available: https://www.gartner.com/en/information-technology/glossary/aiops-artificial-intelligence-operations.

[4] Available: https://www.cio.com/article/3625580/top-aiops-platforms.html?upd=1626871907417.

[5] P. Prasad, P. Byrne, J. Chessman, Market Guide for AIOps Platforms, Published 6 April 2021 - ID G00735577 - 2021. Available: https://www.gartner.com/en/documents/4000217/market-guide-for-aiops-platforms.

[6] Available: https://www.ericsson.com/4a03c2/assets/local/mobility-report/documents/2021/june-2021-ericsson-mobility-report.pdf.

[7] T. McElligott, 5G future: Targeting the enterprise 9 2019. Available: https://inform.tmforum.org/research-reports/5g-future-targeting-the-enterprise/.

[8] E. Finegold, AIOps: From Automation to Autonomous Networks 12 2020. Available: https://inform.tmforum.org/research-reports/ai-ops-from-automation-to-autonomous-networks/.

[9] 6 2021. Available: https://rakuten.today/blog/rakuten-mobile-edge-computing-hub.html.

[10] R. Mobile, «Evolving Autonomous Networks,» Available: https://netlab.mobile.rakuten.co.jp/.

[11] T. Forum, «Autonomous Networks Project,» Available: https://www.tmforum.org/collaboration/autonomous-networks-project/.

[12] «ITU (FG-AN),» Available: https://www.itu.int/en/ITU-T/focusgroups/an/Pages/default.aspx.

[13] «ETSI - ENI,» Available: https://www.etsi.org/technologies/experiential-networked-intelligence.

[14] «ITU-T (FG-AN),» Available: https://www.itu.int/en/ITU-T/focusgroups/an/Documents/FG-AN_Terms_of_Reference.pdf.

[15] U. Gasser, V. A. F. Almeida, *A Layered Model for AI Governance,* IEEE Internet Computing 21 (6) (November): 58–62. doi:10.1109/mic.2017.4180835., 2017.

[16] D. D. Clark, C. Partridge, J. C. Ramming, J. T. Wroclawski, *A Knowledge Plane for The Internet,* New York: Conference on Applications, Technologies, Architectures and Protocols for Computer Communications (SIGCOMM), 2003.

[17] K. R. Sollins, «An Architecture for Network Management,» de *Workshop on Re-Architecting the Internet (ReArch)*, New York, NY, 2009.

[18] A. Mestres, A. Rodriguez-Natal, J. Carner, P. Barlet-Ros, et.al., *Knowledge-Defined Networking,* SIGCOMM Computer Communications, vol. 47, nº 3, pp. 2-10., 2017.

[19] Z. Zhu, *Knowledge-Defined Network Orchestration in a Hybrid Optical/Electrical Datacenter Network,* Dublin, Ireland: Conference on Optical Network Design and Modeling, 2018.

[20] S. S. Mwanje, C. Mannweiler, Towards Cognitive Autonomous Networks, Wiley, 2020.

[21] C. Righetti, M. Fiorenzo, E. Gibellini, et.al., *Network Capacity and Machine Learning,* SCTE Cable Tec Expo 2017, 2017.

[22] J. Hu, et.al., *CableMon: Improving the reliability of cable broadband networks via proactive network maintenance.,* 17th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 20), p. 619-632., 2020.

[23] J. Watson, R. Brooks, A. Colby, Christian, P. Kumar, A. Malhotra, M. Jain, *Predicting Service Impairments from Set-top Box Errors in Near Real-Time and What to Do About It,* SCTE Cable-Tec Expo 2018, 2018.

[24] N. V. Chawla, et.al., *SMOTE: synthetic minority over-sampling technique.,* Journal of artificial intelligence research, 2002, vol. 16, p. 321-357., 2002.

[25] C. NORTHCUTT, L. JIANG, I. CHUANG, *Confident learning: Estimating uncertainty in dataset labels.,* Journal of Artificial Intelligence Research, 2021, vol. 70, p. 1373-1411, 2021.

[26] C. Righetti, E. Gibellini, et.al., *Augmented Intelligence: Next Level Network and Services Intelligence,* SCTE NCTA CableLabs 2020 Fall Technical Forum, 2020.

The Path to 100 Gbps DAA Nodes

Analyzing DOCSIS Bandwidth and its Impact on the CIN

A Technical Paper prepared for SCTE by

John T Chapman
CTO Broadband Technologies & Fellow
Cisco Systems
jchapman@cisco.com

https://www.linkedin.com/in/john-t-chapman/

1. Introduction

This white paper will look at DOCSIS bandwidth as it relates to a DAA (Distributed Access Architecture) node. Although similar to an Integrated CMTS (Cable Modem Termination System), there are some additional design considerations to be aware of. The first section looks at various design considerations. The second section then goes right into the bandwidth calculations.

Specifically, the following goals will be addressed with this white paper:

Network Design Considerations

- Review channel and port differences between an integrated CMTS (I-CMTS) and a DAA node
- Discuss a new concept of DS:US bandwidth ratio (DUCR) and peak versus average capacity

Bandwidth Studies of DAA Nodes and its CIN requirements

- What is possible today (with some DOCSIS 3.1 and video)
- What is possible with DOCSIS 3.1
- What is possible with DOCSIS 4.0
- What is possible after DOCSIS 4.0

This white paper will establish a set of metrics and then use them along with an extensive bandwidth analysis of DOCSIS to provide guidelines on what technology direction operators may decide to go.

All calculations are supported by a publicly available spreadsheet as show in Figure 20 by contacting the author. The numbers capture in the while paper are supportive of general conclusions. As time goes on, the spreadsheet may contain more exact bandwidth numbers for a given upstream and downstream channel configuration.

2. Network Design Considerations

2.1. Review of I-CMTS channel/port configurations

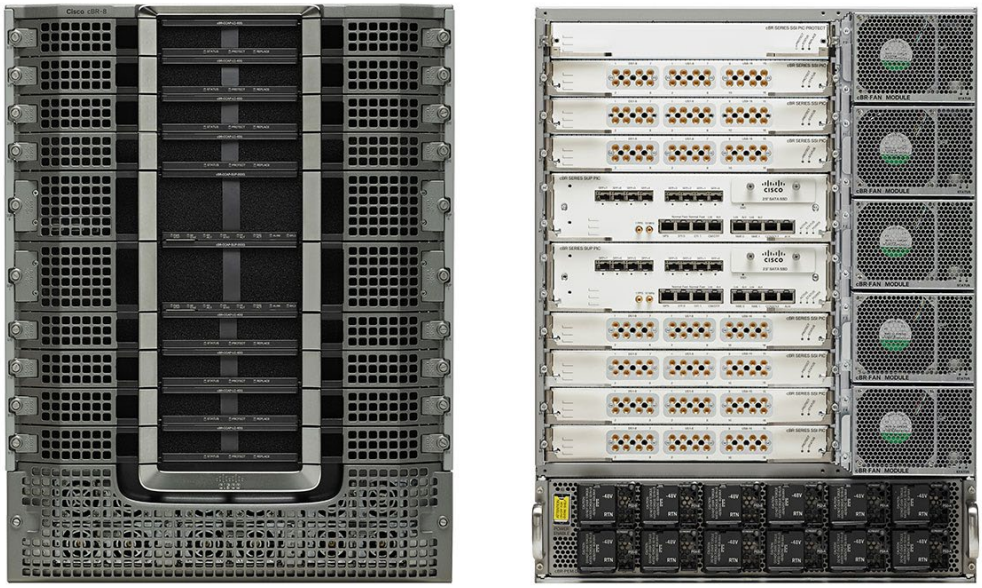


Figure 1 - Example of an I-CMTS

To start the analysis, we first begin with the Integrated CMTS (I-CMTS). A typical I-CMTS is shown in Figure 1. The I-CMTS was designed to connect to separate radio frequency (RF) downstream (DS) combining and upstream (US) combining network. As such, the DS and US RF ports are separate and do not contain a diplexer.

The I-CMTS also contains RF redundancy. In this example, there are eight line cards (8 LCs) but only seven RF physical interface cards (PICs). Internally, there is a 7+1 redundancy scheme that allows one LC to replace any of the other seven line cards without changing the external RF path.

Here are some example I-CMTS capacity values

CMTS Ports

- 7 LC @ 8 DS ports x 16 US ports = 56 DS ports and 112 US ports per I-CMTS chassis

Channels per I-CMTS Port

- DS: 96 SC-QAM and 2 OFDM
- US: 8 A-TDMA and 2 OFDMA

The SC-QAM and A-TDMA channels are for DOCSIS 3.0 and earlier versions. OFDM and OFDMA channels are added for DOCSIS 3.1 and DOCSIS 4.0

Ethernet Capacity

- Dual 100 Gbps

The Ethernet connectivity represents the total aggregate capacity of the chassis. The DOCSIS protocol allows for over subscription on the HFC (Hybrid Fiber-Coax) plant. In addition, the I-CMTS allows for some oversubscription on the CMTS chassis. Thus, the total RF capacity may be greater than the Ethernet capacity. Over-subscription is important as each Ethernet port needs to connect to a router port, so each port has a cost associated with it. Over-subscription both keeps costs down and represents proper traffic patterns.

Internal Constraints

- Bus and backplane bandwidths, various memory sizes, packets-per-second (PPS) limitations of data and control planes, CPU core count and clock speeds, software architecture, etc.

These internal constraints lead to DOCSIS performance constraints such at the number of service groups (SGs), classifiers, and data throughput.

2.2. DAA Node channel/port configurations

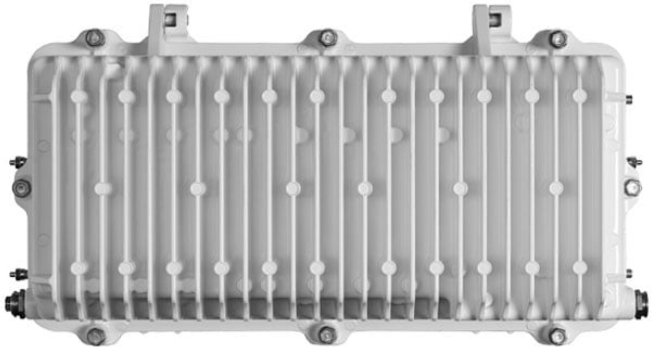


Figure 2 - HFC DAA Node

There are two types of DAA nodes defined at CableLabs.

- The first type is the Remote PHY (RPHY) system. The RPHY design takes the PHY chip out of the I-CMTS and puts it in the RF node. The MAC chip remains in the I-CMTS and is connected to the PHY chip in the node with a 10 Gbps (or higher) link. [1][2][3]
- The second approach is a Remote MAC and PHY (RMACPHY) system where the entire layer 1 and layer 2 CMTS is placed in the node. Layer 3 connectivity is supplied by a leaf router. Aggregation of RMACPHY nodes replies on system software known composed of a MAC Manager and a PacketCable Aggregator.

Both RPHY and RMACPHY are part of the Flexible MAC Architecture (FMA), although FMA commonly refers to RMACPHY as RMACPHY is its first system deliverable. Additional non-DOCSIS features such as MPEG-TS video, narrowband digital forward (NDF), narrowband digital return (NDR), out-of-band (OOB) channels, are all managed using the RPHY protocols and add to the Ethernet bandwidth requirements.

For the analysis in this white paper, which focuses on RF channels, RF ports, and Ethernet ports, both the RPHY and RMACPHY systems are identical. Hence forth, the term DAA node equally applies to an RPHY node or a RMACPHY node. A typical RF node is shown in Figure 2. A breakdown of the connectivity of that node is shown in Figure 3.

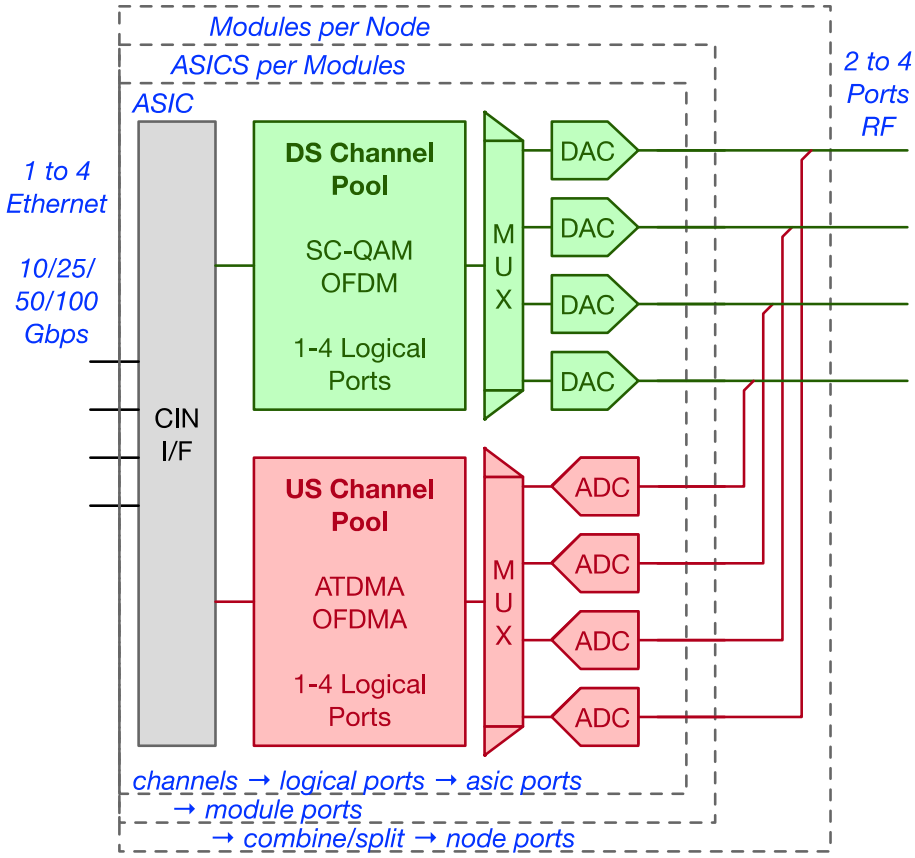


Figure 3 - DAA Node channels and ports

A DAA node has two to four RF ports. These RF ports after a diplexer, so they contain both DS and US spectrum and channels. That is different than an I-CMTS. There is also no RF redundancy which simplifies things.

A DAA node is composed of:

- One or more RF ASICS that fit into a module
- One of more modules that fit into a node

The DAA ASIC contains a pool of RF channel for SC-QAM, OFDM, ATDMA, OFDMA as well as resources for NDR, NDF, OOB 55-1, OOB 55-2, and the upstream spectrum burst receiver. This pool of resources can be organized into one or more port groups which are then mapped into DS digital-to-analog (DAC) and US analog-to-digital (ADC).

Now, here is a tricky part. Say that the chip supports a 1x2 port config (so 1 DS and 2 US). That could map to one DAC and two ADC. Or, it could map to four DAC and four ADC. Now, why would that be interesting? There is an additional function that DAA silicon may provide called digital pre-distortion (DPD).

The downstream power amp is a class A amplifier which means it does not run anywhere near saturation. DPD pre-distorts the DS signal in the opposite manner that the DS power amp may distort it. When combined, the linearity of the power amp is increased. This allows the output power amp in the node to be biased at a lower voltage and thus the power amp runs at a lower power level for a given output level. To make this all work, the ideal DPD circuit has to monitor all four DS power amps. It does with an extra four set of full range ADC converters.

In the upstream direction, a node would normally combine RF inputs from four down to two or one and then connect to an ASIC. This creates noise funneling where the noise floor from multiple inputs is combined. Instead, if each RF upstream input is received on a separate ADC, and then digitally processed, then the impact of noise funneling can be reduced.

The DAA module acts like a small CMTS. There can be one or two modules in DAA node. The modules may be split across downstream ports or directly connected. The upstream ports may be combined or directly connected to the US ports on the modules.

2.3. DAA Node CIN connectivity

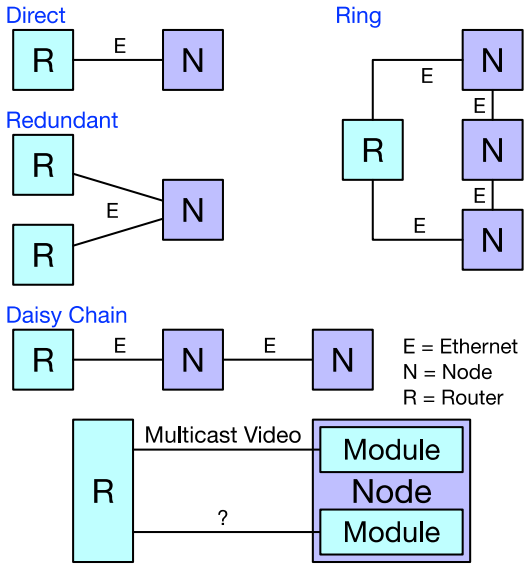


Figure 4 - DAA Node Ethernet Connectivity

The aggregate of all bits from all the DS and US RF channels pass through the Ethernet ports and through the Converged Interconnect Network (CIN). The definition of the CIN is in part defined by the use of the Ethernet ports.

The various Ethernet configurations are shown in Figure 4.

- Direct is most common
- Daisy chain allows multiple nodes to share a common 10 GE link. This may become a less popular option as node bandwidth increases
- Rings are rare but they do provide a redundant path on a common 10 GE link. Rings may become less popular use as the node bandwidth requirements increase. Or, the rings could just move to higher bandwidth
- Redundant connectivity is rare today, but may become useful in the future for a CIN.

Imagine a DAA node that connects to separate hub sites. If one hub site fails, the DAA nodes remain connected to the other hub site. In fact, since the DAA nodes are now IP connected, each DAA node connects to all hub sites, not just the next hop hub site. This is a powerful concept as it allows DAA nodes that might not fit on one hub site to be connected to another hub site.

Backhaul may be one or more Ethernets per Node, in addition to the above configurations. This presents both an opportunity and a dilemma for MPEG-TS video. Linear video which can consume 64 SC-QAM channels is usually sent using IP Multicast.

- Multicast video may be shared across network ports of the same module if designed to do so
- Multicast video cannot be shared across network ports that go to separate modules

So, if you have to 1x2 modules to make a 2x4 node, then the multicast video will have to be sent separately to each. If you have one 2x4 module, and if the ASICs permit sharing of video, then you may only need to send one multicast video stream to the node.

This is a very important consideration for DAA node planning as it could double the number of Ethernets you require. For denser DAA node configurations, like for DOCSIS 4.0, it may be necessary to move to 25 Gbps Ethernet or higher.

Make sure to work out these considerations on the DAA bandwidth spreadsheet.

2.4. DUCR – Downstream to Upstream Capacity Ratio

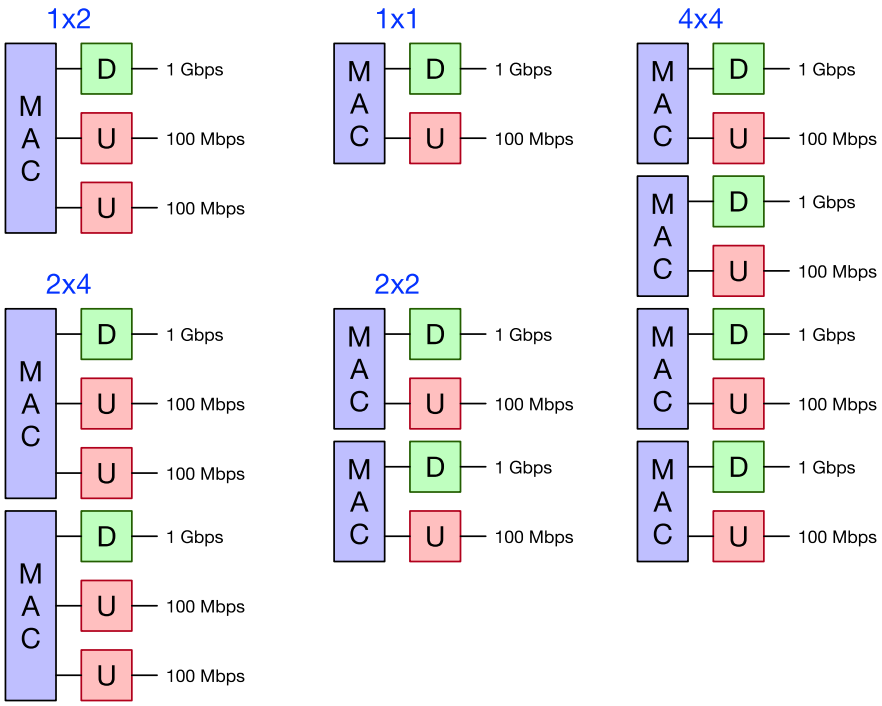


Figure 5 - DAA Node Ethernet Connectivity

As mentioned before, even though a DAA node may have four physical DS and US ports, it has an effective number of ports depending upon the internal configuration of resources. Figure 5 shows these various combinations with example bandwidth numbers. Here are the important points to keep in mind.

1x1 modules

- One module creates a 1x1 node
- Two modules create a 2x2 node
- The 1x1 module may be a 1x2 module with one port disabled.

1x2 module

- One module creates a 1x2 node
- Two modules create a 2x4 node
- This is the current state of the industry in 2021

2x2, 2x4 module

- Next generation chipsets contained in a single module
- Targeted at one module per node.

4x4 module

- Not available or deployed yet, but could be with future ASIC densities. ADC and DAC density is already here.
- Also, two 2x2 modules could form a 4x4 node.

These port configurations can be further sorted into two classes:

- Single-return: 1x1, 2x2, 4x4
- Dual-return: 1x2, 2x4
- Quad-return: 1x4 (not in actual use)

A single-return system has the same number of DS and US ports; a dual-return system has twice the upstream ports as downstream ports. Quad-return is not in use but is included for consistency. (the usage of these terms are defined here)

So, when should you use a single-return system versus a dual-return system?

The first reason may be cost. If you as an operator are paying a license per channel or per port, a dual-return system will cost more money. However, if you are paying a license per customer, then a single-return and dual-return system may be the same price if the hardware is the same.

The second reason is peak capacity versus total capacity.

The peak capacity is set by the port capacity, and any one particular CM is only connected to one physical port of a DAA node and thus one DS port and one US port. Thus, a single-return system and a dual-return system have the same peak capacity.

The total capacity is set by the aggregate bandwidth of all ports and how they fit within the Ethernet port bandwidth. In this scenario, a dual-return system has twice the upstream capacity as a single-return system, although they both have the same amount of downstream bandwidth.

Here is a potential conclusion.

- If a significant number of your customers are running at an upstream rate that is higher than say 50% of the upstream bandwidth (say 1 Gbps service rate on a 1.4 Gbps port), there are a low number of CMs on a node, and sales are based upon good speed-test performance, and then you may want to go with a single-return system.
- If a significant number of your customers are running at an upstream rate that is lower than say 50% of the upstream bandwidth (say 300 Mbps on a 1.4 Gbps upstream), there are a high number of CMs on a node, and sales are based on connectivity and overall throughput, then you may pick the dual-return system.

The third reason is if there are not enough OFMDA channels. For example, the silicon may support two upstream ports at 2 OFDMA each, but only one upstream port at 4 OFDMA each. Both solutions have the same average capacity, but the single upstream port has higher peak capacity.

The fourth reason is to get to a desired DS:US capacity ratio (DUCR) for a DAA node. DUCR (*pronounced "duck-r"*) is a proposed definition originating in this white paper.

- Average DUCR takes into account the total bandwidth from the total number of DS ports and US ports. This is a measure of the capacity of the HFC plant from the CMTS viewpoint.
- Peak DUCR only takes into account the bandwidth of one DS port and one US port. This is a measure of the capacity of the HFC plant from a CM viewpoint.

A single-return system will have the same value for average and peak DUCR. A dual-return system, the average DUCR will be half the peak DUCR.

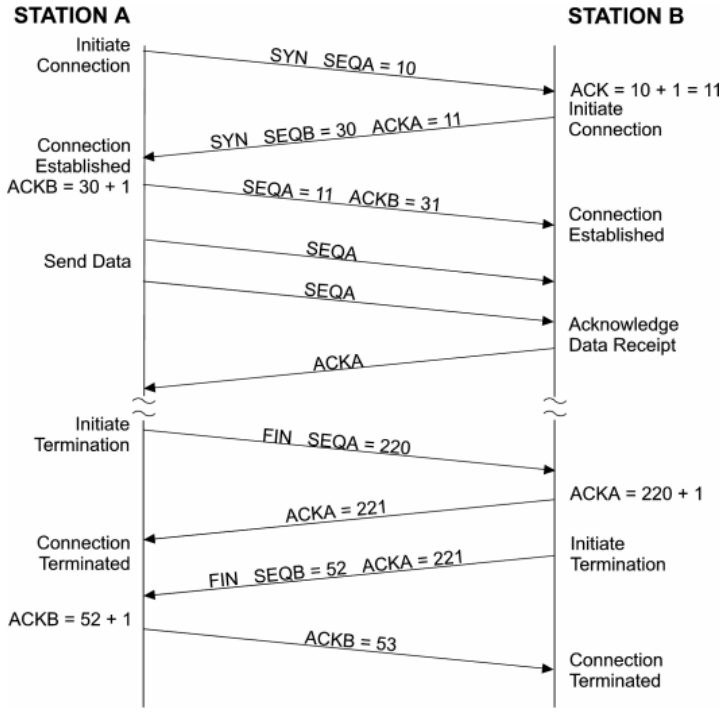


Figure 6 - TCP Handshakes

So, what is a good value DUCR? That depends upon the nature of traffic being run over the DOCSIS interface. About 75% of Internet traffic is TCP (transport control protocol). To understand this, let's do a quick re-hash on how TCP works.

A set of TCP signaling actions are shown in Figure 6. Typical TCP stack implementations try to send one ACK for every two data packets. TCP will fragment files into max MTU (about 1500 bytes) packet sizes with the last packet being a remainder. Note that Ethernet packets on an Ethernet interface have a 12-byte inter-packet gap (IPG). When the Ethernet frames are put on DOCSIS, that IPG is removed by replaced with DOCSIS framing overhead which is 13 bytes or more.

- DS: (1518 + 12 bytes IPG) x 2 = 3060 bytes
- US: 64 bytes + 12 bytes IPG = 72 bytes
- TCP DS:US ratio 3060 / 72 ~= 40 (most extreme case)

Shorter TCP transfers from web pages can dramatically reduce this ratio. The remaining UDP traffic and upstream TCP traffic also reduces this ratio. The DUCR value of the DOCSIS system should be less than the TCP ratio in order to all the downstream to fill with TCP traffic and still allow room in the upstream for non-ACK traffic.

Cable Modems (CM) contain a TCP ACK suppression mechanism that removes redundant TCP ACKs, which in term allow for a higher ratio to work. Ack suppression will remove an older ACK from a TCP flow if a new ACK comes along, if there is upstream queuing in a CM, and if the ACKs are in a burst. A nice side effect is if the upstream queue begins to build up due to upstream compression, the probability of a redundant ACK being present goes up.

5G TDD (time division duplex) systems share the same spectrum for downlink and uplink transmissions. The proportion of bandwidth between downstream and upstream is set by a frame structure configuration. The typical profile used in 5G is the so called "DDDSU" where D = downlink, U = uplink, and S = 10:2:2 (D:G:U) where G is guard time. Thus, there is a time ratio of about 3.25:1 for DL:UL. Ignoring overhead, that would correspond to a DUCR value of 3.25 if the modulation was the same. If the mobile downlink has a higher modulation that the mobile uplink, then the mobile DUCR value could be closer to 4 or 5.

Table 1 - Proposed DUCR Guidelines

DUCR		Usage Notes
Average	Peak	
40	80	• Downstream may not get fully utilized as there is not enough upstream bandwidth to support TCP ACK traffic. • Not recommended • ACK suppression in the CMs allow this extreme condition to work and this does represent some
20	40	• At the edge of working for DS and US. • This is what is often deployed today but really only works because of ACK suppression.
10	20	• Good balance between DS and US bandwidth for asymmetric traffic.
5	10	• Better balance between DS and US bandwidth for asymmetric traffic that also allows for upstream originated TCP transfers • Average DUCR on DOCSIS matches 5G TDD systems
< 2.5	< 5	• Upstream may not get fully utilized in the presence of asymmetrical traffic • Allows high symmetrical SFs to be sold. • Peak DUCR on DOCSIS matches 5G TDD systems

Based on upon these data points, Table 1 provides some guidelines for picking the right DUCR value.

A high DUCR means the downstream may not get full utilized. A low DUCR means the upstream will may get fully used (unless there is a lot of upstream TCP transfers).

Note that the asymmetry of a CM service flow may be different that the asymmetry of the DOCSIS channels. For example, say there is a 1 Gbps downstream and a 100 Mbps upstream with a 1x2 node. Thus, the upstream capacity is really 200 Mbps, the average DUCR is 5 and the peak DUCR is 10. Meanwhile, the CM has been configured to use 200 Mbps downstream and 10 Mbps upstream which is a DUCR of 20. How does that play out?

The CM will follow its individual CM DUCR, and either the downstream or upstream service flow will max out first. The DOCSIS channel, however, is an aggregate of all CMs, and of CMs are different rates. Thus, the aggregate traffic will ultimately shape to the channel DUCR.

3. DOCSIS Bandwidth Basics

3.1. Baseline assumptions

Summary	Explanation
2 x 4 Node Capacity	Number of DS and US ports on Node
Scenario	Reference number
DS End MHz	DS Spectrum upper limit
DS Start MHz	DS Spectrum lower limit
US End MHz	US Spectrum upper limit
VOD/SDV MPEG-TS	VOD used in calculation
Linear Video MPEG-TS	Linear channels used in calculation
DOCSIS DS port Gbps	DOCSIS Downstream Port
DOCSIS US port Gbps	DOCSIS Upstream port
Ethernet DS Gbps	Ethernet Downstream (DOCSIS + Video)
Ethernet US Gbps	Ethernet Upstream (DOCSIS)
DUCR, Avg	Node Ratio for DS:US capacity
DUCR, Peak	Port Ratio for DS:US capacity
ODFM ch per Node	DS OFDM channels needed per node
OFDMA ch per Node	US IFDMA channels needed per node
DOCSIS DS BW MHz	Total DS DOCSIS RF Bandwidth
Cross-over MHz	Cross-over band
DOCSIS US BW MHz	Total DS DOCSIS RF Bandwidth

Figure 7 - Spreadsheet Inputs/Outputs Explained

The study in this paper is contained within a spreadsheet. The outputs of the spreadsheet and what they represent are summarized in Figure 7 and the common set of input assumptions is in Figure 8.

32	VOD/SDV MPEG-TS	1794	DS Stop MHz for D4.0	ESD	ESD or FDX for D4.0	4096	OFDM Mod
64	Linear Video MPEG-TS	16.4	US Start MHz	YES	Video in FDX Trans Band	2048	OFDMA Mod
2	DS ports per Node	32	ch SC-QAM @ 6 MHz	120	MHz FDX Trans Band	256	SC-QAM Mod
4	US ports per Node	4	ch ATDMA @ 6.4 MHz	24	MHz DS unused < 108	64	ATDMA Mod

Figure 8 - Spreadsheet Common Inputs with default values

VOD/SDV MPEG-TS

- Unicast video services that are reserved on each downstream. The analysis is that VOD is carried on IP unicast on Ethernet and delivered to one downstream. SDV may be on IP multicast on the Ethernet, but is also delivered to only one downstream. For 6 MHz video spacing, 32 channels of video will be 192 MHz.

Linear Video MPEG-TS

- Also known as broadcast video. The analysis assumes that linear video is carried on IP multicast so there is one copy on a shared Ethernet for two DOCSIS downstreams.

Ports per node

- The number of unique downstream and upstream ports per DAA node.

DS Stop MHz for D4.0

- This allows the ending frequency for DOCSIS 4.0 to be manually set higher or lower

US Start Frequency

- Common starting point for the upstream. The spreadsheet will subtract out ATDMA channels and then fill the remaining upstream spectrum with OFDMA. To prevent OFDMA being calculated below 42 MHz, specify 4 channels of ATDMA (default is 6.4 MHz each), and a start frequency of 16.4 MHz (42 – 4*6.4).

Channels of SC-QAM

- These are downstream channels. DOCSIS 3.0 CMs have/had configurations of 8x4, 24x8, and 32x8. SC-QAMs are retained for legacy CMs.

Channels of ATDMA

- Typically, four upstream ATDMA channels were used for DOCSIS 3.0 and this occupied all available bandwidth in the upstream.

ESD or FDX for DOCSIS 4.0

- Changes the downstream starting frequency. You can tell FDX is being used when the "cross-over MHz" is negative.

Video in the FDX transition band

- FDX excludes 684 MHz to 804 MHz. That band can be used for MPEG-TS video or for legacy CMs. This is usually a yes.

FDX Transition Band

- This is defined as 120 MHz in the specification and is included here in case operators need to change it.

DS MHz skipped below 108 MHz

- 72 MHz to 76 MHz is used for the legacy OOB channel, and 88 MHz to 108 MHz are used for the FM band. Typically, these bands are not available for DOCSIS downstream or MPEG-TS video. This totals to 24 MHz.

Modulation

- The average QAM modulation level used for DOCSIS 3.0 and DOCSIS 3.1
- ODFM and OFDMA have profiles which can be configured to optimize modulation levels per subcarrier [4][5]

Note that if the upstream was configured for no ATDMA channels, and there would be seven full 96 MHz channels of OFDMA between 12 and 684 MHz, and if the modulation of each data subcarrier was 4K QAM, then the theoretical throughput of the upstream channel would be 6 Gbps. This matches the 10G

vision in the Cable industry [6]. The analysis in this whitepaper, however, assumes four ATDMA channels for backwards compatibility below 42 MHz.

3.2. Acceptance criteria

Scenario	1	2	3	4	5	6	7	8	9	10	11
DS End MHz	1002	1002	1002	1218	1218	1794	1794	1794	1794	1794	1794
US Rtn Path MHz	42	42	85	85	204	85	204	300	396	492	684
1) DS Path ≥ 5 Gbps	54%	54%	49%	92%	62%	207%	177%	154%	130%	108%	62%
2) 2 < avg DUCR < 20	148%	148%	131%	246%	52%	181%	149%	91%	56%	36%	14%
3) US SF 1 Gbps, 0.4 K	7%	7%	33%	33%	106%	33%	106%	151%	209%	268%	385%
4) 100% > 85 MHz US	-78%	-78%	0%	0%	217%	0%	217%	350%	526%	701%	1051%
Combined Results:	-78%	-78%	0%	0%	52%	0%	106%	91%	56%	36%	14%

Acceptance Criteria	1)	5	Gbps DS path min					Success Critieria	
	2)	2	DUCR min	20	DUCR max	avg	based	success if >	100%
	3)	1	Gbps US SF with	40%	of additional headroom (K)			borderline	
	4)	100%	more BW than a	85	MHz return path			reject if <	90%

Diplexer ratio %	29%	29%	27%	27%	26%	27%	26%	24%	24%	23%	22%

Figure 9 - Acceptance Criteria

One of the extended features of the spreadsheet is to evaluate each scenario against a set of criteria. These criteria so far are.

How close is:

- the downstream above a desired speed?
- DUCR between two level for peak or avg?
- the headroom of a service flow?
- how much better is the proposed return path compared to another return path solution?

The results can then be ranked within an upper and lower percentage.

There is also a diplexer ratio calculator. Negative values indicated FDX operation.

3.3. QAM/ATDMA calculator

Table 2 - QAM Configuration

Country	USA	USA	Europe	Europe
J.83 Annex	**Annex B**	**Annex B**	**Annex A**	**Annex A**
bandwidth (MHz)	6	6	8	8
constellation size	64	256	64	256
symbol rate (Msps)	5.056941	5.360537	6.952	6.952
alpha	0.186	0.119	0.151	0.151
bits per symbol	6	8	6	8
PHY Layer Raw BW (Mbps)	**30.34**	**42.88**	**41.71**	**55.62**
FEC Frame Sync	0.08%	0.05%	0.00%	0.00%
FEC Parity bytes	4.69%	4.69%	7.84%	7.84%
Trellis Coding Overhead	6.67%	5.00%	0.00%	0.00%
PHY Overhead	11.11%	9.50%	7.84%	7.84%
PHY Payload BW (Mbps)	**26.97**	**38.81**	**38.44**	**51.25**
MPEG Header	2.13%	2.13%	2.13%	2.13%
MPEG Pointer Byte	0.09%	0.09%	0.09%	0.09%
MPEG Total Overhead	2.22%	2.22%	2.22%	2.22%
PDU Layer BW (Mbps)	**26.37**	**37.95**	**37.59**	**50.12**
PDU after DOCSIS OH	**26.03**	**37.46**	**37.11**	**49.47**
Adjusted b/s/Hz	**4.34**	**6.24**	**4.64**	**6.18**

Ethernet Avg Frame Size	1000	bytes
DOCSIS Overhead	13	bytes

This calculation is the bit rate for an Ethernet frame rate.

3.4. OFDM Configuration

Table 3 - OFDM Configuration

Downstream			
Scenario	**A**	**B**	
Size of channels (MHz) - 24-192 MHz[1]	192	192	MHz
FFT size (4K or 8K FFT)	4096	8192	subcarriers
Subcarrier spacing	50	25	kHz
Cyclic prefix (Ncp)	192	192	samples[2]
Roll-off (Nrp) - must be less than Ncp	128	128	samples[2]
Ncp overhead	4%	2%	
Guard band override (leave blank if not used)[3]	1	1	MHz
Guard band on upper and lower edge (MHz)[4]	1.000	1.000	MHz
Number of active subcarriers	3800	7600	subcarriers
PLC overhead (number of subcarriers)	8	16	subcarriers
Continuous Pilot Scaling (48 - 120 subcarriers)	48	48	subcarriers
Continuous Pilots (include pilots for PLC)	56	56	subcarriers
Scattered Pilots (estimate)	29	59	subcarriers
Num of NCP - must be >0 (estimate)	4	4	
QAM order of NCP (QPSK, 16QAM, 64QAM)	6	6	bits / sym
NCP overhead (including CRC)	40	40	subcarriers
FEC overhead	12%	12%	8/9 code
Data QAM order (bits per symbol)	12	12	bits / sym
Data Rate (Mbps)	**1844**	**1911**	**Mbps**
Overhead % based on active subcarriers	19.1%	16.2%	
Efficiency	80.9%	83.8%	
Adjusted bits/second/Hz	9.61	9.95	

Notes

1. If using exclusion bands, reduce channel size by amount of spectrum excluded for data rate
2. Sampling rate is 204.8 MHz (based on OFDM spectrum - FFT size x subcarrier width)
3. Note that maximum active OFDM spectrum is 190 MHz so must have 1 MHz or higher guard band if channel is 192 MHz
4. Note that guard bands are based on Appendix V of D3.1 PHY spec based on roll-off period samples shown below (rounded to whole number of subcarriers)

OFDMA for DOCSIS was first introduced for DOCSIS in [6]. For a great discussion on OFDM and OFDMA settings, refer to [8].

3.5. OFDMA Configuration

OFDMA Upstream			
FFT size (2K or 4K FFT)	2048	4096	subcarriers
Subcarrier spacing	50	25	kHz
Size of channels (MHz) - 7.4 - 96 MHz[1]	96	96	MHz
Exclusion bands (or unused bands)			MHz
Guard band on upper and lower edge (MHz)[3]	0.5	0.5	MHz
Number of active subcarriers	1900	3800	subcarriers
Cyclic prefix (Ncp)	96	96	samples[2]
Roll-off (Nrp) - must be less than Ncp	64	64	samples[2]
Ncp overhead	4.48%	2.29%	
Maximum K values based on channel size	18	9	
OFDM symbols per OFDMA Frame (K)[4]	18	9	
Minislot size (400 kHz)	8	16	subcarriers
Minislots per OFDMA frame - (minimum 25 or 16 to ASIC)	237	237	
Edge minislots per OFDMA Frame (estimate)[5]	23	23	
Unused subcarriers (assumes continuous exclusion band)	4	8	
OFDMA frame duration	376.88	368.44	µsec
Pilot pattern	1	8	pattern
Pilots per minislot (body)	2	2	pilots
Pilots per minislot (edge)	4	4	pilots
complementary pilots per minislot (body)	2	2	comp pilots
complementary pilots per minislot (edge)	4	4	comp pilots
Total data carriers per OFDMA frame	33088	33088	
Total complementary pilots per OFDMA frame	520	520	
Data QAM order (bits per symbol)	10	10	bits / sym
bits per minislot (body) - no FEC overhead	1412	1412	bits
bits per minislot (edge) - no FEC overhead	1384	1384	bits
FEC overhead (long codeword)	11.11%	11.11%	8/9 code
FEC overhead (medium codeword)	15.15%	15.15%	28/33 code
FEC overhead (short codeword)	25.00%	25.00%	3/4 code
bits per OFDMA frame	296889	296889	bits
Data Rate (Mbps)	**788**	**806**	**Mbps**
Overhead % based on active subcarriers	17.08%	15.18%	

Table 4 - OFDMA Configuration

Notes

1. Valid channel sizes 11 - 96 MHz for 50 kHz; 7.4 - 96 MHz 25 kHz (includes 0.5 MHz guard band per edge)
2. sampling rate is 102.4 MHz (based on OFDM spectrum - FFT size x subcarrier width)
3. Note that guard bands are fixed at 0.5 MHz on cBR-8
4. K values must fall in range from table on right
5. Edge minislots occur at start of OFDMA frame, after excluded or unused spectrum, and at start of modem burst (10% might be a good estimate - has minimal impact on estimate)

4. DAA Bandwidth Studies

4.1. DOCSIS 3.1 with 1002 MHz, low split and 2x4 DAA Node (today)

(a)

2 x 4 Node Capacity		
Scenario	**1**	**2**
DS End MHz	**1002**	**1002**
DS Start MHz	**54**	**54**
US Rtn Path MHz	**42**	**42**
VOD/SDV MPEG-TS	32	10
Linear Video MPEG-TS	64	48
DOCSIS DS port Gbps	2.8	5.0
DOCSIS US port Gbps	0.10	0.10
Ethernet DS Gbps	10.3	12.6
Ethernet US Gbps	0.4	0.4
DUCR, Avg	14	25
DUCR, Peak	27	50
ODFM ch per Node	2	4
OFDMA ch per Node	0	0
DOCSIS DS BW MHz	348	576
Cross-over MHz	12	12
DOCSIS US BW MHz	26	26

2	DS ports per Node	4096	OFDM Modulation
4	US ports per Node	2048	OFDMA Modulation
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation

24	DS MHz skipped below 108 MHz
1794	ESD Stop Frequency (MHz)
16.4	US Start Freq (MHz)

(b)

1	**2**
1002	**1002**
54	**54**
42	**42**
10	0
48	0
5.0	8.5
0.20	0.20
12.6	17.0
0.8	0.8
13	21
25	43
4	8
4	4
576	924
12	12
32	32

2	DS ports per Node	4096	OFDM Modulation
4	US ports per Node	1024	OFDMA Modulation
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation
2	ATDMA 6.4 MHz ch	64	ATDMA Modulation

24	DS MHz skipped below 108 MHz
1794	ESD Stop Frequency (MHz)
10	US Start Freq (MHz)

Figure 10 - Bandwidth today with low-split

Deployment scenarios today are generally with a 1x1 or 1x2 DAA node with newer ones using a 2x2 or 2x4 DAA node. This scenario looks at a 2x4 node on an existing 1002 MHz HFC plant, with and without MPEG-TS video, and with a 42 MHz return path.

DOCSIS 3.1 CMs are capable of receiving 32 6 MHz SC-QAM channels (192 MHz total) plus two OFDM channels (192 MHz each), for a total of 576 MHz of DOCSIS spectrum. That is over half of the available spectrum on the HFC plant today. At this time, few deployments are fully utilizing that available DOCSIS capacity on their HFC plant.

In scenario 1 (a) with video and on a 1002 MHz plant, there is 96 SC-QAM carriers (576 MHz) of MPEG-TS video is present, resulting in only 348 MHz of remaining spectrum available for DOCSIS. The resulting DOCIS downstream bandwidth is about 3 Gbps which can be implemented a partial OFDM channel (132 MHz) and 32 SC-QAM channels (or a full 192 MHz OFDM channel and 26 SC-QAM channels).

The total Ethernet DS bandwidth, including video, is about 10 Gbps. If the multicast video cannot be shared across the RF downstream ports, there is an additional about 2.5 Gbps (64 x 38.5 Mbps) load on the Ethernet. This exceeds a single 10 Gbps backhaul, so two 10 Gbps backhauls may be needed or one 25 Gbps backhaul.

The upstream is 100 Mbps and the resulting average DUCR is around 14 and peak DUCR is around 27. These are reasonable ratios.

In Scenario 2 (a), only 58 channels of video are deployed, so DOCSIS can now be allocated a full 576 MHz. That equates to two OFDM channels (2 x 192 MHz) and 32 SC-QAM channels (192 MHz). The DOCSIS downstream increases to about 5 Gbps per DS port. The avg/peak DUCR jumps to about 25/50. This is high but in alignment with today's practices.

In scenario 1 (b), the upstream is changed from the default four ATDMA channels below 42 MHz to only 2 A-TDMA channels, while ODFM fills in the gaps down to 10 MHz (19.2 MHz). The average modulation was dropped to 1024-QAM. The resulting theoretical throughput is 200 Mbps, although in practice it may be slightly less than this due to plant noise. Bear in mind with a 2x4 node, that is the equivalent of 400 Mbps of capacity for each 5 Gbps downstream. This shows up with slightly high but good DUCR values.

In scenario 2 (b), all MPEG-TS video is removed, and 948 MHz (54 MHz to 1002 MHz) is used for DOCSIS. The DOCSIS downstream is about 8.5 Gbps per port. This could be used to drive two banks of CMs that share a common set of 32 SC-QAMs but separate pairs of OFDMA channels. Or, the 32 SC-QAM could be used for legacy CMs and there would be two banks of DOCSIS 3.1 CMs, each connected to two OFDMA channels but not the SC-QAM channels.

The doubling of the US throughput has kept the DUCR values from getting too high which would mean the upstream may saturate and the downstream may not get fully utilized. The DAA silicon should be able to support this scenario. Without the increase in upstream bandwidth, this may not make sense. Instead, it shows that the upstream return path needs to be increased, and that is the subject of the next section.

The observation from these scenarios is that video could be reduced to 58 or so channels on a 1002/42 MHz plant to allow the full potential of a DOCSIS 3.1 CM to be used. Further, if OFDM is placed below 42 MHz, then the upstream can be almost doubled. If video is entirely removed, DOCSIS can be increased three-fold from about 2.8 Gbps to about 8.5 Gbps. There is a lot that can be done for DOCSIS without changing the HFC plant.

4.2. DOCSIS 3.1 with 1218 MHz, mid/high split and 2x4 DAA Node

2 x 4 Node Capacity	DOCSIS 3.1			DOCSIS 3.1		
Scenario	**3**	**4**	**5**	**3**	**4**	**5**
DS End MHz	**1002**	**1218**	**1218**	**1002**	**1218**	**1218**
DS Start MHz	108	108	258	**108**	**108**	**258**
US Rtn Path MHz	**85**	**85**	**204**	**85**	**85**	**204**
VOD/SDV MPEG-TS	32	32	32	0	0	0
Linear Video MPEG-TS	64	64	64	0	0	0
DOCSIS DS port Gbps	2.5	4.6	3.1	8.2	10.3	8.8
DOCSIS US port Gbps	0.47	0.47	1.48	0.47	0.47	1.48
Ethernet DS Gbps	9.7	14.0	11.0	16.4	20.7	17.7
Ethernet US Gbps	1.9	1.9	5.9	1.9	1.9	5.9
DUCR, Avg	2.6	4.9	1.0	8.7	11.0	3.0
DUCR, Peak	5.2	9.8	2.1	17.5	22.1	6.0
ODFM ch per Node	2	4	2	8	10	8
OFDMA ch per Node	4	4	8	4	4	8
DOCSIS DS BW MHz	318	534	384	894	1110	960
Cross-over MHz	23	23	54	23	23	54
DOCSIS US BW MHz	69	69	188	69	69	188

2	DS ports per Node	4096	OFDM Modulation	24	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1794	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 11 - Bandwidth with DOCSIS 3.1 mid-split and high-split

This section exams the move to an 85 or 204 MHz return path. There are three deployment plans suggested and are shown in Figure 11.

Scenario 3, looks at the impact of moving to an 85 MHz return path but retaining the 1002 MHz downstream. This can be a reality of the downstream RF limit that is imposed by the HFC amps and/or the CMs, rather than the DAA node. The return path now has four times the capacity with over 450 Mbps rather than the previous 100 Mbps.

Scenario 3 with video has a decent 2.5 Gbps downstream with low DUCR values. Low DUCR values in theory allow for symmetrical bandwidth, but in practice may mean that the US is under-utilized. When MPEG-TS video is completely removed from the HFC plant, the DOCSIS downstream bandwidth can go to over 8 Gbps. The DUCR values are almost ideal at 10 and 20. With no video, the Ethernet backhaul barely fits into two 10 Gbps links. With video, it fits.

Scenario 4 increases the downstream to 1218 MHz with an 85 MHz return path. The downstream bandwidth goes up to about 4.5 Gbps and the DUCR values increase to the ideal zone where both downstream and upstream can be fully utilized. With no video, the downstream goes to about 10 Gbps and the DUCR values are still excellent. Note that the OFDM channel count is getting high. The Ethernet bandwidth crosses 20 Gbps which is not good as it does not allow headroom for signaling and there could be the multicast video hit. This is a good place for a 25 Gbps Ethernet serving both downstream ports.

Scenario 5 increases the upstream return path to 204 MHz. This allows a 1 Gbps upstream service tier to be sold. But it also ends up decreasing the DS spectrum and throughput somewhat. The downstream is about 3 Gbps with video and 9 Gbps without video, which is quite decent. The DUCR values are on the low side so the upstream may not be fully utilized, and the Ethernet bandwidth falls back into a dual 10 Gbps target.

The observation here is that DOCSIS 3.1 provide a huge amount of bandwidth growth in both the downstream and upstream, especially if video is removed.

4.3. DOCSIS 4.0 with ESD 1794 MHz and 2x4 DAA Node

2 x 4 Node Capacity	D4.0 ESD with video						D4.0 ESD with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**
DS Start MHz	108	258	372	492	606	834	108	258	372	492	606	834
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	10.3	8.8	7.7	6.5	5.4	3.1	16.1	14.6	13.4	12.2	11.1	8.8
DOCSIS US port Gbps	0.47	1.48	2.11	2.93	3.75	5.39	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**25.5**	**22.5**	**20.2**	**17.8**	**15.6**	**11.0**	**32.1**	**29.2**	**26.9**	**24.5**	**22.2**	**17.7**
Ethernet US Gbps	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	11.0	3.0	1.8	1.1	0.7	0.3	17.2	4.9	3.2	2.1	1.5	0.8
DUCR, Peak	22.1	6.0	3.7	2.2	1.4	0.6	34.3	9.8	6.4	4.2	3.0	1.6
ODFM ch per Node	10	8	8	6	6	2	16	14	14	12	12	8
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	1110	960	846	726	612	384	1686	1536	1422	1302	1188	960
Cross-over MHz	23	54	72	96	114	150	23	54	72	96	114	150
DOCSIS US BW MHz	69	188	261	357	453	645	69	188	261	357	453	645

2	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1794	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 12 - Bandwidth with DOCSIS 4.0 ESD at 1794 MHz

Six scenarios, each with a different upstream split (85, 204, 300, 396, 492, 684 MHz) but a common downstream cap of 1794 MHz are shown in Figure 12, with and without MPEG-TS video. This version of DOCSIS 4.0 is referred to as extended spectrum DOCSIS (ESD). [6] The 85 MHz return path is not an actual DOCSIS 4.0 use case, but is included for comparative purposes.

The DUCR values are very telling here. For almost all the higher splits, the ratios are so low that it is likely that the upstream will get fully utilized, unless all the traffic was symmetrical, which it is not. In fact, for a 684 MHz 1x2 or 2x4 node, DUCR drops below one which means there is more upstream capacity than downstream capacity. For 1x1, it is still low. So, 684 MHz may not be a good choice. The 492 MHz upstream also has low ratios.

Note the impact on the Ethernet backhaul per node. In general, two 10 Gbps backhauls for a 2x4 node are no longer enough. 25 Gbps is likely needed or even 40 Gbps. At 25 Gbps and a 2x4 node, both downstream paths could not be saturated at the same time, although that is probably acceptable. Not that this analysis is not included downstream bandwidth for OOB and NDF, so some headroom is required.

The 396 MHz upstream is right in the middle. It has twice the upstream bandwidth of 204 MHz. But remember 204 MHz already did a 15x increase over 42 MHz, so is another 2x worth it? The driving goal would have to be the ability to offer a 2 Gbps upstream service. Note that the OFDM channel count without video is high and may not fit some silicon implementations.

With no video, and at lower return paths, the DOCSIS downstream goes past the 10 Gbps mark. With a 204 MHz upstream, and no video, the DOCSIS downstream is about 15 Gbps, which could allow a 10 Gbps x 1 Gbps service to be sold to the home. That's interesting for competitive purposes.

The observation here is that 204 MHz looks like a great choice for asymmetrical traffic and for a potential 10 Gbps x 1 Gbps service offering. 396 MHz is a second choice for more symmetrical traffic and a 2 Gbps upstream service offering.

4.4. DOCSIS 4.0 with ESD 1602 MHz and 2x4 DAA Node

2 x 4 Node Capacity	D4.0 ESD with video						D4.0 ESD with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**	**1602**
DS Start MHz	108	258	372	492	606	834	108	258	372	492	606	834
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	8.4	6.9	5.8	4.6	3.5	1.2	14.2	12.7	11.5	10.3	9.2	6.9
DOCSIS US port Gbps	0.47	1.48	2.11	2.93	3.75	5.39	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	21.6	18.7	16.4	14.0	11.7	7.2	28.3	25.3	23.1	20.7	18.4	13.9
Ethernet US Gbps	1.9	5.9	8.4	11.7	15	22	1.9	5.9	8.4	11.7	15	22
DUCR, Avg	9.0	2.3	1.4	0.8	0.5	0.1	15.1	4.3	2.7	1.8	1.2	0.6
DUCR, Peak	18.0	4.7	2.7	1.6	0.9	0.2	30.2	8.5	5.5	3.5	2.5	1.3
ODFM ch per Node	8	6	6	4	4	0	14	12	12	10	10	6
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	918	768	654	534	420	192	1494	1344	1230	1110	996	768
Cross-over MHz	23	54	72	96	114	150	23	54	72	96	114	150
DOCSIS US BW MHz	69	188	261	357	453	645	69	188	261	357	453	645

2	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1602	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 13 - Bandwidth with DOCSIS 4.0 ESD at 1602 MHz

What if the HFC plant can't make it to 1794 MHz and only gets to say 1602 MHz? or even 1506 MHz or 1554 MHz? All is not lost. The results are shown in Figure 13 for 1602 MHz. Both the 85 MHz and 204 MHz cases have good DUCR values. Above that, the DUCR values are too low.

The observation here is that the DUCR values suggest that the 204 MHz upstream in this scenario is the way to go. DOCSIS 4.0 CMs would still be required.

4.5. DOCSIS 4.0 with FDX and 2x4 DAA Node

2 x 4 Node Capacity	D4.0 FDX with video						D4.0 FDX with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**	**1218**
DS Start MHz	108	108	108	108	108	108	108	108	108	108	108	108
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	4.6	4.6	4.6	4.6	4.6	4.6	10.3	9.1	9.1	9.1	9.1	9.1
DOCSIS US port Gbps	0.47	1.29	2.11	2.93	3.75	5.39	0.47	1.29	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**14.0**	**14.0**	**14.0**	**14.0**	**14.0**	**14.0**	**20.7**	**18.3**	**18.3**	**18.3**	**18.3**	**18.3**
Ethernet US Gbps	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	4.9	1.8	1.1	0.8	0.6	0.4	11.0	3.5	2.2	1.6	1.2	0.8
DUCR, Peak	9.8	3.6	2.2	1.6	1.2	0.9	22.1	7.1	4.3	3.1	2.4	1.7
ODFM ch per Node	4	4	4	4	4	4	10	10	10	10	10	10
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	534	534	534	534	534	534	1110	990	990	990	990	990
Cross-over MHz	23	-96	-192	-288	-384	-576	23	-96	-192	-288	-384	-576
DOCSIS US BW MHz	69	165	261	357	453	645	69	165	261	357	453	645

2	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1218	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 14 - Bandwidth with DOCSIS 4.0 FDX

DOCSIS 4.0 full-duplex (FDX) with a downstream limit of 1218 MHz, a maximum upstream return path of 204 to 684 MHz, and with and without video is shown in Figure 14. [9][10][11][12][13]. The 85 MHz return is not an FDX use case an dis included here for comparative purposes.

The DUCR values are low and suggest that a 684 MHz return will not fill the upstream with asymmetrical traffic. A 1218 MHz downstream with two 85 MHz upstream has very nice DUCR values, although it is not FDX. The FDX choices would be a 204 MHz return path for 1 Gbps subscriber service or 396 MHz return path for a 2 Gbps subscriber service.

The downstream bandwidth is about 4.5 Gbps with video present and about 9 Gbps with all video removed.

The OFDM channel requirements are high and care should be taken on silicon choices. Alternatively, using more than 32 channels of SC-QAM for DOCSIS is an alternative.

The Ethernet bandwidth should be doable with two 10 Gbps interfaces or one 25 Gbps. Again, be careful of how multicast MPEG video is shared as this could make two 10 Gbps Ethernet saturate and the DOCSIS downstream would not be fully utilized.

4.6. Post DOCSIS 4.0 with ESD + FDX and 2x4 DAA Node

2 x 4 Node Capacity	D4.0 FDX with video						D4.0 FDX with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**
DS Start MHz	108	108	108	108	108	108	108	108	108	108	108	108
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	10.3	10.3	10.3	10.3	10.3	10.3	16.1	14.9	14.9	14.9	14.9	14.9
DOCSIS US port Gbps	0.47	1.29	2.11	2.93	3.75	5.39	0.47	1.29	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**25.5**	**25.5**	**25.5**	**25.5**	**25.5**	**25.5**	**32.1**	**29.7**	**29.7**	**29.7**	**29.7**	**29.7**
Ethernet US Gbps	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	11.0	4.0	2.5	1.8	1.4	1.0	17.2	5.8	3.5	2.5	2.0	1.4
DUCR, Peak	22.1	8.0	4.9	3.5	2.8	1.9	34.3	11.5	7.1	5.1	4.0	2.8
ODFM ch per Node	10	10	10	10	10	10	16	16	16	16	16	16
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	1110	1110	1110	1110	1110	1110	1686	1566	1566	1566	1566	1566
Cross-over MHz	23	-96	-192	-288	-384	-576	23	-96	-192	-288	-384	-576
DOCSIS US BW MHz	69	165	261	357	453	645	69	165	261	357	453	645

2	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	1794	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 15 - Bandwidth with ESD + FDX, 2x4 DAA Node

The following post DOCSIS 4.0 scenarios are not official CableLabs projects, and are thus theoretical and for discussion and planning purposes.

The scenarios in Figure 15 combines FDX and ESD from DOCSIS 4.0. The upstream return path would be 204 to 684 MHz and the downstream path would extend to from 108 MHz to 1794 MHz. This could be an achievable design goal by combining the best of ESD and FDX. With video, there is enough downstream spectrum for a shared 10 Gbps and without video, there is about 15 Gbps of bandwidth. The upstream is about 5 Gbps. This could allow service offerings of 10 Gbps x 4 Gbps.

The DUCR values are low suggesting that all the upstream bandwidth may not get used. The OFDM channel count is high and could be a challenge for ASICs. The Ethernet bandwidth also exceeds the 25 Gbps limit, so either a 25 or 40 Gbps Ethernet would be needed.

4.7. Post DOCSIS 4.0 with 3 GHz ESD and 2x4 DAA Node

2 x 4 Node Capacity	Post D4.0 ESD with video						Post D4.0 ESD with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**
DS Start MHz	108	258	372	492	606	834	108	258	372	492	606	834
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	21.8	20.3	19.2	18.0	16.8	14.6	27.5	26.0	24.9	23.7	22.6	20.3
DOCSIS US port Gbps	0.47	1.48	2.11	2.93	3.75	5.39	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**48.4**	**45.4**	**43.1**	**40.8**	**38.5**	**33.9**	**55.1**	**52.1**	**49.8**	**47.4**	**45.2**	**40.6**
Ethernet US Gbps	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	23.3	6.8	4.5	3.1	2.2	1.4	29.4	8.8	5.9	4.0	3.0	1.9
DUCR, Peak	46.6	13.7	9.1	6.1	4.5	2.7	58.8	17.5	11.8	8.1	6.0	3.8
ODFM ch per Node	22	20	20	18	18	14	28	26	26	24	24	20
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	2262	2112	1998	1878	1764	1536	2838	2688	2574	2454	2340	2112
Cross-over MHz	23	54	72	96	114	150	23	54	72	96	114	150
DOCSIS US BW MHz	69	188	261	357	453	645	69	188	261	357	453	645

2	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	2946	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 16 - Bandwidth with 3 GHz ESD, 2x4 DAA Node

The scenarios in Figure 16 is a 3 GHz version of ESD [14]. The value of 2946 MHz is derived from taking 1794 MHz from DOCSIS 4.0 ESD and adding six more 192 MHz OFDM channels.

The downstream bandwidth is nice and high and is around 25 Gbps with no video. What is interesting with the DUCR values is that 204 MHz is a good choice, especially if it is a dual-return upstream. However, this would limit service offerings to 1 Gbps. Alternatively, a single 396 MHz would work which could allow a 20 Gbps x 2 Gbps service to be sold.

The OFDM channel count is very high indicating yet another generation of silicon, which would be needed anyway to hit 3 GHz. The Ethernet bandwidth is hitting the 50 Gbps level now.

4.8. Post DOCSIS 4.0 with 3 GHz FDX and 2x4 DAA Node

2 x 4 Node Capacity	Post D4.0 FDX with video						Post D4.0 FDX with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**
DS Start MHz	108	108	108	108	108	108	108	108	108	108	108	108
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	21.8	21.8	21.8	21.8	21.8	21.8	27.5	26.3	26.3	26.3	26.3	26.3
DOCSIS US port Gbps	0.47	1.29	2.11	2.93	3.75	5.39	0.47	1.29	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**48.4**	**48.4**	**48.4**	**48.4**	**48.4**	**48.4**	**55.1**	**52.7**	**52.7**	**52.7**	**52.7**	**52.7**
Ethernet US Gbps	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	23.3	8.5	5.2	3.7	2.9	2.0	29.4	10.2	6.2	4.5	3.5	2.4
DUCR, Peak	46.6	16.9	10.3	7.4	5.8	4.0	58.8	20.4	12.5	9.0	7.0	4.9
ODFM ch per Node	22	22	22	22	22	22	28	28	28	28	28	28
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	2262	2262	2262	2262	2262	2262	2838	2718	2718	2718	2718	2718
Cross-over MHz	23	-96	-192	-288	-384	-576	23	-96	-192	-288	-384	-576
DOCSIS US BW MHz	69	165	261	357	453	645	69	165	261	357	453	645

2	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	2946	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 17 - Bandwidth with 3 GHz FDX, 2x4 DAA Node

The scenarios in Figure 17 are a 3 GHz version of FDX, with and without video, on a 2x4 DAA node.

The bandwidths are higher than the 3 GHz ESD case since there is more downstream spectrum on each of the downstream ports (606 MHz per port). However, the downstream is slightly above 25 Gbps and will likely be held at 25 Gbps due to the Ethernet backhaul which will be dual 25 Gbps or 50 Gbps.

4.9. Post DOCSIS 4.0 with 3 GHz ESD and 4x4 DAA Node

4 x 4 Node Capacity	Post D4.0 ESD with video						Post D4.0 ESD with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**
DS Start MHz	108	258	372	492	606	834	108	258	372	492	606	834
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	21.8	20.3	19.2	18.0	16.8	14.6	27.5	26.0	24.9	23.7	22.6	20.3
DOCSIS US port Gbps	0.47	1.48	2.11	2.93	3.75	5.39	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**94.4**	**88.4**	**83.9**	**79.1**	**74.6**	**65.5**	**110.1**	**104.2**	**99.6**	**94.8**	**90.3**	**81.2**
Ethernet US Gbps	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	46.6	13.7	9.1	6.1	4.5	2.7	58.8	17.5	11.8	8.1	6.0	3.8
DUCR, Peak	46.6	13.7	9.1	6.1	4.5	2.7	58.8	17.5	11.8	8.1	6.0	3.8
ODFM ch per Node	44	40	40	36	36	28	56	52	52	48	48	40
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	2262	2112	1998	1878	1764	1536	2838	2688	2574	2454	2340	2112
Cross-over MHz	23	54	72	96	114	150	23	54	72	96	114	150
DOCSIS US BW MHz	69	188	261	357	453	645	69	188	261	357	453	645

4	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	2946	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

Figure 18 - Bandwidth with 3 GHz ESD, 4x4 DAA Node

The scenarios in Figure 18 take the 3 GHz ESD scenarios and expand them to a 4x4 node.

We now see the capacity of the DAA node hit 100 Gbps. The DUCR values are the same for average and peak as the DAA node is single-return. This implementation packs around 50 OFDM channels into a node housing.

4.10. Post DOCSIS 4.0 with 3 GHz FDX and 4x4 DAA Node

4 x 4 Node Capacity	Post D4.0 FDX with video						Post D4.0 FDX with no video					
Scenario	**6**	**7**	**8**	**9**	**10**	**11**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**	**2946**
DS Start MHz	108	108	108	108	108	108	108	108	108	108	108	108
US Rtn Path MHz	**85**	**204**	**300**	**396**	**492**	**684**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	32	32	32	32	32	32	0	0	0	0	0	0
Linear Video MPEG-TS	64	64	64	64	64	64	0	0	0	0	0	0
DOCSIS DS port Gbps	21.8	21.8	21.8	21.8	21.8	21.8	27.5	26.3	26.3	26.3	26.3	26.3
DOCSIS US port Gbps	0.47	1.29	2.11	2.93	3.75	5.39	0.47	1.29	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**94.4**	**94.4**	**94.4**	**94.4**	**94.4**	**94.4**	**110.1**	**105.4**	**105.4**	**105.4**	**105.4**	**105.4**
Ethernet US Gbps	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**	**1.9**	**5.2**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	46.6	16.9	10.3	7.4	5.8	4.0	58.8	20.4	12.5	9.0	7.0	4.9
DUCR, Peak	46.6	16.9	10.3	7.4	5.8	4.0	58.8	20.4	12.5	9.0	7.0	4.9
ODFM ch per Node	44	44	44	44	44	44	56	56	56	56	56	56
OFDMA ch per Node	4	8	12	16	20	28	4	8	12	16	20	28
DOCSIS DS BW MHz	2262	2262	2262	2262	2262	2262	2838	2718	2718	2718	2718	2718
Cross-over MHz	23	-96	-192	-288	-384	-576	23	-96	-192	-288	-384	-576
DOCSIS US BW MHz	69	165	261	357	453	645	69	165	261	357	453	645

4	DS ports per Node	4096	OFDM Modulation	27	DS MHz skipped below 108 MHz
4	US ports per Node	2048	OFDMA Modulation	YES	Video in FDX Transition Band
32	SC-QAM 6 MHz ch	256	SC-QAM Modulation	2946	D4.0 Stop Frequency (MHz)
4	ATDMA 6.4 MHz ch	64	ATDMA Modulation	16.4	US Start Freq (MHz)

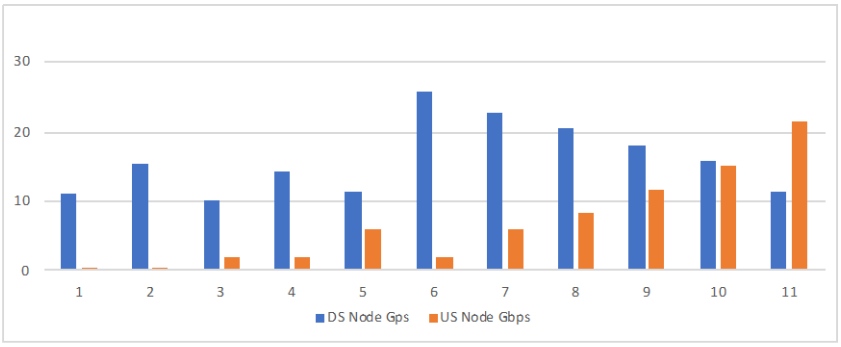
Figure 19 - Bandwidth with 3 GHz ESD, 4x4 DAA Node

The scenarios in Figure 18 is a 3 GHz version of FDX with a 4x4 DAA node.

The downstream DOCSIS speeds are a healthy 25 Gbps per port which adds up to 100 Gbps downstream for the node. FDX has an extra 606 MHz per port which adds up across four ports. The DUCR levels are the same for average and peak since the number of downstream and upstream ports are the same.

100 Gbps per DAA node is impressive and an indication of where the CIN network may need to scale to in the future.

4.11. Summary of Calculations



Summary

2 x 4 Node Capacity	User	DOCSIS 3.1 with MPEG video					DOCSIS 4.0 with MPEG video					
Scenario	**Calc**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
DS End MHz	**1002**	**1002**	**1218**	**1002**	**1218**	**1218**	**1794**	**1794**	**1794**	**1794**	**1794**	**1794**
DS Start MHz	714	54	54	108	108	258	108	258	372	492	606	834
US End MHz	**42**	**42**	**42**	**85**	**85**	**204**	**85**	**204**	**300**	**396**	**492**	**684**
VOD/SDV MPEG-TS	0	32	32	32	32	32	32	32	32	32	32	32
Linear Video MPEG-TS	0	64	64	64	64	64	64	64	64	64	64	64
DOCSIS DS port Gbps	2.3	3.2	5.3	2.6	4.8	3.3	10.5	9.0	7.9	6.7	5.6	3.3
DOCSIS US port Gbps	0.10	0.10	0.10	0.47	0.47	1.48	0.47	1.48	2.11	2.93	3.75	5.39
Ethernet DS Gbps	**4.7**	**11.1**	**15.4**	**10.1**	**14.4**	**11.4**	**25.8**	**22.8**	**20.6**	**18.2**	**15.9**	**11.4**
Ethernet US Gbps	**0.4**	**0.4**	**0.4**	**1.9**	**1.9**	**5.9**	**1.9**	**5.9**	**8.4**	**11.7**	**15**	**22**
DUCR, Avg	11.6	16	26	2.8	5.1	1.1	11.2	3.0	1.9	1.1	0.7	0.3
DUCR, Peak	23.1	31	53	5.6	10.2	2.2	22.5	6.1	3.7	2.3	1.5	0.6
ODFM ch per Node	2	4	6	2	6	4	12	10	8	8	6	4
OFDMA ch per Node	0	0	0	4	4	8	4	8	12	16	20	28
DOCSIS DS BW MHz	288	372	588	318	534	384	1110	960	846	726	612	384
Cross-over MHz	-	12	12	23	23	54	23	54	72	96	114	150
DOCSIS US BW MHz	26	26	26	69	69	188	69	188	261	357	453	645

32	VOD/SDV MPEG-TS	1794	DS Stop MHz for D4.0	ESD	ESD or FDX for D4.0	4096	OFDM Mod
64	Linear Video MPEG-TS	16.4	US Start MHz	YES	Video in FDX Trans Band	2048	OFDMA Mod
2	DS ports per Node	24	ch SC-QAM @ 6 MHz	120	MHz FDX Trans Band	256	SC-QAM Mod
4	US ports per Node	4	ch ATDMA @ 6.4 MHz	24	MHz DS unused < 108	64	ATDMA Mod

Acceptance Criteria

Scenario	Calc	1	2	3	4	5	6	7	8	9	10	11
DS End MHz	1002	1002	1218	1002	1218	1218	1794	1794	1794	1794	1794	1794
US End MHz	42	42	42	85	85	204	85	204	300	396	492	684
1) DS Path ≥ 5 Gbps	47%	63%	106%	53%	96%	66%	210%	180%	158%	134%	111%	66%
2) 2 < avg DUCR < 20	173%	127%	76%	141%	255%	55%	178%	152%	94%	57%	37%	15%
3) US SF 1 Gbps, 0.4 K	7%	7%	7%	33%	33%	106%	33%	106%	151%	209%	268%	385%
4) 100% > 85 MHz US	-78%	-78%	-78%	0%	0%	217%	0%	217%	350%	526%	701%	1051%
Combined Results:	-78%	-78%	-78%	0%	0%	55%	0%	106%	94%	57%	37%	15%

								Success Criteria	
Acceptance Criteria	1)	5	Gbps DS path min						
	2)	2	DUCR min	20	DUCR max	avg	based	success if >	100%
	3)	1	Gbps US SF with	40%	of additional headroom (K)			borderline	
	4)	100%	more BW than a	85	MHz return path			reject if <	90%

Diplexer ratio %	1600%	29%	29%	27%	27%	26%	27%	26%	24%	24%	23%	22%

Figure 20 - DAA and DOCSIS Bandwidth Calculator

5. Conclusions

The future of DOCSIS is inter-twined with the DAA Node. The bandwidth and port density calculations that used to be done for a I-CMTS are now done for the DAA Node.

For any frequency plan, DAA has benefits over a conventional CMTS.

1. By placing the modulation in the node where there is a lower noise floor, high modulations can be run.
2. The number of nodes that a single physical CMTS can support increases, sometimes 4x or more
3. The fiber network is converted to a Ethernet based CIN which can now support PON, business services, and mobile backhaul.

Care must be taken that the DAA node has enough OFDM channels to support a particular configuration. The CIN network has to be designed to support the Ethernet capacity of the Node. DOCSIS 4.0 breaks through the 10 Gbps Ethernet capacity boundary. In the future, DAA nodes could approach 100 Gbps in downstream capacity.

DUCR was defined in this white paper as the ratio of downstream to upstream capacity of a node. The average DUCR took into account the number of ports and looks at the HFC plant from the CMTS perspective. The peak DUCR only looked at one downstream port and one upstream port and looks at the HFC plant from the CM perspective. High DUCR values indicate the downstream may not get fully utilized while low DUCR values indicate the upstream may not get fully utilized.

DUCR is a good tool for evaluating the usefulness of a deployment plan and is a measure of asymmetry. 5G uses a DUCR value of 5. A good target for cable would be 5 to 10 for average capacity. Too low a value indicates the upstream may not be fully used. Too high a value indicates that the downstream may not get fully used. CM ACK suppression allows this value to be artificially high.

Removing MPEG-TS video from the downstream spectrum works well to increase DOCSIS bandwidth. The example used in this white paper showed an increase of 3x in bandwidth, from 3 Gbps to 9 Gbps, on a 1002 MHz plant by just removing video. The 1218 MHz plant for DOCSIS 3.1 was useful for retaining the same downstream spectrum size as 42 to 1002, but with a 204 MHz lower limit. DOCSIS 4.0 ESD added even more bandwidth, but went over the 10 Gbps Ethernet backhaul ceiling.

The 42 MHz return path can be increased by 50% to 100%, from 100 Mbps to 200 Mbps, by removing two of the A-TDMA channels, adding OFDMA, moving the modulation to 2048-QAM, and starting at a lower frequency. An 85 MHz return path provided excellent return path bandwidth with great DUCR values. The 204 MHz return path added more bandwidth than needed on DOCSIS 3.1 and worked well for DOCSIS 4.0. DOCSIS 4.0 would benefit from two 204 MHz return path ports.

For 204 MHz to be deployed, all OOB signaling for video has to be removed from the plant, and OUDP updates have to be done to accommodate signal leakage detection. Return paths greater than 204 MHz have diminishing returns in a 1x2 node configuration, but are more attractive in a 1x1 configuration. Note that two 204 MHz return paths are roughly equivalent to one 396 MHz return path from an average capacity viewpoint.

It was also demonstrated that future silicon solutions could sub-split a node to 4x4 and if the plant went to 3 GHz, then nodes could approach 100 Gbps of capacity.

Abbreviations

ADC	analog-to-digital converter
ATDMA	advanced time division multiple access
CIN	converged interconnect network
CMTS	cable modem termination system
DAA	distributed access architecture
DAC	digital-to-analog converter
DPD	digital predistortion
DS	downstream
DUCR	downstream-to-upstream capacity ratio
ESD	extended spectrum DOCSIS
FDX	full duplex
FMA	flexible MAC architecture
HFC	hybrid fiber coax
I-CMTS	integrated CMTS
LC	line card
MUX	multiplexer
NDF	narrowband digital forward
NDR	narrowband digital reverse
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
OOB	out of band
PPS	packets per second
QAM	quadrature amplitude modulation
RF	radio frequency
RMACPHY	remote MAC PHY
RPHY	remote PHY
SC-QAM	single carrier QAM
TCP	transport control protocol
TDD	time division duplex
US	upstream

Bibliography & References

[1] John T. Chapman, "DOCSIS Remote PHY", *SCTE Cable-Tec Expo Fall Technical Forum*, Atlanta, GA, Oct 2013.

[2] John T. Chapman, "Remote PHY for Converged DOCSIS, Video and OOB", *NCTA/SCTE Technical Forum*, Los Angeles, Jun, 2014. [link][link2]

[3] Pawel Sowinski, Andy Smith & Tong Liu, "Remote PHY 2.0: The Next Steps For Remote PHY Technology", *SCTE Cable-Tec Expo Fall Technical Forum*, New Orleans, Sep, 2019. [link]

[4] John T. Chapman, "Managing DOCSIS 3.1 Profiles", *SCTE DOCSIS 3.1 and Wireless Symposium*, Denver, Oct, 2014.

[5] John T. Chapman, "The Power of DOCSIS 3.1 Downstream Profiles", *NCTA Spring Technical Forum*, Jun, 2013. [link]

[6] CableLabs, "DOCSIS 4.0 Technology", [link]

[7] John T. Chapman, Mike Emmendorfer, Robert Howald, Shaul Shulman "Mission is Possible: An Evolutionary Approach to Gigabit-Class DOCSIS", NCTA Spring Technical Forum, May 2012. [link]

[8] Jason Miller, "OFDM Capacity Optimization", Cisco Knowledge Network, May 2021 [link]

[9] John T. Chapman, Hang Jin, "Full Duplex DOCSIS", *SCTE/NCTA Spring Technical Forum*, Boston, May, 2016. [link]

[10] Tong Liu, John T. Chapman, Hang Jin, "Interference-Aware Spectrum Resource Scheduling for FDX DOCSIS", *SCTE Journal of Network Operations*, Vol 1, No 2, Sept, 2016. [link]

[11] Hang Jin & John T Chapman, "Echo Cancellation Tecniques for Supporting Full Duplex DOCSIS.", *SCTE Cable-Tec Expo Fall Technical Forum*, Denver, October, 2017. [link]

[12] John T Chapman, Hang Jin, "FDX DOCSIS Line Extender: Deploying FDX DOCSIS Beyond N+0", *SCTE Cable-Tec Expo Fall Technical Forum*, Atlanta, Oct, 2018 [link]

[13] Hang Jin, John T. Chapman, "FDX Amplifier for Supporting N+M Network", *SCTE Cable-Tec Expo Fall Technical Forum*, New Orleans, Sep, 2019. [link]

[14] John T. Chapman, Hang Jin, Thushara Hewavithana; Rainer Hillermeier, "Blueprint for 3 GHz, 25 Gbps DOCSIS," *SCTE Cable-Tec Expo Fall Technical Forum*, New Orleans, Sep, 2019 [link]

The Road to 10G

Migrating Today's HFC Network to Meet Tomorrow's Demand

A Technical Paper prepared for SCTE by

Mike Cooper
Principal
Cox Communications
6305 (CTECH-B) Peachtree Dunwoody Rd, Atlanta, GA 30328
404-269-4133
Michael.Cooper4@cox.com

David Job, Principal, Cox Communications
David.Job@cox.com

Bill Wall, Principal, Cox Communications
Bill.Wall@cox.com

1. Introduction

The hybrid fiber coaxial (HFC) network has served as both a reliable and flexible architecture which has enabled cable operators to continue to migrate their networks to meet the exponential growth in bandwidth demanded by both commercial and residential customers. This architecture has supported cost effective capacity increases with minimal disruptions to the network and end-users. While more recent efforts have focused on building fiber-to-the-home (FTTH) networks, they are generally quite expensive with cost estimates often nearing $1000 per household passed (HHP) or greater. Given the long-term capacity of an all-fiber network, it may be considered the obvious choice for green field deployments. However, for brown field areas, the cost of a FTTH network upgrade drives the continued evolution of the existing HFC architecture as a better choice.

The introduction of extended spectrum Frequency Division Duplex (FDD) Data Over Cable Service Interface Specification (DOCSIS) within the DOCSIS 4.0 specifications enable that continued evolution by defining standards that for Cox would support capacities as high as 12 Gbps on the downstream and 3 Gbps on the upstream across the legacy HFC node and amplifier (N+X) architecture.[CM-SP-PHYv4.0] The strategy for operational implementation of this upgrade will be a critical component to the success of multi-Gig deployments over HFC.

This paper will explore several considerations for upgrading the HFC network to 1.8 GHz FDD DOCSIS 4.0 including a drop-in amplifier approach, intermediate upgrade steps, amplifier gains/tilts, and other operational aspects.[1]

2. Requirements and Constraints

A key challenge for cable system operators is the ability to transform the HFC network to meet the growing bandwidth demands while simultaneously continuing to provide reliable services to a large existing customer base. In fact, many of these customers have service level agreements (SLAs) in place which prohibit significant down time in the network. As a result, it is almost always the case that an operator cannot upgrade an entire market or, in some situations, even a single node within a maintenance window of a few hours. The ability to execute an upgrade to a network in incremental steps while continuing to support legacy devices and services is critical to the success of the operator, and the evolution to 1.8 GHz extended spectrum DOCSIS will be no exception to this requirement.

2.1. Tap Output Radio Frequency (RF) Levels

The DOCSIS 4.0 specification was developed with an assumption that extended spectrum DOCSIS 4.0 FDD signals operating up to 1.8 GHz would be communicating with a single point of entry (POE) device in the home. That is, there would be no splitters, drop amplifiers, or pass-through devices between the tap output and the DOCSIS 4.0 customer premise equipment (CPE) device. For extended spectrum FDD, this assumption was based upon the need to provide adequate forward signal levels at the CPE to support up to 1.8 GHz in spectrum with higher order quadrature amplification modulation (QAM) levels. Historically, operators have modeled drops with 2 or 4-way splitters and a 100-150 ft of cabling. With a POE device, at a minimum, the 2 or 4-way splitters are no longer necessary providing an additional 4 to 8 dB of signal level to a DOCSIS 4.0 device. However, that signal level increase does not carry-over to

[1] The authors wish to acknowledge and express their appreciation to others in the Cox Outside Plant Engineering team who also contributed to the technical work presented in this paper

legacy devices in other homes. It will take significant time for operators to migrate their in-home CPE to DOCSIS 4.0, and, as a result, the network will be required to support both single DOCSIS 4.0 POE device homes and legacy homes with multiple CPE devices including legacy video.

2.1.1. Legacy Levels

Video service offerings were the driver for the birth and growth of the cable industry during the 70's, 80's and 90's. The incorporation of data services (DOCSIS) was first introduced in the late 90's long after video services provided in the form of set tops were prolific within the network. While many operators, including Cox, are developing roadmaps to replace proprietary video CPE devices with Internet Protocol Television (IPTV), it is both cost prohibitive and infeasible to do a wholesale replacement of millions of these legacy devices within a short time. As a result, operators are required to continue to support both legacy video and data services to set tops and cable modems at their current location within the home as they migrate the network. This requirement constrains the architecture as tap output power levels in the legacy spectrum must be kept close to what they had been prior to the upgrade. When tap levels change, legacy CPE devices within the home are at risk of no longer receiving the minimum signal level required, driving the potential for a poor customer experience and truck rolls.

During Cox's initial experimentation with node plus 0 amplifier (N+0) architectures in 2017, early attempts were made at minor reductions in tap output levels to increase the reach or HHP per node. (See Figure 1) Early trials revealed a more than 20% increase in truck rolls and while there are a number of factors which can contribute to truck rolls when migrating from a N+X to a N+0 architecture, subsequent trials with restoration of design tap output level targets to more closely match legacy levels brought the number of truck rolls back in line with historical levels.

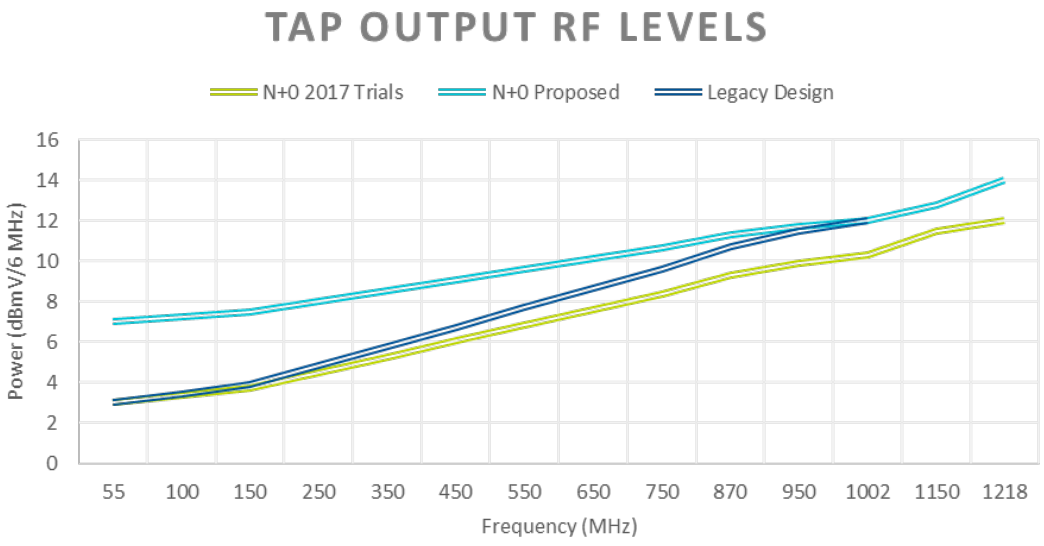


Figure 1. Previous Cox Trials Revealed the Importance of Maintaining Legacy Levels

2.1.2. In-House Conditioning

In addition, a more subtle dependency in tap output levels also exists. While HFC plant design constraints dictate a minimum tap output level, fielded implementations generally yield levels that are

above that minimum. See Figure 2 for an example of output levels for a series of 5 consecutive taps on a cable run between amplifiers. Actual output levels are provided in blue and minimum design requirements are provided in red. If a future change or upgrade to the plant results in the output levels changing, for example, if any of the actual output levels are reduced and brought closer to the design minimum levels, then any homes/devices connected to taps whose output levels were reduced might be at risk of no longer functioning properly.

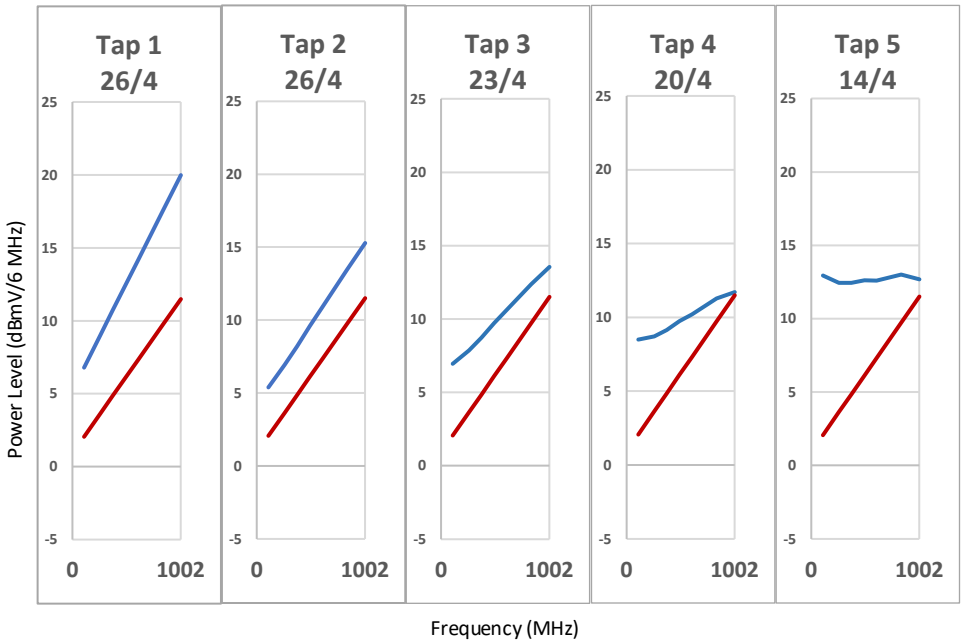


Figure 2. Example Tap Cascade Output Levels (Blue) versus Minimum Design Spec (Red)

While the red curve meets minimum design constraints, there is a risk that in-house wiring may have leveraged the additional power previously provided at that tap to deliver adequate levels perhaps deeper into the home. That is, either the operator or the homeowner may have placed devices within the home that leveraged this additional power. Alternatively, technicians may have installed in-home attenuators to force levels closer to what they considered to be optimum. If the levels are subsequently lowered (but still at or above spec), adequate signal level may not be present for those devices, driving the need for possible in-house amplification, splitter reconfigurations, or attenuator removals that were not required before. From the homeowner's perspective, his service was working before the upgrade but no longer functions properly afterward. This situation can be referred to as in-house conditioning and is quite common.

2.2. Mainline Amplifier Considerations

During most prior HFC downstream bandwidth upgrades, a "drop-in" approach was used. The drop-in term was meant to convey that legacy amplifier locations were maintained and that new amplifier modules/stations with increased downstream bandwidth were dropped in as replacements at those locations. Mainline passives (taps, splitters, and directional couplers) also maintained their legacy locations and were only replaced if needed to support the increased downstream bandwidth. The drop-in approach minimized the need for extensive plant construction/redesign that would have been required if amplifier locations were moved, which in turn helped minimize downtime during the upgrade.

When considering a 1.8 GHz upgrade, the drop-in approach continues to be attractive in that upgrades to passives (which far outnumber actives) could be accomplished regardless of whether the amplifiers in the area have been upgraded or not. This increases the flexibility to efficiently deploy resources for that part of the upgrade and enables small sections of plant (down to individual feeder legs) to be upgraded without the long duration outages inherent in upgrades where actives move locations, requiring the passives to be changed along with the actives.

Historically, both downstream amplifier gains and downstream amplifier RF output levels were increased to help facilitate drop-in upgrades while maintaining amplifier and tap RF output levels in the existing legacy band. Each are discussed next in more detail.

2.2.1. Downstream Amplifier Gain

To facilitate a drop-in bandwidth upgrade, the downstream gains of the mainline amplifiers were increased to offset increased cable losses at the higher frequencies. Table 1 below shows the downstream amplifier gains for many of the amplifiers commonly used in North America, and how they increased as downstream bandwidths increased.

Note that downstream amplifier gain generally refers to the gain at the downstream upper band edge (highest rated frequency). It is important to keep in mind that sufficient gain is required to offset the coaxial and passive losses between amplifiers (to achieve unity gain) and that gain requirements are not directly linked to amplifier RF output levels.

Table 1 – Amplifier Gains for 750, 870, 1000, and 1218 MHz Amps

Manufacturer	Amplifier Type	Operational Gain at Rated Frequency (dB)			
	Amplifier Rated Frequency (MHz):	**750**	**862/870**	**1000**	**1218**
Commscope - Motorola	MB-AGC	36	38	42	47
	4 Port BT–AGC	40	40	42	-
	BLE–AGC/Therm/Man	28	30	34	38
Commscope – C-COR	FMB–AGC	37	40	43	47
	FMT–AGC–Trunk Port	28	30	33	36
	FMT–AGC–Bridger Ports	37	39	43	47
	FM LE–AGC/Therm	30	32	35	38
	FM LE-Manual	33	35	38	42
Cisco - SA	HGBT–AGC/Therm		38	41	46
	HGD–AGC	35	40	43	48
	HGD-Therm	39	40	43	48
	UBT–AGC–Main Port	28	31	34	-
	UBT–AGC–Bridger Ports	37	40	43	-
	LE-AGC	24.5	31	33	38
	LE-Thermal	29	32	34	38
	LE-Manual	34.5	37	39.5	-

For a full drop-in type of upgrade to 1.8 GHz to be supported, the downstream gains for the amplifiers would need to increase even higher.

Figure 3 shows the historical gains for Line Extenders plotted against cable loss increases. The frequency axis was extended to show the associated gains that would be required for drop-in upgrades to 1.8 GHz. A cable length and type (1360 ft of PIII .500 cable) whose loss matches the legacy gains at various lower

frequency points was selected to graphically show the increased gain versus increased cable loss relationship. This shows the gains required if the span loss preceding the amplifiers were all coax, and if the amplifiers were at maximum spacing.

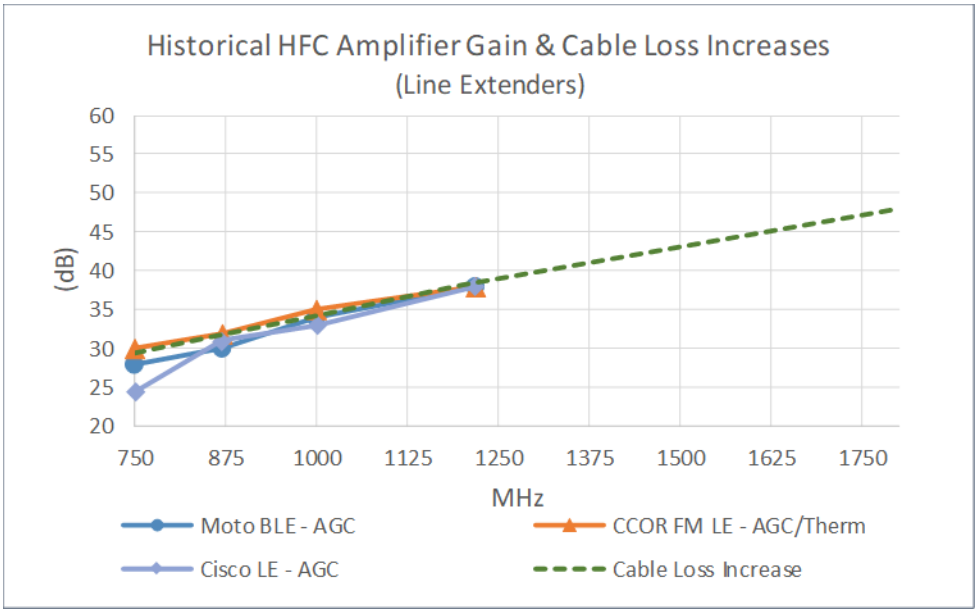


Figure 3. Line Extender Gain & Cable Loss Increase

For Line Extenders, the gains shown are for automatic gain control (AGC)/Thermal types. Based upon this information, gains at 1.8 GHz of approximately 48-49 dB may be required for full drop-in capability for these type amplifiers. There appears to be some consensus in the industry that this amount of gain is feasible for 1.8 GHz Line Extenders.

An internal Cox study was also performed to estimate what Cox's actual 1.8 GHz spacings would be for AGC/Thermal Line Extenders. Information from Cox's design data base was used to determine the existing 1 GHz coax and passive losses preceding each Line Extender. Based upon vendor data sheets, coax losses were increased by 40 percent and passive losses by 20 percent and summed to estimate the expected total loss at 1.8 GHz. The results are shown in histogram format in Figure 4.

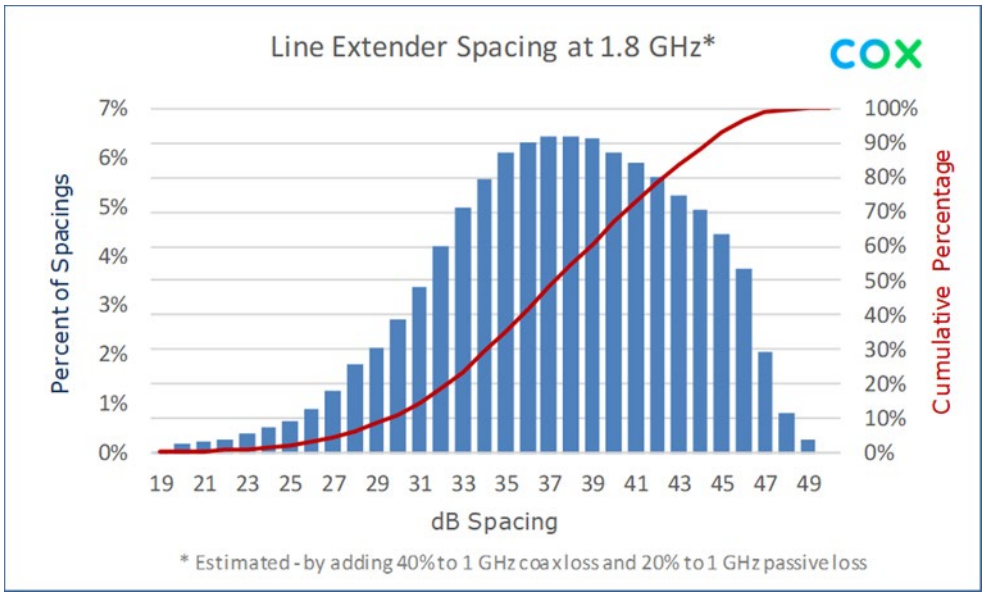


Figure 4. Estimated 1.8 GHz Line Extender Spacing for Existing Cox Spans

The histogram correlates the above assumption that AGC/Thermal Line Extenders with 48-49 dB of gain at 1.8 GHz should cover the vast majority of spacings for these type amplifiers using a drop-in upgrade approach.

Figure 5 shows the historical gains for high gain Multiport Bridger amplifiers plotted against cable loss. The frequency axis was extended to show the associated gains that would be required for drop-in upgrades to 1.8 GHz. A cable length and type (2450 ft of PIII .750 cable) whose loss matched the legacy gains at various lower frequency points was selected to graphically show the increased gain versus increased cable loss relationship. This show the gains required if the span loss preceding the amplifiers were all coax, and if the amplifiers were at maximum spacing.

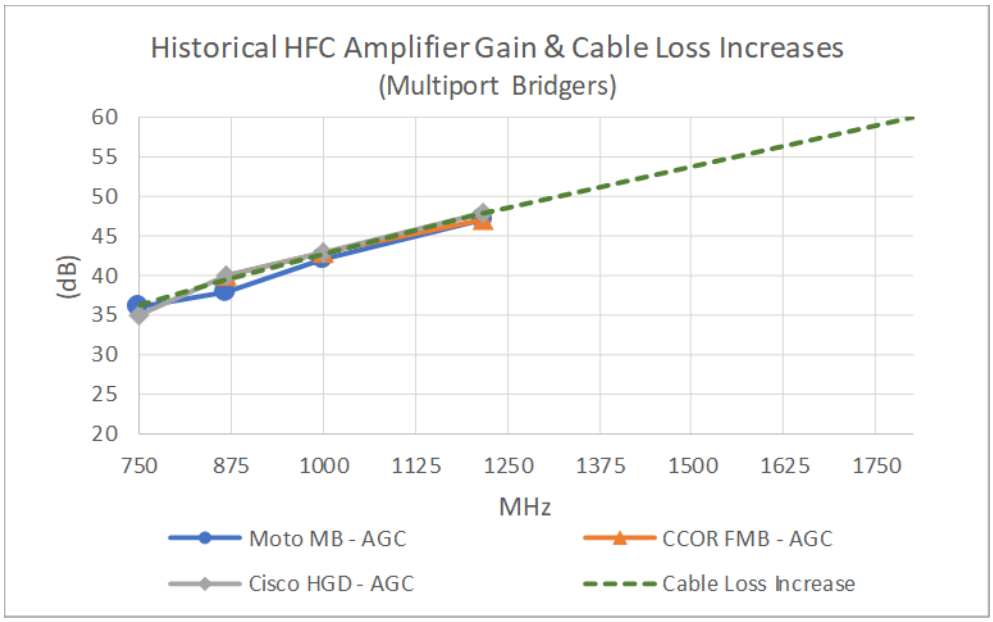


Figure 5. Multiport Bridger Amplifier Gain & Cable Loss Increase

For Multiport Bridgers, the gains shown are for AGC types. Based upon this information, gains at 1.8 GHz of 59-60 dB may be required for full drop-in capability for these type amplifiers. There appears to be consensus that this amount of gain would not be feasible for these types of amplifiers. For one, Carrier to Noise Ratio (CNR) would suffer due to the extremely low input levels that would need to accompany amplifiers with such high gains. Additionally, it would be very difficult to achieve sufficient closed loop isolation in a station with so much downstream gain combined with an increased upstream gain. There appears to be some consensus in the industry that 48-50 dB gain would be feasible for these types of amplifiers, but in systems that had made use of maximum spaced amplifiers with high coax losses in their legacy plant that would not be enough gain to overcome losses at 1.8 GHz. This led to the consideration of ways to address this concern.

The Booster Amplifier concept was conceived based upon the realization that there could be a 10 to 12 dB gain shortage in some cases where high gain Multiport Bridgers had been deployed at or near full spacing where most of the loss was coaxial. At Cox, this is called an express application, where the RF from the node was fed deeper into the network to efficiently reach tapped feeder areas by using lower loss express cable in conjunction with high gain amplifiers. In scenarios where the 1.8 GHz drop-in amplifier would not have enough gain to cover the high express losses, Cox initially considered adding a Line Extender, but realized a preference to use a smaller, lower gain, lower cost amplifier placed in the express cable route to supplement the drop-in amplifier's gain. Currently several amplifier manufacturers are planning to make such amplifiers which have come to be known as Booster Amplifiers. Cox recognizes that installing these type amplifiers will be much easier in aerial plant than in underground, but even in underground there are often locations along long express runs where the cable run is interrupted to accommodate a splitter, coupler, or splice block.

An internal Cox study was also performed to estimate what Cox's actual 1.8 GHz spacings would be for AGC Multiport Bridger amplifiers, to determine what percentage of those amplifiers might need Booster Amps added. Information from Cox's design data base was used to determine the existing 1 GHz coax and passive losses preceding each amplifier. Based upon vendor data sheets, Cox increased those coax

losses by 40 percent and passive losses by 20 percent and summed them to estimate the expected total loss at 1.8 GHz. The results are shown in histogram format in Figure 6.

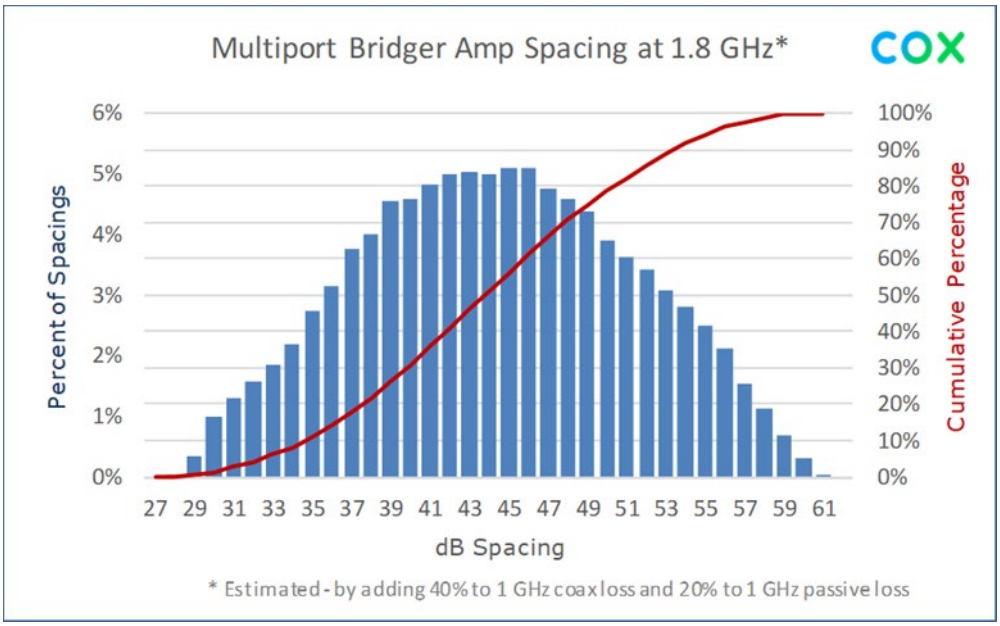


Figure 6. Estimated 1.8 GHz Multiport Bridger Spacing for Existing Cox Spans

The histogram shows that Multiport Bridgers with 48-50 dB of gain at 1.8 GHz should cover roughly 75 to 80% of Cox's spacings for these type amplifiers using a drop-in upgrade approach. The remaining 20 to 25% would require the addition of a Booster Amplifier. With Multiport Bridger amplifiers representing about 32% of the network, Cox's expects to need booster amplifiers for approximately 8% of the network spans.

2.2.2. Downstream Amplifier RF Output Levels

When considering downstream amplifier RF output levels for plant upgrades to 1.8 GHz, two important aspects must be considered. One is how closely the tap RF output levels after the upgrade should match the current tap RF output levels in the legacy band, and the other is the total composite power (TCP) constraints that exist when operating amplifiers with significantly greater channel loading at high output levels. When considering a drop-in upgrade approach these two factors are inherently linked.

The expectation that both legacy CPE and the new extended spectrum D4.0 CPE will need to co-exist in the same network for years to come should be well understood. With that comes a requirement to consider both the downstream RF levels that will feed the legacy CPE and the new extended spectrum CPE. As was mentioned prior, Cox learned that in order to minimize the potential for service complaints and associated truck rolls when performing a plant upgrade, it is best to keep the downstream RF levels feeding legacy CPE very close to what they were.

If the assumption is to leave the downstream RF output levels of the taps nearly unchanged in the legacy band, and to replace the existing taps with 1.8 GHz taps using the same tap values (which looks possible based upon 1.8 GHz tap insertion losses), then it follows that the downstream RF output levels of the replacement amplifiers should remain nearly unchanged in the legacy band. At Cox, the legacy band is 54 or 108 MHz to 1002 MHz for most of the network. Existing typical node/amplifier RF output levels in

that band are shown in Figure 7. Note that actual signal level per 6 MHz is shown (not analog equivalent).

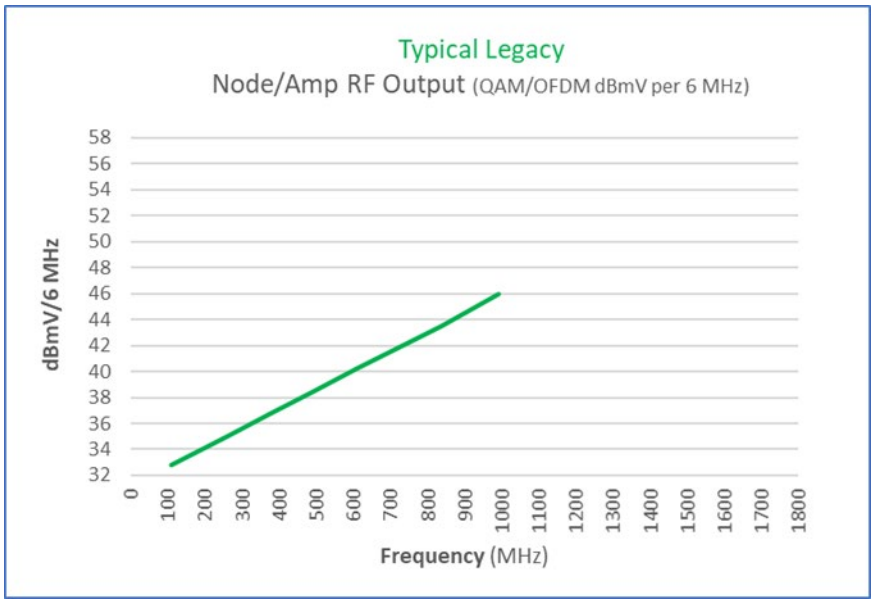


Figure 7. Typical Node/Amp RF Output Levels

In drop-in plant upgrades from 750 or 870 MHz to 1002 MHz, traditionally the amplifier RF output levels were increased along a common output tilt line. This trend also applied to some earlier plant upgrades that started at 550 MHz and has recently been incorporated for upgrades to 1218 MHz. Evidence of this approach is shown in Table 2 and Figure 8. Table 2 shows the reference output levels on the published data sheets for two major manufacturer's amplifiers. Figure 9 shows the same information in graphical format. While not all operators ran the exact reference RF output levels that were specified on published data sheets, most did tend to use similar RF output tilts as those shown.

Table 2 – Amplifier Data Sheet Reference Output Levels/Tilts (dBmV/Channel – analog equivalent)

Mfr	Rated Frequency (MHz)	Amp Type	54 MHz	750 MHz	870 MHz	1002 MHz	1218 MHz
S-A	**750**	**LE, HGD, HGBT**	36	46			
S-A	**870**	**LE, HGD, HGBT**	35	45.7	47.5		
S-A	**1002**	**LE, HGD, HGBT**	35	45.7	47.5	49.5	
Moto	**750**	**BLE, MB, BT**	37	47			
Moto	**870**	**BLE, MB, BT**	37	47	49		
Moto	**1002**	**BLE, MB, BT**	37	47	49	51	
Moto	**1218**	**BLE, MB**	38	48.2	49.9	51.8	55

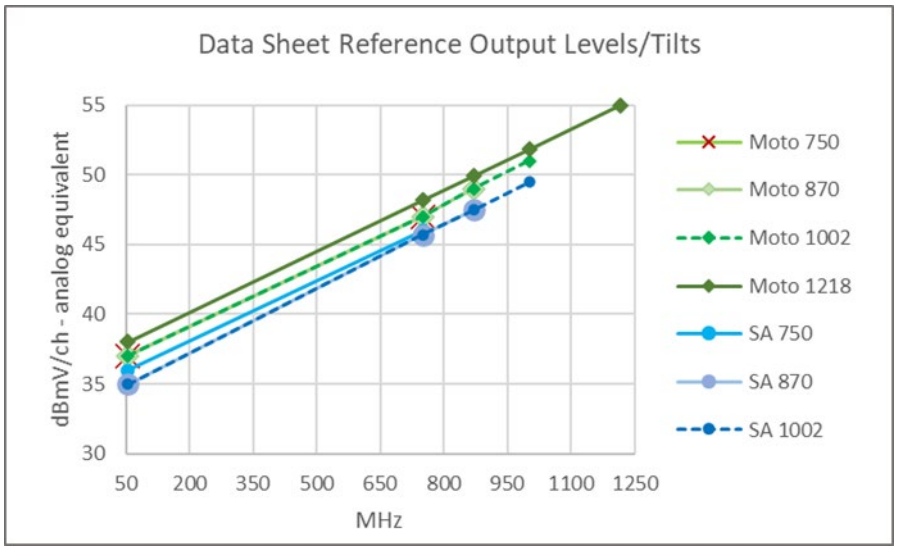


Figure 8. Amplifier Data Sheet Reference Output Levels/Tilts

Cox followed the trend of increasing the RF output levels along a common tilt line during upgrades to 1002 MHz; however, using a similar approach for upgrades to 1.8 GHz introduces a problem. Figure 9 shows what that would look like.

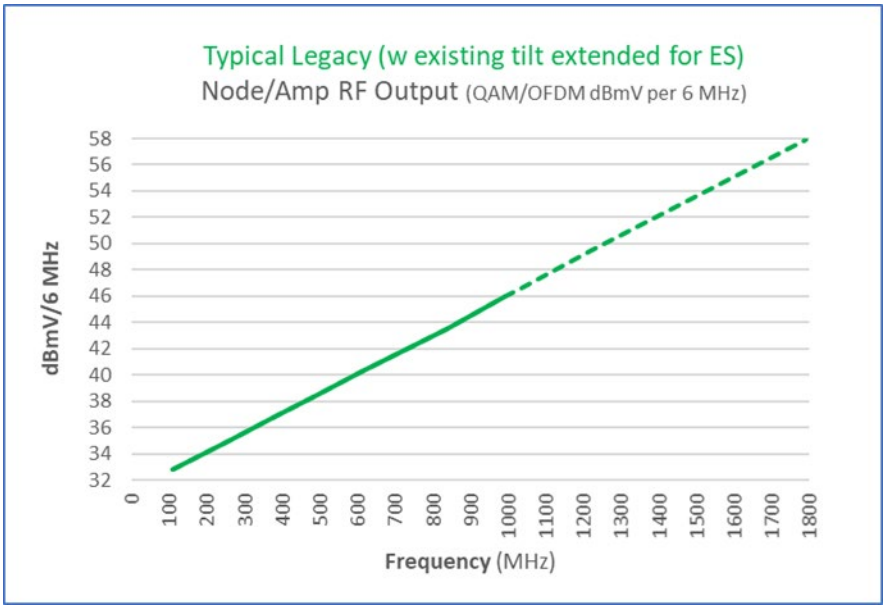


Figure 9. Typical Node/Amp RF Output Levels - if extended to 1.8 GHz

If the tilt line representing Cox's typical amplifier RF output levels in the legacy spectrum was extended out to 1.8 GHz and the band was fully loaded at those levels, the total composite power (TCP) at the amplifier output would be excessively high reaching 74.8 decibels relative to one millivolt (dBmV). It is worth noting that the associated TCP for the legacy band in this case is only 62.7 dBmV. If you consider that there may be roughly 4 dB of loss from the output of the final amplifier gain stage (commonly referred to as the power amplifier) to the station RF output port, 74.8 dBmV TCP at the port would equate

to ≈ 78.8 dBmV TCP at the power amplifier output. That is much higher than any power amplifier can currently support while outputting usable signals.

The TCP limitations of state-of-the art broadband power amplifiers were recognized early on during the extended spectrum discussions, as similar discussions had gone on previously in association with the very high RF output powers required for N+0 architectures. All broadband amplifiers have limitations on how high they can run before distortion becomes excessive. The question of how high the new amplifiers will be able to run with extended spectrum loading is under investigation but gain block manufacturers are diligently working to optimize and characterize their output power capabilities. While beyond the scope of this paper, it is worth mentioning that there's more than one aspect to consider, as in some cases better performance may come at the cost of higher power consumption.

For reference, amplifier output TCP can be calculated by converting the output signal level (in dBmV per 6 MHz for each 6 MHz of bandwidth that will be occupied) to power (typically expressed in milliwatts), summing those powers, and converting the summed power to dBmV.

When the limitations due to TCP were discussed in conjunction with the desire to keep RF amplifier output levels near to what they have been in the legacy band, the idea of reducing or stepping down the RF power in the extended spectrum became a widely discussed topic. Various output power profiles for the extended spectrum were suggested, such as making use of multiple step downs; up to one per 192 MHz orthogonal frequency division multiplexing (OFDM) channel (referred to as zig-zag or lightning bolt), or shaping the amplifier RF output power in the extended spectrum to be flat while keeping the legacy spectrum tilted (referred to as tilt-flat).

After much discussion, most operators coalesced around the idea of having one or possibly two step downs in the extended spectrum, with all output power levels referenced to a virtual linear tilt line from lowest to highest frequency (referred to as tilt-tilt). This concept was not new, as prior to conversions to "all-digital" signal carriage, cable operators ran amplifier RF outputs referenced to what was called an "analog equivalent" tilt line - with analog video carriers ran at the levels shown on the tilt line and digital (SC-QAM256) signals typically ran 6 dB lower relative to the tilt line.

The ability for the DOCSIS 4.0 remote PHY device (RPD) / remote MACPHY device (RMD) in a distributed access architecture (DAA) node to make use of a step down (or multiple step downs) in the extended spectrum was written into the D.4.0 PHY specification. Up to 10 dB of step-down capability was mandated. The step-downs, once created at the RPD/RMD module, would pass through amplification and tilt stages in the node so that the tilted RF output with associated extended spectrum step-downs would be outputted at the node RF output ports. With this approach, the broadband amplifiers following the node in N+X architectures would not have any part in producing the RF profile with the step downs, and would just need to make use of equalization, amplification, and internal tilt to allow the duplication of the tilted RF output with associated extended spectrum step downs at their RF output ports.

With the extended spectrum step-down capability, operators will need to work to optimize a variety of factors. If it is desired to maintain legacy band amplifier RF output levels near to what they have been, the amount of step down for the extended spectrum can be adjusted to align the output TCP of the amplifier to a place that the operator deems to be optimum. That will require interaction with amplifier station manufacturers familiar with the output level capabilities of the amplifiers.

Some of the trade-offs to consider are that increasing the amount of extended spectrum step down lowers the TCP and increases the RF power margin between design operating levels and the point where significant ill effects from amplifier overdrive occur – such as 4-5 dB reduction in modulation error ratio

(MER) with only 1 dB increase in levels, or the onset of significant bit errors. That margin needs to be considered because real world amplifiers and networks will have some inherent factors that could cause amplifier output levels to be higher than targeted. Those factors include the accuracy of the actual RF output levels versus the programmed RF output levels, the AGC's ability to hold output levels over temperature, and the fact that frequency response build-up in the network can increase the TCP even when high and low frequency output levels are held tightly at target levels. On the flip side, increasing the amount of extended spectrum step down reduces the RF input levels to amplifiers and the DOCSIS 4.0 modems in that spectrum, which can reduce overall carrier to noise and potentially reduce the order of modulation that the extended spectrum OFDM signals can support.

At Cox, based upon vendor feedback concerning power amplifier performance and on Cox's plant performance modeling, Cox identified two potential initial target profiles for amplifier RF output levels, shown in Figure 10 below. The green trace represents holding Cox's RF output levels in the legacy spectrum where they are now, extending that virtual tilt line, and using a 6 dB step down in the extended spectrum from 1 to 1.8 GHz. This represents a total virtual tilt of 25.1 dB from 111-1791 MHz. The blue trace shows what is Cox's preferred option, where the legacy RF output level at 1 GHz is held, but the overall virtual tilt is reduced and again a 6 dB step down is used in the extended spectrum. This represents a total virtual tilt of 21 dB from 111-1791 MHz.

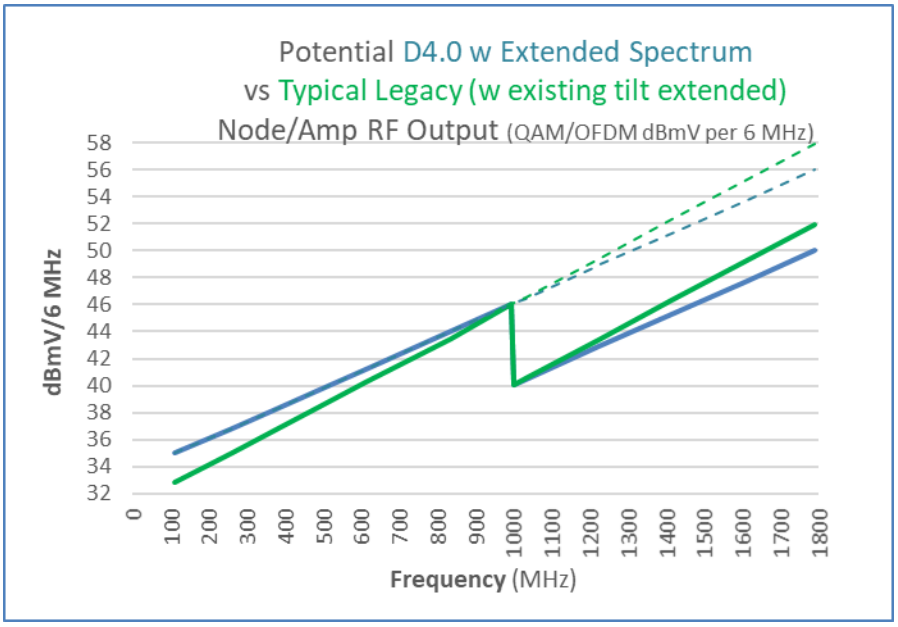


Figure 10. Potential Amplifier RF Output Profiles for 1.8 GHz

There are three potential advantages to the reduced tilt option:

1. Cox learned from amplifier gain block manufacturers that not all TCP is created equal. While it is true that conceptually, amplifiers with lower TCP also produce lower distortion, it has been found that the RF power at the highest frequencies has a disproportionate effect on distortion that is not captured via TCP differences alone. The RF energy at those very high frequencies can cause significant effects on the overall distortion, and it appears that keeping the RF energy lower at those frequencies is beneficial. In the blue reduced tilt trace the RF output power per 6 MHz at 1791 MHz is 2 dB lower than in the green trace.

2. There is also an improvement in TCP for the blue reduced tilt trace (68.6 dBmV) compared to the green higher tilt trace (69.5 dBmV).

3. The first incarnations of 1.8 GHz taps available on the market have "flatter" insertion losses than legacy taps, i.e., less tilt from low to high frequency across the band. Since the losses at 1 GHz are no greater than the losses of the legacy taps they will replace, and they are flatter, that means that the losses at the lower frequencies for the new taps end up being somewhat higher. With the blue reduced tilt profile, the higher outputs from the amplifiers in the lower frequency range help to offset the effect of a tap cascade having higher insertion losses in that band.

Based upon modeling, the slight reduction in RF power per channel at the higher frequencies with the detilted blue trace profile does not appear to be enough to alter the supported order of OFDM modulation in that region; however, Cox will continue to work with amplifier manufacturers to monitor measured performance attributes and further adjustments to the extended spectrum step down may be needed.

One final input on this topic. Since Cox's initial efforts to develop a target amplifier RF output profile, a new wrinkle came up, which was the need to plan to support 1.2 GHz high-split in portions of the network. Understanding that RF output levels in the new 1 to 1.2 GHz spectrum will become a legacy condition to deal with on the road to 1.8 GHz, Cox put careful thought into what should be done in that spectrum. Many of today's DOCSIS 3.1 RPD's and the software applications used to configure them do not currently support the 6 dB step down shown above while continuing to provide quality overall performance. Cox also realized that if that spectrum was deployed with no step down, it would be painful to later reduce those levels significantly to match the 6 dB step down with the migration to 1.8 GHz. After careful consideration, Cox's current plan is to introduce a 3 dB step down, which is more readily available in RPD products today, in the new 1 to 1.2 GHz spectrum. With respect to how that will play into the eventual migration to 1.8 GHz, the target RF output profile (with the new double step approach), is shown in Figure 11 below. This did increase TCP slightly (by 0.3 dB to 68.9 dBmV) relative to the initial target shown above. There is a benefit in that OFDM signal in the ≈ 1 to 1.2 GHz spectrum is now expected to have better MER than anticipated in the original single stepdown profile. The virtual tilt from 111-1791 MHz remains unchanged at 21 dB.

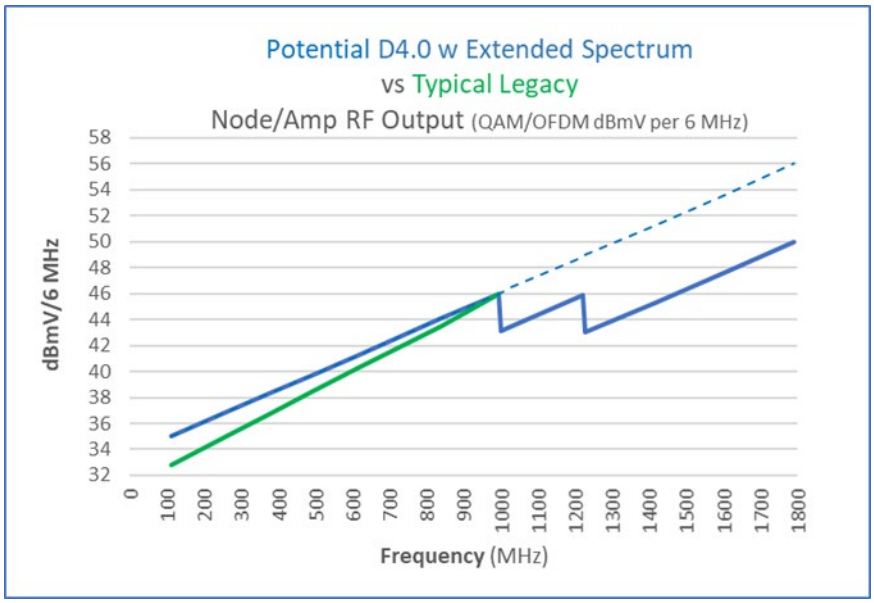


Figure 11. Current Potential Amplifier RF Output Profile for 1.8 GHz

2.3. Impacts of Spectrum Changes

Cable operators have a great deal of experience over the years with frequency expansion upgrades, i.e., increasing the top end frequency of the network (e.g., 550, 750, 870, 1000 MHz); however, these upgrades have always targeted increasing the downstream bandwidth and have rarely modified the upstream/downstream frequency split until recently. The transition from sub-split (42/54 MHz) to mid-split (85/108 MHz) represents one of the first times that operators are moving that upstream/downstream frequency split. Movement to mid-split is a minimal impact upgrade as most services that are locked into a particular frequency do not fall in the affected range or, if they do, the services were agile enough to allow movement beyond the band of change. With the transition to a high-split configuration (204/258 MHz), operators can no longer expect minimal impact as several downstream services are fixed in the spectrum below 204 MHz which will transition to upstream.

2.3.1. Service Impacts

One cannot say that DOCSIS has not been a resounding success. Cable was once a video only service that is now a majority data service. The compounded annual growth rate (CAGR) of data services continues to be in the 30% range, putting pressure on the need for more and more bandwidth devoted to data. This has led to an unprecedented number of node splits annually and the move to mid-split to relieve upstream congestion. Mid-split and the move to high-split and eventually ultra-high-split reduces the available downstream bandwidth below 1 GHz. While the higher spectral efficiency of DOCSIS 3.1 has helped significantly, it is not enough by itself.

2.3.1.1. QAM Video

To compensate for the reduced downstream bandwidth, Cox has converted QAM based video from MPEG2 to MPEG4, reducing the number of QAM video channels from approximately 70 to 42. Cox has already made extensive use of switched digital video (SDV) to reduce the number of QAM channels devoted to video. The MPEG4 conversion allowed the move to mid-split and the addition of another 96 MHz DOCSIS 3.1 OFDM block with no reduction of video services. Eventually, over a several year

period, Cox plans to eliminate QAM based video services in favor of IPTV. Approximately 10% of our video base is already IPTV only. Based on current utilization, an all IPTV 250 HHP node would consume approximately 450 Mbps of IPTV traffic at peak busy hour. That is equivalent to ¼ of a DOCSIS 3.1 OFDM block, or 48 MHz (8 QAM video channels). However, the conversion to all IPTV comes at a significant price, the retirement of QAM only STB's and a much larger population of DTA's. The newest generation of QAM enabled STB's include bonded DOCSIS 3.0 modems and can be converted to IPTV.

There is also significant interest in being able to offer a symmetric gigabit service prior to the availability of DOCSIS 4.0. There are two paths using DOCSIS 3.1 to achieve this. The first is to move to high-split and increase the upper end of the downstream spectrum to 1.2 GHz to recover the lost downstream bandwidth. This requires that not only actives in the plant be upgraded, but passives will have to be upgraded beyond their current 1 GHz limitation. A second approach would be to accelerate the move to IPTV, removing QAM video from the network and devote that bandwidth to DOCSIS. This would require retiring all QAM-only CPE including the millions of DTA's in the network. The plant upgrade would then touch only actives and would be similar in cost to a mid-split upgrade. The tradeoff is then the cost of the passive upgrade vs. the cost of retiring CPE. Both must be done eventually to meet the final DOCSIS 4.0 ultra-high split (UHS) end state. A "halfway" solution is also possible; Cox dedicates 16 of the QAM channels to SDV/VOD. Retiring the QAM only CPE would allow removing the SDV/VOD QAMs from the network, keeping the broadcast QAMs and DTAs. This would allow high-split on 1 GHz plant with only a slight limitation in DOCSIS capacity while minimizing the amount of retired CPE.

2.3.1.2. Legacy OOB STBs

The upstream move to high split or ultra-high split introduces several issues. The first is the legacy out-of-band (OOB) downstream for STB's. Downstream OOB is limited to a highest frequency, by specification, of 130 MHz, which now falls in the upstream spectrum. Cox has retired all legacy OOB STB's in favor of DOCSIS Set top Gateway (DSG) STB's, however customer owned UDCP devices, such as TiVo, that use CableCards are required by regulation to be supported using legacy OOB. Several approaches to a solution have been proposed, such as moving the OOB to a higher frequency in the downstream and providing an in-home down converter for devices needing legacy OOB. Cox has chosen another approach, along with other operators, in developing a DSG to legacy OOB converter to be placed in the homes of UDCP customers. This eliminates the need for OOB support within the coaxial network.

2.3.1.3. Impacts of HS and UHS Devices

Large numbers of sub-split and mid-split STB's and modems will continue to be in the network, potentially for years, after high-split or ultra-high-split (UHS) networks are enabled. A high-split DOCSIS 3.1 modem or an UHS DOCSIS 4.0 modem could be transmitting upstream with levels bursting as high as 65 dBmV, with a significant portion of that energy falling within the downstream passband of the sub-split or mid-split devices. For a STB or modem, the total composite power (TCP) received from the network may be on the order of 20 dBmV. The automatic gain control (AGC) on the front end of those devices reacts relatively slowly, and the bursting upstream from the high-split or UHS modem, even through the isolation of a splitter, may overload the front end of the older device causing dropped packets or macro blocking. For this reason, Cox is recommending that all homes that are high-split or UHS enabled, should be single point-of-entry, and any video services should be IPTV only using wired or WiFi connectivity to the gateway. A potential similar problem exists for adjacent homes off the same tap of a high-split or UHS home. However due to drop cable loss and better isolation in the tap, analysis

estimates this to be a problem in less than 5% of adjacent homes. Remediation could be through conversion of that home to HS/UHS or the installation of upstream blocking filters in that home.

2.3.1.4. MoCA Networks

Cox, along with other operators, has a significant footprint of Multimedia over Coax Alliance (MoCA) enabled homes used for multi-room DVR or in some cases WiFi extensions. Nominally, MoCA point-of-entry (POE) filters are used to isolate MoCA signals from the network and from interfering with adjacent MoCA networks. Once the downstream spectrum is extended above 1 GHz, to 1.2 or 1.8 GHz, MoCA will be incompatible with the network. Homes that have only 1 GHz CPE and MoCA, can continue as is if the MoCA POE filter provides adequate isolation. Homes using spectrum above 1 GHz cannot use MoCA, and any existing POE filters will have to be removed and the MoCA network replaced with WiFi or other compatible network technology within the home.

2.3.2. The Downstream Bandwidth Squeeze

In order to execute the migration process of increasing both the upstream and downstream frequency spectrum, two options are available as illustrated in Figure 12 and Figure 13. Figure 12 illustrates one option which provides for the lowest cost and lowest network downtime, as it benefits from completing the construction process (upgrading the node, and amplifiers for both upstream and downstream with the desired upstream/downstream frequency split) as a single step. However, by executing the entire construction change as a single step, the network maybe temporarily forced to operate with reduced downstream bandwidth, referred to here as the “squeeze”. The operator can avoid the squeeze by deploying an adequate number of CPE supporting the higher downstream frequency prior to the upgrade to absorb the lost bandwidth incurred by raising the split. Some limited modeling has indicated upgrading 20% of the CPE deployment should provide adequate congestion relief. However, the fact that these devices will need to be single point of entry (and drive the need to swap other sub-split and mid-split in-home devices) may limit early customer acceptance and thusly, the ability for an operator to deploy CPE early.

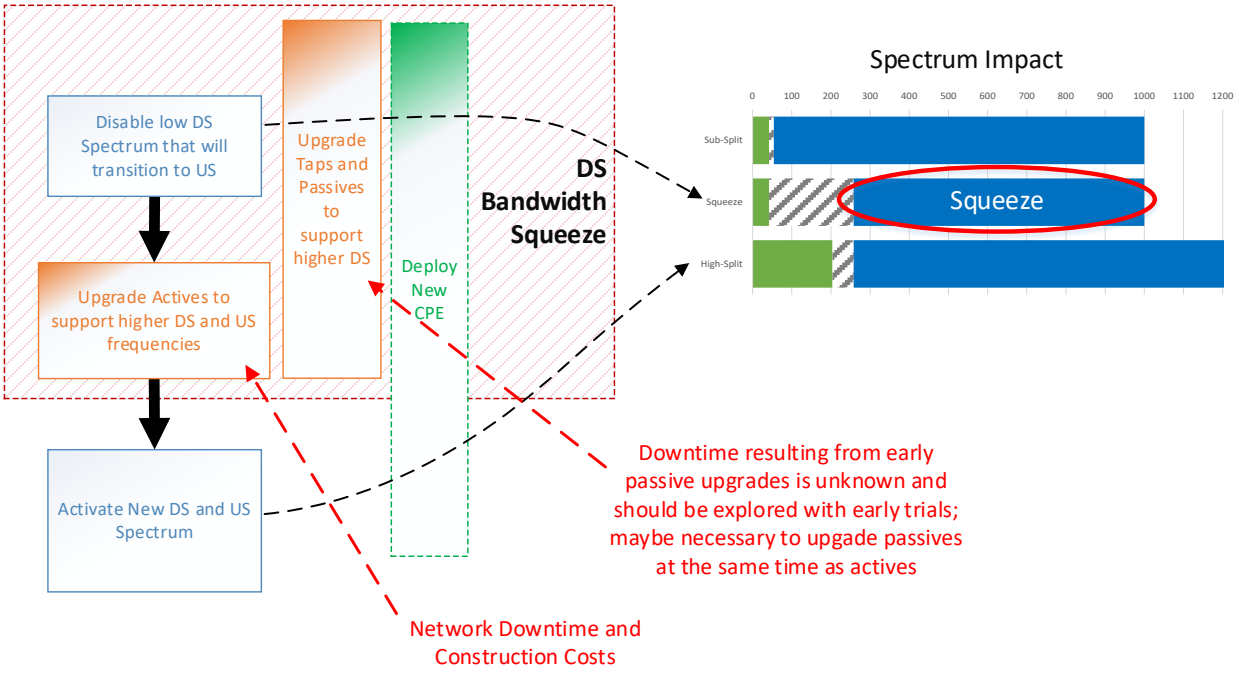


Figure 12. Migration Option to Support both DS and US Frequency Expansion with Minimal downtime but possibly incurring a Bandwidth Squeeze

Figure 13 illustrates a second option which eliminates the downstream bandwidth squeeze by executing the construction process as two steps. First the operator replaces the active and passive components to handle the higher downstream frequencies and activating that new downstream prior to changing the upstream/downstream frequency split in the actives. In parallel, the operator would be upgrading the CPE population to accommodate the increased downstream spectrum until an adequate number of high frequencies devices are deployed to eliminate the possibility of congestion when lower downstream spectrum is disabled to make room for upstream. Once an adequate number of new CPE were deployed, the operator would execute a second step to upgrade the active components a second time with the desired upstream/downstream frequency split and activating the new expanded upstream bandwidth.

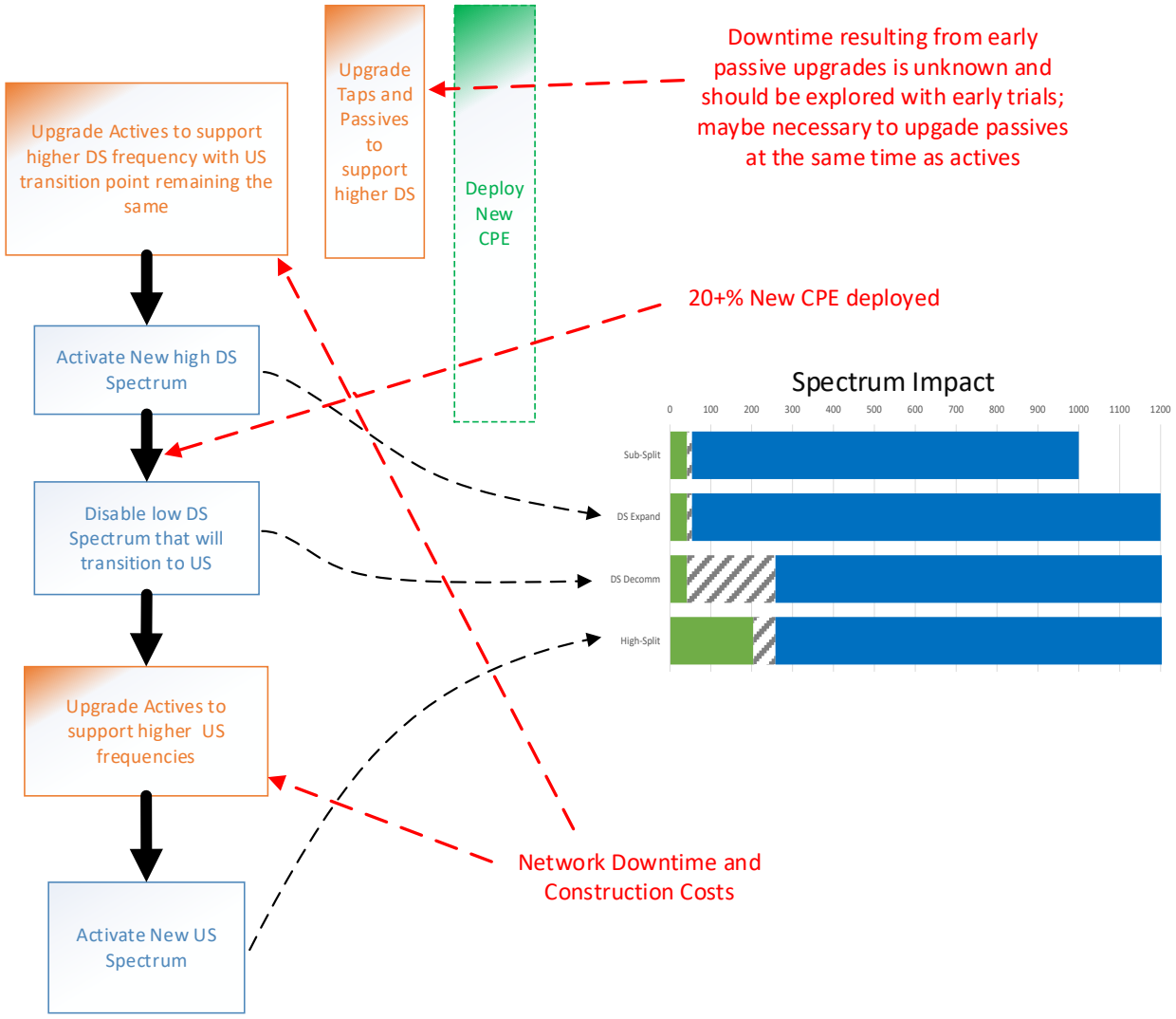


Figure 13. Migration Option to Support both DS and US Frequency Expansion with Increased Costs and Downtime but without a Bandwidth Squeeze

The increased costs and extended customer impacts in network down time, make the second option an undesirable choice. As a result, cable operators are strongly incentivized to identify ways to temporarily absorb the lost downstream bandwidth during the migration process. In the case of Cox, efforts are currently underway to look at accelerating the retirement of legacy video spectrum and replacement with IPTV which will provide some downstream spectrum relief during this transition period. Alternatively, operators may want to investigate accelerating their high-split upgrades prior to downstream congestion problems to also enable some absorption of this spectrum squeeze.

2.3.3. FCC Requirements for Leakage

Signal leakage is another factor which introduces significant challenges for operators as they move to DOCSIS 4.0 FDD, and, in particular, increased upstream spectrum. In the United States, the Federal Communications Commission (FCC) mandates that operators actively monitor their networks for signal leakage to assure that HFC network signals don't interfere with wireless services. While the FCC's requirement extends to many wireless services and frequencies, aircraft navigational signals (ranging

from 108 to 137 MHz) are of particular concern and as such, operators are required to provide documentation annually to the FCC to assure compliance.

Historically, this frequency band overlapped with what is the downstream spectrum on the cable plant (sub-split or mid-split spectrum plans). (See Figure 14) As such, vendor solutions relied on reference leakage test signals either generated by specialized equipment feeding the combining network in a headend or by RPDs in the case of DAA. With the consideration of high-split and ultra-high-split spectrum configurations, the lower edge of the downstream moves to 258 MHz or higher. The leakage reporting region of the aeronautical band is then clearly in the middle of the upstream frequency spectrum.

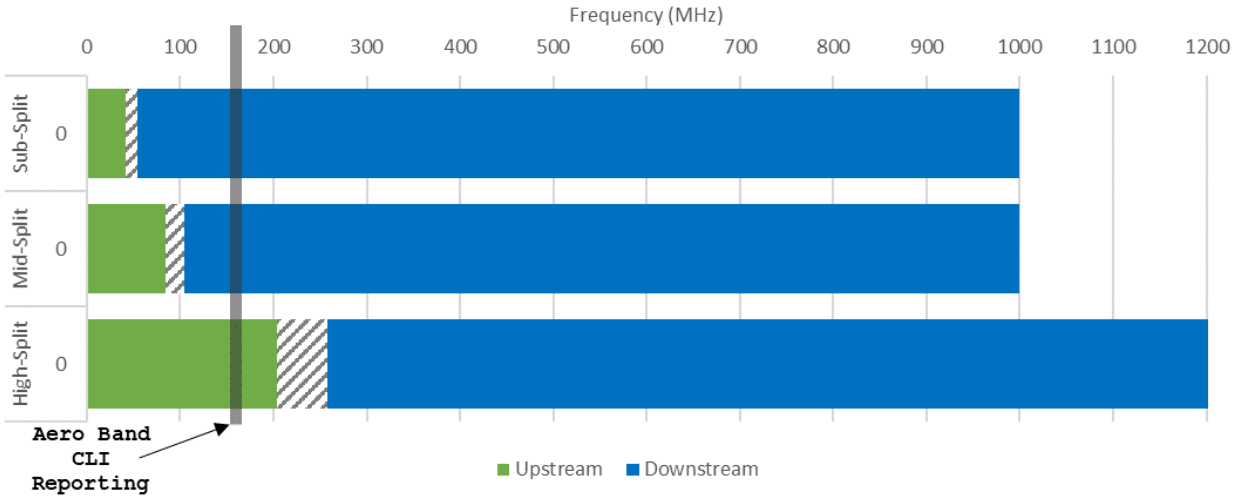


Figure 14. Leakage Reporting Spectrum Relative to Upstream/Downstream Split Configurations

This is a fairly dramatic paradigm shift for the leakage problem in that historically downstream signals are at their peak at the outputs of nodes and amplifiers in the network and are weaker as they progress through the tap outputs and onto the drop and in-house wiring. Conversely, upstream signals are near their peak at the cable modem output (on the order of 40 dB stronger than the downstream signals at the same point in the network). As a result, the signals in the aeronautical band are stronger in the house and on the drop, which is also the most common location of leaks.

High-split leakage solutions require a reference signal sourced from within the home to manifest these leaks within the network. An upstream signaling protocol called OUDP, that was part of the initial DOCSIS 3.1 standard, was leveraged to generate this reference signal. With some minor enhancements to the DOCSIS specifications, the industry enabled any DOCSIS 3.1 or 4.0 CPE device to function as a leakage reference signal source.[MULPIv3.1][CCAP-OSSIv3.1][MULPIv4.0][CCAP-OSSIv4.0] While requiring only firmware updates to CPE and CMTS devices, this solution does require hardware upgrades to leakage meters in order to detect this new signal. As a result, operators will need to lead their network upgrades with leakage meter upgrades to their truck fleet prior to enabling high and ultra-high-split configurations. It is generally expected that leakage meter vendors will provide solutions that function in both high-split and mid/sub-split networks enabling operators to continue to meet FCC reporting requirements during the transition period of upgrading from sub/mid-split to high/ultra-high-split on their network.

3. DOCSIS 4.0 Migration Strategy

The optimal strategy for upgrading a cable operator's network depends upon several factors including initial architectural state, network congestion conditions, competitive positions, as well as planned future product offerings. In addition, standardization of an architecture is highly desirable providing an operator with economies of scale in all aspects of business operations including volume pricing, flexibility in supply chain management, market product offerings, organizational knowledge, network maintenance, and technical expertise. However, the diverse set of factors affecting the strategy means that operators must plan the execution at a global level while realizing that they will often need to respond in a local market to unique competitive situations.

With Cox's EON (Extended Optical Network) program which began in 2007, Cox became one of the first cable operators to migrate their network to a 1 GHz design; however, Cox chose to remain at a sub-split configuration (42/54 MHz upstream/downstream frequency split) during that upgrade primarily due to its large base of legacy video CPE requiring significant amounts of downstream QAM video spectrum. As time progressed, due to other efficiency upgrades such as MPEG4, Cox remained fairly well positioned relative to downstream capacity; however, in more recent years, Cox has begun to experience the pressure of upstream bandwidth demand with more and more nodes encroaching upon upstream congestion.

The most common tool readily available to operators to address congestion problems is a node split. (See Table 3) Node splits are a rather coarse tool that divides the serving area in half providing the same capacity to both serving areas, each with half as many users. It effectively provides a doubling of capacity for both the downstream and upstream. However, node splits are a congestion relief tool only as they cannot provide for an increase in individual maximum customer rates since the total spectrum allocation remains the same.

Table 3 – Relief Levers Available to Operators to Address Bandwidth Challenges

Bandwidth Challenge	Target Issue	Relief Action
Congestion Relief	Upstream and Downstream	Node Split
Congestion Relief	Upstream	Mid-Split
Product Competition	1 Gbps Symmetric	High-Split
Product Competition	2 Gbps Symmetric	Ultra-High-split

In recent years, Cox has experienced an enormous growth in necessary node actions with more than 90% of those driven to address upstream congestion issues. Initially, Cox leveraged node splits, however, beginning in 2020, Cox turned toward mid-split plant upgrades as the preferred course of action. Mid-split upgrades provide an additional 150% of upstream spectrum (from 30 MHz of usable upstream spectrum to nearly 75 MHz) providing greater upstream congestion relief while also providing for a greater maximum sustained data rate.

3.1. Capacity Milestones

While a mid-split configuration provides for significant runway for operators dealing with the tactical problem of upstream node congestion, it fails to address some of the key upstream data through-put milestones such as 1 Gbps upstream which may be desired due to competitive threats. Table 4 provides the expected maximum sustained data rates for various split configurations under consideration by Cox. Similarly, Figure 15 provides for possible Cox frequency allocations for these split configurations. Table 4 illustrates that the capacities provided by DOCSIS 4.0 enable HFC to remain competitive with all-fiber networks. High-split and ultra-high-split configurations enable operators to offer 1 Gbps and 2 Gbps

symmetric speeds respectively through their networks. As such, it is the DOCSIS 4.0 1.8 GHz ultra-high-split configuration that is the end game for this plan.

Table 4 - Split Configuration Maximum Sustained Data Rates

Split Config	Upstream	Downstream
Sub-Split (42/1002 MHz)	0.1 Gbps	5 Gbps[1]
Mid-Split (85/1002 MHz)	0.5 Gbps	5 Gbps[1]
High-Split (204/1002 MHz)	1.5 Gbps	6.5 Gbps[2]
High-Split (204/1218 MHz)	1.5 Gbps	9 Gbps[2]
Ultra-High-Split (396/1794 MHz)	2.5 Gbps	12 Gbps

1 Assumes D3.1 limit of 2 4k-OFDM blocks and 48 D3.0 channels DS
2 Assumes all DOCSIS 4.0 4k-OFDM DS

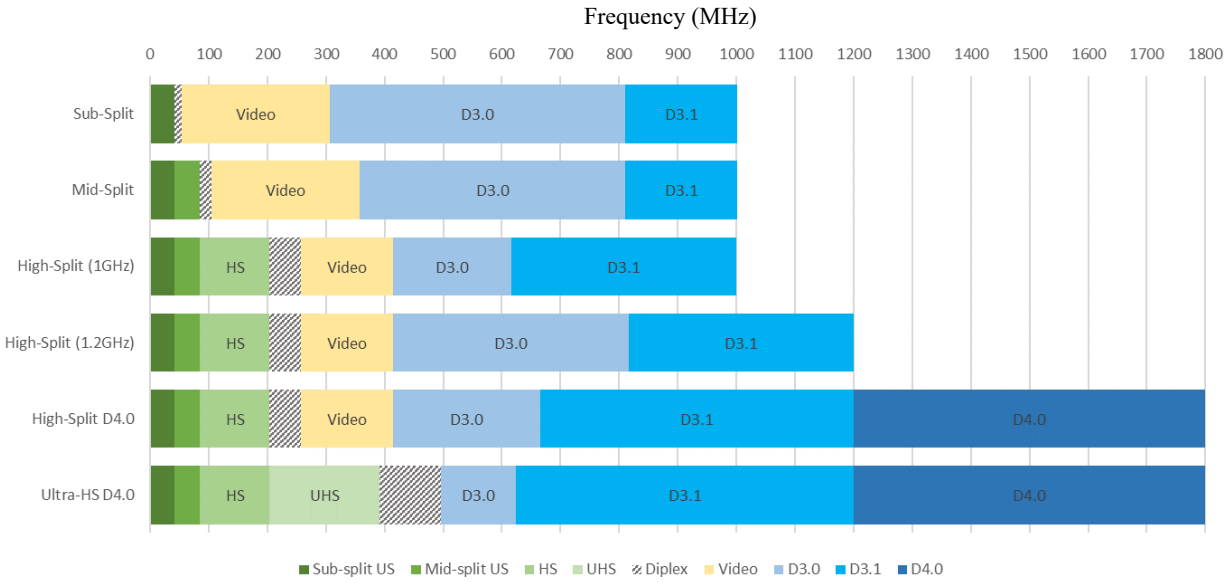


Figure 15. Possible Cox Spectrum Allocations for Various Split Configurations

3.2. Cost Implications

One important element to consider when evaluating migration paths is cost, and for the purposes of this analysis, Cox considered 8 alternatives for reaching a 1.8 GHz ultra-high-split target. (Figure 16). Each colored column in Figure 16 represents a network capacity as identified in Table 4. The circles in Figure 16 represent the network state of the plant where the letters represent the upstream/downstream frequency split (SS is sub-split, MS is mid-split, HS is high-split, and UHS is ultra-high-split) and the numbers represent the highest frequency supported by the actives and passives respectively. For example, the orange circles represent a network state with a high-split configuration and with nodes/amplifiers supporting 1.2 GHz, and taps/passives supporting 1.0 GHz (referred to as 1.2/1.0). Such a network would only run signals up to 1 GHz. While such a combination may seem odd to the reader, such states are incorporated into the plan to represent transition milestones that also minimize regrettable spend. Cox's network is primarily a sub-split 1.0/1.0 plant; however, as mentioned above, Cox is in a multi-year process of upgrading the network to mid-split 1.2/1.0. With Mid-split representing approximately 12% of the Cox network, starting points of both sub-split and mid-split were used as migration path options. In addition, remaining states were selected based upon vendor guidance concerning likely future product offerings and timelines.

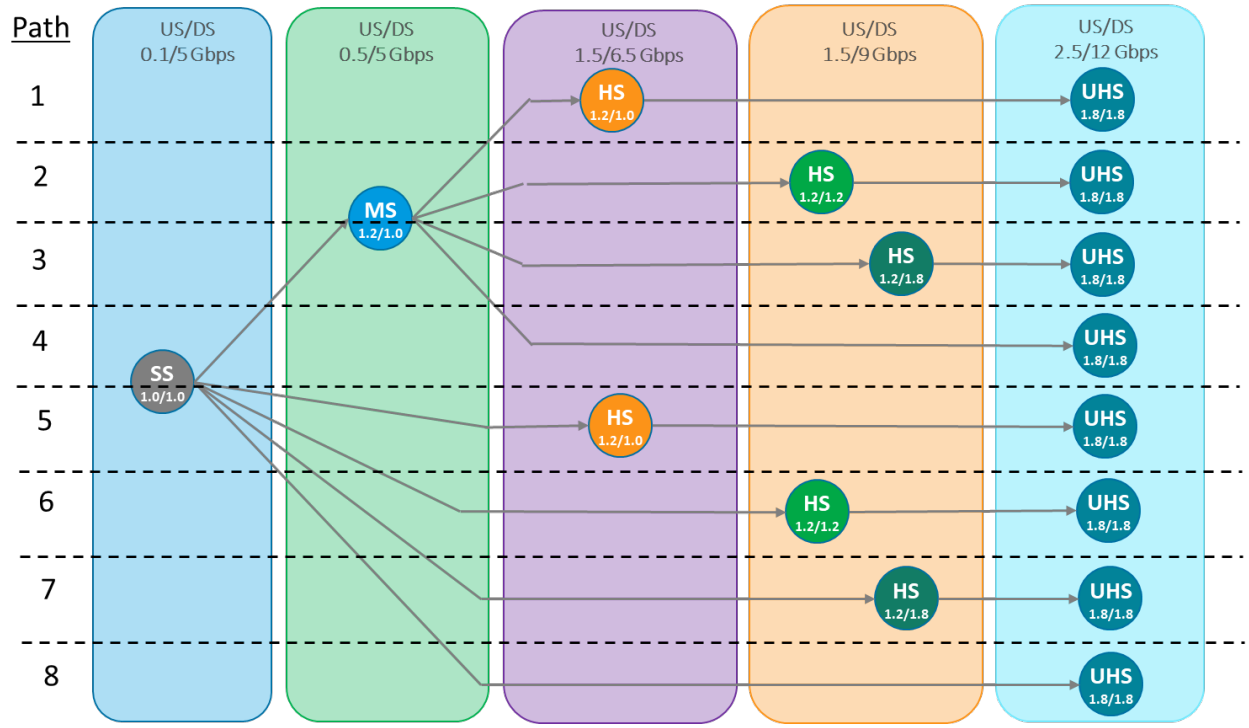


Figure 16. Cox Progression Path Options to Achieve 1.8 GHz UHS

Cox then developed an outside plant cost model for each of the transitions between states in Figure 16. It was then possible to sum those transitions to estimate total cost for each of the path options. (In order to remove proprietary cost data within this paper, all cost numbers were scaled relative to the cost of the transition of Sub-split 1.0/1.0 to Mid-split 1.2/1.0. The expectation is that other operators, equipped with their own unique costs for a mid-split upgrade, may then leverage the cost model results for their own unique circumstances.) Figure 17 provides the estimated total cost relationships for each of the 8 migration paths. In addition, the primary driver for high cost is provided for the more expensive paths.

As shown in Figure 17, paths 2 and 6 are significantly more expensive driven by the need to revisit all passive elements in the network twice for upgrades; first to 1.2 GHz and subsequently to 1.8 GHz. In addition, path 3 is quite expensive primarily because of the need to revisit the actives three times for bandwidth changes and twice for sweep and rebalancing to accomplish the upgrades; however, depending upon competitive circumstances, path 3 may still be necessary in some competitive market situations.

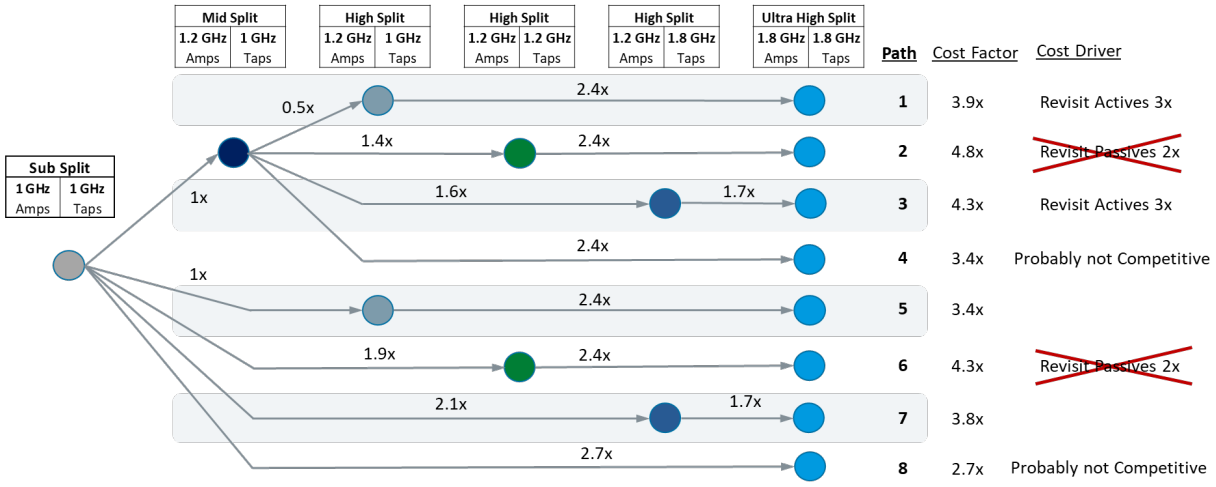


Figure 17. Estimated Cox Cost to reach DOCSIS 4.0 UHS for Eight Migration Paths

3.3. Product Availability and Path Implications

In addition to cost, commercial availability of products within the timeline demanded by competitive situations is also an important factor. As of the release of this paper, mid-split 1.2 GHz nodes, amplifiers, and passives are readily available. High-split 1.2 GHz nodes and amplifiers are expected to be readily available in 2022. However, as described above in sections 2.3.1 and 2.3.3, legacy video services and leakage monitoring need to be addressed in order to field a high-split network. Each operator must estimate their unique time implications to swap CPE to address service impacts; however, at a minimum, leakage aspects are likely to slow any large scale high-split deployments until at least late 2022 or early 2023 in order to deal with hardware upgrades of leakage meters and cumulative leakage index (CLI) reporting applications as well as CMTS modifications required in the enhanced DOCSIS specifications. 1.8 GHz nodes and amplifiers will likely not be available much earlier than late 2023 and more than likely middle of 2024 driving deployments into late 2024 to early 2025.

Relative to 1.8 GHz taps and passives, these are commercially available today with a cost premium over existing 1.2 GHz taps. However, this cost premium is more than offset by the high labor costs associated with revisiting passives twice in migration paths 2 and 6, driving a preference for a strategy where passives are only updated once in any timeline. In addition, based upon vendor feedback, 1.8 GHz passives are expected to require new housings. As such, there is some debate within Cox as to how disruptive a 1.8 GHz passive upgrade will be on network downtime, especially when considering the effects of power passing on other portions of the network. Given the raw number of passives within a node (Cox averages ~150 passives/node), it seems desirable for operators to begin passive upgrades significantly earlier than the rest of the plant upgrade; however, excessive network downtime might drive operators to a single cutover time for all actives and passives for a given node. Cox will need to pursue early trials using both a precursor and a concurrent upgrade process to better understand this impact and drive the scaled approach.

Migration paths 4 and 8, which basically capture jumping straight to a fully ultra-high-split 1.8 GHz configuration, are likely not competitively viable as it is quite likely that significant deployment penetration of symmetric 1 Gbps products beyond trials might be required earlier than 2025.

As a result, 4 recommended paths (1, 3, 5, and 7) (Figure 18) are left for consideration, costing between 3.4x and 4.3x the cost of a mid-split upgrade.[2] Cox is continuing to investigate the possibility of accelerating the replacement of legacy video CPE with an IPTV-based solution which may be required for paths 1 and 5, both of which provide substantial cost savings. Depending upon the competitive urgency to move to high-split, there may not be adequate time to replace CPE, and as a result, Cox would need to pursue paths 3 and 7. However, paths 3 and 7 drive the need for accelerating and targeting 1.8 GHz passive upgrades as activating a high-split for these paths requires these new taps and passives. Selection between the options 1 and 5 or 3 and 7 is driven by whether the node faces more immediate upstream congestion in which case a mid-split upgrade would be needed and paths 1 or 3 would be taken.

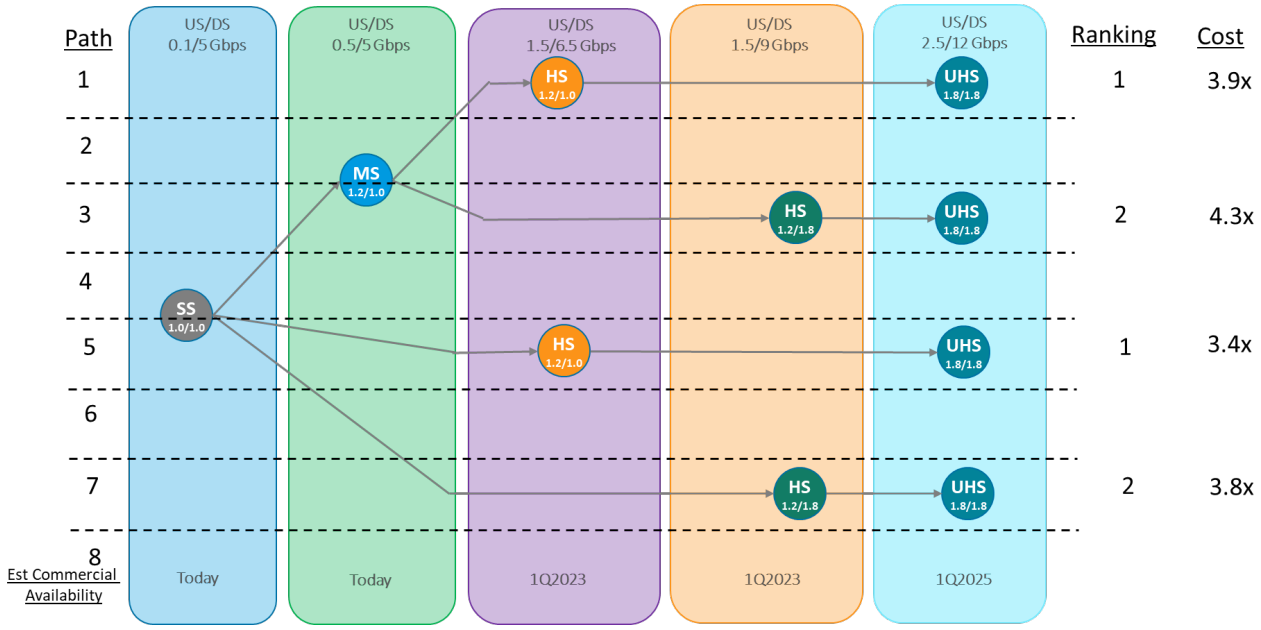


Figure 18. Recommended Migration Path Options to 1.8 GHz UHS

Figure 19 provides a possible timeline for paths 1 and 5. Key highlights of this process include the need to retire SDV/VOD or execute an IPTV upgrade to provide adequate downstream spectrum to achieve a high-split on a 1 GHz network as well as the need to upgrade homes with high-split service to a POE CPE device to eliminate in home interference with legacy devices. As mentioned earlier, there are some questions that will need to be resolved during trials as to the impact of earlier passive deployments on network downtime.

Figure 20 provides a possible timeline for paths 3 and 7. One key requirement in this process is the need to upgrade to 1.8 GHz passives much earlier in the process during the actual 1.2 GHz upgrade step; whereas that was a flexible option in Figure 19 assuming it didn't contribute to significant downtime.

[2] A key parameter in estimating the overall upgrade cost is cable replacement. The multipliers provided represent a 2% cable replacement assumption; however, Cox is currently performing tests in various markets to attempt to better refine that estimate. Any change in that percentage will impact the overall cost multiplier for each path but the relative differences will remain the same as the total cable replacement cost would be constant for all eight paths.

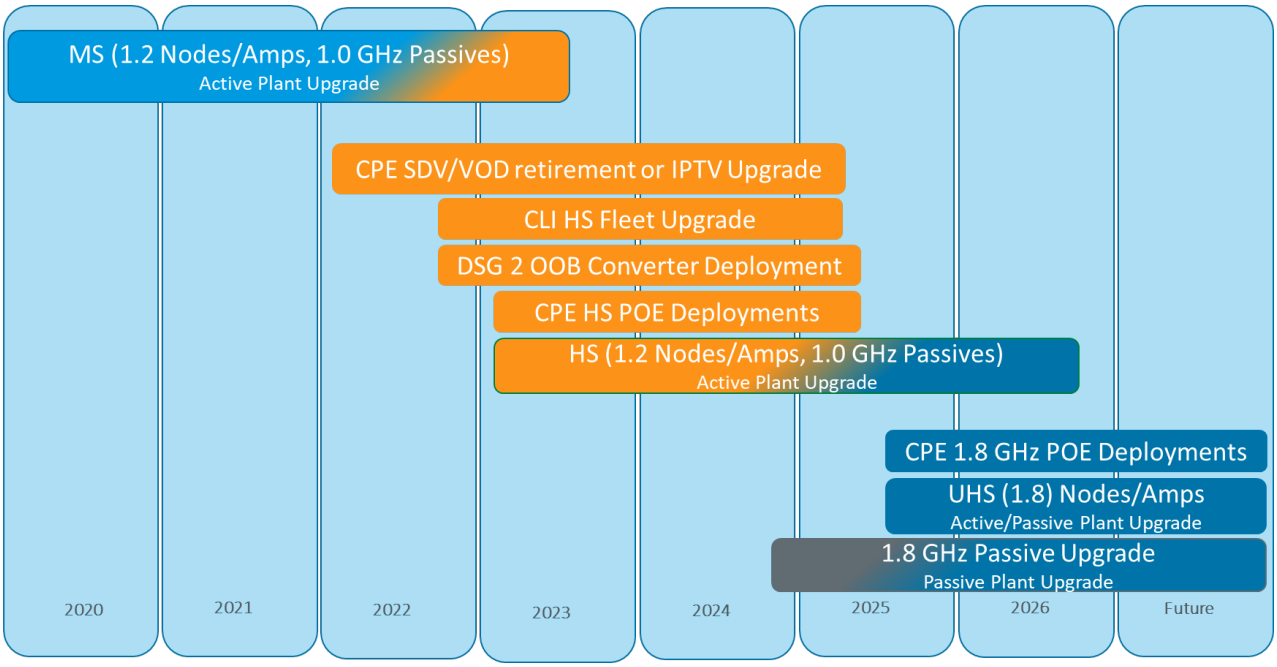


Figure 19. Possible Cox Migration Roadmap (Paths 1 and 5) with a 1.0 GHz Intermediate Spectrum Limit

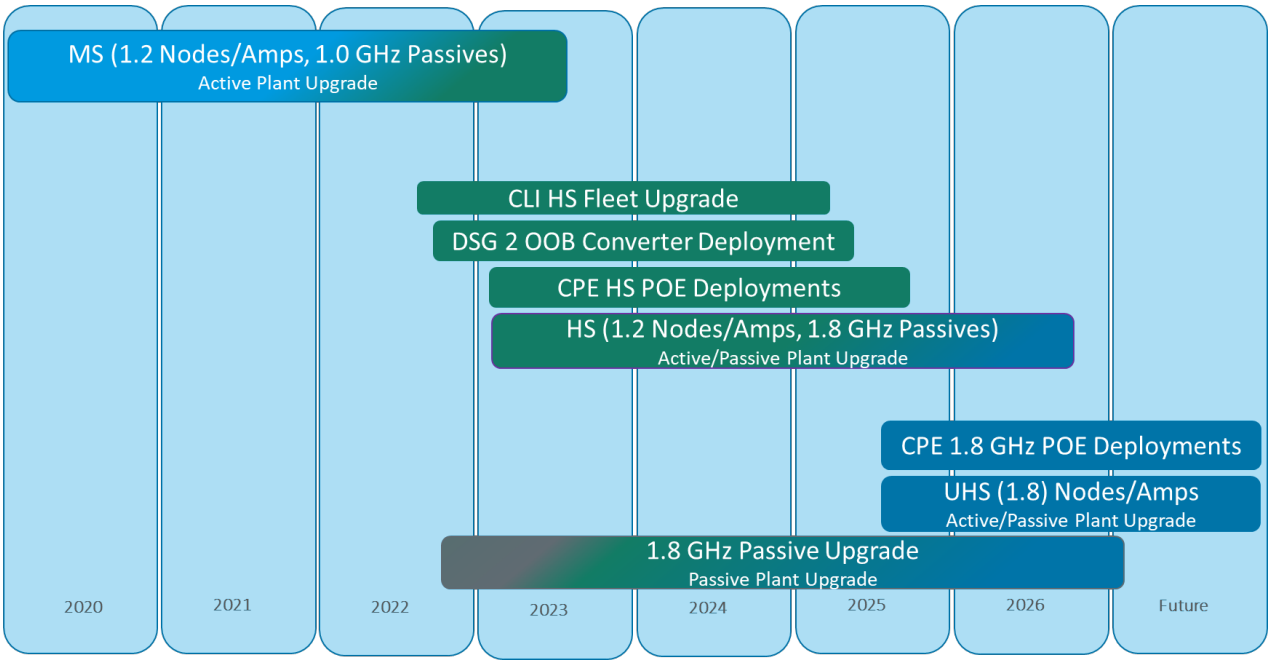


Figure 20. Possible Cox Migration Roadmap (Paths 3 and 7) with a 1.2 GHz Intermediate Spectrum Limit

4. Conclusion

In summary, there are many important aspects to consider in association with potential bandwidth upgrades made possible by FDD DOCSIS 4.0. In the paper we explored many of those aspects and:

- Pointed out the rationale for maintaining legacy band tap RF output levels during plant upgrades.
- Described the historically used drop-in approach to plant bandwidth upgrades, and its benefits.
- Explained likely 1.8 GHz amplifier downstream gain requirements for drop-in upgrades and why Booster Amps may be needed.
- Provided for consideration, proposals for amplifier RF output profiles with associated rationale.
- Described some of the major items to consider associated with spectrum changes, with potential methods for handling them, including:
    - Reducing QAM Video
    - Meeting CableCard Regulations and Addressing Legacy OOB (55-1, 55-2)
    - Requiring POE for HS and UHS homes
    - Retiring MoCA
    - Meeting FCC Leakage Regulations
- Pointed out the potential requirements for a downstream spectrum “squeeze” associated with the plant upgrades and offered alternative mitigation approaches to consider.
- Shown a variety of potential approaches making use of progressive steps leading up to ultra-high-split 1.8 GHz and presented associated rationales and cost factors.

Abbreviations

AGC	automatic gain control
Amp(s)	amplifier(s)
BAU	business as usual
BLE	Motorola line extender
bps	bits per second
BT	Broadband Telecommunications
CLI	cumulative leakage index
CMTS	cable modem termination system
CNR	carrier to noise ratio
CPE	customer premises equipment
DAA	distributed access architecture
dB	decibels
DOCSIS	data over cable system interface specification
DS	downstream
DSG	DOCSIS set top gateway
ES	extended spectrum
FCC	Federal Communications Commission
FDD	frequency division duplex
FEC	forward error correction
FM	CCOR FlexMax line extender
FMB	CCOR FlexMax bridger
FMT	CCOR FlexMax trunk/bridger
FTTH	fiber to the home
Gbps	Gigabits per second
GHz	Gigahertz
HFC	hybrid fiber coax
HGBT	high gain balanced triple
HGD	high gain dual
HHP	households passed
HS	high-split (204/258 MHz split)
Hz	Hertz
IPTV	internet protocol television
ISBE	International Society of Broadband Experts
LE	line extender
Man	manual
MB	minibridger
Mfr	manufacturer
MHz	Megahertz
MoCA	Multimedia over Coax Alliance
Moto	Motorola
MS	Mid-split (85/108 MHz slit)
N+0	node plus 0 amplifier architecture
N+X	node plus X amplifier architecture
NCTA	National Cable Television Association

OFDM	orthogonal frequency-division multiplexing
OFDMA	orthogonal frequency-division multiple access
OOB	out of band
OSP	outside plant
OUDP	OFDM Upstream Data Profile
POE	point of entry
RF	radio frequency
RPD	remote phy (physical layer) device
RMD	remote mac device
SA	Scientific Atlanta
SCTE	Society of Cable Telecommunications Engineers
SLA	service level agreements
SS	sub-split (42/54 MHz or 40/52 MHz splits)
STP	set top gateway
TiVo	television input video output
TCP	total composite power
Therm	thermal
UBT	Scientific Atlanta unbalanced triple
UDCP	Uni-directional cable product
UHS	ultra-high-split (Cox tentative target 396/500 MHz split)
US	upstream
WiFi	wireless fidelity

Bibliography & References

CM-SP-PHYv4.0 – DOCSIS 4.0 Physical Layer Specification, Version I03, 12/02/2020

SCTE Cable-Tec Expo 2021 – "Leakage Detection in a High-Split World: Industry Progress Toward a Viable Solution", Rex Coldren, et al.

CM-SP-MULPIv3.1 - DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification, Version I21, 10/20/2020

MULPIv3.1-N-21.2165-2 – DOCSIS 3.1 MULPI ECN, 5/27/2021

CM-SP-CCAP-OSSIv3.1 - CCAP Operations Support System Interface Specification, Version I20, 4/19/2021

CCAP-OSSIv3.1-N-21.2179-3 – DOCSIS 3.1 OSSI ECN, 7/1/2021

CM-SP-MULPIv4.0 - DOCSIS 4.0 MAC and Upper Layer Protocols Interface Specification, Version I03, 12/02/2020

MULPIv4.0-N-21.2166-2 – MULPI ECN, 5/27/2021

CM-SP-CCAP-OSSIv4.0 - CCAP Operations Support System Interface Specification, Version I04, 5/21/2021

CCAP-OSSIv4.0-N-21.2168-1- DOCSIS 4.0 OSSI ECN, 6/10/2021

The Scheduler and the Tap: The Odd Infrastructure Couple

A 100 Gbps Coaxial Future Story

A Technical Paper prepared for SCTE by

L. Alberto Campos
Fellow
CableLabs
858 Coal Creek Circle, Louisville CO, 80027
303 661 3377
a.campos@cablelabs.com

Lin Cheng
Lead Architect
CableLabs
858 Coal Creek Circle, Louisville CO, 80027
303 661 3335
l.cheng@cablelabs.com

Zhensheng (Steve) Jia, CableLabs

Jing Wang, CableLabs

Chris Stengrim, CableLabs

1. Introduction

The death or end of coaxial cable transport has been predicted in the past on more than one occasion to give way to fiber-to-the-home. The resiliency of coaxial transport, however, has proven to be quite enduring. Coaxial cable itself has not yet been used to its full potential and operators have demonstrated that with the always improving operational practices in addition to the robustness and flexibility provided by the evolving transport technologies, there are still effective means to get value out of our coaxial infrastructure investment. A key factor in this equation is that under several operator starting point scenarios, fiber-to-the-home (FTTH) deployment requires significant investment and coaxial based evolution alternatives can still address most subscriber requirements.

As it appears that coaxial based transport in the access will still be with us for quite some time, it behooves us to think how coaxial transport would look like 10+ years from now. This paper explores how a coaxial based future could evolve in the long run, what would be the implications on topology, spectrum, and technologies that we could see taking shape under the pressures of deployment costs, service offerings, operational efficiency, reliability and market competition.

Today demand for capacity and consumption patterns from subscribers is quite varied. Unfortunately, in deployments, where serving area sizes until recently consisted of about 400+ households per fiber node, there has been little flexibility to design based on the traffic characteristics that we experience. Instead to maintain acceptable levels of customer experience, we have been designing these serving groups driven by peak user speeds and competitive forces. The diversity of our subscribers' usage patterns indicates that peak-capacity based design can leave a good amount of resources and investment on the table. Service elasticity and resulting in network elasticity to flexibly support our subscriber diversity becomes more relevant in our environment that not only caters to residential subscribers but also small businesses and wireless connectivity sharing the same coaxial infrastructure.

1.1. Capacity Improvement Mechanisms

Our industry has leveraged 3 ways of improving HFC capacity. First is by improving efficiency, second through segmentation and third by increasing the amount of spectrum we use. We have been using all three tools available at our disposal as we have evolved our cable networks.

In cable, we have used several efficiency improvements tools. We have improved the efficiency of how we carry video leveraging more efficient compression techniques. We have introduced profiles in our DOCSIS 3.1 systems so that if a CM is closest to an amplifier or a fiber node, this CM could transmit at greater speeds/efficiencies and not at the lowest common denominator as was the case in earlier implementations. Cable has done a very good job at cleaning the HFC plant, improving channel conditions so that we could carry higher order modulations, reaching up to 4096 QAM in the upstream and up to 16384-QAM in the downstream. Distributed architectures have helped in that they eliminated noise and distortion contributions of the analog optical link, thereby facilitating higher modulation transport. Figure 1 shows a 16384 QAM DOCSIS constellation example.

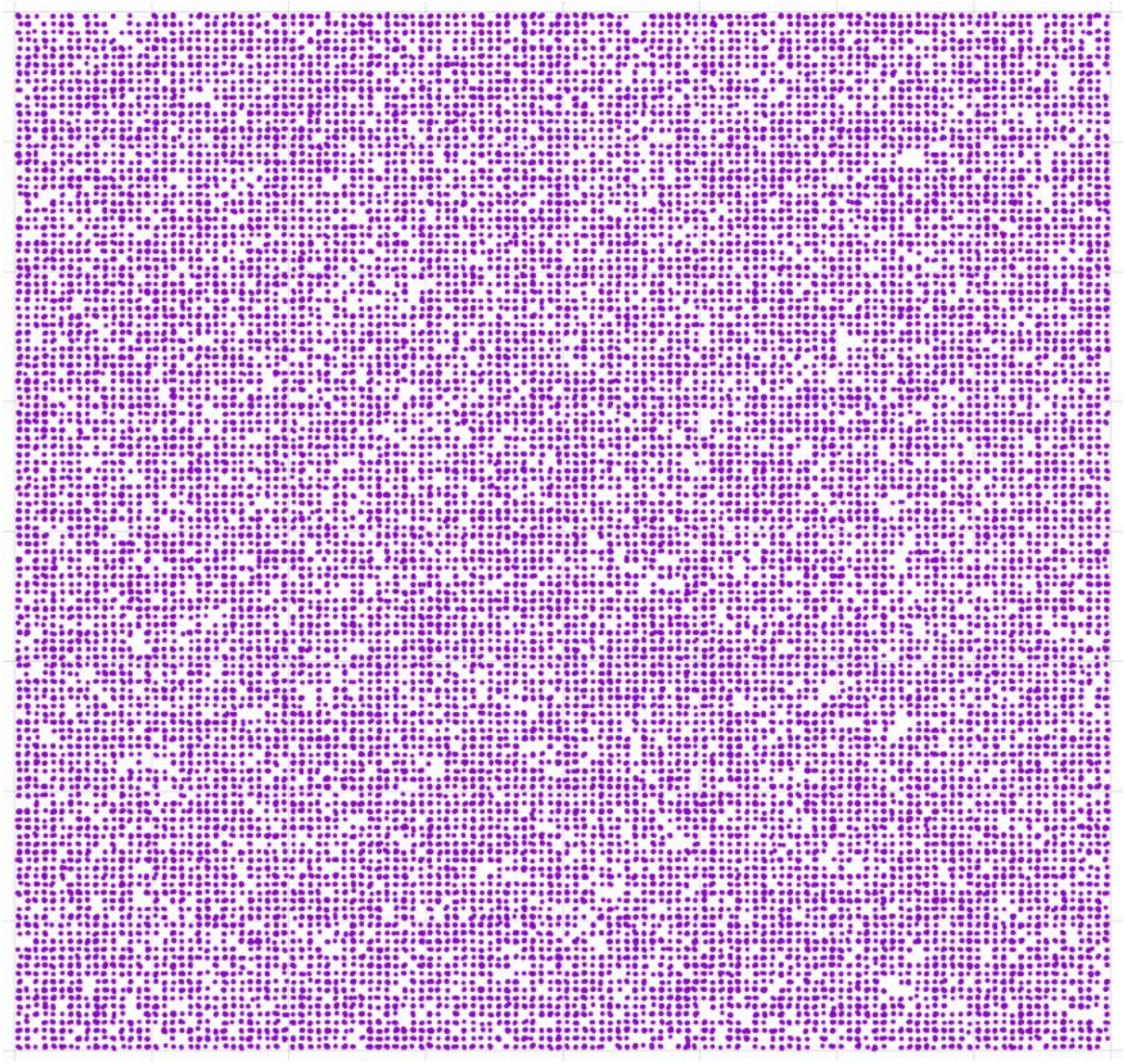


Figure 1 - Highest Efficiency 16384-QAM DOCSIS Downstream Constellation

We have an RF domain peak capacity limitation imposed by the amount of spectrum available. Still by dividing this RF domain into smaller serving areas or segments the amount of capacity per user can be increased. Segmentation in cable has been used on an "as-needed-basis" for quite some time, relying on node splitting to address capacity shortages. The growth in demand for capacity has reached a point in which a surgical approach to segmentation may not always be sufficient as this increase in capacity is more widespread and the rapid growth in demand may require longer term solutions. These solutions consisted of an HFC architecture migration from the original Node+4/Node+5 500 HHP architectures to the smaller Node+2 to Node+0 architectures. In addition to increases in aggregate capacity, they also bring improved performance and higher reliability. Figure 2 shows a Node +0 network migration from a legacy 500 HHP node.

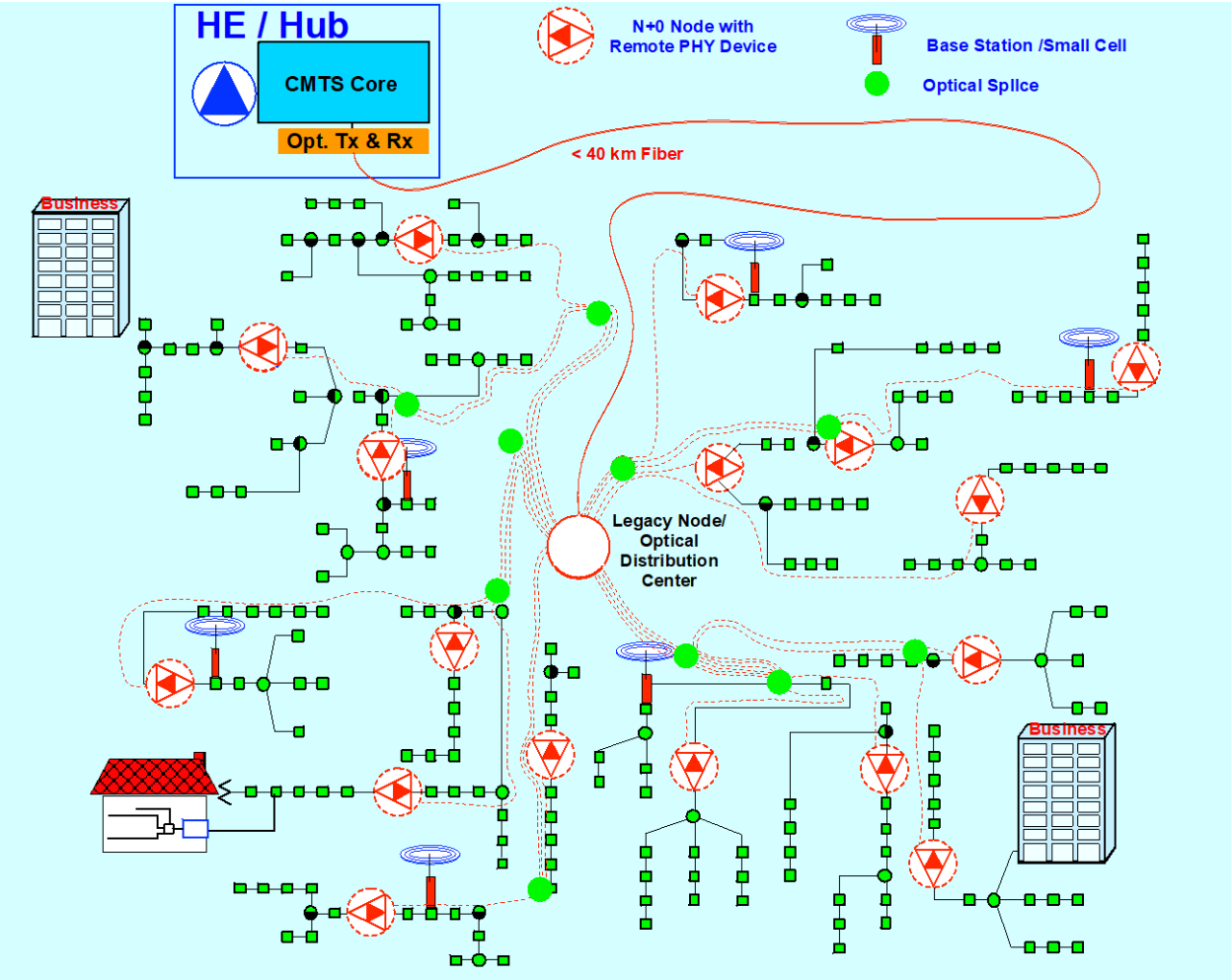


Figure 2 - Legacy/Original Node Segmented Into Node+0 Child Nodes

The changes that have been taking place regarding segmentation and efficiency improvement are quite transformational. The area that perhaps has shown a steady improvement is spectrum enhancement. In the past, it has been somewhat limited by regulatory forces but also by demand, operational complexity and ultimately by cost.

We mentioned earlier how spectrum could enable peak service capacity. In the past peak capacity has been increased by aggregating/bonding more channels within the overall available spectrum. Now however, we have reached the point that the maximum number of channels used by CMs is quickly approaching the maximum amount of spectrum we have available. Therefore, in order to increase service tiers, we need to increase our coaxial RF spectrum. DOCSIS 4.0 addresses this need but at a service tier CAGR of 25% to 40%, we need to start thinking about longer term options. We will discuss the paths we could take to address this demand.

2. Ultimate Coaxial Spectrum Resources

In the early days of cable, only video was carried over the network, video content was limited and the demand to support higher frequencies was also limited. Support to carry the highest practical frequency was also impacted by the attenuation of coaxial cable, the need for multiple amplification stages and the noise and distortion amplification introduced. In this early all-coaxial transport, cable networks evolved from maximum frequencies of 250 MHz to 350 MHz, 450 MHz up to 550 MHz. It was in the 1990's when the transformation of cable networks into a Hybrid Fiber Coaxial (HFC) architecture took place. At that time the HFC network typically consisted of fiber serving areas of 500 households passed (HHP). Instead of having up to 30 amplifiers in cascade, many fiber nodes serving areas were upgraded to 4 to 5 amplifiers in cascade before reaching the furthest subscriber. This network transformation also included the transport of digital video and bidirectional transport to carry data and video at a highest downstream frequency of 750 MHz. With IP-data becoming the dominant use of cable networks, DOCSIS system capabilities evolved along with plant frequency upgrades. DOCSIS versions included 860 MHz, 1002 MHz, 1.2 GHz and with DOCSIS 3.1 and DOCSIS 4.0 specifications 1.794 GHz became the next high frequency target.

The questions that we want to ask ourselves are: How high in frequency can we leverage our coaxial infrastructure? What are the challenges that we need to consider and are there potential approaches to address these challenges?

Coaxial transport relies primarily on the transverse electromagnetic (TEM) mode of propagation. In coaxial cable, TEM mode is radially symmetric and propagates along the direction of the center conductor (Figure 3).

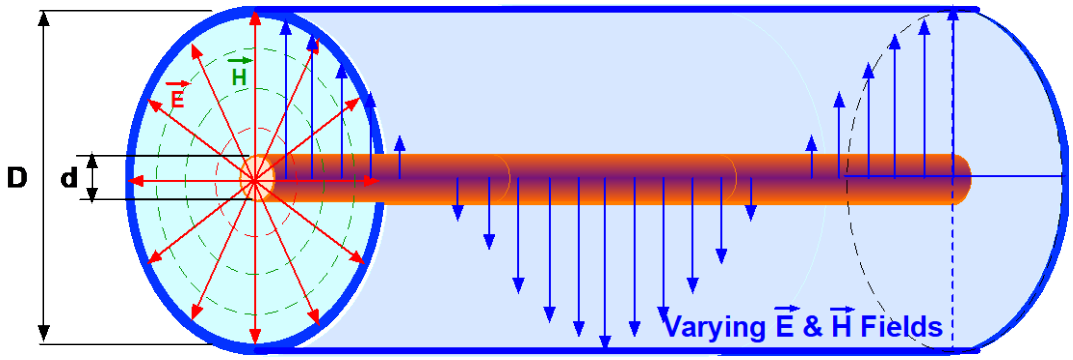


Figure 3 - TEM Propagation Mode in Coaxial Cable

As the frequency increases and the wavelength approaches the dimensions of the radius of the coaxial cable, other modes of propagation are excited. The transverse electric TE_{11} mode is the first to appear. When these modes are allowed to propagate with the TEM mode, they interfere with each other. The frequency at which TE_{11} appears is called the cut-off frequency (f_c) which for TE_{11} mode is given by:

$$f_c = \frac{c}{\lambda_c} = \frac{c}{\pi\left(\frac{D+d}{2}\right)\sqrt{\mu_R \varepsilon_R}}$$

where c is speed of light, D is the inner diameter of the outer conductor, d is the outer diameter of the center conductor and μ_R and ε_R respectively are the relative permeability and relative permittivity of the

dielectric material. Table 1 shows the estimated cut-off frequency for different cable types we use in our industry.

Table 1 – TE11 Coaxial Cable Cut-Off Frequencies

Cable Type	Cut-Off Freq
RG6	29.1 GHz
RG11	18.6 GHz
0.5"	11.5 GHz
0.625"	9.3 GHz
0.75"	7.7 GHz
0.875"	6.6 GHz

As you can see with 0.625” and 0.5” hardline cables which are common in the distribution portion of the network, have cut-off frequencies of 9.3 GHz and 11.5 GHz providing significant opportunities for higher bandwidths. This opens up attractive potential resources we could leverage. Nevertheless attenuation, implementation, operational costs, service tier and aggregate capacity are also factors that need to be considered.

3. Legacy of Analog Video Transport

Transport of analog video was required to transmit at a very a high carrier-to-noise ratio. The downstream plant required very clean maintenance and other services had to coexist well with analog video. The channel spacing used by analog video was adopted by newer data services, and out-of-band emissions of the new tenants of cable’s coaxial spectrum were also tightly controlled. There have been numerous decisions that have been made in cable based on analog video transport, not only in the design of services that share the same spectrum as analog video, but also decisions in the transport infrastructure itself so that it can optimally support analog video.

The original cable industry plant was designed around one type of service, mainly the delivery of broadcast analog video services. A Cable Television (CATV) service provider frequency-multiplexed a lineup of analog video channels from a central location such as a hub or headend. It transmitted the video signals to subscribers connecting to the coax network within a fiber node serving area. In order to have suitable reception of analog video, each home had to ideally receive the video channel signal at about the same target power level. Cable accomplished this with an RF distribution network where taps coupled RF energy out of the hardline into drop ports to connect to the subscribers’ homes via drop cables. Each successive tap following a fiber node or an amplifier, has a specific coupling loss to the drop port so that even after the attenuation of coaxial cable, the power reaching the end-device, such as a set-top-box, is about the same for all subscribers. Figure 4 shows a schematic representation of a coaxial segment with taps of decreasing values.

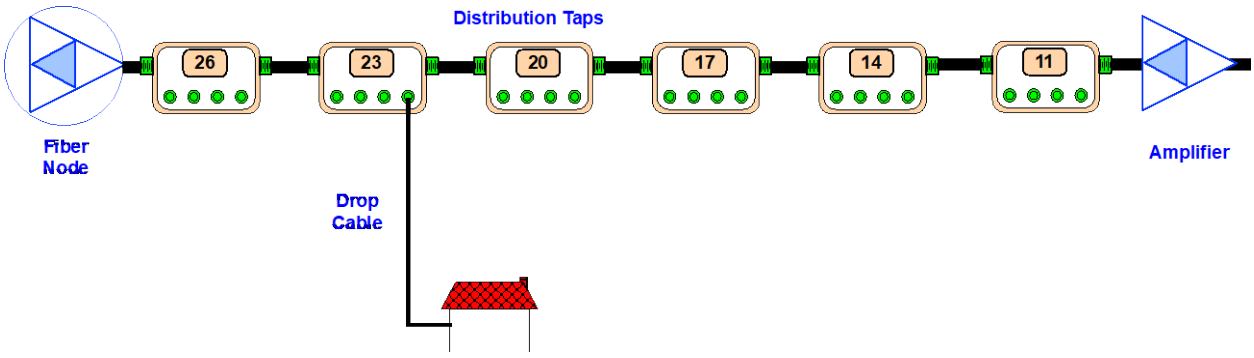


Figure 4 - Fiber Node-to-Amplifier Coaxial Segment With Tap Values To Provide Similar Power Levels Per Video Channel

Over the relatively narrow band that the upstream occupies, differentiation in channel conditions can depend on the amplifier cascade value and specific impairments which predominantly impact one or more cable modems. There is also some frequency dependence when comparing the edges of the upstream band with the middle of the upstream band. The downstream also imposes larger frequency-dependent behavior due to cable attenuation, however this behavior is typically ameliorated by an up-tilted downstream signal and by the decreasing tap values which effectively equalizes the power levels for all customers along the coaxial segment.

4. Evolving HFC Environment

While early in Cable, the focus was on having every analog video channel be received by every Set Top Box (STB) and TV receiver at approximately the same power level, an allowance to deviate from that philosophy of operation has been enabled with the introduction of channel conditions' dependent profiles in the DOCSIS 3.1 specification. Another trend worth noting is that more traffic in cable networks has steadily moved from broadcast to unicast. This has been due to the way services are delivered such as video on demand as well as IP video services. Therefore, the need of having one stream having to be received by all receivers is disappearing.

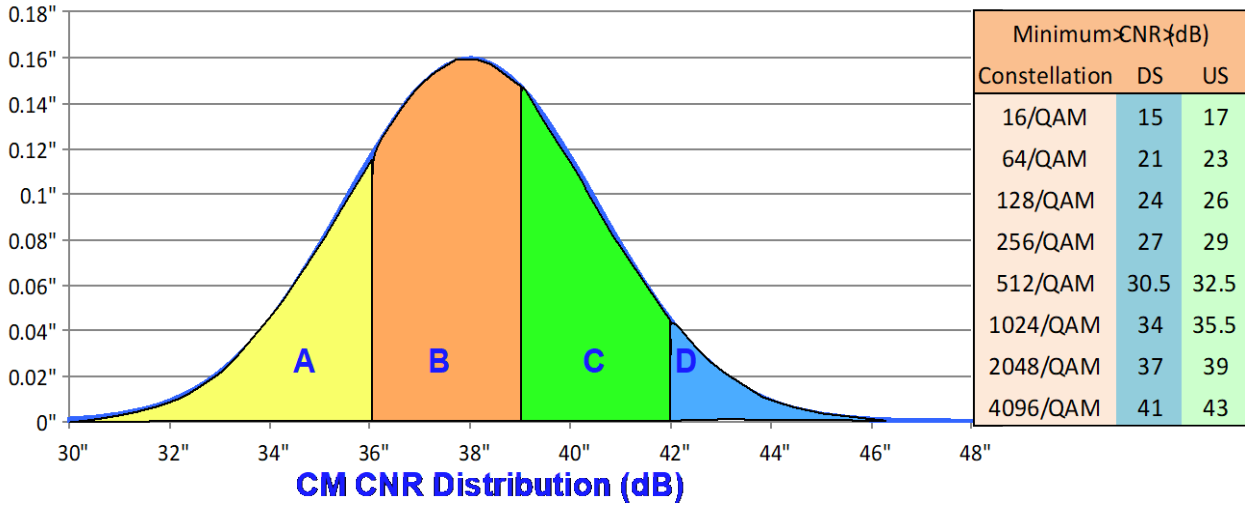
About a decade ago, the transition to digital video broadcast started in the United States which, except for low power TV broadcast stations, has now been completed. This transition also impacted "must carry" rules. When analog is no longer transmitted, there is no longer the requirement to carry broadcast channels at the over-the-air frequencies. For example, Channel 2 does not need to be carried at 50 MHz in our cable networks when only carrying digital video channels. For all practical purposes, we live today in an all-digital world. Not only video but also voice and data are carried over bit streams. We are no longer required to optimize our networks to distribute video signals so that their service endpoints are at about the same power levels. This is a brand-new ball game; the change in conditions allows us to break free from the restrictions of analog video distribution and everything it entails. The consequences of analog video distribution are not just limited to reclaiming coaxial RF spectrum by replacing analog video with the more efficient digital video. The change in conditions is significantly more encompassing—. Think about revisiting all the network design decisions that have been made since the early days of community antenna television.

4.1. A Step Beyond DOCSIS Profiles

As we move to an all-digital transport, our criterion needs to shift from the optimal transport of analog video to the optimal transport of bits. No longer are we required to deliver every analog channel at about

the same power level. Now we need to focus on how to transport the highest number of bits across our entire spectrum.

When we introduced the DOCSIS 3.1 version of the specification, instead of treating all CMs the same, the concept of transmission profiles was incorporated. This meant that CMs that could perform at a higher efficiency level due to a more benign channel would be placed in a profile where higher order modulation would be used. Figure 5 highlights the variation in CNR of a population of CMs and an example of how they could be grouped in different efficiency buckets.



Minimum CNR (dB)		
Constellation	DS	US
16/QAM	15	17
64/QAM	21	23
128/QAM	24	26
256/QAM	27	29
512/QAM	30.5	32.5
1024/QAM	34	35.5
2048/QAM	37	39
4096/QAM	41	43

Figure 5 - CM CNR Distribution and Modulation Order Thresholds

The assignment of profiles is adaptable to the current conditions of a CM. Today, a CMTS is specified to support up to 16 downstream profiles and 7 upstream profiles. These profiles are defined within the downstream channel with subcarrier granularity and in the upstream with minislot granularity since the modulation order is defined on a per minislot basis. CMs could go to a higher or lower efficiency profile depending on the channel conditions and leveraging the ranging response mechanisms and MER and codeword error metrics.

4.2. Higher Frequency Off First (HFOF)

In a DOCSIS OFDM downstream, resources exist in symbols (time) and subcarriers (frequency). In the upstream, minislots consisting of symbols and subcarriers are also allocated in time and frequency. As we use higher frequencies, the variation of CNR versus frequency is more evident. The attenuation in hardline and drop cables versus frequency becomes the dominant factor determining the MER and CNR in a CM. Figure 6 shows coaxial attenuation versus frequency and Figure 7 highlights how such losses may impact CNR at higher frequencies on a CM attached to a coaxial segment.

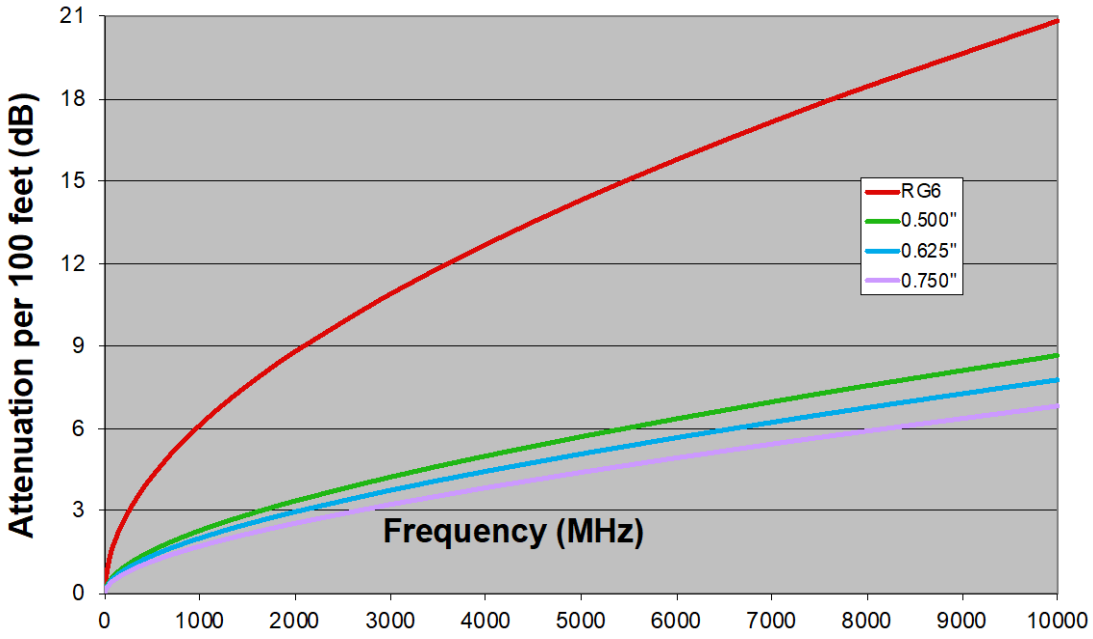


Figure 6 - Hardline And Drop Cable Attenuation Versus Frequency

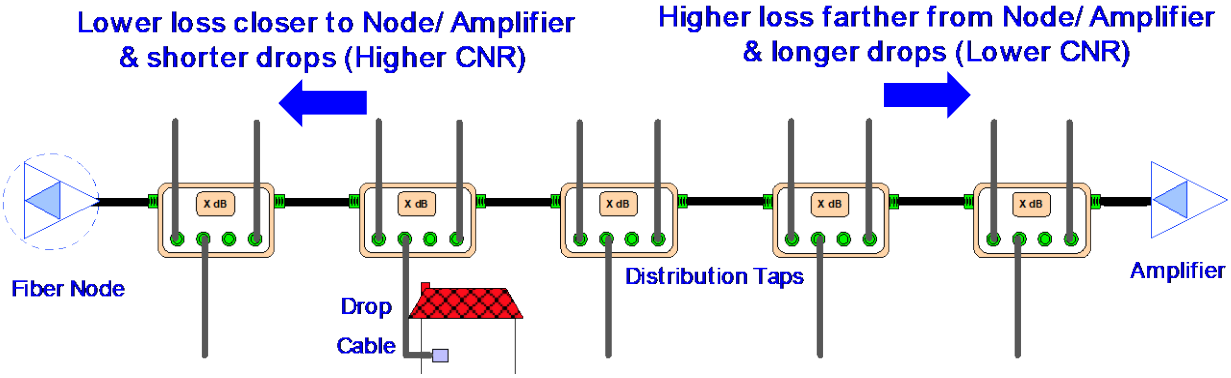


Figure 7 - CM Higher-Frequency CNR Based on Topology Location and Drop Length

Figure 7 shows that, at higher frequencies, CMs that are closer to the fiber node and with shorter drops will enjoy better CNR while CMs that are further out from the node and with longer drops cables will have lower CNR. This effect is due in large part to the fact that Total Composite Power for the transmitted signal from any Node/Amplifier is limited to "reasonable" levels on the order of ~70 or so dBmV, so the received power at higher frequencies is reduced by the combination of limited transmit power at the higher frequencies and increased attenuation at the higher frequencies. This qualitative assessment is quantified using the topology example of Figure 8. This coaxial segment consists of 4 taps in cascade and 12 CMs attached to these taps through drop cables. The drop lengths have been selected on purpose with great diversity of lengths so that we can better observe the significant effect that drop cable attenuation can have on performance. In this analysis, a hardline feeder cable of 0.5" and an available bandwidth of 11 GHz are assumed.

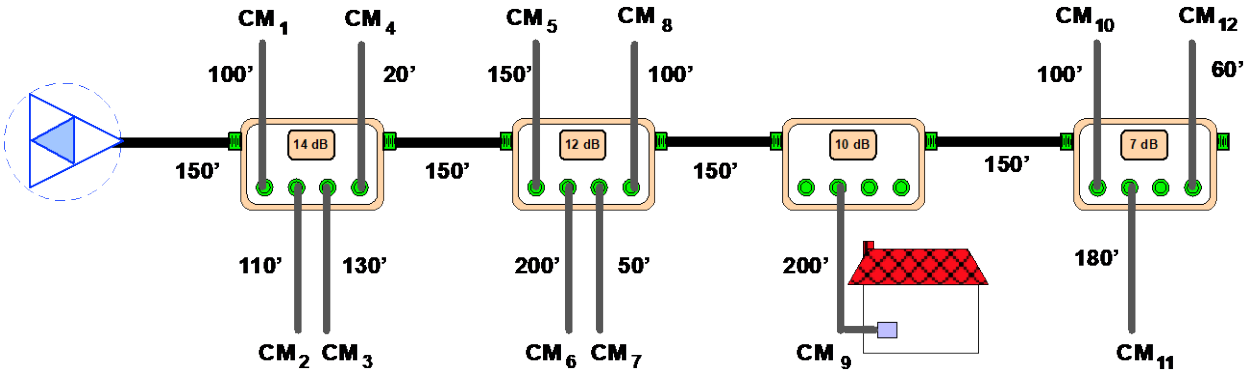


Figure 8 - Sample Coaxial Segment For Ultimate Capacity Estimation

Table 2 shows the total capacity that the CMs within the sample coaxial topology in Figure 8 could have if each CM has access to the entire 11 GHz spectrum.

Table 2 – Full Bandwidth Capacity Of CM Within Sample Topology

Modem	Full Bandwidth (Gbps)
CM_1	134.3592
CM_2	129.5459
CM_3	118.1095
CM_4	146.1113
CM_5	105.6579
CM_6	75.2112
CM_7	145.609
CM_8	134.6757
CM_9	74.7192
CM_{10}	134.053
CM_{11}	85.2012
CM_{12}	144.8434

As you can see the capacity that each CM could obtain varies widely from a lower 74.7 Gbps value to a higher 146.1 Gbps rate. In general, CMs that are close to the Node/Amp and CMs with short drop lengths tend to experience much higher capacities than the others. The average aggregate capacity is 112.14 Gbps which results in an average capacity per CM of 9.345 Gbps if it would be equally shared. If we assign spectral efficiency according to the downstream CNR table in Figure 5 and assuming a 3 dB receiver noise figure, the resulting efficiency versus frequency for the different CMs within the sample topology is shown in Figure 9.

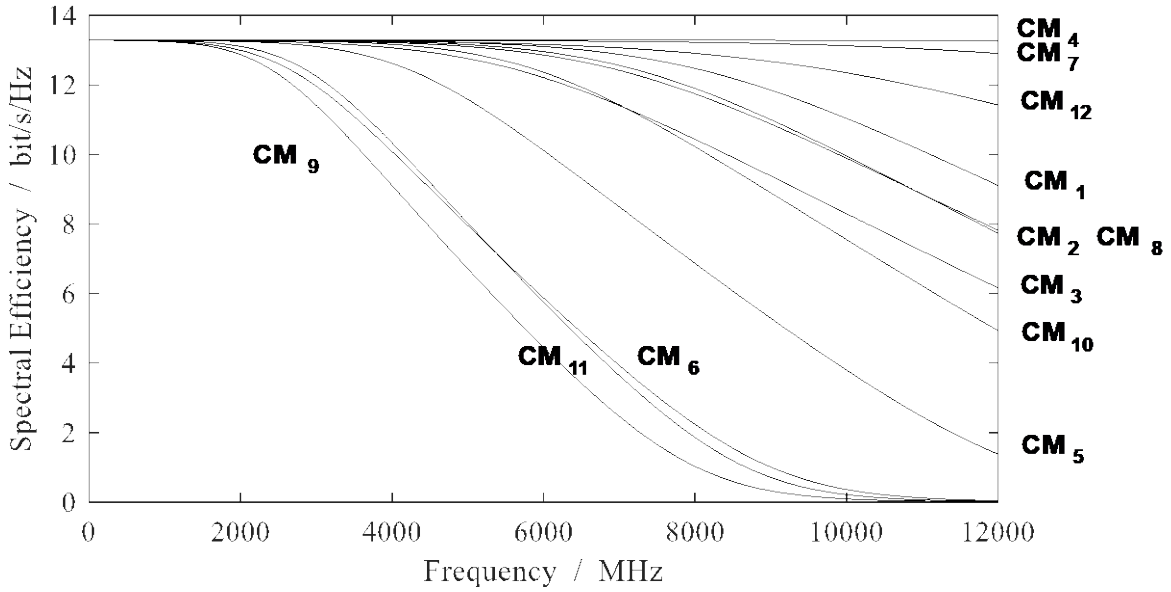


Figure 9 - Spectral Efficiency Versus Frequency Of CMs Within Sample Topology

As observed in Figure 9, CM_4 with a short drop and closest to the fiber node enjoys the best performance while CM_9, which has the longest drop and is located after 3 hardline segments, has the most limiting performance. The performance of CM_{12}, that is located in the last tap but has a short drop highlights the impact that drop cable can have as well as the fact that there is plenty of capacity left for the amplifier that sits behind the cascaded hardline segment.

If one would have the capability of flexibly assigning capacity on a frequency basis to the different CMs that enjoy different channel performance the overall aggregate capacity could be optimized. This mechanism is what we call "higher frequency off first" (HFOF) mechanism, which assigns the higher frequencies to the CMs that enjoy best higher-frequency CNR performance and leaves the lower frequencies for the CMs that have limited higher-frequency CNR performance. Figure 10 shows the allocation of capacity according to frequency bands. The capacity allocated to each CM is the same so there is some variation in the bandwidth allocated to each CM. An overall aggregate capacity of 142.81 Gbps is obtained which represents an improvement of 27% compared to the traditional approach calculated by averaging values in Table 2. (Note: If a particular CM is not utilizing its assigned spectrum, then the CMTS scheduler would be able to re-assign that spectrum to one or more other CMs. This intelligent scheduling would likely be typical in real-world scenarios).

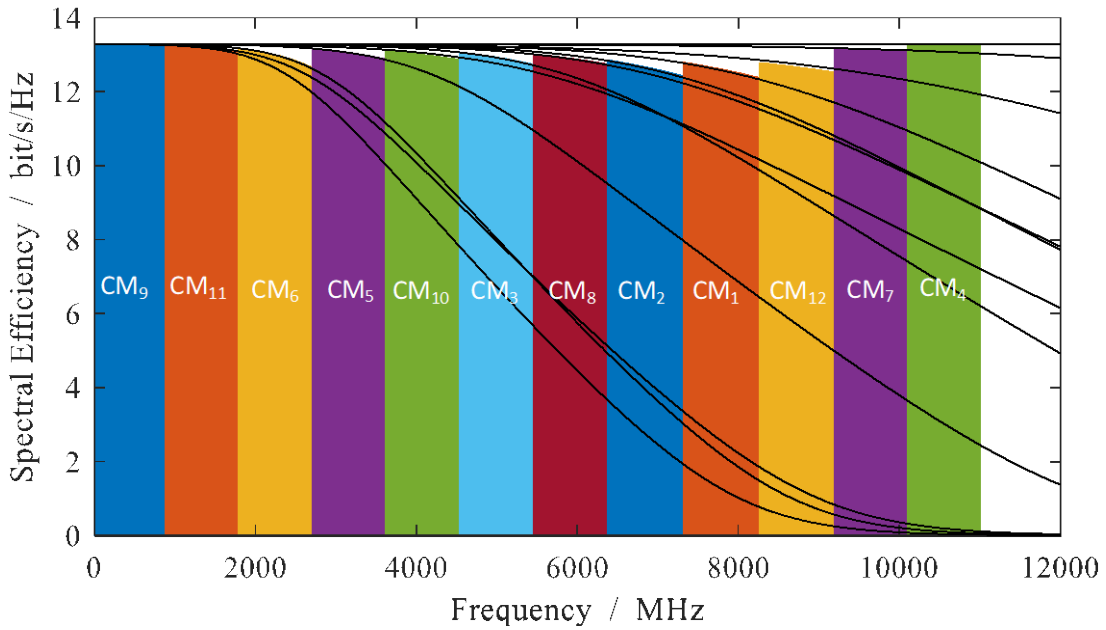


Figure 10 - CM Capacity Allocation Following HFOF Approach

A second simulation was conducted using 200' coaxial segments for a total of 800' total segment length. In that scenario using the traditional approach an average aggregate capacity of 105.888 Gbps was achieved or an average capacity per user of 8.824 Gbps. When leveraging the HFOF approach an aggregate capacity of 139.992 Gbps or 11.666 Gbps per user was obtained. This represents a 32% improvement of HFOF over the traditional approach.

4.3. Dynamic Range

The attenuation of hardline and drop cable versus frequency as shown in Figure 6 can be significant. In Figure 11, we combine the frequency response of all CMs along the coaxial segment. As you can see the loss across the entire 11 GHz bandwidth can be significant, but the loss across the portion of the spectrum allocated to each CM according to HFOF is bounded, resulting in a significant relaxation of dynamic range requirements. The blue and red segments on the CM loss curves stay within a limited loss range, highlighting the dynamic range benefits of the HFOF approach.

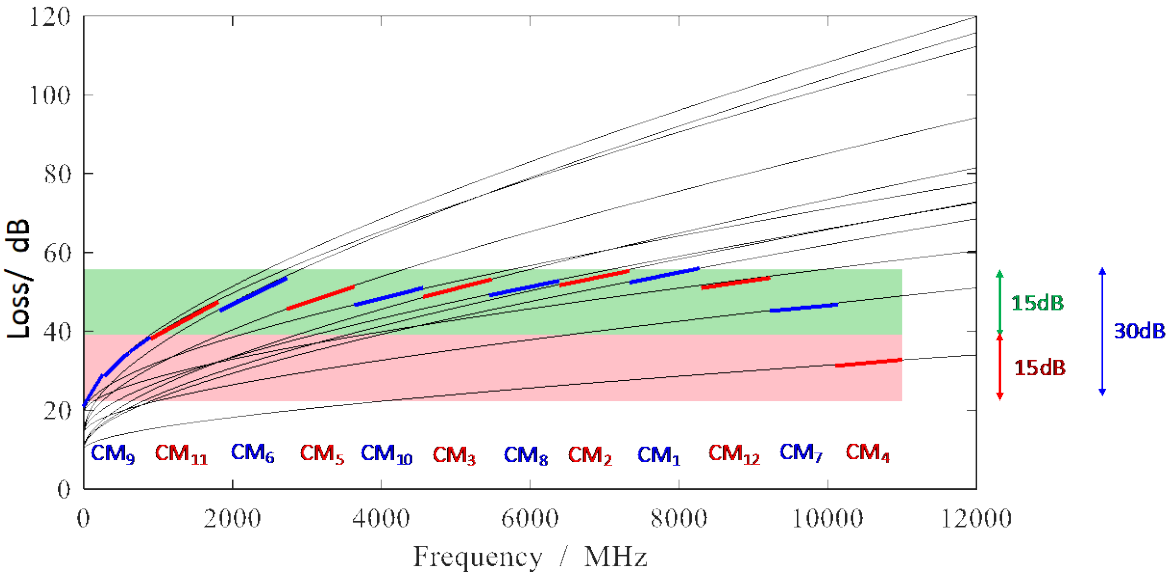


Figure 11 - End-to-end CM Attenuation Within Allocated Frequency Band

4.4. Implementation Implications Of Peak And Aggregate Rates

In the earlier sections, we have seen how a large amount of coaxial spectrum can be made accessible to CMs. In this section, we explore techniques in making this accessibility cost effective. In WiFi and mobile applications, we have systems with limited amounts of bandwidth available out of a diverse selection of spectrum bands. The accessibility to the many options of spectrum bands is achieved through the tuning capabilities of the receiver. How much bandwidth can a receiver simultaneously capture, process and aggregate, is an indication of the peak capacity a handset could reach.

A similar approach could be followed in cable where CMs could have accessibility to a wide spectrum while the bandwidth capture capabilities would indicate the potential peak bandwidth a CM could reach. Figure 12 shows an example of the cable analogy where 10.8 GHz of coaxial spectrum is available. This amount of spectrum is consistent with the cut-off frequency of 0.5" hardline cable. In this example a CM capture bandwidth of 1.8 GHz is assumed. The total amount of spectrum available for the downstream in this example is 10.2 GHz. Leveraging HFOF techniques and assuming a clean plant, a modulation order of 2048 QAM can be reached which leads to a 17 Gbps capacity per 1.8 GHz capture bandwidth and an aggregate capacity out of the entire 10.2 GHz of spectrum of approximately 100 Gbps assuming DOCSIS 3.1 level overhead. The CM capture range could be adjusted based on the target peak rates and implementation cost complexity criteria. Figure 12 depicts the scenario just described where some CMs, depending on where they are within the coaxial segment topology, are assigned certain frequency bands.

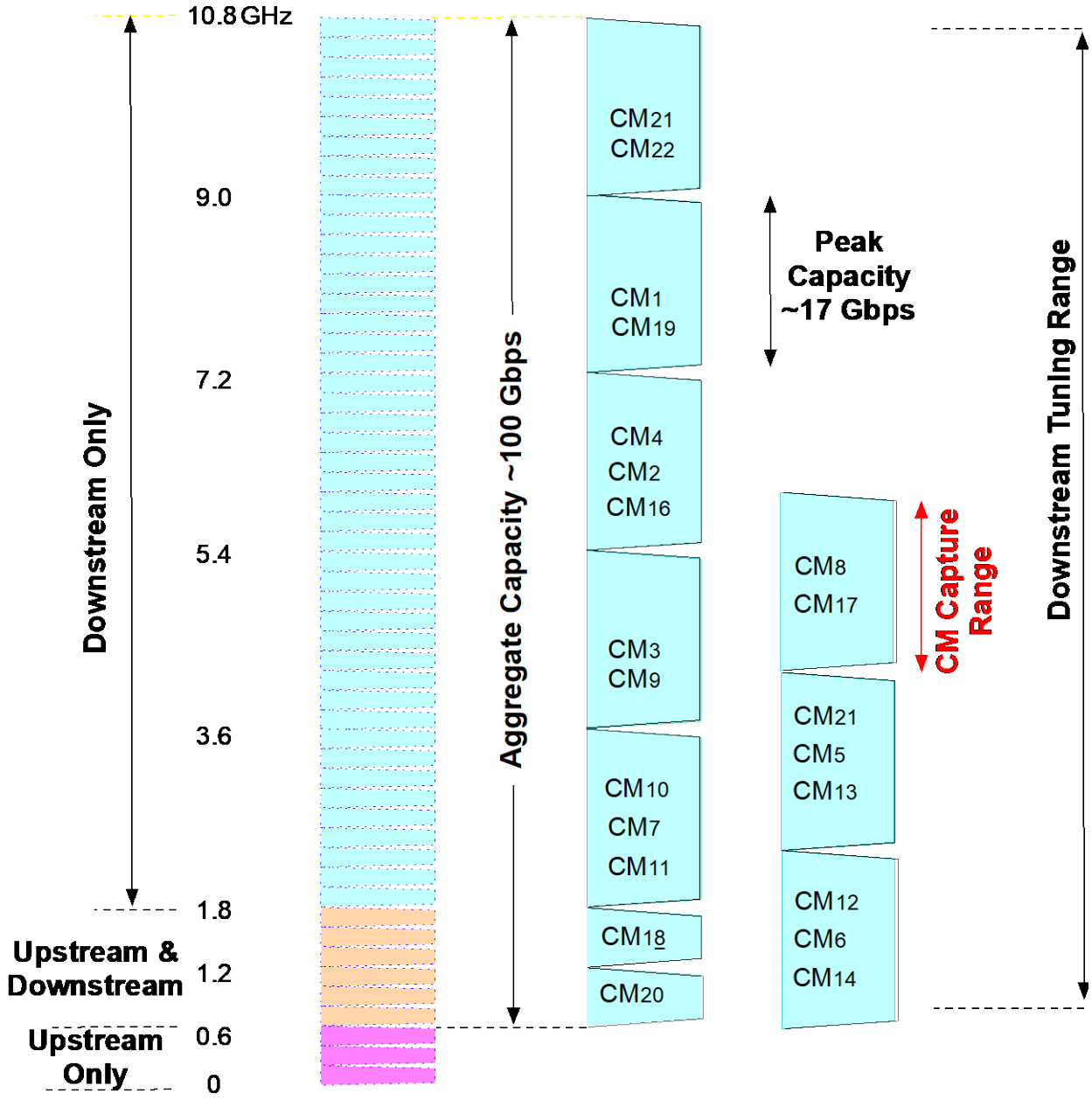


Figure 12 - CM Spectrum And Bandwidth Allocation Example

5. An Evolved Scheduler

In an environment such as the one depicted in section 4 where much higher frequencies are used, such as the case of coaxial spectrum above 3 GHz, subscribers' frequency-dependent performance becomes more noticeable. This behavior becomes highly dependent on which tap they are connected to and how long the drop cable is. In this environment, the importance of having an evolved scheduler that can allocate resources based on frequency and Modulation Error Ratio (MER) is very important.

We introduced profiles in the DOCSIS 3.1 specification which allows custom mapping of subcarrier modulation order versus frequency for groups of CMs associated to a profile. Figure 13 shows an example describing its implementation.

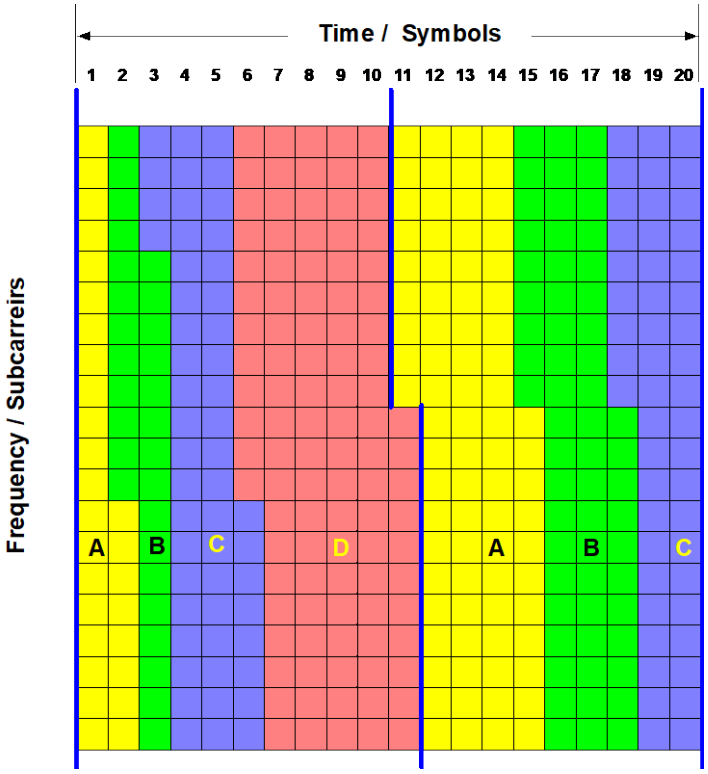


Figure 13 – Conventional Implementation of Downstream Modulation Profiles, A,B,C,D

The profile, however, covers the entire range of frequencies within a channel. In an environment with strong frequency dependent behavior, having the capability of limiting users under the same profile within a certain frequency range would be advantageous to more flexibly implement the HFOF concepts and improve overall system performance. If the channel could be flexibly split in two or three frequency segments as shown in Figure 14, one could optimize overall network performance.

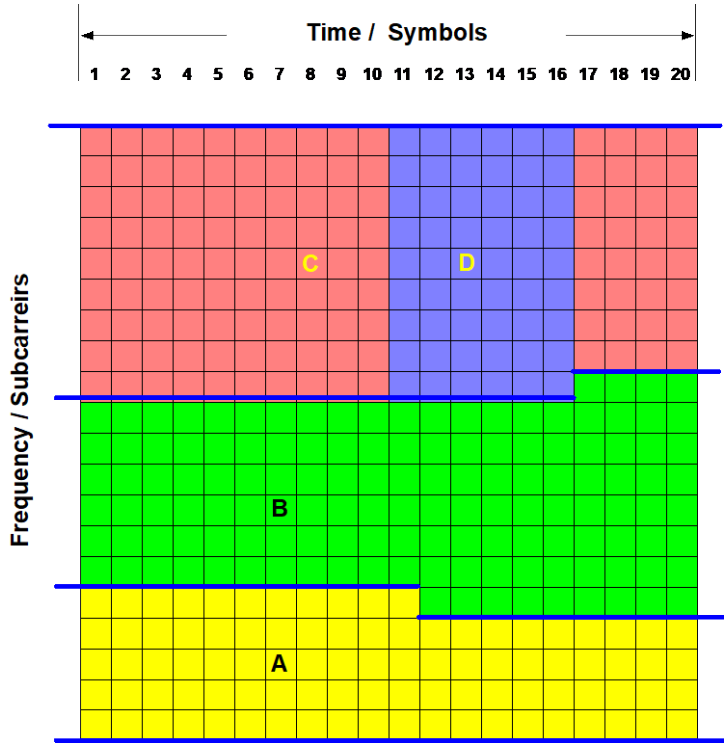


Figure 14 – Frequency Dependent Implementation of Downstream Modulation Profiles, A,B,C,D

Figure 14 shows a downstream frequency-dependent implementation of profiles. A similar capability can be implemented in the upstream

5.1. Additional Frequency Aware Scheduling Benefits

In addition to capacity optimization, having resource allocation control based on frequency and SNR provides other operational benefits. With efficient allocation, a user's capacity is less sensitive to the length of cables and location. Therefore, we could have a system implemented with more flexible coaxial reaches. As operators migrate to N+0 topologies, it is advantageous to split the original HFC node into as few child nodes as possible. To that effect operators can smartly reposition child node locations, not necessarily coincident with former amplifier locations leveraging longer coaxial segments.

Leveraging frequency dependent resource allocation of a next generation scheduler, you can have, for example, a 1.8 GHz coaxial segment with 6 taps supporting 9 192 MHz DOCSIS 3.1 channels. In such a scenario, you can have CMs connected to the first two taps allocated to use channels 5 through 9, CMs attached to the next two taps could be allocated resources from channels 3 through 7 and the CMs attached to the last two taps could be allocated resources from channels 1 through 5. Since the CMs that are farther away use the lower frequency channels one can afford much longer coaxial segments. Another way of looking at this is requiring only lower gain amplifiers resulting in overall lower power consumption.

6. Conventional Taps and Connectors

We mentioned earlier that in order to distribute analog video channels such that they arrive at approximately equal power levels to our subscribers, our coaxial networks were designed with decreasing

tap values so that the attenuation of coaxial cable is somewhat compensated by the tap value. In this original approach, a tap in proximity and following a node or an amplifier would have a higher coupling loss than a tap that is farther away from the node or amplifier so that the impact of longer cable attenuation is compensated.

According to the old paradigm, every end device such as a TV set, a set-top box or a cable modem would receive about the same power level per channel. Operators would use RF distribution taps that consisted of a housing structure and a faceplate. Typically, the housing included ports to connect to the hardline cables or the rigid portion of the coaxial network. The faceplate, on the other hand, included ports that connected to the flexible portion of the coaxial network, the drop cables (Figure 15). The faceplate also included coupling and splitting circuitry to provide specific coupling loss values to the drop ports. These faceplates are removable and designed with different coupling loss values to reach the subscriber's premises at the target power level. If different coupling values or tap values are desired, the faceplate is replaced by another one with the desired coupling values.

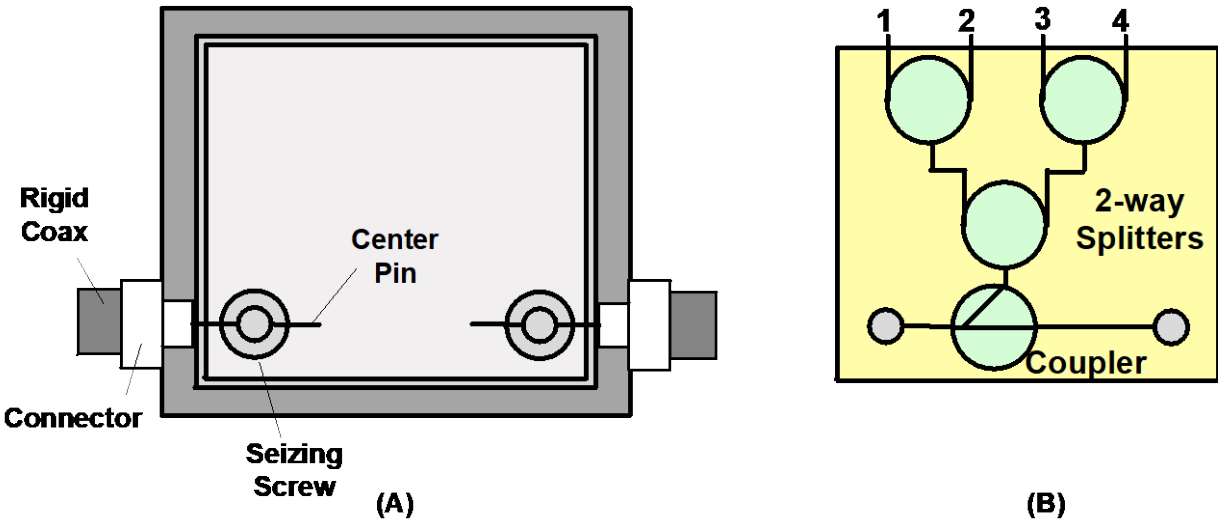


Figure 15 – Tap Housing (A) and 4-Drop-Port Removable Faceplate (B)

One reason to have the tap consist of two components, the housing and the faceplate, is so that faceplates with different tap values can be easily interchanged.

Another reason to have removable faceplates is so that during installation, technicians could have access to the internal structure of the tap. This would allow them to set/configure the tap to receive the center pin of an external KS connector attached to the hardline cable. The connector could attach from a vertical direction when used in pedestals with an underground coaxial distribution network, or from a horizontal direction when the transmission cables are inline as is the case of an aerial distribution network (Figure 16).

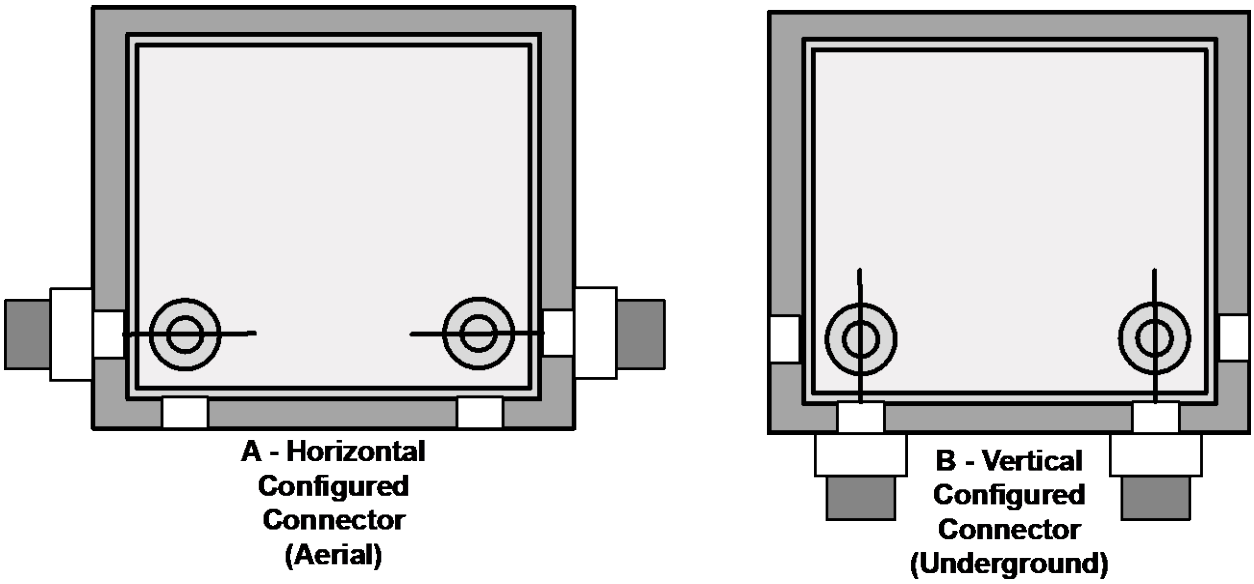


Figure 16 – Horizontal/Aerial (A) and Vertical/Underground (B) Configuration Of Traditional Tap Housings

With a faceplate removed, one can also verify that the long center pin of the KS connector is trimmed to the right length (different tap vendors require different center pin lengths and technicians adjust by manually cutting the center pin) and adjust the seizure screw to make sure a good contact is made with the center conductor of the KS connector.

A fourth reason is to change faceplates with a larger or smaller number of drop ports. This occurs when new customer premises are built and/or a greater number of ports are required.

One challenge that comes from having removable faceplates is that when the faceplate is removed the RF transmission to the elements downstream from the tap is interrupted. The industry solved this by including a conductive path that switches in place enabling an alternate path between the taps' input and output ports, therefore avoiding interruption of AC and RF transport. This alternate path is often implemented using a metal strip which has suboptimal performance at higher frequencies (Figure 17).

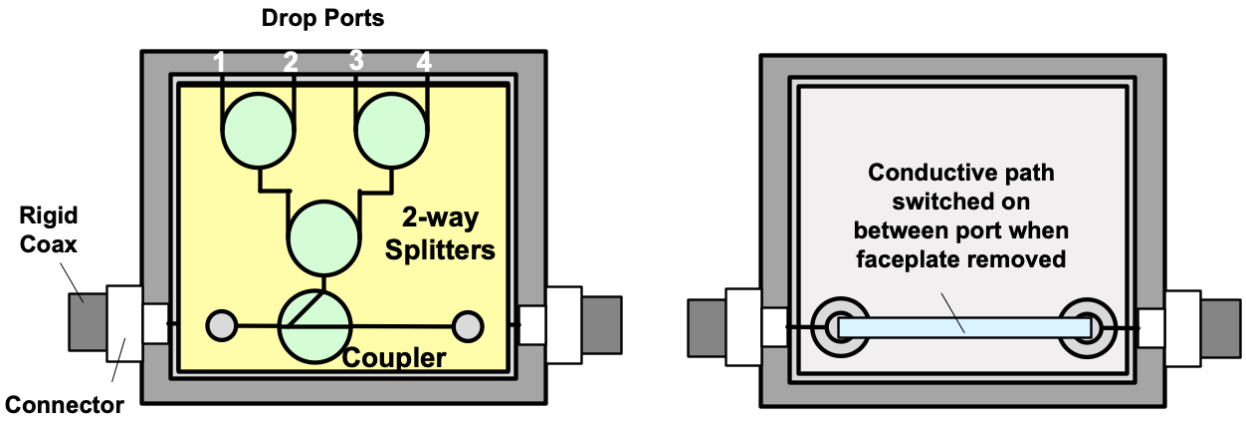


Figure 17 – Tap With And Without Faceplate Showing Conductive Path Switched On When Faceplate Is Removed

7. Revisiting Cable Distribution and Network Components Design

In today's digital age, if we want to control capacity that each subscriber could ultimately consume, we would ideally leverage digital tools that control resource allocation, mainly the CMTS scheduler. If possible, we should avoid using infrastructure means that impact resource allocation when digital means are available. It has been a long time since we moved away from using inline RF notch filters to control premium content access. We have evolved to using digital encryption tools instead. For resource allocation tasks we must also leverage as much as possible our digital domain tools.

Therefore, for the taps that are closer to the node or amplifier we can use a much lower value tap than what is conventionally used (Figure 4). Figure 18b, shows an alternate coaxial segment to the one presented earlier in Figure 4 and Figure 18a, with tap values adjusted so that subscribers leverage the channel conditions and performance they have available in their transmission medium to the node or amplifier.

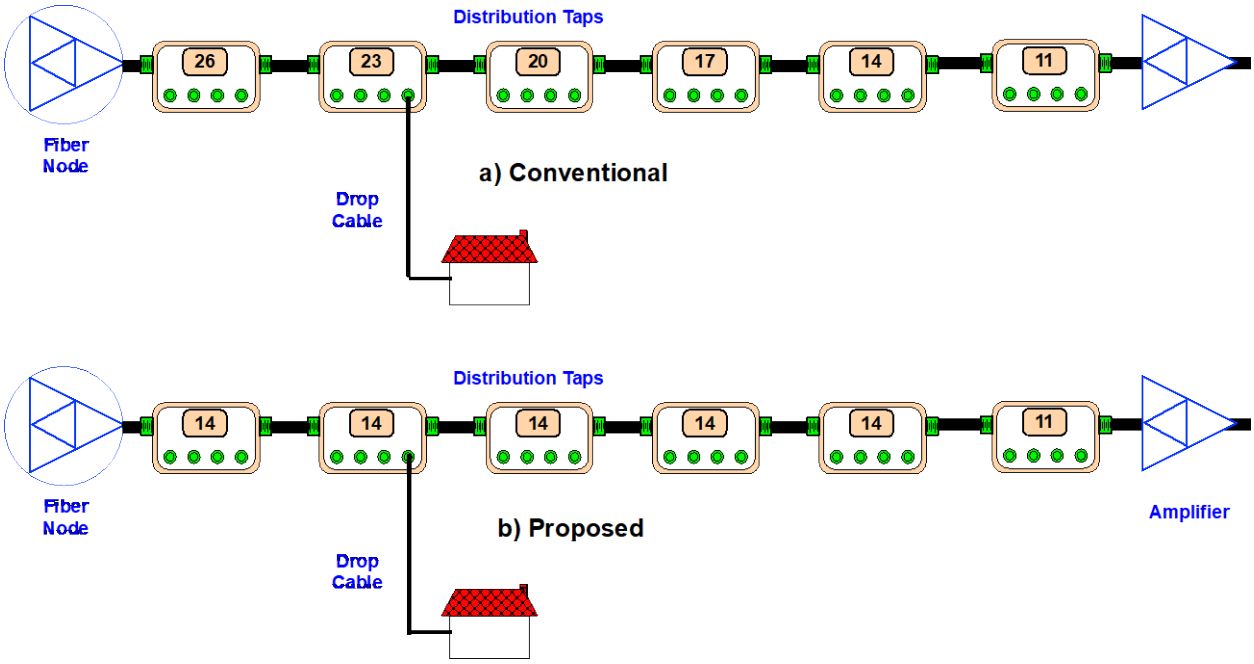


Figure 18 – Coaxial Segment With Adjusted Tap Values Optimize Subscriber Capacity

In Figure 18, the first four taps have decreased their tap values to 14 dB, the 5th tap remains at 14 dB and the last at 11 dB. At lower frequencies, the insertion loss of taps with values 17 dB or higher is dominated by the implementation or excess loss. The insertion loss value for the 14 dB four-port tap is still below 2 dB even after adding excess loss. At higher frequencies the tap implementation becomes more complex and there is a small, gradual excess loss that increases with frequency.

7.1. Single Value Tap

In Figure 19, all the taps for this 4-drop-port scenario, have the same coupling loss value of 14 dB. It does not represent a drastic change, even for the last tap where only a small capacity penalty is incurred. An

exception can be made with the end-of-line "tap" when we are actually dealing with a splitter. Keep in mind that not all end-devices or cable modems need to have the same RF power level. The scheduler is in charge of controlling the capacity that the CMs receive, even though the RF receive power may vary among end-devices. Figure 19 shows the implementation of the coaxial segment using only one tap value.

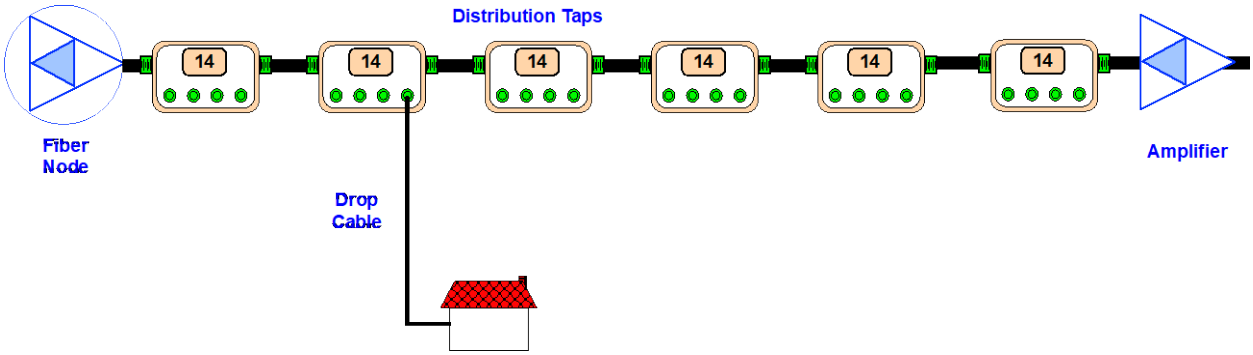


Figure 19 – Single Value Tap Segment For A 4-Port Tap

Using only one tap value for a type of tap significantly simplifies operations. The stocking of taps is much easier. Table 1 shows typical options for tap types and values commercially available.

Table 3 – Typical Tap Types And Coupling Loss Values in dB

2-port	29	26	23	20	17	14	11	8	4
4-port	29	26	23	20	17	14	11	8	
8-port	29	26	23	20	17	14	11		

The 14 dB tap coupling value shown as an example for a 4-port tap has not yet been optimized. Its optimization will depend on the spectrum bandwidth, lengths of coaxial cables and the number of taps that the operators are targeting in a coaxial segment. However, a suboptimal tap value should not significantly affect performance, according to our recent study.

With this in mind, it is safe to assume that there could be a single tap value for 2, 4 and 8-port taps in addition to an end-of-line splitter. This approach would result in the types of taps shown in Table 4.

Table 4 – Modified Tap Types And Coupling Values – Last Column For End-Of-Line Taps

2-port (dB)	11	4
4-port (dB)	14	8
8-port (dB)	17	11

The number of spare taps that a technician would need to stock has been greatly reduced. Without counting the end-of-line splitters, we are dealing with the implementation of a single-value tap for each number of port types.

In this new environment, can the tap performance be improved? One of the challenges when implementing taps is its port-to-port isolation. A simplified traditional tap circuit can be represented by a coupler followed by a splitter. We represent a 2-port tap by a coupler followed by a 2-way splitter, a 4-

port tap by a coupler followed by a 4-way splitter and an 8-port tap by a coupler followed by an 8-way splitter.

This configuration has a potential isolation issue between splitter ports. Isolation between drop ports is not optimal and degrades at higher frequencies. Figure 20 shows a design alternative for a 4-port tap.

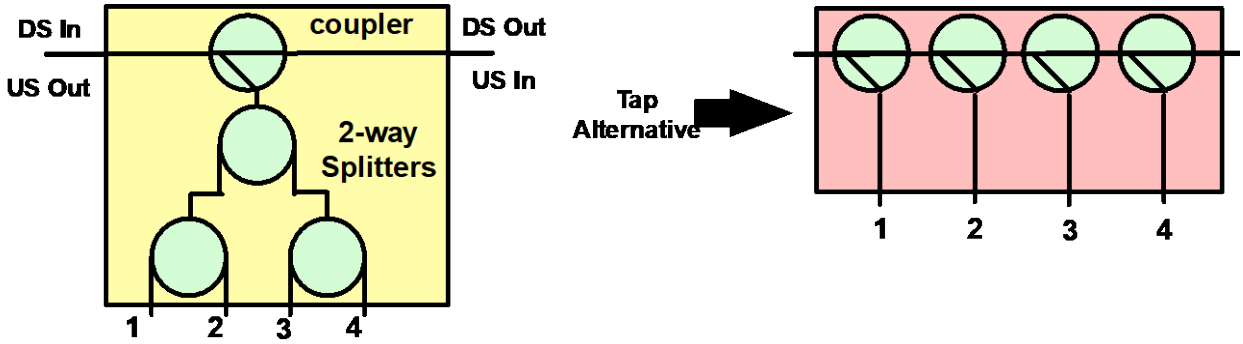


Figure 20 – High Isolation 4-Port Tap Design Alternative

7.2. Revisiting Need For Removable Tap Faceplates

Now that we have an approach to drastically reduce the inventory of taps, we need to ask the question again. Do we really need removable faceplates? Earlier, we discussed some of the reasons why we have removable faceplates. We will now explore the impact of not having them.

Properties of taps with removable faceplates include:

1) Changing tap values

2) Changing the number of drop ports

3) Switching the tap configuration between vertical and horizontal

4) Verifying proper length of center pin

5) Verifying proper contact of seizure screw

6) Maintaining connectivity to the elements downstream during faceplate removal

Changing tap values may no longer be necessary in an environment with a reduced number of tap types. The burden on inventory has been reduced with single-value taps for different numbers of ports. A reduced number of tap types also facilitates the development of a vertical-only housing and a horizontal-only housing. This provides the opportunity to explore using two types of tap housings, mainly horizontal and vertical connector entry tap housings (Figure 21).

Separate vertical and horizontal housings eliminate the need for technicians to mechanically configure a horizontal or a vertical connector entry option. The mechanism that enables dual connector entry options makes operation at higher frequencies more challenging.

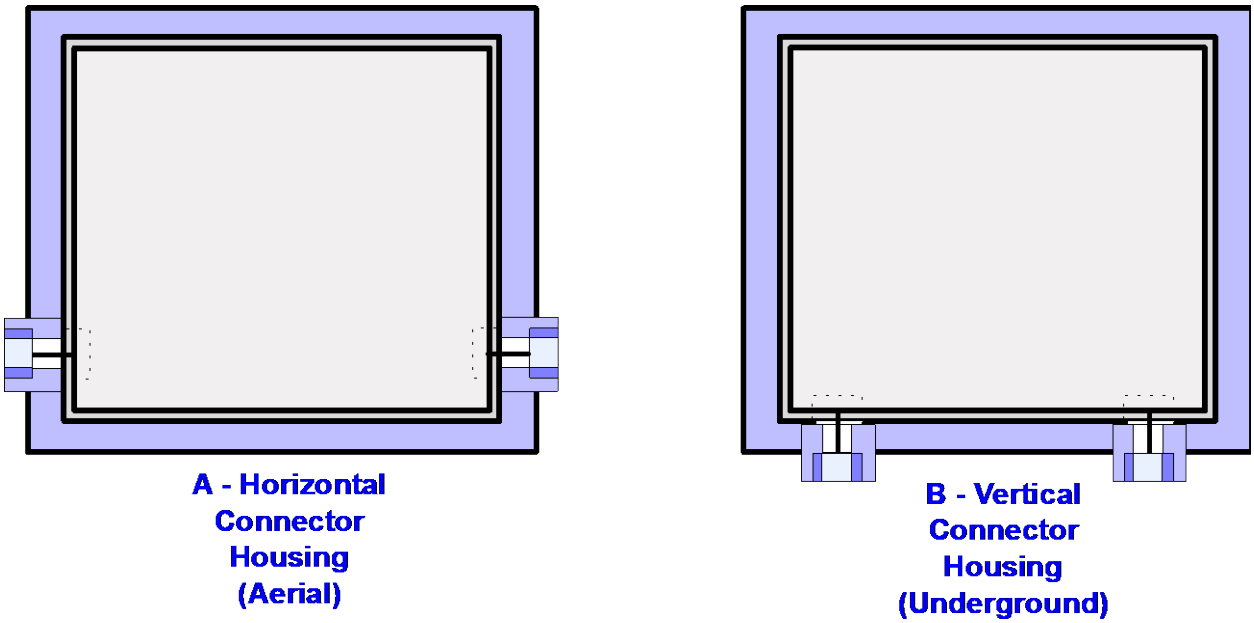


Figure 21 – Separate Housing SKUs For Horizontal (A) And Vertical (B) Taps

Specific connector entry (horizontal or vertical) implementation enables a mating tap connector permanently attached to the tap circuit board. This means that technicians do not need to trim the center pin at vendor specific lengths. It actually enables the use of a connector with fixed length center pin that screws to establish connectivity without the need for seizure screws. You can use a female connector attached to the tap and the board inside the tap and a male connector that attaches to the hardline cable (Figure 22).

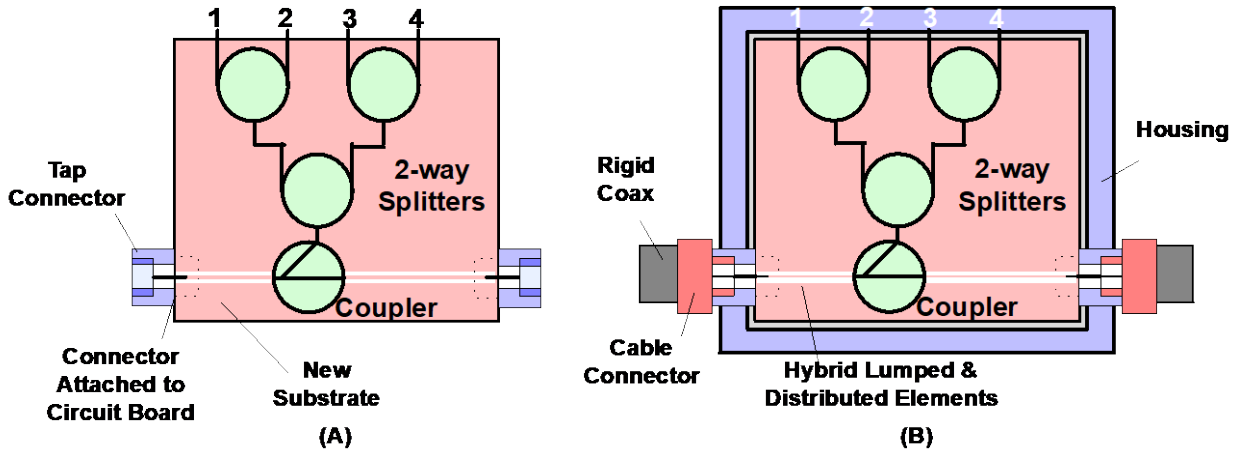


Figure 22 – Updated 4-Port Tap Design

There is no longer a need to verify contact and center conductor length by removing a faceplate. The tap can be implemented as a closed unit. Water leakage and radio frequency ingress problems would be drastically diminished. Challenges in implementing the switchable conductivity path when faceplates are removed are avoided. All circuitry can reside in one board and the assembly of a permanent connector that mates to a hardline connector would enable much higher frequency implementations.

One drawback of this proposed approach is that tap replacement results in a service outage for subscribers downstream from the tap. Reasons to require the changing of taps are diminished with the use of enclosed taps. In the rare event when a new premise needs connectivity. which was not anticipated during the original network design, and no spare drop port is available, you would need to replace the tap. Tap failure would also require a tap replacement, but in an environment where the taps are never opened this would be rare. In a fiber deep architecture, tap replacement is less of an issue as the network affected area is smaller. To replace a tap, two of these next generation hardline connectors are unscrewed to remove the whole tap, estimated at less than 1 minute of interruption.

As we move to higher frequencies, we need to be mindful of the natural tendency to hang on to traditional approaches. We have an inertia to continue using techniques that may no longer be the most efficient. Our environment is changing, we have been gradually pushing to higher and higher frequencies and the way we manufacture taps needs to evolve as well.

In the past, tap circuitry has been implemented using lumped circuit elements. However, as we look to support higher frequencies, incorporating distributed elements may be necessary. Similar questions must be asked regarding the tap circuit substrate. Fiberglass-based substrate, FR4, has been used in the past. Its low permittivity may result in higher loss and leakage at higher frequencies. Ceramic and PTFE (Polytetrafluoroethylene) based substrates should be explored. Their higher permittivity helps confine the RF energy and reduce leakage. While cable's traditional support of lower frequencies (< 1GHz) takes advantage of lumped element circuit components, hybrid lumped and distributed circuits next-generation designs could provide good performance at both higher and lower frequencies.

8. 100+ Gbps Experimental Setup and Demonstration

At this point, the discussion has focused on theoretical aspects of using higher portions of the coaxial spectrum leveraging simulations and modelling of a coaxial segment with cascaded taps. This section discusses experimental results obtained from an actual coaxial segment that has been built using cascaded taps linked by rigid coax cable. This network is actually a 50-ohm network, although the cables that were selected, have the exact attenuation versus frequency behavior of 0.54" rigid coax and RG6 flexible coaxial cable. This network leverages two types of transmitters. From 500 MHz to 3000 MHz a single-frame DOCSIS signal was composed in MATLAB and generated from an arbitrary waveform generator (AWG). This signal was configured with DOCSIS 1024 QAM subcarriers and the output was received in the last tap by a Rohde & Schwartz DOCSIS analyzer. A raw rate of 20 Gbps was estimated at the DOCSIS receiver.

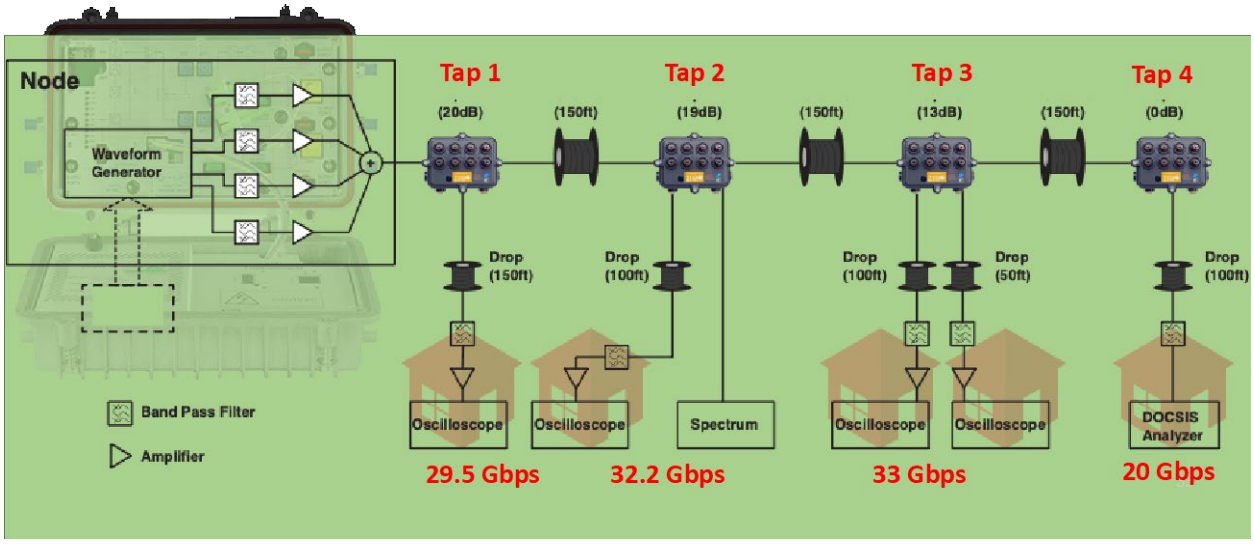


Figure 23 – Experimental Setup of > 100 Gbps Coaxial System

The rest of the spectrum covering frequencies of up to 12 GHz was occupied by re-designed DOCSIS-like OFDM symbols generated from AWGs (Figure 24).

Figure 24 – Spectrum Allocation Of Experimental Setup

A transmission band from 3 GHz to 6 GHz was assigned to Tap3, a band from 6.5 GHz to 9.5 GHz was assigned to Tap 2, and a band from 9.5 GHz to 12 GHz which was assigned to Tap 1 which enjoys the best CNR at higher frequencies. The portion of the spectrum between 6.0 GHz and 6.5 GHz was not used due to the aliasing signal from the AWG module operating at 12 GS/s. An optimized custom design could avoid that frequency gap. Capacity was estimated at 29.5 Gbps at tap 1, 32.2 Gbps at tap 2, 33 Gbps at tap 3 and 20 Gbps at tap 4. An aggregate capacity of 114.7 Gbps was obtained.

9. Node+1 Architectures

The analysis in previous sections examined a single coaxial segment. This capacity is therefore accessible to a Node + 0 architecture. Nevertheless, the use of higher frequencies doesn't have to be limited to N+0. In fact, as you further examine the performance of CMs that connect to the last tap and have a shorter drop length, such as in the case of CM_{12}, the spectral efficiency is quite high. An equivalent longer hardline segment with no drop connecting to an amplifier that follows can take advantage of an N+1 architecture. To highlight that higher cascade use, the spectral efficiency in CM_{12} is shown in red in Figure 25.

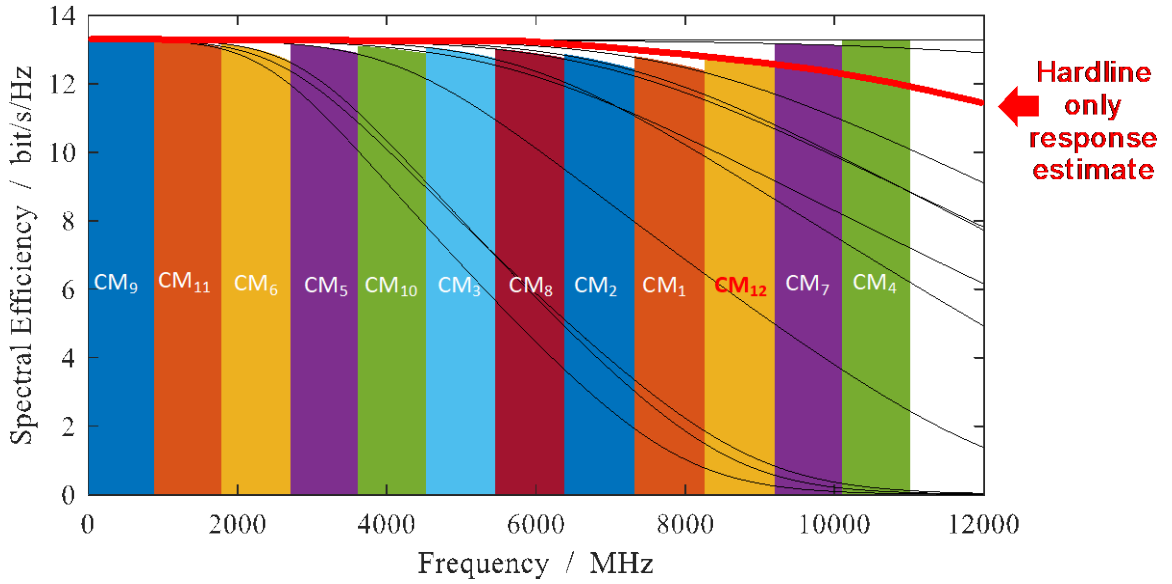


Figure 25 – Spectral Efficiency Available At End Of First Coaxial Segment N+1 HFOF Implementation

Figure 26 shows a N+1 concatenated coaxial segment with some sample scenarios regarding the bandwidths that could be implemented in the segment that follows the fiber node and the one that follows the N+1 amplifier. Figure 26 includes an aggressive scenario with 10 GHz bandwidth following the fiber node (Option 1), in addition to a more conservative 6 GHz bandwidth implementation (Option 2). Keep in mind that higher bandwidths are feasible if lower efficiencies are allowed or in shorter drop length scenarios.

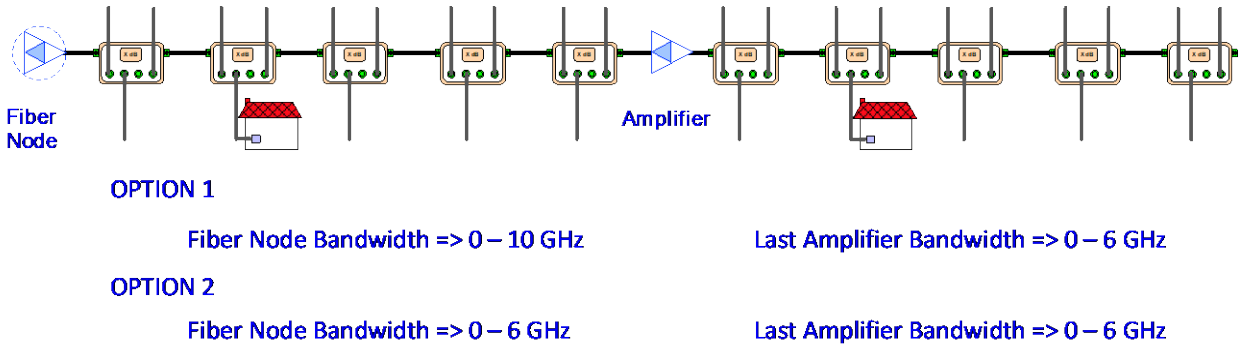


Figure 26 – Coaxial Segments in N+1 HFOF Implementation

9.1. Longer Coaxial Segments

As HFC networks evolve to fiber deeper. An alternative to Node+1 is to use longer coaxial segments. In order to avoid having too many optical nodes in a N+0 architecture, we have been pushing very high gains and power levels out of our optical nodes. We are doing this so that the furthest home can access all the resources that the CMTS or the RPD or RMD makes available. If we submit to the HFOF philosophy, you don't have to expect that every CM can handle the entire spectrum available. It is OK not to be able to consume all the resources. A consequence of this is that you could afford longer passive segments. We

are calling this topology “N+0-Long”. The intelligence of the evolved scheduler will help you with the appropriate resources based on the capabilities of each endpoint.

10. New Kind Of “Hybrid” Fiber Coax

As we have seen how we can still achieve significant capacity leveraging coaxial resources, in order to balance the cost-complexity of the end devices or CMs, one has to determine how much coaxial aggregate capacity and how much peak capacity are practical. Even though 100 Gbps is achieved as an aggregate from a coaxial serving area, the complexity of a CM capturing significant RF bandwidth has to be assessed when determining its practical peak rate. We estimate that 25 Gbps could be a practical peak rate target or 50 Gbps when stacking two receivers. An alternative to coax is the use of coherent PON or CPON. The emergence of coherent optical innovations in the access environment along with the cost reduction trends of coherent optical components, make CPON an attractive long term access solution. The cost of deploying fiber deeper and fiber-to-the-home varies significantly among operators and even within an operator, which may lead an operator to different coax versus fiber deployment strategies. This will depend; on the specifics of the starting point scenario to evolve towards FTTH, on the availability of conduits, on the cost of deploying fiber drops, on economical and operational aspects for extending the life and frequency range of coax and many other dependencies. For some operators, it may make sense to migrate directly to FTTH and for others leveraging the existing coax may make economical sense. The evolution of fiber deeper and FTTH could be made on an as needed basis. A technology like CPON which allows users to reach 100 Gbps on typical Hub-to-subscriber lengths, supporting split ratios of up to 512, could be leveraged to support subscribers demanding higher peak rates. There are many users that may not need peak rates above 25 Gbps for quite some time. Now that fiber is penetrating much deeper in cable networks, the high-end users requiring high peak rate services are a long optical drop away which could be implemented on a success basis. Operators can design the ultimate fiber-to-the-home network but only deploy it partially based on where the high-end customers are. This would result in a gradual transition towards FTTH depending on where the demand is. In some places, there may not ever be such a demand. A CPON network could feed “Extreme Coax” nodes, base stations, enterprise and residential high-end users. Figure 27 shows such a “Hybrid” network where CPON and an N+1 “Extreme Coax” network are jointly leveraged to address subscribers’ long term demand of peak and aggregate data rates.

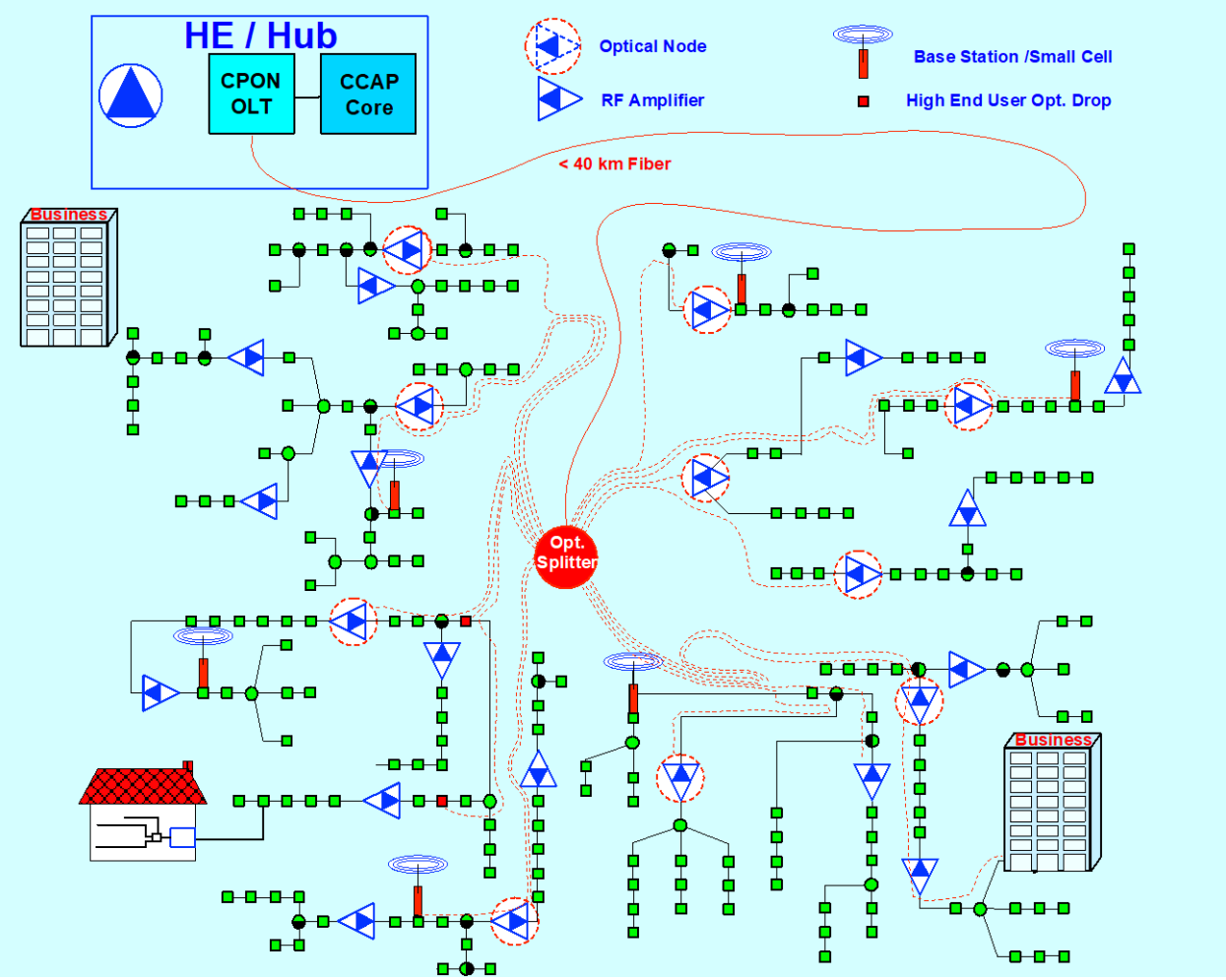


Figure 27 – Ultimate CPON/Extreme Cable "Hybrid" Fiber Coax Network

11. Conclusion

We have finally said goodbye to analog video and we need to fully embrace the digital era with all of its benefits, including the opportunity it provides in re-designing our coaxial network. The CMTS being the device that controls resource allocation still has plenty of room for improvement in this new environment.

As Cable entertains the support of 1.8 GHz, 3 GHz and even higher frequencies, the coaxial cable medium exhibits greater dependency in frequency. Having frequency-aware resource allocation provides great strategic advantage, helping enhance our data delivery capabilities over coax. Being free from analog video restrictions provides Cable the opportunity to drastically simplify the implementation of its coaxial infrastructure while preparing it to evolve to higher frequencies. Except for end-of-line taps, Cable can follow a single value tap for each tap type with the same number of ports. This reduction in inventory makes attractive horizontal- and vertical- specific taps, as well as taps without removable faceplates, avoiding many of the challenges in the evolution to higher frequencies.

A Higher Frequency Off First (HFOF) approach to allocating bandwidth has been proposed to optimize how we can use higher frequency resources as well as to facilitate the extension of coaxial segment lengths. This approach is not limited to N+0 architectures but can also be used with N+1 and higher

cascade scenarios. Coaxial bandwidths greater than 100 Gbps have been demonstrated over a coaxial segment using HFOF and leveraging frequencies approaching the cut-off frequencies of the hardline cable (11 GHz). Balancing the capture bandwidth of the CM versus its tunability allows the optimization in the system's cost-complexity through peak versus aggregate rate assessment. The proposed HFOF approach also bounds the system's dynamic range. Frequency aware scheduling, HFOF, single value tap and high frequency tap redesign are key ingredients to this Extreme Cable approach. Together these concepts are powerful, but they could also be used independently and provide benefit to the evolution of our coaxial environment.

A new "Hybrid" Fiber Coax environment where CPON and Extreme Cable join forces to deliver data services is considered as a gradual, success-based transition to FTTH in the areas where it is needed.

Abbreviations

AWG	arbitrary waveform generator
CAGR	compound annual growth rate
CATV	cable television
CCAP	converged cable access platform
CM	cable modem
CMTS	cable modem termination system
CNR	carrier to noise ratio
CPON	coherent passive optical network
dB	decibels
DOCSIS	data over cable service interface specification
DS	downstream
FEC	forward error correction
FR4	flame retardant 4 circuit
FTTH	fiber to the home
Gbps	gigabit per second
GHz	gigahertz
HE	headend
HFC	hybrid fiber coax
HFOF	higher frequency off first
HHP	household passed
KS	klemmschrauben (clamp screw)
MAC	medium access control layer
MER	modulation error ratio
MHz	megahertz
OLT	optical line terminal
ONU	optical network unit
PHY	physical layer
PON	passive optical network
PTFE	polytetrafluoroethylene
QAM	quadrature amplitude modulation
RF	radio frequency

RG	radio grade
RPD	remote PHY device
RMD	remote MAC-PHY device
Rx	receiver
SNR	signal to noise ratio
STB	set-top-box
TE	transverse-electric
TEM	transverse-electromagnetic
TV	television
Tx	transmitter
US	upstream

Bibliography & References

DOCSIS 3.1 Physical Layer Specification. CM-SP-PHYv3.1-I18-210125, January 25, 2021, Cable Televisions Laboratories, Inc.

Microwaves101.com, https://www.microwaves101.com/encyclopedias/coax-cutoff-frequency

DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification. CM-SP-MULPIv3.1-I21-201020, October 20, 2020, Cable Televisions Laboratories, Inc.

DOCSIS 4.0 Physical Layer Specification. CM-SP-PHYv4.0-I03-201202, December 02, 2020, Cable Televisions Laboratories, Inc.

Z. Jia and L. A. Campos, "Coherent Optics Ready for Prime Time in Short-Haul Networks," in IEEE Network, vol. 35, no. 2, pp. 8-14, March/April 2021, doi: 10.1109/MNET.011.2000612.

J. Zhang, Z. Jia, M. Xu, H. Zhang and L. Alberto Campos, "Efficient preamble design and digital signal processing in upstream burst-mode detection of 100G TDM coherent-PON," in IEEE/OSA Journal of Optical Communications and Networking, vol. 13, no. 2, pp. A135-A143, February 2021, doi: 10.1364/JOCN.402591.

J. Zhang, Z. Jia, M. Xu, H. Zhang, L. A. Campos and C. Knittle, "High-Performance Preamble Design and Upstream Burst-Mode Detection in 100-Gb/s/λ TDM Coherent-PON," 2020 Optical Fiber Communications Conference and Exhibition (OFC), 2020, pp. 1-3.

The Tooling Abyss

A Technical Paper prepared for SCTE by

Joann Shumard
Vice President –Engineering Operations
Comcast Cable
1800 Arch street, Philadelphia, PA 10103
770-652-3836
Joann_Shumard@cable.comcast.com

1. Introduction

System support tools can be a vast ocean of internally-developed and externally-purchased groups of software solutions. This paper will focus on drivers and solutions to an organizational problem, the tooling abyss, that centers around the dual impacts of pace and sprawl of tools created and utilized to keep in stride with the innovation growth of the technology of our industry. This document will clarify the definition of what a tool is for our industry, identify the business impacts of our decisions, and recognize the perspectives of the user community.

"Just build a new one" is a typical response to developing tools to support new technology. This mantra has defined many aspects of the transformative technology growth in the industry. The "build a new" strategy, without a plan for the old, has led to an exponential growth of complex tools required to operate the business. This methodology increases user complexity and causes operational inefficiency. Technology innovation has unlocked an expansive sea of data and knowledge, but this simultaneously created a myriad of tools making it more complex for the user communities. Increasing complexity can result in inefficiency, such as when a single team member must reference multiple interfaces to troubleshoot a single problem. This complexity creates and exacerbates perceptions of information overload. The concept of "adding new to old" has also shaped new challenges, like a sense of searching through an abyss of information when trying to find critical troubleshooting data. The influx of added information and operating multiple interfaces causes challenges for teams responsible for managing the customer experience, maintaining the plant, and monitoring system events.

The persistent launch of new technologies in the industry has driven a constant change in the tools needed to operate them. This paper will focus on building a strategy around making instrumental tooling decisions, concentrating on when to develop new, when to integrate, and when to sunset tools. Tooling culture will be discussed to delve into the psychology of change, and the navigation of managing evolution, including the emotional attachment to tools team members, have historically known. This paper and presentation will review strategies on tool development that are based on operational efficiency while supporting new technology and driving overall business strategies.

The purpose of this paper is to understand how more effective development methods can lead to improved team performance across multiple technology tooling systems. Evaluating real-world organizational problems provides a pathway to investigate a specific business problem and seek opportunities that shape future business goals or strategies. The information will enhance the understanding of how business evolution, financial requirements, and consumer needs have led to the multitude of systems utilized in today's operations. The paper will focus on the history of how the tooling abyss has been created and the impacts on the user community. It will propose new opportunities in managing tooling systems to support the development of plausible and implementable solutions.

2. History Of Industry Technology

To help understand today's issues of the complexity of the tooling abyss, we can review the pace of innovation of the industry because every "chapter" brought with it a corresponding increase in tools. Cable television dawned in the 1940s when cable became the connection into the homes that could not be easily reached by standard over-the-air broadcast capability [2] (Cable History, 2014). Through the years, the technology evolved through amplifiers and the continuing focus to reach more people. The 1960s was an innovation milestone with the introduction of the Jerrold Channel Commander headend signal processor, the first pay-for-view (PPV) event offering, solid-state technology, and the enhancement of aluminum-shielded foam dialectic distribution cable. In 1969 the Society of Cable Television Engineers (SCTE) was established, focused on training and educating the industry's technology professionals. The innovations of the 1970s brought new concepts of two-way addressability and enhancements in drop

cable to improve in-home quality. Video evolution continued to grow through the years, and the expansion of the industry continued through compression technologies, industry channel expansions, and innovation in delivery systems.

The 1990s accelerated the pace of technology innovation with the introduction of the fiber to the node methodology or Hybrid Fiber Coax (HFC) [2] (Cable History, 2014). The innovation of the return spectrum opened the door for two-way communication that would support PPV revenues, telephony voice, and the ability to deliver the internet to homes. In 1995 copper-wire phone lines provided 14.4 kbps internet, while the then-new cable modems would provide 4000 kbps which, at the time, was groundbreaking. In 1997 CableLabs emerged as a governing presence to align broadband standards for manufacturers and operators to ensure consistency in the business. Since the beginning of DOCSIS 1.0 in 1997, these standards and the industry have grown exponentially. DOCSIS 1.0 became 2.0, 3.0, 3.1, and laid the future for 10G and beyond. Analog delivery systems have transformed into fully digital architectures, providing a stream of innovative technical data. All of this evolution transformed the experience for all internal users.

3. Tool Development Influence

The historic pace of innovation is a significant factor in recognizing how the tools needed to operate them have also transformed. The first 50 years of evolution centered primarily on video technology. The catalyst of the evolution of the cable modem created a pace of innovation in the next 25 years that has dramatically altered the amount of software development and tooling innovation required to keep pace with the accelerated growth. In early history, basic measurements of levels were the primary performance indicators compared to today's expansive data capabilities that provide insight into the customer premise equipment in the home. This history of acceleration is a primary contributor to the exponential growth of support tooling in the industry.

4. Industry Tooling

Some may ask: What is considered a cable tool? The focus of this paper is on the software developed "electronic toolbelt" [1] (Breymier, 2020) rather than the hands-on physical tools used by field teams. If starting from the catalytic point of the introduction of cable modems, that question can be answered by the desire to understand the state and levels of the cable modems when troubleshooting. Early troubleshooting would require logging into the CMTS (Cable Modem Termination System) and looking for a specific modem that required its MAC (Media Access Control) address. The use of the "show cable modem" feature and directly logging into the CMTS became quickly identified as a non-scalable way to gather and interpret information for troubleshooting.

As the technology and customer base have grown at an accelerated pace over the past 25 years, so has the technology of the tools that support it. Tools in the business can be those used for monitoring the systems, databases that store information, back-office systems, or tools used by front-line team members who work in the home or the plant. New user interfaces (UIs) are tool interfaces created using data pulled from the CMTS and developed into a more consumable view. As the technology evolved and new data became available, in many cases, new tools were created to support it. The answer to the question about what constitutes a necessary tool must acknowledge certain contributing factors. The next sections describe the vast amount of tools utilized in the course of supporting business needs to identify opportunities to optimize strategy for future development.

5. Organizational Importance

The purpose of this review of industry tooling strategy is to investigate the impact development decisions have on team performance, tooling development environment, and intra-organizational communication. The problem identified is that the methodology of creating tools in multiple departments to meet only specific needs, and potentially without a long-term strategy, can negatively impact team performance. When tools are created without a centralized, defined strategy, the organization is impacted by wasted investment tech debt due to software teams duplicating work effort. Team performance is negatively impacted when information does not directly aid internal users to efficiently complete their tasks. Cross-department team conflict leads to organizational impacts on performance and job satisfaction.

There are multiple inputs into what is needed for system tooling, such as new technology, process change, and internal operational-user requests. When internal departments do not have aligned priorities, this can lead to challenges with professionals and software-development engineers working with conflicting priorities resulting in resource inefficiencies. Understanding the drivers for tooling development within the organization and the various business departments are imperative when recognizing what influences the tooling strategy.

6. The Tooling Abyss

To support today's technology, the parallel rise in tooling can cause many internal users to feel like they are swimming in an unending deluge of information. There is an endless stream of new data driving constant change. New tools are developed to support new technology, while legacy tools are sometimes not decommissioned. Some methodologies add new data or widgets to existing tools, making them large and complex to operate. There are three notable areas where the growth of technology creates the pathway for this number to continue to grow:

- Centralized Tooling System Management
- Decentralized Tooling System Management
- Tool Ownership Conflict

Centralized Tooling Systems

Many organizations promote the centralization of tool development to control the development lifecycle. However, it is essential to recognize that the technology in and adjacent to our industry is advancing at an exponential rate. It is critical to understand that a centralized approach to tooling development may not result in one specific strategy. In essence, even centralized organizations may not always have a consistent strategy for tool development.

Commonly, organizations focus on launching new technologies and products to their internal users from a central viewpoint. However, not all have a centralized strategy for tooling, nor a software development plan for operating tool readiness. An increasingly common challenge that illustrates this is a new technology that becomes available to external users, but the tooling is not fully developed for the internal operators that support them. One practice commonly used is "just add new," where engineering organizations feel compelled to not only build the new capability but also create a tool to support it. New architectures and components may take years to completely roll out throughout a large system, which leaves the operators and technicians with no choice but to use multiple sets of tools. In this situation, one tool supports the new architecture, while one remains to operate the legacy. Further complicating this issue, by the time one "new' technology rollout is nearing completion, the following new technology is waiting to ramp up. Tooling never ends; the parallel strategic goal that is continuous integration/continuous development (CI/CD) assures it.

The negative impact to the teams responsible for operations and the technicians is this constant swiveling and cross-referencing of tooling information to complete everyday tasks. Advocates of the centralized model tend to focus on security standards as well as enhancing and growing existing operational toolsets. The risk and challenge with this methodology are that without discipline and long-term strategy, a system's code-stack can become obsolete. It can become so monolithic that continuing development becomes less agile and stability questionable. The thought in this method is that maintaining a centralized toolset reduces the swivel, but as technology data increases, the tool can become more complex to operate, and the amount of information contained in it can become overwhelming. Overall, technology and innovation are great things that drive growth, but there is an impact on the operations when the tools needed to support it are not available.

Decentralized Tooling Systems

Most of today's largest service providers are a product of acquisitions [2] (Cable History, 2014). That creates an increased number of tools integrated into the organization, some developed in-house and some contract acquired through business acquisition. The challenges consolidating organizations face include what to do with the newly acquired tool. Is its function duplicative? What is the cost to sunset the old one? Without a consistent strategy on organizational tooling systems, many choose to simply keep everything and operate in segments.

The broadband communications industry was started by and continues to attract an amazing talent pool that is capable of innovating new ideas, especially when it comes to supporting software-developed tools. Many talented teams have taken it upon themselves to create "gap solvers" where there may be centralized or common systems, but to operate them more effectively, they develop secondary enhancements. These individual solutions developed across the organization can create quick short-term solutions to a current business problem. They are also one of the highest contributors to the tooling sprawl. This methodology creates challenges in operational consistency, security, and system consistency, as well as hinders the ability to make future system changes.

Tool Ownership Conflict

Pride in ownership is a valued characteristic. People involved in the ideation and development of new tool ideas are no exception. Being competitive is a cornerstone of our business. Metrics commonly drive our teams to work hard and strive to be the best. In many business areas, this is a practice that has determined constant operational improvement through the years.

However, in the tooling and development space, this mindset can also be a contributor to conflict. It can perpetuate the keeping of obsoleting systems because of emotional attachments from those who created and used them. In organizational terms, dominant or competitive conflict culture is a winner versus loser strategy that can lead to negative, organizational influences [3] (Choi & Ha, 2018). The competitive conflict culture applied in the software development space leads to inefficiency and negative employee experience. In contrast, the collaborative conflict culture is a partnership-oriented culture focusing on coordinating the best outcomes for the business.

In both centralized and decentralized examples, leaders are responsible for the technology that overlaps with other teams. This overlap can result in continuous conflict, even when attempting to be effective and productive. This dynamic creates both a task and a relationship conflict for the teams that manage the tooling systems. When there is a lack of a clear strategy, negative team impacts can include job dissatisfaction and team inefficiency [3] (Choi & Ha, 2018). Negative individual behaviors are caused by a lack of direction as well as a competition where people are diligent in keeping what they own. It also means some tools may stay in operation because of a lack of willingness to let go of obsolete and antiquated electronic software systems.

Cost of Inconsistency

Job dissatisfaction is not the only business challenge created by competing or unclear tooling strategies. Another is duplication. Tooling duplication is a common result of the decentralization of operational support tools, where the same data is used in multiple ways. For example, different groups may download data on customer premise equipment, then use that data to create a plan of action, or evaluate a system's status. In doing this, developers create duplicate data stores for the information. This duplication creates a cost to the business in multiple ways.

- Multiple data stores mean replication on multiple platforms and systems, which take up much more storage space
- Multiple support teams are required to maintain the data and UI, or one support team is required to support multiple similar tooling systems
- Replicated system upgrades and security requirements
- A technical debt of overlapping and duplicated operational systems
- Inconsistent user experiences

Technical user inefficiency is an increasing cost to the business, especially when the abyss of tools and complexity requires team members to swivel to multiple interfaces to troubleshoot system issues. Team and individual metrics are impacted. As an example, time-per-task increases when the tooling lacks stability, speed of response, or contains more information than is needed for the task. When the utilization of the tool or tools to complete the work is inefficient, its users will avoid adoption. When complex information flows disrupt a tool user's ability to work swiftly and accurately, they will disregard it.

The pace of new consumer and system architecture technology has reached extraordinary heights. This timeline requires system tooling to innovate at the same pace to support the operational needs of the business. When there is system tooling duplication, lack of a focused strategy or extremely large monolithic systems, the ability to maintain pace is reduced. Reduced agility, when launching new features and technology, can result in products becoming available to consumers before they can be effectively supported, from an operations perspective.

7. Information Overload And The Human Factor

Understanding the user experience is an essential concept in any tool's development. The accessibility of data through the internet and system technology has created a rich opportunity to connect and collaborate in extraordinary ways. Recognizing the power of data, however, needs to be balanced with the human factor. Essentially, the cognitive value in the information relies on the design, structure, and quantity of the information shared with the user [6] (Voinea, Vică, Mihailov, & Savulescu, 2020). Cognitively people are likely to take mental shortcuts because, at any given time, the human brain can only process so much complex information. Information overload is a phenomenon that emerges when the amount of information and choices become so expansive that they impact the user's cognition. The psychology of how people process information is a critical understanding point in tool-development concepts. That is because simply adding new data, new widgets, and new information can ultimately have a significant impact on how the internal users adapt to the technology change.

The constant evolution of tooling and strive to drive more data to the internal users also create challenges in the best way to train new tooling technology. In an interview to inform this paper with Walter Breymier, a veteran technician at Comcast, he observed that the constant pace of change and volume of information impacts team members in different ways [1] (Breymier, 2020). The technician perspective must be a key factor in evaluating how a tool will be adopted:

- There is not always a clear understanding of who and what the software tool is for
- Technicians are not given time to truly absorb the best way to utilize the tools
- As a direct result, the tendency is to "stick to what you know" and avoid utilizing the new information.

It is worth noting that it can be the team members with the longest tenure who are the most resistant to change. They may also need more time to comprehensively understand the newest technology. Understanding the human factor is critical to ensure the information in the tooling is comprehended and utilized by all team members.

A study in information-avoidance behavior substantiates this view by modeling a stressor-strain-outcome framework. The intent is to qualify this avoidance as a consideration of network fatigue in comparison with social networking [4] (Guo, Lu, Kuang, & Wang, 2020). This concept evaluates the thresholds in which the time pressures, data irrelevance, information overload, and data fatigue indicate the predicted result of information avoidance. The study illustrates that stressors are created through information irrelevance, information overload, and social overload. These stressors become a strain when the individual reaches the point of network fatigue. When time pressures are factored in, the rate of information avoidance also increases; the calculation is noted in Figure 1 below.

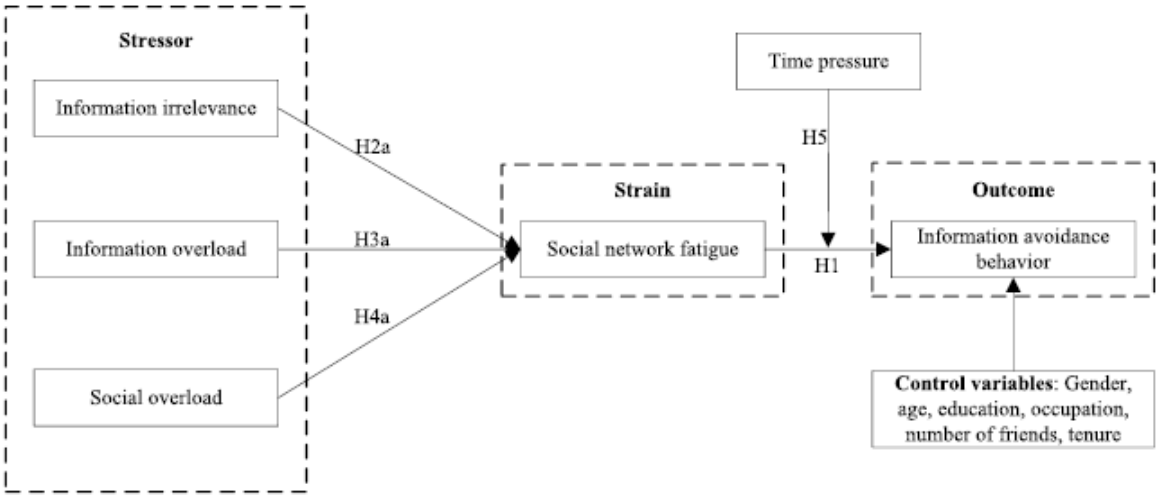


Figure 1 - Information Avoidance Calculation

This study is relevant to the concepts of exposing large quantities of information through system tooling with added time pressures. As an industry, the teams that utilize the system tools are typically held accountable to a very high level of time pressures, intrinsic with responding to system or customer issues. There is likely to be a broad and deep spectrum of data available to our team members that may not be the most relevant to the issue they are working on at that moment. The constant interaction with the networking data can conceptually be an aspect of network fatigue. All of these elements are factors in the introduction of tooling technology to ensure they are not creating even higher incidents of information avoidance.

8. The Solution

Creating a solution to the complex world of supporting tools for advanced technologies relies on multiple factors. The initial assessment comes from a survey of operational teams to identify how many system tools the team members are using and what information they value. Depending on the organization's

history and tooling strategy, it is entirely possible that a single department team could use up to 100 different electronic system tools to complete their work.

As a result, organizations must not only invest to advance network and consumer technology, but must also commit to the system tools needed to support it. Definition of the development culture, blended with leadership influence, creates the foundations of agility and success. Figure 2 shows a view of a tool development ecosystem. It begins with the need to ensure that regardless of what changes are being made, the tenet of operational reliability and stability are foundational to the business.

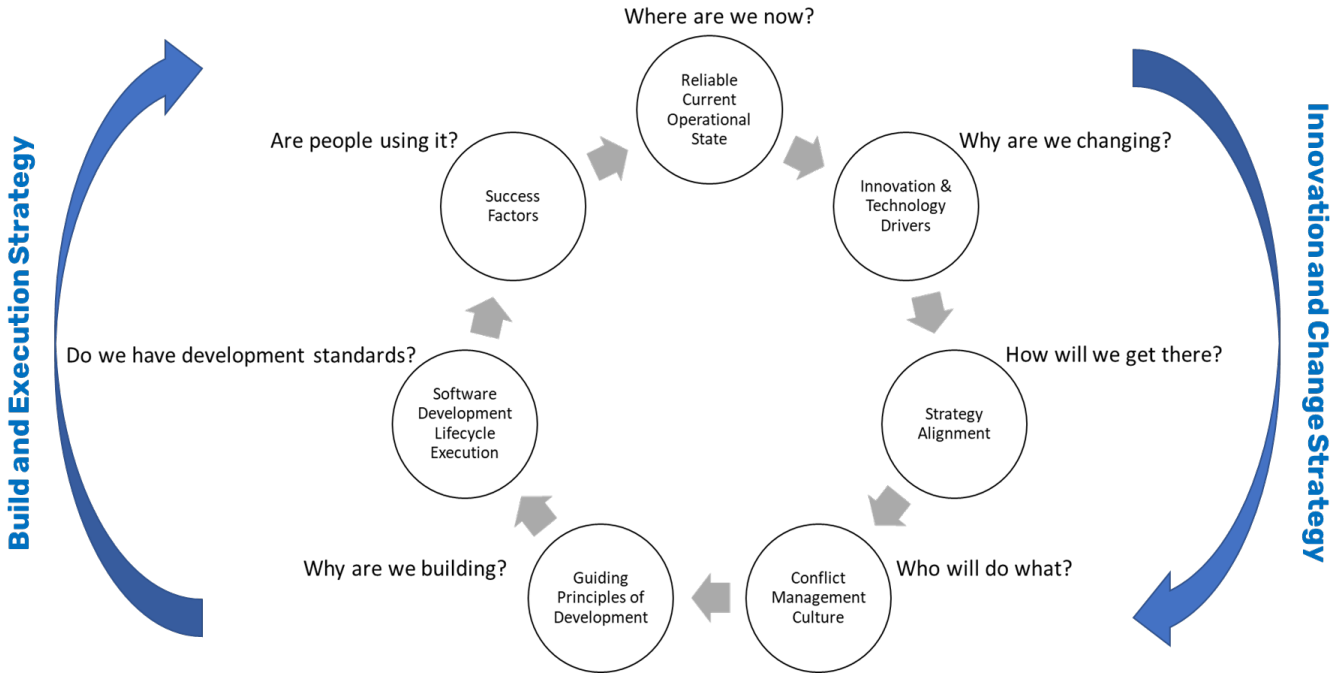


Figure 2 - Software Tooling Development Strategy

Change Drivers

The pace of innovation has primarily driven the need for tool evolution. The history of the technology evolution creates the basic drivers for much of the need to evolve the tooling systems. Comcast colleague and Engineering Fellow Larry Wolcott (2021) [8], a recognized industry expert on the topic of Proactive Network Maintenance (PNM), shared his insights about tooling for this paper. He noted that it is critical to recognize the drivers of the development changes and to be sure to integrate ideation and capability requirements when operationalizing a tooling system or update.

There are typically four primary drivers of change in tooling evolution: new industry technology, tool infrastructure technology, organizational process improvement, and enhancing the user experience.

Figure 3 - Drivers of Tooling Change

Strategy Alignment

The first step in optimizing tooling development for both the development teams and the teams that ultimately use the tools is to agree on a unified strategic planning process, specifically for software-developed systems and electronic tooling. Simply deciding if the organization will be centralized or decentralized will aid strategic alignment. As previously discussed, methodologies can and do differ from department to department, and there are various visions on what is valuable in the data that is presented to internal users.

Coordinating on any strategy, including tooling, starts with leadership alignment, both top-down, and ground-up. Department and area leaders need to be committed to a discipline that outlines the requirements for tooling development. Governance matters for the organizational accountability that integrates all of the change drivers into a cohesive strategic plan. The strategic plan should address standards for tooling technology, coding languages, reliability requirements, financial payback, and business value determination.

The strategy and structure for governance require a collective of experts representing the knowledge wheelhouses of engineering, software tooling development, finance, and operations. All must be equally represented to ensure that the expertise of each discipline is integrated, which creates a balanced business strategy. Engineering experts are responsible for the architectural vision of the technology that supports products. The software development experts are responsible for maintaining tooling technologies that support the monitoring and operations of the engineering systems. The operations experts utilize the software-developed systems and electronic tools needed to maintain and operate the engineering systems. The finance team is responsible for maintaining overall business financial objectives. Each of these groups is critical. Collaboratively, they create a comprehensive business strategy that directly influences the creation of operational support tools.

When the strategy is well balanced, the tools that are used to operate these systems are fully adopted, truly optimized, and streamlined. Each team has a critical role in creating a highly collaborative and focused strategy that optimizes overall business strategy and value. When they are not in balance and the strategy is not aligned, the abyss of tools grows, complexity increases and conflict can slow down the pace of innovation and development.

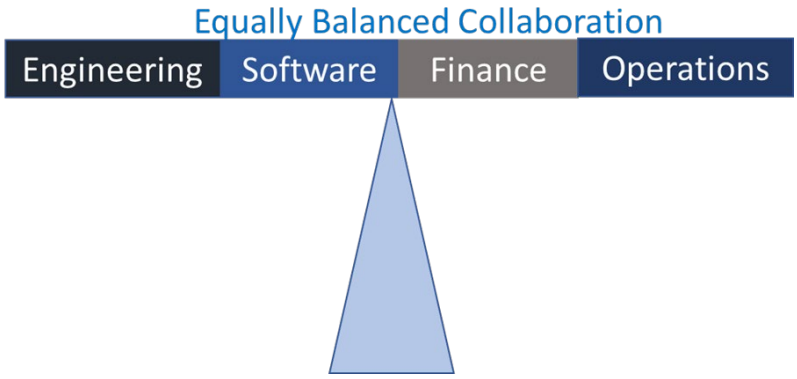


Figure 4 - Balanced Business Strategy

Conflict Management Culture

To maintain a balanced strategy, an organization needs to define expectations for resolving inevitable conflicts so that working teams have a clear path to work together positively. A collaborative conflict-management culture can have a positive impact on organizational effectiveness and improve job satisfaction [3] (Choi & Ha, 2018). Successful conflict navigation relies on cognitive flexibility, self-awareness, and commitment to collaboration. When the teams have more clarity on aligned goals and operate collaboratively, there is an opportunity to enhance innovation and create healthy tension. In this application, teams working in technology need healthy tension to innovate and strive for constant improvement. When there is trust and willingness to be more flexible, individuals feel that their contributions are valued. In turn, job satisfaction will be higher, and internal users will adapt to change and resolve conflict most effectively.

Executive leaders are tasked to define organizational responsibilities and create a process to entrench an inclusive and collaborative conflict-management culture into the entire organization. A collaborative conflict-management culture relies on individual cooperation to create a compromising versus competitive method for deriving agreements [3] (Choi & Ha, 2018). As a metric-driven industry, this concept for tooling development represents a notable shift from the competitive dynamic. An innovation strategy that is only focused on what is best for the technology, for purposes of a competitive result, can create conflict with financial standards, or challenges for operations. A compromise-oriented solution provides the potential to create a healthy tension in the organization that can lead to innovative breakthroughs and positive organizational influence.

When there is a non-competitive, collaborative environment, innovation is maximized, and integration is optimized. Larry Wolcott, mentioned previously, notes the engineering concept that "just because we can, doesn't mean we should" mindset. This concept means that even groundbreaking technology can be available, but if operational teams find it too complex or there are conflicts with operational efficiency, it may not reach its full potential. To ensure that balance is maintained between engineering and operations, there must be a focus on collaborative problem solving and conflict management.

Development Principles

The technical sprawl of tooling is expanded or contained by the software and electronic tool development strategy. Eric Wall (2021) [7], a leader of the support teams that develop several of Comcast's widely utilized operational tools, shared the perspective that the growth of tool development must have a clear, collaborative strategy. Without recognizing business benefits and how operators utilize them, the software developed may not have the intended impact. To aid in ensuring the strategic direction of development, below are a few guiding questions to drive the most efficiency in software tool development.

- Is this requirement strategically aligned and governed?
- What is driving the change or the problem that needs to be solved?
- Who and what development team is responsible for solving it?
- How will the user behavior be influenced, and what is the behavioral expectation?
- How does this fit in the business financial model?
- Should this be developed internally or purchased?

The answers to those questions will clarify new tooling initiatives and align with the expressed strategies of the business. As noted previously, the drivers for creating a large set of tools are driven by many factors: Building new without a plan for the old, "because my team can," and gap functionality all contribute to the steady growth in tools volume. To resolve this continuous sprawl, discipline and commitment adherence to guiding principles are paramount.

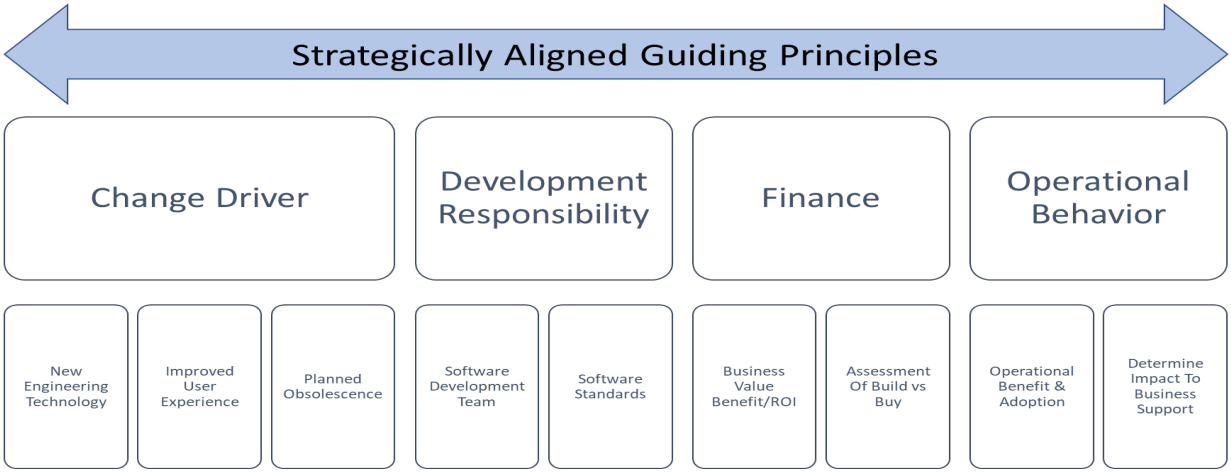


Figure 5 - Strategically Aligned Guiding Principles

Lifecycle Development Execution

Lifecycle execution for tools and tooling should be a structured approach based on the resources available and the overall business requests. The concepts of people, process, and technology are critical components to prioritizing work for development [7] (Wall, 2021). Development standards are set by software engineering experts to ensure a commitment to security and development standards. The challenge that developers face is that of a large number of requests to complete work versus a set amount of dollars to accomplish them.

"Cost modeling can lead strategy" is a common axiom in determining business value drivers for development [7] (Wall, 2021). Cost modeling is a strong influencer of development prioritization, along with engineering, tooling software, operations, and finance, all are essential parts of a comprehensive strategy. With this in mind, the recommendation is to create a foundational basis for development work prioritization by clarifying how each influences the lifecycle.

- Engineering Deployment Of New Technology
  - If tooling systems are not built or updated for the new technology, what experiences would fail?
- Software Development
  - If there was no investment into the systems to maintain reliability, modernize, and grow in capacity, would the system tools become unreliable or unusable?
- Operational Requirements

    - How will features and functionality improve the experience and drive business value?
    - Will there be improved efficiency or saved time per task?
- Financial Requirements
    - Is there a positive return on investment (ROI) analysis of the cost to create the tool or enhancement compared to the value benefit for the business?

Overall, even when the answers to some, or all, of these questions, are yes, the development work still needs to be prioritized to optimize developer time. All of these factors require equal consideration when planning the development work. Discipline in the business strategy and execution lifecycle is the best way to ensure the internal users of the tools are utilizing them most effectively.

Success factors

In the overall business strategy, success factors can rely on many considerations. The first leading indicator of successful development is whether or not the tool has been adopted and utilized by its internal users. Simply put: Are people using it beyond the break-in period? Positive feedback would indicate that the work completed has made their job simpler or better in some way. The success criteria of adoption and utilization is a leading indicator of success, because if people are not using it, then there was no value in building it.

"Success equals adoption. If they adopt it quickly, they value it, therefore, they're less likely to work around it."

-Larry Wolcott, Engineering Fellow, Comcast

The second leading success factor is recognizing how the development of the tool is benefiting the business. This factor requires evaluating that the desired impact was achieved and is positively influencing user behavior. Too much information at the wrong time, or the creation of swivels, can cause team members to take more time to accomplish a task. Therefore, success comes from the reduction of time per task by improving the efficiency of the person using the tool. This reduction in time also creates a component of the payback model by illustrating the cost of build versus the improvement in technician efficiency. A hypothetical example of the influence of calculating task time is in Table 1, below. If in pre-development a task takes 5 minutes to complete, and post-development it takes 3 minutes to complete, the savings are demonstrable.

Table 1 - Business Benefit Calculation (Average Cost Is Illustration Only, Not Actuals)

	Task Quantity	Time Per Task	Average Hourly Cost	Task Cost
Pre Development	100,000	5 min	$60	$500,000
Post Development	100,000	3 min	$60	$300,000
Total Business Benefit				**$200,000**

The third success factor is confirmation that the tool developed achieved the intended expected results and business behaviors without creating more tools for the technicians and other internal users to use in the course of completing their work. The strategy to innovate and create new must be balanced with the roadmap of the legacy technologies and not perpetuate information overload for the users. The success could be measured by a metric to create more technician efficiency, such as reducing the time per task and thus improving operational metrics.

"The gauge of success is when the tool accomplishes a meaningful task and the percentage of users is high".

-Walter Breymeir, Field Operations Comtech, Comcast

The perspective of the end-user is the final success factor. The tool must be easy to use and provide the best information at the right time, without overwhelming, and be quickly adopted. The tool must be viewable in a consolidated way, to 1) minimize swivel, 2) reduce confusion on when to use the tool in what situation and 3) be easy to learn for its users.

As well, and from a good sense perspective, the tool must be cost-effective to be available to the broadest number of users for the least amount of cost. Choosing the right balance of information and finding the most cost-effective way to reach the largest number of users provides an avenue to optimize adaptability and usability. The information provided is a vital factor in preparing the training for new technology and supporting tools. The tooling systems must be intuitive and easy to learn. Training must be provided to the end-users to not only teach them new technology but to provide the team members time to absorb and learn.

"A tool could be great, but without training and it being easy to learn, it will likely fail to be used by the majority."

-Robert Snare, Field Operations Comtech, Comcast

9. Conclusion

The tooling abyss is a real and reasonably universal consequence of industry consolidation and the after-effects of growth-by-acquisition. There are the tools traditionally used by the field, then there are the "new" tools that come in through the different (acquired) doors. The tooling abyss is similarly fed by the technological advancements in adjacent and over-arching technology sectors, thus creating a myriad of information that needs to be quantified and managed to operate the business.

It is a top-down and bottom-up challenge: Large companies that became larger because of industrial and geographic consolidation are particularly susceptible to performance challenges associated with duplicative tooling. And, strong emotional attachments to favored or long-used tools, at the individual and small-group level, as well as the earnestly-developed patches and workarounds developed over time, can make it difficult to "centralize" tooling and related support.

The outcome is "tool sprawl," which is evidenced by the amount of new swiveling teams need to do across multiple applications, which serves only to add time to task completion and raise frustration. While it is true that as service providers, we "sit on" mountains of actionable data, the definition of "actionable" requires continuous attention. Too much information at the wrong time can lead to information avoidance and overload, both of which hinder operational performance.

When cross-department technology teams are not aligned, tool development can overlap, which also drives inefficiency for the business. There is a business cost when there is competitive conflict, inadvertently causing developers to work against each other, usually to achieve a desired or published performance metric. Strategic alignment is crucial to ensuring a collaborative instead of conflicting culture exists for the teams.

Therefore, a clear and comprehensive business strategy needs to provide clarity to streamline both the overall technology development process and the tooling process. Collaboration within a cohesive strategy is what generates an environment of optimized innovation. When the development lifecycle is optimally

based on a balance of experiences and wisdom representing engineering, software, operations, and financial elements, the resulting tooling products will drive business benefit most effectively.

Success factors hinge on meeting the needs of the internal users and business operations. Development teams need to ensure that system reliability is a foundational cornerstone because unreliable systems will not be adopted and utilized by users. Listening to feedback provides insight into why or why not the tool is being used. Two-way communication provides a conduit for information sharing for both the developers and the users. This avenue allows developers to aid in training and learning while receiving critical feedback on improvement opportunities.

"The measurements of reliability, adoption, and feedback determine why or why not tools are utilized and successful in meeting the needs of the business."

-Eric Wall, Executive Director Software and Development Engineering, Comcast

The tooling abyss can be avoided through an overall alignment in business strategy that integrates the views of multiple teams. Engineering, tooling software developers, operations, and finance teams must work together to commit to a business strategy that is effective for all. When there is a strategic commitment, innovation is increased, conflict is positive, complexity is reduced, and the overall health of the business improves. The tooling abyss ultimately becomes a manageable pool of critical resources widely adopted and utilize to optimize business initiatives.

Abbreviations

CI/CD	Continuous Innovation/Continuous Development
CMTS	Cable Modem Termination System
DOCSIS	Data Over Cable Service Interface Specification
HFC	Hybrid Fiber-Coax
MAC	Media Access Control
PNM	Proactive Network Management
ROI	Return On Investment
UI	User Interface

Bibliography & References

[1] Breymier, W. (2021, July 15). Technician Perspective [Telephone interview].

[2] Cable History. (2014). The Cable History Timeline. Retrieved from https://www.cablecenter.org/images/files/pdf/CableHistory/CableTimelineFall2015.pdf

[3] Choi, Y., & Ha, J. (2018). Job satisfaction and work productivity: The role of conflict-management culture. *Social Behavior and Personality, 46*(7), 1101-1110. doi:http://dx.doi.org.libauth.purdueglobal.edu/10.2224/sbp.6940

[4] Guo, Y., Lu, Z., Kuang, H., & Wang, C. (2020). Information avoidance behavior on social network sites: Information irrelevance, overload, and the moderating role of time pressure. *International Journal of Information Management, 52*, 102067. doi:10.1016/j.ijinfomgt.2020.102067

[5] Snare, R. (2021, July 16). Technician Perspective [Telephone interview].

[6] Voinea, C., Vică, C., Mihailov, E., & Savulescu, J. (2020). The internet as cognitive enhancement. *Science & Engineering Ethics*, *26*(4), 2345–2362. https://doi.org/10.1007/s11948-020-00210-8
[7] Wall, E. (2021, July 9). Developer Perspective [Telephone interview].
[8] Wolcott, L. (2021, July 8). Industry Expert [Telephone interview].

The WiFi Happiness Index (WHIX)

A Technical Paper prepared for SCTE by

Krithika Raman
Vice President Product Engineering
Comcast India Engineering Center LLP,
Chennai One SEZ, 5th Floor, North Block in Phase II,
Module 7 & 8 - Chennai One,
Pallavaram-Thoraipakkam 200 Feet Road,
Thoraipakkam, Chennai - 600 097,
Tamil Nadu, India
+919152007903
Krithika_raman@comcast.com

Charles Moreman
Executive Director of System Architecture
Comcast Cable
1800 Arch Street, Philadelphia, PA 19103
Charles_ Moreman@comcast.com

1. Introduction

During the 2018 SCTE Cable-Tec Expo, panelists in a "birds of a feather" RDK session introduced the concept of a "Wi-Fi Happiness Index," abbreviated WHIX. Described as a weighted, realistic view of Wi-Fi behavior in a home, the WHIX index may involve more than 50 parameters, and very large quantities of machine-derived data. Since then, the WHIX concept has grown in breadth and depth, and is the topic of this paper.

The paper will cover what it takes to deliver per-device WiFi metrics, normally gathered across the industry, that contribute to an aggregate / whole house WiFi health assessment; what factors shape high and low WiFi performance; algorithmic detail regarding the non-linear aspects of whole-home "happiness;" and the value of customer feedback in informing health-centered WiFi metrics.

The paper will also cover best practices that can be proactively applied to WHIX assessments, in terms of recommended next steps. For instance, in situations where the WHIX score may indicate a suboptimal WiFi experience, a series of complementary algorithms can be run to evaluate different recommendations for improving the WiFi connection. For example, an algorithm can be run to determine if Wi-Fi extender "PODs" are recommended for this WiFi network. Attendees will also learn how WiFi health assessments can be applied to triage situations, and for engineering-level troubleshooting of new software releases.

1.1. The WiFi Happiness Index

WiFi is complex: The user experience of someone using WiFi is affected by multiple different WiFi conditions, occurring simultaneously. Even a WiFi subject matter expert can find it difficult to consider all the various WiFi conditions, assess the health of the WiFi network and determine how these conditions interact to affect the user experience. WHIX can be designed to process this complexity, with a goal of making it easy to understand whether a customer is having a good or bad experience. WHIX can also provide details on what is affecting each customer's WiFi experience.

WHIX can measure how well a WiFi network is working by modelling the WiFi user experience. It can provide details on specific WiFi conditions and how they correlate to good or bad WiFi connectivity. WHIX can assess the WiFi user experience at two levels: Per client device and for each WiFi network as a whole. Specifically, it can measure the WiFi connection on every client device in every network, every hour. It may also use this client device information to measure the user experience of each WiFi network. As API-based software, it may run in the cloud and can be integrated into other internal tools and UIs, including those used by field technicians to find and fix customer problems.

This may be used to assess the user experience for each specific WiFi client device. WHIX may also make it possible to understand the overall user experience provided by that WiFi network. Together, the per-device and per-network assessments may inform and model the WiFi user experience in a way that is intuitive and comprehensive.

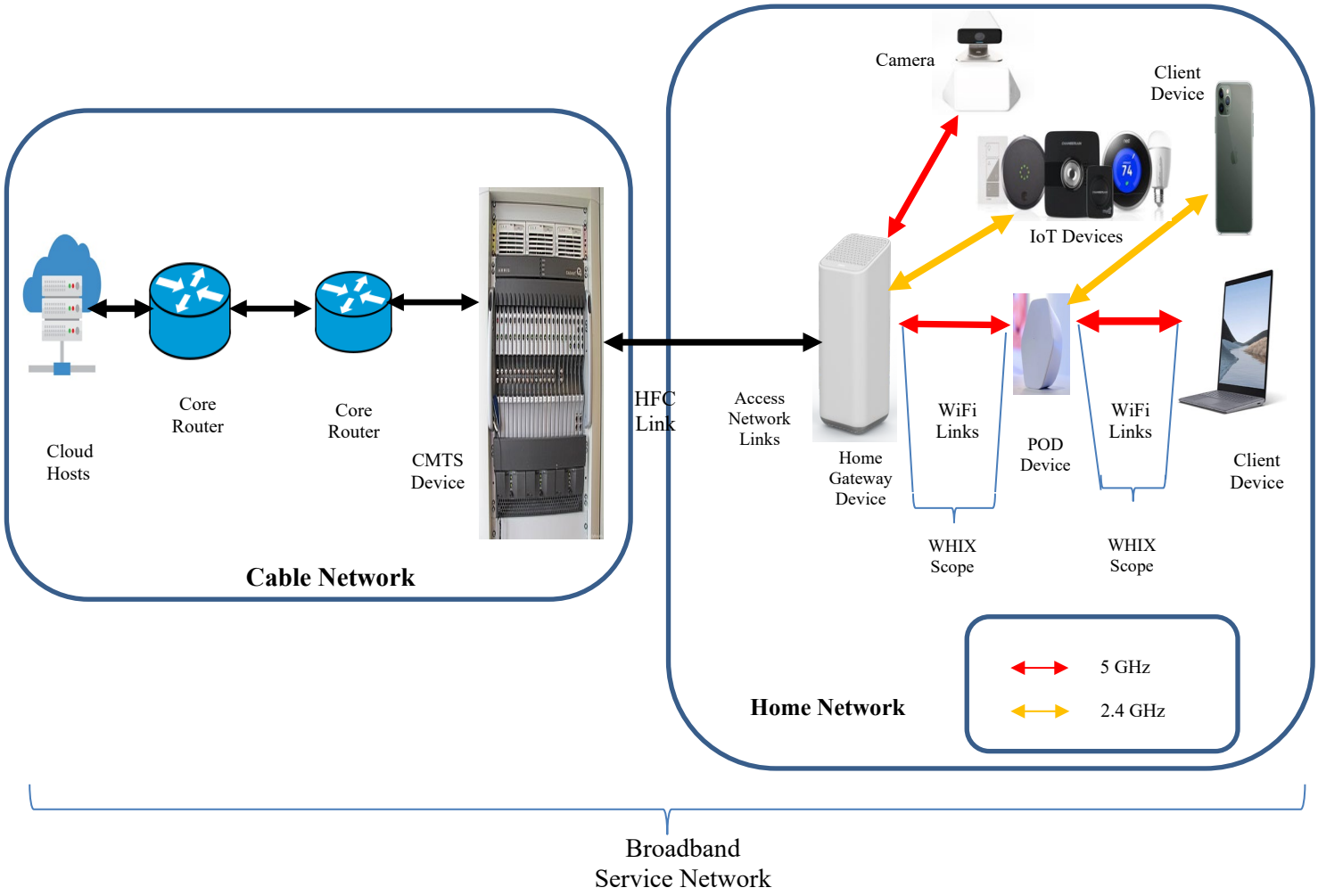


Figure 1 - The scope of the WHIX in an end-to-end broadband network

Figure 1 shows a typical end-to-end residential WiFi network. The cloud hosts the WHIX server components. The cable network component shows the typical network components, such as Core Routers and Cable Modem Termination Systems (CMTSs.) The "last mile" is over Hybrid Fiber-Coax and is connected to the Home Gateway Device (a WiFi router). Together, the Home Gateway and the WiFi extender "PODs" provide the WiFi service.

The WiFi router and WiFi extender PODs may use either dual-band (2.4 GHz + 5 GHz) or triple-band (not shown; 2.4 GHz + 5 GHz + 6 GHz) WiFi technology to provide high-speed connectivity. The WiFi extender (POD) is optionally used to improve the signal strength of WiFi in specific locations of the house. The various client devices shown in Figure 1 include a mobile phone, laptop, camera, and home automation devices, including a touch screen and various sensors. The Home Gateway device is

connected to the POD, client laptop and to the mobile device(s) on the 5GHz channel; the touch screen and the camera over the 2.4 GHz channel.

The scope for the WHIX WiFi assessment is all the WiFi connections (both the 2.4 GHz and the 5GHz) that form the end-to-end path between the WiFi gateway and each WiFi client device. Other links, including the core network and access network, are not in scope for the WHIX WiFi assessment.

WHIX uses a machine learning (ML) model to process raw telemetry data and create WHIX "scores" that can model the user experience. Each score is based on a 0 to 100 range. 100 may indicate the best possible user experience, and 0 may indicate the worst possible user experience. In practice, most scores fall on a sliding scale between these two extremes. Different industry implementations could use different score ranges based on the granularity of the score that is desired and the data available.

Figure 2 shows a sample of score interpretations.

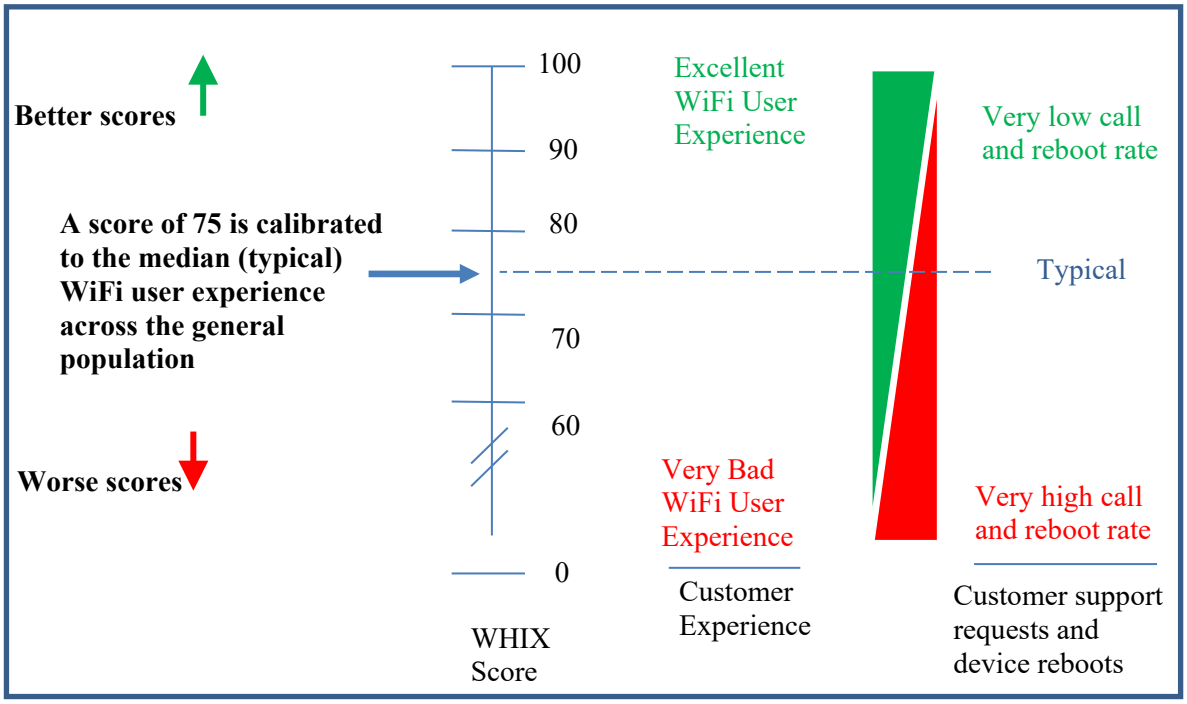


Figure 2 - The WHIX Uses Human-Readable Labels

The WHIX machine learning (ML) model can be trained using billions of historical WiFi data points collected over time from millions of WiFi networks. The ML model can be calibrated to provide a score of 75 when the user experience matches the median (typical) user experience. Scores greater than 75 may indicate a better than typical WiFi user experience. Scores less than 75 may indicate a worse than typical WiFi user experience.

A low WHIX score generally indicates a poor user experience and often results in a higher rate of customer support requests to the connectivity provider.

WHIX Scores may be mapped to human-readable labels. These labels may include:

- Excellent (scores >= 90 and <= 100)
- Very good (scores >= 80 and < 90)
- Good (nominal) (scores >= 70 and < 80)
- Bad (scores >= 60 and < 70)
- Very Bad (scores >= 40 and < 60)
- Extremely Bad (scores >= 20 and < 40)
- Worst (scores >= 0 and < 20)

Multiple factors simultaneously affect the user experience on a WiFi network. The WHIX assessment can consider both positive and negative factors to create these scores. When the positive factors are more significant than negative factors, the WHIX assessment score may be greater than the median score of 75. Conversely, when negative factors outweigh the positive factors, the WHIX score may be less than 75.

2. Telemetry data used by WHIX

One of the most common problems seen with WiFi is the change in the signal strength at different locations within a residence. WiFi signal strength can be a problem when the signal is too weak -- and can also be a problem when it is too strong. Weak signal strength can be caused by longer distances between WiFi devices and by obstacles that impact signal strength. Weak signals can be difficult for the receiver to reliably discern from background noise and interference. Very strong signals can sometimes occur when WiFi devices are too close to each other. Strong signals can overwhelm the dynamic range of some WiFi receivers and cause errors. WiFi tries to adjust for poor signal strength by changing how the WiFi signal is modulated. A lower signal modulation index can improve the reliability of the WiFi reception, but can have adverse effects, like lower throughput and higher channel utilization for a given bit rate.

Another common WiFi problem is high WiFi channel utilization. This is often expressed as the percentage of airtime used on the WiFi channel. The available airtime on a channel is finite, yet is used for multiple reasons. Airtime is used by the WiFi stations and access point(s) (APs) in the WiFi network being assessed; by WiFi networks in neighboring homes/businesses; and by non-WiFi RF sources such as baby monitors and microwave ovens. When a greater percentage of the available airtime is consumed, WiFi devices must contend with other WiFi devices to access the WiFi channel. As a result, the WiFi channel utilization increases. High WiFi channel utilization can cause the following undesirable scenarios:

1) *It becomes more difficult for WiFi stations and access points to send traffic.* During periods of high channel utilization, WiFi devices often wait for other WiFi devices to finish transmitting before the begin sending their data.
2) *There's a greater chance that multiple WiFi devices will attempt to begin transmission at the same time*. This can cause collisions, such that none of transmissions can be received.
3) *It can cause higher latency*. Higher latency means that packets incur longer delays. This reduces the user experience of delay sensitive applications like interactive games.

It follows that lower WiFi channel utilization can cause improved latency, improved throughput, fewer dropped packets -- and an improved customer experience.

Many other factors can affect someone's WiFi experience. These include downlink and uplink PHY rates, MCS (Modulation and Coding Scheme) levels, the background RF level (also called the "noise floor"), and Signal to Noise Ratio (SNR). Also impactful: The rate of packet re-transmissions on each client device; the rate of lost packets (when all retransmissions fail) on each client device; service availability (percentage uptime) of the access points in the WiFi network; and the ability (or inability) of WiFi clients to sustain the WiFi connection (measured as "rapid reconnects"). The total number of WiFi clients on each WiFi channel can impact performance, as can the rate of WiFi password failures, the configured maximum channel width, the operating channel width for each client device, media access delay for transmitted packets, the ability (or inability) of the network to steer each client device to a closer AP or steer that client device to a better WiFi band, QOS configuration/usage -- and more. It's a long list!

These WiFi factors tend to affect different WiFi client devices in different ways. This is because different types of client devices use WiFi very differently. For example: WiFi cameras use more uplink bandwidth (in the direction toward the access point) than downlink bandwidth. This makes WiFi cameras more susceptible to problems that impair uplink data transmission. WiFi media players and set-tops use more downlink bandwidth, so these are more susceptible to downlink-related problems. Other devices, like IoT sensors and thermostats, use very little bandwidth and can successfully operate at low PHY rates -- but must have extremely reliable connections. Still other devices, like game consoles, tend to be more sensitive to latency. WHIX may use a mapping of client device "fingerprint" data to categorize client devices into one of multiple WiFi device categories. WHIX may consider the WiFi device category when analyzing the WiFi user experience for each client device.

These factors can work in combination to affect multiple WiFi links used to create the end-to-end connection for each client device. For example: when a client device is connected to a gateway that provides both a WiFi Access Point and the home's WAN service connection, then only a single WiFi link is used in the end-to-end path between that client device and the WAN connection. In the case where a client device is connected to an access point provided by a WiFi extender POD, additional WiFi link(s) can be used to backhaul traffic to the WiFi gateway that includes the WAN service connection. In this case telemetry data from multiple WiFi links may be analyzed to understand the user happiness for that client device.

This telemetry data may be collected by both main WiFi gateway access points and WiFi extender POD access points. WHIX may import this data and use it to measure how well WiFi is working for each client device and for each WiFi network as a whole.

3. WHIX algorithm overview

The algorithms that can be used to create a WiFi Happiness Index may use a supervised ML training model based on this WiFi-related telemetry data. Telemetry data may be sourced from each of the Home Gateways and WiFi extender PODs, in millions of homes, and aggregated as anonymized high level statictics to train the ML model.

Client types spanning nine WiFi categories may be derived from device fingerprint data (see Table 1). In a typical iteration, the training model may examine >4K specific WiFi conditions over these nine different client device types. This model may form the core of the WHIX analysis and feed the outputs of the WHIX API.

Table 1 - Types of WiFi client device categories assessed by WHIX

WiFi Client Device Types Categories Assessed by the WiFi Happiness Index
1. Broadband Devices: Smartphones, laptops, tablets
2. Media Players: Smart TVs, streaming video devices, streaming dongles
3. Media Sources: Cameras, streaming servers
4. IoT / Printers: Thermostats, lightbulbs, sensors, and networked printers
5. Managed Set-tops
6. Managed WiFi Extender PODs
7. Audio Devices: Smart speakers, connected audio devices
8. Network Devices: Customer-owned WiFi extenders, etc.
9. Game Consoles

4. Training the Machine Learning Model

Multiple feedback mechanisms can be used to determine WiFi conditions associated with user happiness (or pain) with WiFi connectivity. These feedback mechanisms can include User Contact Rate statistics and Unscheduled Reboot Rate statistics.

Customers contact their service provider for many different reasons. Some contacts are for questions about billing or if the customer wants to upgrade/downgrade service. Other contacts occur when the customer needs help to resolve a WiFi-related problem. As an indication of WiFi-related user happiness, customer contacts may be filtered for contacts related only to connectivity problems. After filtering, connectivity-related contact rates may be compared with connectivity-related contact rates for the general (total) population of all customers. A higher rate of connectivity-related contacts may be associated with a poorer WiFi experience. A lower rate of connectivity-related contacts may be associated with a better WiFi experience.

Gateway reboots can occur for a variety of reasons. These may include “scheduled” reboots that coincide with software upgrades during a late-night maintenance window. Reboots can also occur during “unscheduled” times. Unscheduled reboots can occur as a result of power outages, software “self-heal” mechanisms and because a customer reboots the device. Many customers may reboot their WiFi gateway when experiencing a WiFi related problem. A higher rate of unscheduled gateway reboots may be associated with a poorer WiFi experience. A lower rate of unscheduled gateway reboots may be associated with a better WiFi experience.

WHIX ML models may be trained using statistical feedback from both user contacts rates and unscheduled reboot rates. Different ML models can be trained for different WiFi client device categories.

First, each WiFi-related condition (WiFi criterion) may be mapped into a range of values called a "bin". Bins may be constructed based on the distribution of telemetry data, so that enough Boolean "matches" occur within a bin range (where possible) to have sufficient data to avoid noise. For example: a bin may be used to characterize channel utilization (CU) criteria in the ML model might include all CU telemetry values greater than 22% and less than or equal to 25% for the client device category of "Media Player" devices.

Next, correlation may be measured for each WiFi criteria bin with customer contacts related to connectivity. For each bin (range of values for a single WiFi criterion), the percentage match rate in the population that made contact with the service provider for connectivity-related issues within the previous 24 hours may be calculated. This may be calculated by dividing the number of criteria matches for this bin range (which occur in the customer population that made a connectivity-related contact in the previous 24 hours) by the total number of telemetry reports received for any value of this WiFi criterion.

Then, the "Criteria Match Difference" (CMD) may be calculated to compare this match rate by this criteria match rate in the total general population. For example, if the criteria match rate for this bin range in the population with connectively-related contacts is 10% higher than the criteria match rate for this bin range in the general population, then Criteria Match Difference = 110%. CMD values equal to 100% may indicate that when conditions match this criteria bin range, there may be no effect on customer happiness. CMD values higher than 100% may indicate a poorer customer experience compared to the general population. CMD values lower than 100% may indicate a better customer experience compared to the general population.

The same process may also be applied using feedback from unscheduled reboots. This may establish correlation between each specific WiFi criteria bin range and customer happiness as indicated by unscheduled reboots.

This process may be repeated for more than 4,000 combinations of WiFi criteria bins across the 9 client device types.

Based on this correlation data, an algorithmic weight may be calculated for each of the 4,000 WiFi criteria bins. Positive weights may indicate WiFi conditions that correlate to WiFi user happiness that is better than the general population. Negative weights may indicate WiFi conditions that correlate to WiFi user happiness that is worse than the general population. Zero weights may be used to indicate correlation with the typical WiFi user happiness level (same user happiness as the general population). For each WiFi criteria, the weights for all bins in the ML model may be normalized such that positive weights offset (balance) negative weights across the total population.

The ML models may be re-trained for every WHIX software release. This may recalibrate WHIX to support and track with all new WiFi Gateway models, extender PODs models and software versions. Each new WHIX software release may typically also expand the ML model, to include new WiFi telemetry data that may become available since the previous release.

5. How WHIX uses the Machine Learning Model

After the ML model may be trained and validated, it can become available for use by WHIX. The following is a description of how WHIX could use this ML model.

On each telemetry interval (currently each hour), WHIX may use the ML model to assess the WiFi conditions for every client device in every home. First, all available WiFi telemetry can be collected from

the WiFi gateway and from all extender PODs in that home. Telemetry for device fingerprinting can also be collected and used to map each client device into one of the 9 WiFi device categories. The WiFi telemetry data may be used to create a Boolean match for a specific bin range of each WiFi criteria.

- For example, if fingerprint data maps client device "A" to the WiFi device category of "Media Player" then the section of the ML table for Media Player devices can be used.
- Then all WiFi telemetry may be compared to the bins defined in the ML model. For example: telemetry for 2 GHz Channel Utilization is collected for all access points used in the end-to-end path for client device "A". If the telemetry indicates 10% channel utilization on the WiFi gateway access point and 42% channel utilization on a WiFi extender POD used in the end-to-end path for this client device, the larger (worst) telemetry value from these 2 WiFi link may be used.
- In this case, a channel utilization telemetry value of 42% may match the 2 GHz channel utilization criteria bin may be defined as "greater than 40% and less than or equal to 45%". Once this Boolean match is determined, the weight for this bin may be used to indicate the user impact of this individual WiFi criterion. In this case, the weight may be -2 for this bin range of Channel Utilization for media player devices. A weight of -2 may indicate that this condition may correlate with a very slight negative impact to the user's WiFi experience.

Next, the process may be repeated for all other telemetry data that was collected on all WiFi link(s) used for the end-to-end path to client device "A". Each additional telemetry data value may be processed by the ML model to add an additional weight for each WiFi criterion.

All weights for all criteria matches for client device "A" may then be summed and added to a value of 75 that can be used tonormalize the WHIX score to the typical score in the general population. This calculation may provide the WHIX score for client device A for the most recent hour. If the WHIX score is greater than 75, this may indicate that WiFi conditions on the links used for this client may provide a better-than-typical user experience. If the WHIX score is less than 75, it may indicate that WiFi conditions provide a worse-than-typical user experience. Previous scores may be retained to support historical views and trend analysis for this client device.

Next, the WiFi network for that home as a whole (WHIX account score) may be evaluated. This process may use the client device scores in that home to calculate an overall WHIX account score for each hour. This may involve analyzing the client device scores to determine which client devices provide the worst WiFi experiences in that hour. The set of these worst devices may then be used, along with an adjustment to recalibrate the median (typical) account score to 75, to create the WHIX account score for that hour.

WHIX can be analyzed over time. Various statistical analysis methods may be used to aggregate WHiX time series data and characterize the WiFi user experience over extended time periods. Statistical methods that provide the best predictor that the customer will/will-not contact the service provider and/or reboot their gateway device within the next 24 hours may have the best correlation with user WiFi happiness.

6. Benefits of WiFi Performance Monitoring

The WHIX concept may enable service providers to determine the quality of each customer's WiFi experience, per connected device and for every customer's account as a whole. This can enable service providers to see the most sample interval (based on available telemetry data) and also see a timeline that shows historical views, including any times when the customer's WiFi experience was materially better

or worse. Data may also provide indications of what caused the customer's WiFi experience to be better or worse for each hour. The detailed data may indicate the root cause of what affected the customer's WiFi experience, and can be used to troubleshoot and remediate any problems. WHIX output data may be provided via an API that supports integration into various user interface tools.

Incremental software upgrades dispatched to Home Gateways and WiFi extender PODs may also have an impact on WHIX scores due to changes that have been incorporated into a software version. By monitoring highlevel statistics based on WHIX scores, service providers can determine if the WiFi quality has changed. If the WiFi quality has degraded, coincident with a software upgrade, the problem can be proactively identified, triaged, and remedied with new software that includes the appropriate fix. This may reduce customer contact rates and may help provide a quick fix to a potential customer-impacting problem.

Monitoring the WiFi Happiness Index may help in troubleshooting difficult issues that normally surface over time. It may also help to narrow down whether problems are intermittent or persistent.

WHIX can also be used with other cloud-based tools to create automated recommendations to improve the WiFi user experience for each home. This can include automated recommendations to perform specific targeted action(s) designed to improve the WiFi network in each individual home.

7. Conclusion

This paper illustrates a comprehensive concept for monitoring and assessing the user's WiFi experience at the per-user and per-client device level. The concept provides details that can be used to troubleshoot common customer WiFi problems. WiFi Happiness Index (WHIX) can be based on a Machine Learning algorithm that models how various WiFi conditions can affect user happiness with each of multiple WiFi client device categories, ranging from laptops and phones to IoT devices, for every account holder, every sample interval.

By using this concept, technician and care agents may identify specific client devices and specific accounts that are may experience WiFi-related issues, and, importantly, may recommend solutions to solve those problems.

WHIX can also be used to measure the aggregate user WiFi experience in WiFi gateway and extender POD software releases. This can be measured in early trial groups and/or in lab test environments. By comparing the aggregate WiFi experience on successive software releases, service providers can identify hidden software problems and prevent problematic software releases from being widely distributed. Conversely, it can be used to identify software releases that improve the user's WiFi experience and/or confirm the customer experience benefit from correcting previous software bugs.

Abbreviations

CMD	Criteria Match Difference
CMTS	Cable Modem Termination System
CU	Channel Utilization
HFC	Hybrid Fiber-Coax
ML	Machine Learning
RDK	Reference Design Kit
WHIX	WiFi Happiness Index
WiFi	Wireless Fidelity

Bibliography & References

802.11ax Wiki article: https://en.wikipedia.org/wiki/IEEE_802.11ax-2021

802.11ax spec: https://standards.ieee.org/standard/802_11ax-2021.html

WiFi Alliance: https://www.wi-fi.org

The Zen of Ticketing

Operational Transformation

A Technical Paper prepared for SCTE by

Melissa Wood
Sr. Director Reliability Engineering, Program Management
Comcast
183 Inverness Drive West, Englewood, Colorado
719-491-0457
Melissa_wood@cable.comcast.com

1. Introduction

Enterprise ticketing systems are among the vital subsystems that "keep the lights on" for broadband service providers – or, more accurately, keep service-related lights off. Service-related tickets, not unlike the complex and interwoven landscape of "tools," in general, are foundational for identifying, tracking, and resolving internal and external operations, from scheduled maintenance to incident handling. For those reasons, ticketing systems are often as layered and difficult to detach as multiple layers of wallpaper on a wall -- they got that way through decades of growth, patches, ownership changes, and management shifts (e.g. from centralized to decentralized and back again).

As the title of this paper indicates, this paper is about how to get to "zen" when it comes to incident and trouble ticketing. We define zen as a general state of enlightenment attained by summing the parts, which in this case, are the many different ticketing systems acquired through decades of individual acquisitions that helped defined our industry's consolidation.

In this paper, we will share how and why Comcast pivoted toward a unified ticketing system -- including what constitutes realistic expectations, lessons learned, useful metrics, and best practices. We will characterize what it takes to unify multiple ticketing systems, so as to attain the benefits that come from operational scale.

2. Ticketing Infrastructures: The Backstory

Comcast provides multiple services to customers to help keep them connected, everywhere, all the time, across a large geographical footprint. From traditional cable television, to Internet, phone, security, and wireless, Comcast and the cable industry deliver products every day that consumers expect to 1) be there when needed, which is all the time, and 2) evolve gracefully.

Overall consumer desires for service availability and speed puts extra emphasis on our engineering and operations teams to execute reliability measures in ways that get a "thumbs up" from consumers every time. Moreover, managing changes to the network production environment, so as not to cause self-inflicted outages, is part of what prompted Comcast to move towards a unified and enterprise-wide ticketing infrastructure.

Why does this matter now? The complexity of the network topology, coupled with heightened consumer expectations, requires a scalable solution. We can no longer "throw more people at it," when it means trying to manually understand all the many cross-departmental tickets and alarms, all firing inside an increasingly complex network topology. The answers require people and technology.

Like many of you reading this, Comcast "got here" through multiple acquisitions, as a result of a heavy few decades of industrial consolidation. This coupled with understandable and necessary localized network requirements at the regional and divisional levels, produced differing hardware, software, and – tantamount to this discussion – ticketing systems, especially as it relates to managing outages and network upgrades. As a direct result, processes and tools vary across our different engineering and operations teams.

Collaborating during an upgrade or outage event demands teams chat on different platforms, and to document ticket information in multiple places to keep stakeholders informed. These variations cause unintended delays to resolve the customer's impairment, resulting in their inability to use services they pay for. The ability to scale and reliably manage the network is more important than it has ever been.

Basically, to provide new and highly reliable services to customers, at scale, requires pioneering towards one network, end-to-end -- including ticketing -- across people, processes, and technologies.

Moreover, our employees provide great feedback on a regular basis about on-the-job processes that are difficult or could be approved upon. Now is the time to make those processes straightforward with the same execution regardless of the geographic location. Building a unified enterprise ticketing infrastructure sets the tone and removes assumptions, which gets us closer to full operational transparency.

3. The Challenges of Managing Outages and Network Upgrades

Tools and processes across Comcast vary when it comes to managing outages and performing network upgrades. Any service provider that grew its geographic footprint by acquisition, then resolved to consolidate and cluster its holdings, invariably faces an avalanche of tools, like ticketing systems, that are used regularly, thus are familiar, yet they all do essentially the same thing. Our internal analysis, for instance, revealed numerous siloed applications that existed to support teams doing incident and change management. Such replication of records adds risk to delivering reliable services -- risks that are potentially service-impacting, thus customer-impacting. Like all service-minded broadband providers, we continually upgrade our network to provide more products and services, with the very specific intent to be "always on." This good work can be overlooked, especially if an outage accompanies an otherwise proactive change.

Data or configuration items (CI) tend to be stored in many different tools and can lack a standard way of providing relationships between them. A CI represents service components, infrastructure elements that need to be managed for successful delivery of services. An example is network components, such as, routers, hub, gateway, etc. Technologies vary greatly from group to group, when it comes to how hardware and software is monitored. Teams tend to work outages from the same CI, producing up to 20 tickets for a single impact, which also produces multiple instances of repeat communications.

Impacts in the network drive impairments to applications that both customers and employees use. When individual teams are working in a silo to conduct incident and change management, the number of incident tickets are inflated, and duplicate technical and executive communications are distributed. Change tickets themselves can result in an outage, due to not having robust conflict assessment across the company. While we are in a far better place than we were 5 to 10 years ago, when it comes to "swivel chair" operations, engineers still spend time manually looking up needed information to solve an outage, having to swivel between tools to manually enter data needed either in an outage or change ticket.

Another contributing factor is alarm correlation across topologies and different activities. When alarm correlation is weak, it is often because root cause analysis is often only done on "severity one and two" tickets, which only equates to about a quarter of the workload. Lower-level severity tickets are created only to form documentation to identify patterns and trends. "Parent" and "child" tickets aren't always properly linked, which puts a weak spot again on correlation. The amount of manual work performed by engineering and operations teams can delay outage mitigation, degrade the customer experience, and limit the ability to scale needed changes within the network.

The guiding principle towards any major overhaul to how development and operations teams work together is to create a vision with shared visibility, communication and collaboration that embraces working together in a blameless culture to drive continuous improvements, always.

4. Transparent Ticketing Processes Create Reliability

Developing a transparent ticketing process for higher reliability is easier said than done! Assessing the nuts and bolts of how your organization is operating today to the desired future state will create a shared vision for alignment utilizing best practices others have identified.

A reliability vision spans people, process, and technology. From a people perspective, there is a need to have clear, accurate and timely communications to the appropriate technical, executive and customer audiences. Processes should establish a robust, data-driven culture to drive resource allocation appropriately. The technology should provide a framework to build automation that decreases outage impacts and reduces unnecessary and especially duplicative labor.

There are several steps outlined in this section that helped to organize the Comcast journey. These steps provided a framework that guided teams through design, to the transition into the zen of ticketing, and eventually to operationalizing ongoing practices.

4.1. Analysis Overview

The first step in this zen of ticketing journey was to identify a solid framework (or two!) to manage the enterprise architecture. The three most common that engineering and information technology (IT) leaders lean towards to accelerate automation of things are information technology infrastructure library (ITIL), site reliability engineering (SRE) and development and operations (DevOps).

Our approach involved blended principles from SRE and ITIL to drive innovation through an agile continuous learning model. ITIL has provided common best practices helping across many operational teams, while SRE brings to light self-service tools and automation scripts. The models together increase reliability and performance of applications and services as manual work decreases. With a constant eye toward automation of manual processes brings the ability to auto-mitigate, remediate, and removes manual labor.

First, we will address the ITIL model. The ITIL is a five-step process to measure, report, plan and implement quality improvement cycles for ticketing operations. Its core components, including Service Strategy, Service Design, Service Transition and Service Operation, provide the framework for continual service improvement. The ITIL service management lifecycle provides guidance for implementation of a configuration management database (CMDB), event, incident and change management with policies, guidelines, and streamlined processes. The is extremely important, as it provides the capability to build physical and software configurations items into relationships that serve as the base for incident and change tickets. It also sets the foundation for automation, and aids in understanding customer impact. Figure 1 shows the ITIL core components (Ghadi, 2011).

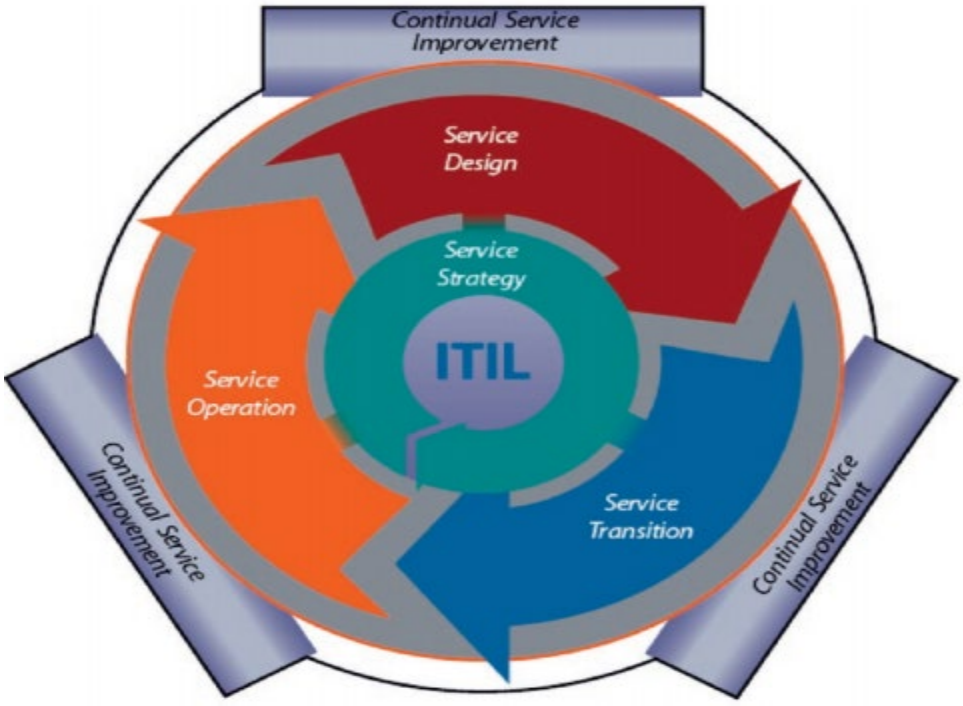


Figure 1 – ITIL Core Components

The ITIL components guided us through the common network enterprise architecture layers of business, data, solution, and technology. The focus on the four common network enterprise architecture layers was intended to reduce duplication, and to set the stage for automating certain repeatable tasks. The layers, informed by the focus needed to achieve a single ticketing platform, included:

- Business architecture or processes to define the strategy for moving people, process and tools to a single ticketing platform. Such as "governance" to inform key business processes.
- Data will include a standardized CMDB to identify physical assets, with relationships that tied to incident, change and problem management tickets.
- A solutions architecture to provide a blueprint of the end-to-end ecosystem, especially with regard to how applications will connect and get deployed.
- A technology architecture to assess what software is appropriate to support the evolution of a single ticketing platform. This view defines applications, databases and how they bridge together.

Each layer provided a different view to achieve the desired ITIL enterprise architecture. The viewpoints consisted of the following components:

- Business
    - A detailed vision to achieve best practices to transform the enterprise to a single ticketing platform

    - An ITIL Service Strategy to guide the design, development, and implementation of service management lifecycle with policies, guidelines, and processes

    - A review of current practices to understand “must have” end user requirements to align to future state
- Data
    - CMDB model with relationships diagramed
    - Schema of Incident, Change and Problem workflows
- Solution
    - Ecosystem design to map out process and software connections
    - Defined service level agreements supporting Incident, Change and Problem processes
- Technology
    - Identification of tools and service delivery platform

The second model we utilized is at the heart and soul of Google as they discovered a need to manage risk and growth called “Site Reliability Engineering” or SRE. SRE goes down the path of having software engineers that have the drive and ability to automate across complex architectures. Scale and reliability are created to manage higher volume of changes to your network with this methodology.

Figure 2 – SRE Key Principles

As shown in Figure 2, SRE drives resiliency into infrastructure and workflows that provide automatic responses with code to what humans have previously done (Site Reliability Engineering, 2020). This practice can transfer operational work into development tasks. As with the ITIL framework, SRE has five base elements to guide towards a software engineering model:

- Business communication for clear alignment on the definition of reliability with service levels that have associated impact.
- Architecture creation of scalable systems that have resilience to reduce outages.
- Build and run using automated toolchains for provisioning and deploying code is the foundation of automating manual work.
- Operating and monitoring aspect to measure everything that matters via service level indicators (SLI) and service level objectives (SLO) to assist with business impact.

- The SRE culture is set up to define a mix around 40% software and 60% administration system capabilities, with each SRE having accountable SLOs.

Implementing an SRE environment provided a higher ability to complete errorless transactions with proactive monitoring that frees the team to work on structural improvements at a more rapid pace. SRE provides not only definition for how to improve reliability, but also provides engineers with the opportunity to build and design instead of just putting out fires. Google has taken lead with the SRE role and methodology to foster sustainability and operational resiliency of all digital assets. Figure 3 provides a sample of what a SRE engineer focuses on (Feoktistov, 2021).

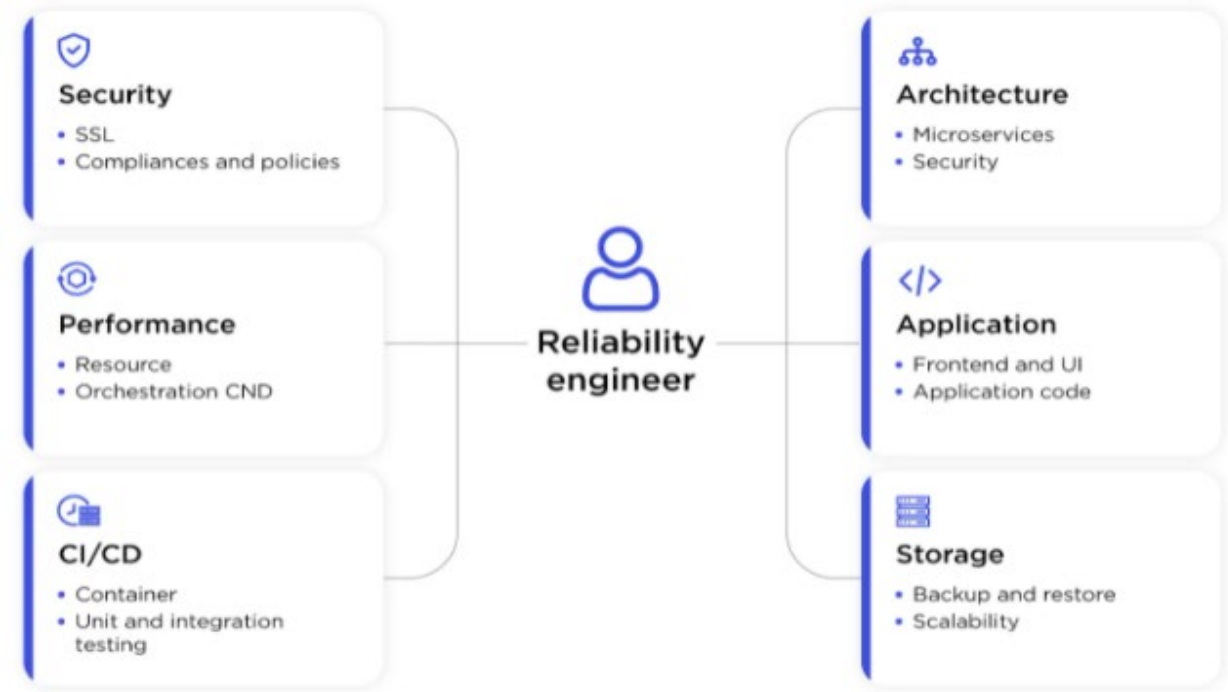


Figure 3 – Site Reliability Engineer Responsibilities

Comcast benefited from a blend of ITIL and SRE to address processes across the organization, as well as the complexity of the CI that existed, with a focus on eliminating manual work. The next step involved mapping out the benefits of this approach, in order to address the question of why any company would break down their current ticketing infrastructure and drive their teams into a single platform. Streamlining incident and change tickets to lower the amount of manual work, improving communications and building an ecosystem to remove toil to help set the stage for scalability are but a few answers to that question. To get to where you are going, you must understand where you have been. Table 1 defines the who, what, why, when, how and where of mapping the current state of ticketing to the future zen state of ticketing. This exercise is intended to understand the different audiences and how they will be impacted.

Table 1 – The Zen of Ticketing Platform Analysis

Classification Names / Audience Perspective	**Who** People, resources	**What** Entities involved in each perspective	**Why** Goals, objectives & business plan	**When** Time and performance criteria	**How** Functions in each perspective	**Where** Locations and interconnections within the enterprise	Classification Names / Model Names
Executive Perspective	List of all teams doing Incident and Change tickets.	List of what each team is responsible for, inventory list and best practices for Incident and Change tickets.	List of benefits for moving to a single ticketing platform.	SLAs for managing the Incident lifecycle from open to mitigation to close; Change ticket lifecycle.	List current & future state to identify transformations to include governance process for future state.	Map of geographical locations to show how they are interconnected.	**Scope Contexts**
Business Management Perspective	Define roles and responsibilities across teams.	Process flow diagrams showing inventory across processes, to include CMDB	Remove duplicate tickets; Create change risk model and conflict assessment.	Determine time needed to get to future state.	Diagram inputs and outputs across process transformations.	Define distribution.	**Business Contexts**
Architect Perspective	Understand system roles to define CMDB for inventory to feed Incident and Change tickets.	Schematic showing inventory across processes, to include CMDB.	Diagram showing system end state CMDB lifecycle.	System timing representation.	System process representation.	System locations.	**System Logic**
Director Perspective	Technology role diagram.	Inventory configuration.	Software selection.	Timing to implement technology.	Software process representation.	Technology locations.	**Technology Physics**
Engineer Perspective	Tool and security roles.	Data definition	Rules for software implementation.	Timing to implement technology.	Process configuration.	Distribution configuration.	**Tool components**
Enterprise perspective	Organization	Data	Strategy	Schedule	Function	Network	

4.2. Future State Network Enterprise Architecture

As stated, ITIL provides a recipe for reliability practices. Figure 4 shows the main areas of focus to improve incident and change management processes. Comcast chose a third-party platform with plugins to align with overall business strategies, including financial planning, applications testing, agile development, and project management, to name a few. The Service Strategy, Service Design, Service Transition and Service Operation guided the team toward agreed-upon implementation plans. An internal SRE team was created to configure the platform towards the aligned future state with a focus in the following areas (IT Service Management, 2019):

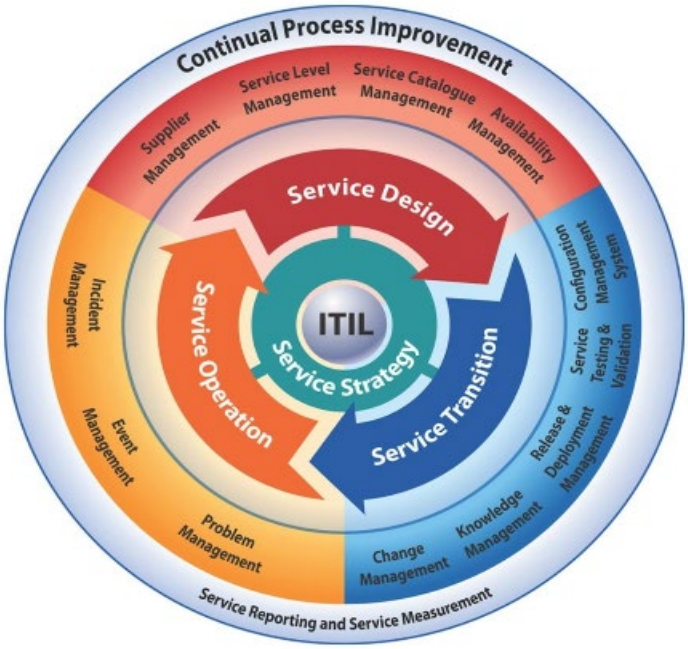


Figure 4 – ITIL Lifecycle Architecture Support

- **Configuration Management Database (CMDB)** – provides the capability to build physical and software configuration information into and between relationships that will be the base of incident and change tickets. This sets the foundation for automation and aids in understanding impact.
- **Incident Management** – provides a single platform for cross-departmental teams to work in unison on outages, with a focus on improving time to engage, time to mitigate and time to restore an outage with strong reporting.
- **On-Call Scheduling** – provides automated escalation from an incident ticket for fix agents to get engaged on an outage.
- **Problem Management –** provides the ability to perform trend analysis, within the same single platform, while tracking internal and third-party vendor action items
- **Change Management** – provides a single platform with a systematic, trackable and controlled approach to the lifecycle of all changes, enterprise-wide, to understand conflicts and customer impact.

4.3. Implementation Plan

Migrating to a zen of ticketing state requires a lifecycle strategy. Collaboration between the team managing the business change, and stakeholders of all impacted teams is key. The zen of ticketing strategy, goals and teams were documented and approved by senior leadership to ensure alignment with the future state. Figure 5 provides an example team structure to build a focused team.

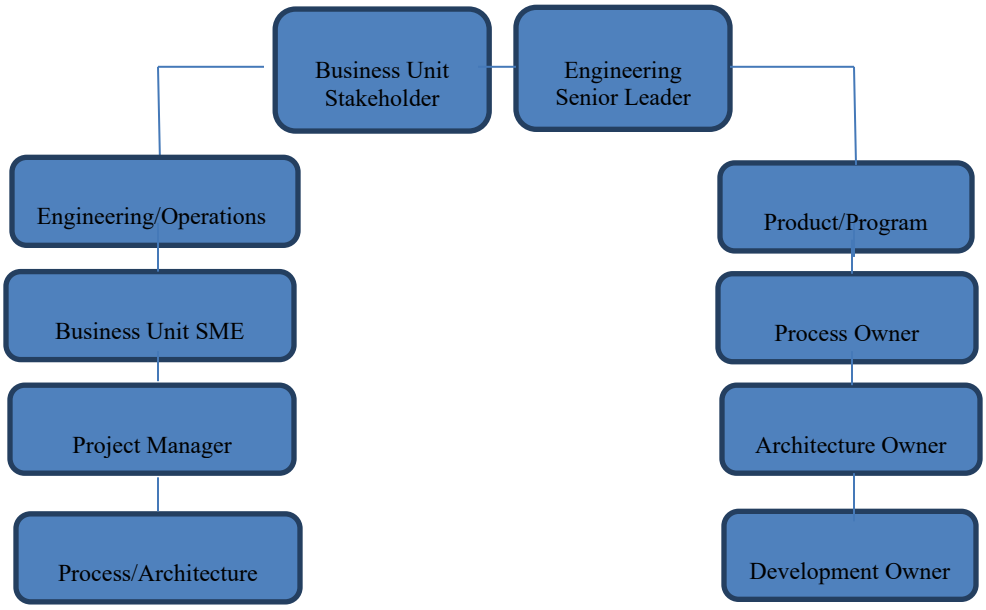


Figure 5 – Implementation Team Structure

Implementing the role of a "product owner" provided a liaison structure across multiple teams to learn and decide how to move through the lifecycle implementation plan. As the discovery process took place, the product owner facilitated the conversation to understand each team's current as-is state, mapping to the future state. With the support of senior leaders, the business goals identified a future visionary state that included the benefits of implementing a new architecture. One might call this a product roadmap and/or charter that is then socialized with teams during a kickoff session. The product manager worked with stakeholders to identify what the "day one" experience would look like, and the architecture team gathered the information it needed to draft different viewpoints to illustrate the alignment.

As the experts on the new technology, the process team mapped current workstream processes to the future state, being careful to identify areas where people would need to change what they were doing. Once the new architecture is approved, the product team will create and submit "stories" for the development team. A story in this context describes what the business is aiming to achieve with incremental code development. The number of stories identified will determine length time needed for development.

The development methodology is necessarily agile and iterative. As functionality is ready to be reviewed, visual demos and user acceptance testing will commence. As each functionality piece is finalized in the lower-level environment, process documentation is created, which will also support training efforts. Training needs may vary by team; therefore, this segment is handled uniquely to ensure the "people readiness" aspects cover awareness, desire, knowledge, and reinforcement. The overarching goal is that teams understand the "why" behind the change. Figure 6 displays the milestone implementation steps for the focused areas each team will need to accomplish.

	Workstreams	Discovery (3 months)			Design (3 months)				Development (3 months)			Onboarding (3 months)		Gating Dependency	Risks	Mitigation
		Kickoff	CMDB Current	Process/ Tools Current	Access Setup	CMDB Future	Process/ Metrics Align	Prod Day 1 New	Req Stories	Arch	Dev UAT	War-game	Training			
Business/People	Vision, Goals & Benefits													-	-	-
	ITIL Service Strategy													-	-	-
Data/Process	Configuration Management Database													-	-	-
	Incident Management													-	-	-
	On-Call Scheduling													-	-	-
	Problem Management													-	-	-
	Change Management													-	-	-
Technology	Core Platform													-	-	-
	Integrations													-	-	-
	Middleware													-	-	-

LEGEND: On track; At risk; Blocked; Completed; Not Started

Figure 6 - Enterprise Architecture Implementation Plan

The mission-critical business drivers encompass improvements that are required in the data and process sections of the workstreams. This, along with developing the overall technology ecosystem, will drive duplicate tickets to one in the single zen platform. Repeatable steps will be identified through discovery, with a focus on automating where it makes sense. The implementation lifecycle can vary depending on the complexity of the network topology and availability of source data. Designing a governance process will help manage the aligned continuous improvements, timeline, and resources required. Finally, regular communications are critical to keep the program team engaged along with leadership across the enterprise. Long term programs can easily succumb to distractions if there is not discipline around program and communication management dwindles.

5. Agile Solutions With Continuous Learning are the Way

Organizations that adopt learning as a practice increase the chances of continuous improvement, while promoting a culture of knowledge and innovation. Leading with a mantra of transparency in all operational and engineering practices also helps to cultivate a blame-free culture. Agility sets boundaries for how to work, in preparation for being able to pivot to new processes and technology. Those pivots may arise because of business needs or competitive threats; in all cases, the customer always plays the primary role. As a result, team environments that achieve shared milestone objectives are more productive, and more emotionally satisfied with their work. A scaled agile framework provides three critical dimensions to help build a learning culture, as shown in Figure 7 (Continuous Learning Culture, 2021). All three dimensions speak to employees at every level to foster growth and transformation.

Figure 7 - The Three Dimensions of a Learning Culture

5.1. Lessons Learned and Pro Tips

As we continue our journey towards the zen of ticketing, we identify a few pitfalls to help our readers avoid those things that can disaffect their single zen ticketing aspirations.

The first lesson learned involves the objective, the goal, and the timeline for a single zen ticketing initiative. Most cable operators already have a ticketing system (or 20!) in place that may have been created over the past 15-20 years. Be sensitive to the time and people that it took to build what worked well until now. Set smaller and realistic timelines that create incremental improvements over time. A governance process will help keep those involved grounded as to what the next quarter's commitments will bring.

A second lesson learned involved the CMDB. As described earlier, the CMDB provides the capability to build physical and software configurations items into relationships that serve as the base for incident and change tickets. It also sets the foundation for automation, and aids in understanding impact.

"Do not underestimate the value in moving towards a single solution that can provide automation and correlation of events, so that engineers can focus on the problem instead of being ticket jockeys," noted Rich Massi, VP, Reliability Engineering Residential Products and Services, who shared his viewpoints during an interview. Build a process to stay current, because the CMDB is never done. Also important is how and who best maintains the CMDB's governance and data quality processes.

Automation is a key aspect of moving into a single zen ticketing platform. Covet all insights about how to perform a task, should automation break or go down -- it happens. Ensure everyone understands manual processes from the beginning, before automating tasks. (You can thank us later for this one!)

Call it what it is. The bigger the organization, the higher the likelihood that big software initiatives like ticketing will come with a fancy name, a logo, stickers and more. Just call it "ticketing"! This will reduce

financial and capital expenditure discussions to come, allowing you to focus on the incremental improvements.

Product and process are alignment upfront. The process should drive the implementation as opposed to the tool trying to drive the process. It is important that there is alignment on the process before the tool is adapted. Otherwise, just another problem is created when multiple groups are trying to modify the tool in competing ways.

Lastly, and speaking of the finances: think about the big picture. Larger companies going through large transformations like moving towards a single ticketing platform should anticipate funding requirements. Considerations like vendor support versus in-house work, i.e., "buy versus build," will depend on what skill competencies exist to start such a major program. A multi-year program, unified ticketing is best approached as an endeavor that is forecasted and budgeted to cover operational and capital expenditures.

5.2. Success Stories

We're starting to gather the success stories that come with moving our teams onto a single ticketing platform. We're not all the way there yet – about halfway, at the time this paper was written. Yet so many wins have been and have yet to be celebrated! Results so far clearly indicate that moving incident and change tickets into the single zen ticketing platform is reducing "swivel" and the time it takes to manage those processes.

The voice service engineering and operations team were the first to embark, diving headfirst into the single ticketing platform. Frances Augustine, Executive Director Reliability Engineering, led the vision to reduce the number of alerts engineers were managing. She was relentless that "eyes on glass" was going to become a thing of past, and that managing with a proactive mindset would be a much healthier work environment. Through the automation associated with a single ticketing system supported by a mature CMDB, the "Voice team" was able to do just that and move from a reactive to proactive operation, focusing on a sustainable level of logged alerts to tackle before an outage occurs. This includes internal alerts, to ensure voice services are always on, to providing automated reporting to the Federal Communications Commission (FCC) for its required outage notifications. This automation enabled engineers to pivot from being reactive to being proactive, finding issues from logs before they became outages. It also enabled a reduction of Tier 1 vendor support.

Another early adopter was in our residential services and products team. Early on, this team aligned with the single ticketing platform, spending hours to build out its CMDB so that all the appropriate software, hardware and respective locations were identified, connected, and correlated when doing incident and change tickets. The extensive time spent to build (while knowing that CMDB work is never really done; see section 5.1) also enabled automation of incident and change ticketing. As a result, the team was able to eliminate an entire vendor team that had provided manual triage support. This team not only tackled the single tools aspect but is also now ensuring that all teams performing incident management are doing it the same way, so as to harvest additional business improvements.

As part of the journey to adapt to a single ticketing platform, having key performance indicators (KPI) per team, across a collective set of teams, will help measure the success in a few sample areas:

1. Mean time to mitigate and resolve incidents
2. Change success percentage (number of changes executed/not executed flawlessly)
3. Number of changes automated
4. Number of incidents created automatically from event management logging and correlation
5. Percentage of time team spends in reactive versus proactive tasks over X Time

6. The Future of Ticketing

The future of the ticketing, zen or otherwise, is really all about scalability. It's about acknowledging industry consolidation, building smart network platforms, and utilizing intelligence in the right way to ticket the things that matter. Tickets are not necessary for every logged event -- but every logged event can provide trend and analysis to move towards managing proactively. Simply put, our networks are smart and intelligent -- but if we're not smart about our networks, we can't let them show their intelligence. Single ticketing is the path to network intelligence.

The nirvana towards moving into a single ticketing platform, besides the obvious benefits of removing unnecessary duplication and effort, is the uplift to overall reliability. "Spend a little bit of time on Incident Management, police Change tickets that fall out from automation, and spend 80% of your time doing problem management to enable even more automation," advises James Manchester, our SVP of Core Platform Technologies.

What's next? Everything that enables reliability to be "always on" and flourishing. As incident and change ticketing puts firm solid practices in place for ticketing, next steps will likely include moving teams into a single problem management space. Doing so will enable further analysis to ensure an outage, if it happens, only happens once.

Cybersecurity is another realm to explore with respect to unified ticketing. Knowing what is happening in the network is always important, and more so in digital times. Professional hackers continue to advance where and how they can breach systems. It's entirely plausible for managed ticketing practices to also advance, to keep would-be interlopers out.

7. Conclusion

Getting to a "zen state" as it relates to change and incident ticketing, or any type of ticketing, is really about "keeping the lights on" and driving reliability as a core service, at a fundamental and strategic level. This is especially true for large companies, like ours, that grew as a result of multiple decades of system acquisitions and geographic clustering, then remained "siloed" through onboarding and corporate integration. As it continues to become mainstream within Comcast, the unified zen of ticketing approach will bring valuable intelligence to understanding topology and automating remedial tasks.

We'll close with this list of tips, when considering a unified ticketing environment where you work:

- Think about a unified enterprise ticketing infrastructure as a way to create alignment on tools, processes, and people.
- Have a vision to steer teams towards.
- Know that transparent ticketing processes create reliability.
- Adopt a framework such as ITIL and/or SRE to design, transition, and operate into a single ticketing platform.
- Analyze the current and future state of tools, processes, and people, to understand what's going on.
- The right team and leadership support structure is crucial for stakeholder support.
- Under-commit and over-deliver wherever possible!
- Go slow to go fast in an agile continuous learning environment.
- Perform retrospectives with a culture of learning; be an advocate for learning.
- Celebrate the successes and continue to look to the future.

Abbreviations

CI	configuration item
CMDB	configuration management database
DevOps	development and operations
FCC	Federal Communications Commission
IT	information technology
ITIL	information technology information library
ITSM	IT service management
KPI	key performance indicator
SLI	service level indicators
SLO	service level objectives
SRE	site reliability engineering

Bibliography & References

Augustine, F. (2021, June 13). Personal interview.

Continuous Learning Culture. Scaled Agile, Inc. (2021, June 24). Retreived from *https://www.scaledagileframework.com/continuous-learning-culture/*

Connors, B. (2021, June). Company meeting.

Comcast. (2019, April 24). Company. Retrieved from *https://corporate.comcast.com/company*

Feoktistov, I. (2021). Why and How to Hire a Site Reliability Engineer (SRE). Retrieved from *https://relevant.software/blog/hire-site-reliability-engineer/*

Ghadi, V. (2011). Adopting ITIL Framework. Retrieved from *https://www.happiestminds.com/whitepapers/Adopting-ITIL-Framework.pdf?sid=5086*

IT Service Management. (2019). Retrieved from *https://docs.servicenow.com/bundle/madrid-it-service-management/page/product/it-service-management/reference/r_ITServiceManagement.html*

John P. Zachman. (2019). The Zachman Framework Evolution by John P Zachman. Retrieved from *https://www.zachman.com/ea-articles-reference/54-the-zachman-framework-evolution*

Manchester, J (2021, June 13). Personal interview.

Massi, R. (2021, June 13). Personal interview.

Minoli, D. (2019). Enterprise Architecture A to Z: Frameworks, Business Process Modeling, SOA, and Infrastructure Technology. Auerbach Publications.

Site Reliability Engineering (SRE) What is it, what are the advantages, and how do I implement it? (2020, March 31). Retrieved from *https://www.linkedin.com/pulse/site-reliability-engineering-sre-what-advantages-how-do-dan-martines/*

Young, C. (2021, June 4). Personal interview.

Tools of the Trade for Supporting Critical Communications of Last Resort

An Operational Practice prepared for SCTE by

Derek DiGiacomo
Senior Director
SCTE
140 Philips Rd, Exton PA 19341
610-594-7310
ddigiacomo@scte.org

1. Introduction – The Why

The cable broadband industry has had a tremendous expansion of products and services since its initial launch. With each expansion of service, the importance of having a plan to help contend with threats has grown. Technology in general has flourished and impacted societies greatly, and cable has been impacting and often driving a lot of that change. Let's take a moment to review some history.

Cable was born to bring better television experience to communities. Reception was poor down in valleys and technology solved that problem of getting to local programming. Next was the expansion of access to more television content, where content providers leveraged the controlled means of signal delivery to the paying customer in an effort to protect their private programming. The third phase of growth marked the first tipping point of criticality of service offering by the industry – voice service. With the advent of voice service, cable began to expand beyond the world of entertainment and increased the importance of the connection. Calling 911 depends on a reliable voice service, so providers needed to pay closer attention to the target of 100% availability. Legacy copper telco phone service was one of the last things that didn't work during a local subscriber power outage and cable voice subscribers desired the same level of service. The fourth phase of cable evolution is high speed Internet access. With the expansion of e-commerce and migration from physical to "digital everything," subscribers could realize an almost unlimited resource that can reshape everyday lives. Today, we are in the midst of the fifth phase of our industry's expansion and continued rise of criticality of service with the promotion of the 10G Platform. With 10G our subscribers will depend on information and services enabled such as telehealth/telemedicine, home/business security monitoring, e-education and a host of undreamed critical Internet based services.

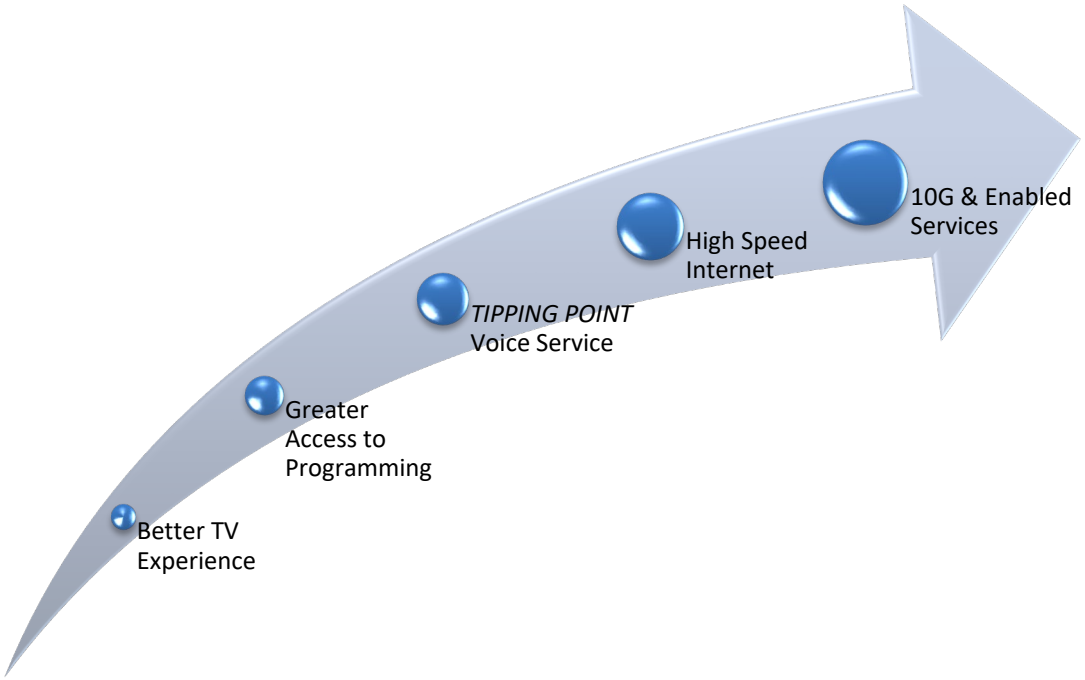


Figure 1 - Cable Industry Importance of Service Milestone Markers

Given this next phase of the cable broadband evolution, let's ask, "What is our game plan when threats to our infrastructure strike?" Natural events, such as earthquakes, hurricanes, fires, floods, winter weather and solar storms; along with manmade threats such as physical attacks, cyberattacks, and electromagnetic

(EM) attacks posing risk to the electric grid could have cascading effects and leave our critical facilities and networks requiring their own power generation and energy storage capabilities for an extended period of time. We are doing a good job at preparing and deepening our resiliency as an industry, however even the strongest plans can and will be tested by unpreventable natural events. This paper will outline high frequency (HF) radio based resources (in particular SHARES) broadband providers can leverage (with proper upfront planning) when all traditional lines of communications are down and we need to reach out for information and request aid typically in the form of security, access, and fuel.

2. What is SHARES?

The SHAred RESources (SHARES) HF radio United States government program administered by Department of Homeland Security (DHS) provides a means for users with a national security and emergency preparedness mission to communicate when all traditional means of communications are unavailable. SHARES members use existing high frequency radio resources to coordinate and transmit messages needed to perform critical functions, including leadership, safety, maintenance of law and order, finance, and public health. SHARES is available on a 24-hour basis to provide an emergency communications link to support intra or inter-sector mission requirements. The use of SHARES requires no prior coordination or activation to transmit messages. A signed non-disclosure agreement (NDA) submitted to DHS is required to obtain proper call sign and access to net frequencies, the SHARES participation directory, and key SHARES resources.

More than 1,400 HF radio stations—representing 104 federal, state, and industry organizations located in all 50 states, the District of Columbia, and several locations overseas—are resource contributors to the SHARES HF radio program also referred to as the SHARES HF network or net for short. Nearly 500 emergency planning and response personnel participate in SHARES. Approximately 200 HF radio reserved channels are available for use by SHARES members. SCTE is the recognized coordinating registration lead for the cable broadband industry. Cable broadband providers interested in deploying a supporting station can contact me for assistance submitting the necessary paperwork and station planning. There is no associated membership or participation fee to get involved.

Membership in the SHARES program is voluntary. As mentioned, the SHARES network is available on a 24-hour basis and requires no prior coordination or activation to transmit messages. Members consult a SHARES handbook hosted at the SHARES private website to find stations, frequencies and/or automatic link establishment (ALE) addresses of participating organizations they need to communicate/coordinate with. Participating SHARES HF radio stations accept and relay messages until a receiving station is able to deliver the message to the intended recipient.

The SHARES program has three operating levels:

- **Operational Level 3** - No emergency exists. The allocated frequencies may be used for training, the weekly SHARES net, on-the-air testing and configuration of SHARES stations, and exercises.
- **Operational Level 2** – An emergency potential exists. Non-essential on-the-air activities on some or all SHARES net channels may be suspended. The SHARES coordinating network monitoring is increased, and regional coordinators for SHARES (RCS) are contacted and advised of potential US Emergency Support Functions (ESF) #2 activation. SHARES net control stations may maintain watch on designated channels to provide stations with an opportunity to test their equipment, and to receive or relay situational awareness messages.
- **Operational Level 1** – An emergency exists, or the potential for an emergency is enough to warrant net activation.

The SHARES Program is organized into several regions and Table 1 represents the geographic breakdown of how the states are organized. The organization is aligned with the Federal Emergency Management Agency (FEMA) regions.

Table 1 - SHARES to FEMA Regions Breakdown

SHARES Region	FEMA Region/States & Territories
Northeast	*FEMA: I, II*, III* CT MA ME NH RI VT NJ NY DC DE MD PA VA WV * PR and VI are in FEMA Region II but are in SHARES Region SE
Southeast	*FEMA: IV* AL FL GA KY MS NC SC TN PR V
South	*FEMA:VI* AR LA NM OK TX
Southwest	*FEMA: IX* AZ CA HI NV American Samoa, Guam, Northern Mariana Islands, FM (Federated States of Micronesia (FM) is not a U.S. Territory but receives certain government services from the U.S.)
Northwest	*FEMA: X* AK ID OR WA
North	*FEMA: V, VII, VIII* IL IN MI MN OH WI IA KS MO NE CO MT ND SD UT WY

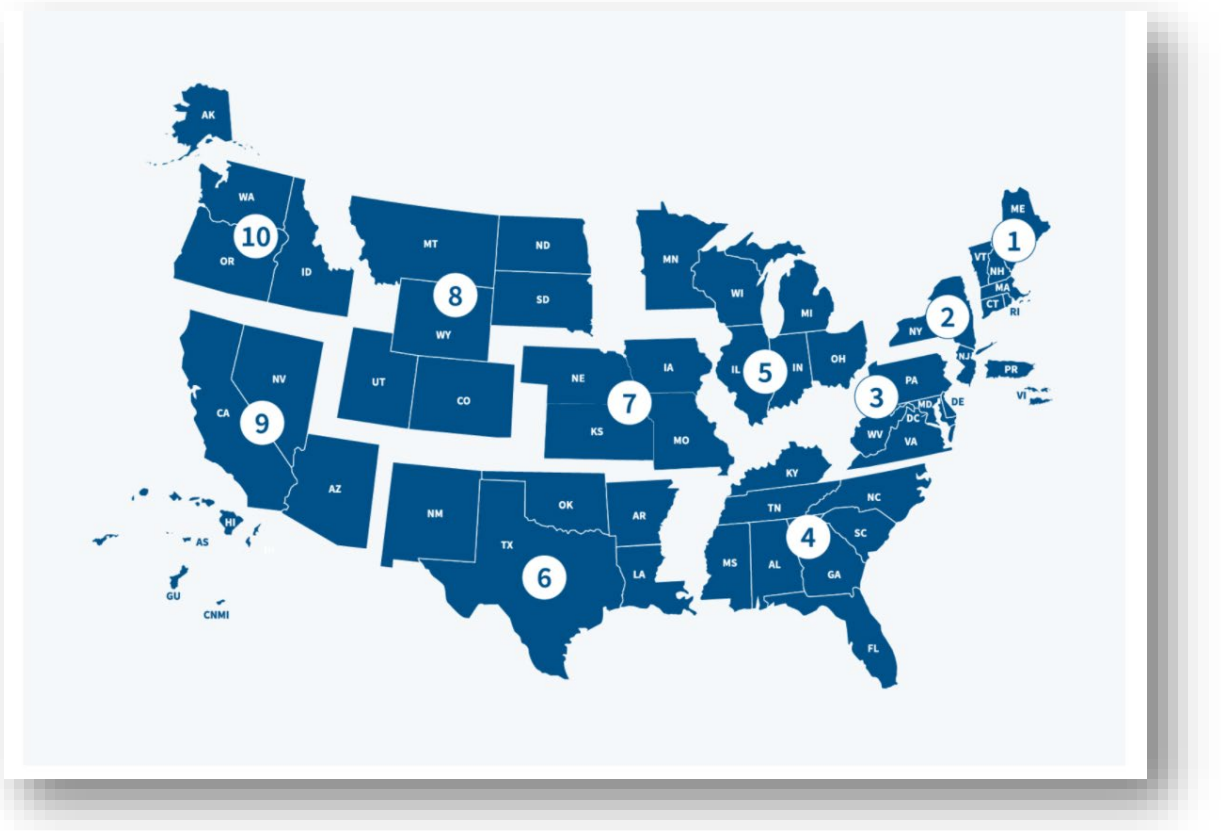


Figure 2 - Federal Emergency Management Agency Ten Regions

Cable broadband providers are present in all ten regions and ideally each operator should evaluate and deploy where they have a need to cover both areas prone to natural disasters as well as zones that could act as critical information relay stations throughout the country.

3. What is HF?

HF operates from 3MHz to 30MHz bands and can be found right above the medium frequency and below very high frequency (VHF) spectrums. Each band has its advantages and use cases. The HF space is used by military, police, emergency services, disaster relief organizations and SHARES. The transmission of communications is called propagation, that is moving of transmissions across the open air. Note there is no physical infrastructure required to carry the message. Up to 1.5 kW single side band (SSB, a form of modulation) and 1 kW carrier wave and data modes can be utilized to get the communication on the net. The national net frequencies often have the capability of transmitting messages coast to coast.

3.1. Propagation Methods

Ground wave propagation takes place when the antenna is configured parallel to the Earth's surface and the range decreases as frequencies increase. The terrain will determine how far the signal will travel. Line-of-sight waves travel point-to-point and are typically found in use for air traffic communications. HF is able to leverage a powerful mode of propagation that takes advantage of the Ionosphere in Earth's atmosphere to refract signals from one point to another. This is called sky wave propagation. Performance using this method will vary by hour of the day, night-day, winter-summer, and allocated frequencies.

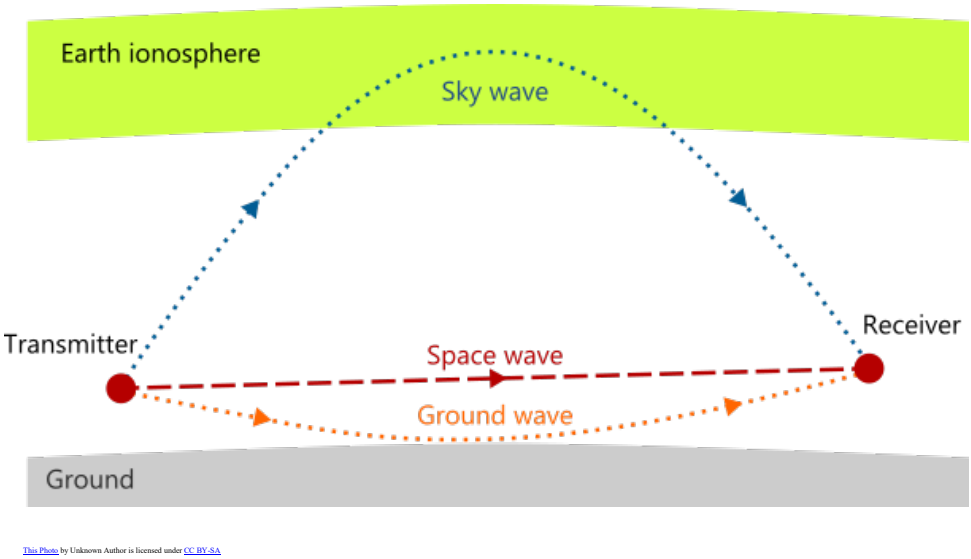


Figure 3 - Common Modes of Propagation

The Earth's magnetic field can play a role in propagation as well. This phenomenon is witnessed in the aurora borealis (northern lights). Space weather or solar winds impact not only the aurora but HF propagation as well. There is a US government service that tracks the space weather conditions like traditional weather forecasting. The United States Space Weather Prediction Center that is part of National Oceanic and Atmospheric Administration (NOAA), has an email alert service that will provide statistics for planetary A and K indexes. These indexes help station operators get a sense of space

weather conditions that could be impacting propagation. It is helpful to subscribe to the service at: https://pss.swpc.noaa.gov/ProductSubscriptionService/RegistrationForm.aspx. The K index is computed every three hours by readings of magnetometers, and conditions tend to be more favorable when K is 3 or lower on a scale of 0-9. The A index is calculated from the previous 8 K index readings and favorable A conditions would be a value of 15 or lower in a recognized range of 0-400. If you like to learn more about space weather, NOAA has released a short video https://www.youtube.com/watch?v=HrloxznL93s.

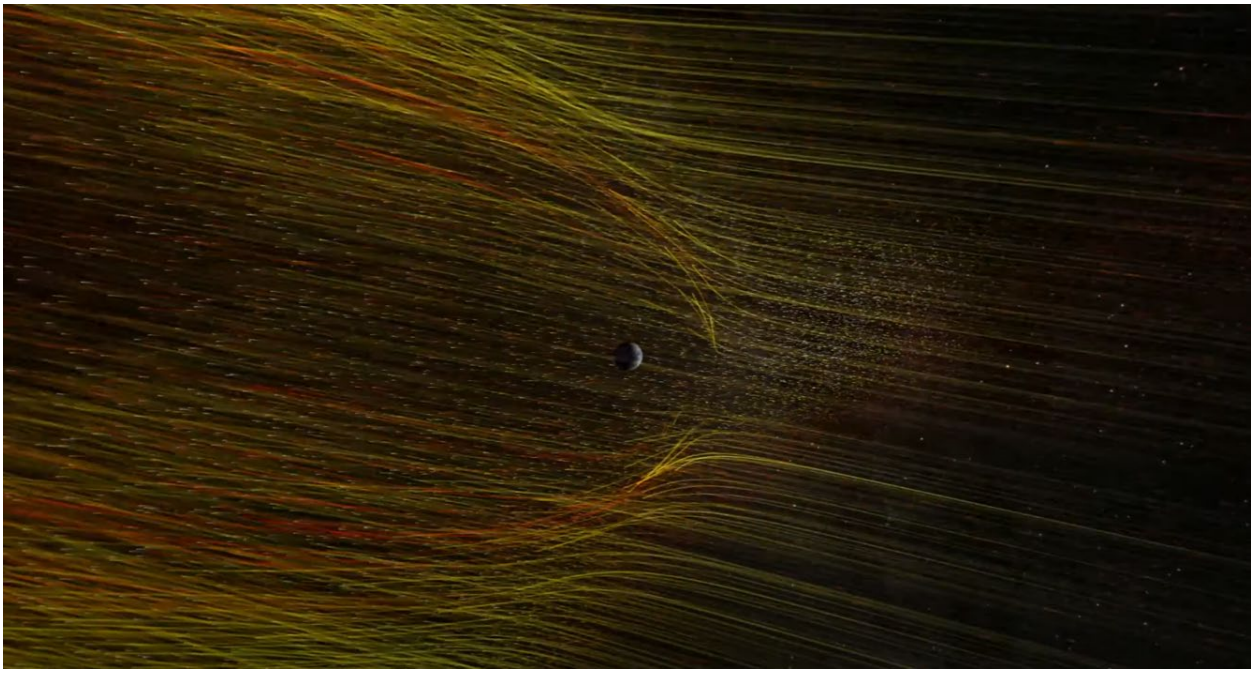


Figure 4 - Space Weather Simulated NOAA Example

When leveraging sky wave propagation, use of an antenna designed for maximizing near vertical incidence sky wave (NVIS used interchangeably with sky wave) will help achieve the receipt of signal. On average NVIS propagation can cover distances between 30-400 miles. The angle of the antenna in comparison to the sky will affect how close/far the signal will return to the earth. It is possible to "bounce" the signal up and down however, strength of signal will diminish with each hop. NVIS frequencies for SHARES will be found between 3 and 12MHz. It is important to note that there is an area between ground wave propagation and the HF hop called the skip zone. In this area, there could be no HF signal received. Alternate means of radio communication could be used to fill in these "HF dead zones." Ionosphere HF conditional monitoring can be found at https://www.sws.bom.gov.au/HF_Systems/6/5.

4. Station Operator Needs

One of the most essential pieces of the SHARES radio station is the operator. Proper planning for the human factor cannot be overlooked or under-planned. This planning begins at home. Before you depart to tend to station needs during times of activation, be sure to have your home plan in order. If you are unable to dedicate focus on your safety and mission of relaying message during incident because your thoughts are trying to react to needs at home then there may be gaps in your own home care planning. Proper home planning is beyond the scope of this paper but is important and worth the mention. Please have a home plan in place to provide peace of mind.

For station deployment needs, simple everyday items should be included in a key assets bag. A search online search for "emergency survival bag/kit –with 72 hours of disaster preparedness" can provide many pre-built options. Suggested items in such a kit can include:

- hammer/mallet
- gloves
- masks
- nylon rope
- bottles of water/water filtering device
- first aid kit
- shade/rain canopy
- rechargeable weather radio
- three days of food and water/water purification system
- sunglasses/safety glasses
- sleeping support needs (bag, tent, etc.)
- personal hygiene support items such as toothbrush/paste, TP, extra socks, undergarments, and showerless cleaning towels
- chair
- small table if possible
- pen and pad of paper
- LED flashlight or evening lighting/battery powered lantern
- waterproof storage bag
- insect repellent
- duct tape and/or electrical tape

5. Station Hardware

As the SHARES Program implies, it depends on access to appropriate hardware to transmit and receive HF signals. Stations generally consist of the person operating the hardware, the HF/RF unit itself and the supporting antenna. When selecting the necessary hardware, it is important to consider the use case and location the station will be deployed. For example, an urban major city with skyscrapers will limit the amount of physical space an antenna can be erected. Please examine the location before making the antenna selection. SCTE has secured both fixed and portable versions of SHARES station equipment to allow for optimum flexibility of deployment. Also, when matching antenna to RF until capability, be sure to analyze the output power of the transmitting equipment to what the highest power capacity of the antenna is rated for. The SCTE equipment can produce transmissions as great at 125 watts; therefore, our antennas should be able to transmit signals output power of at least 125 watts.

5.1. Radio

When selecting a radio for SHARES one should ensure the feature set matches capabilities of other key stations on the network. For example, the ability to store frequencies (software defined radios) in a list format and rapidly change from one frequency to another can save a lot of time and limit frustration especially during challenging times like net activation. Other commonly deployed station qualities include rugged construction (not necessarily military grade but approaching that quality of construction). During activation, having well built equipment will relieve some anxiety about station reliability and allow the operator to perform the key task of relaying information.

Another consideration when choosing an HF radio is how easily can the radio be changed without needing to replace the entire device. Typically, this is accomplished via a software defined method and many of the higher end devices support the ability to manipulate the handset appearance, what features are restricted/enabled, and many more operating functions. A good practice would include a plan to purchase this equipment once and anticipate having this gear in service for many years (unlike laptops or cell phones). Finally, a built-in standing wave ratio (SWR) meter is very handy to help troubleshoot antenna issues to help optimize power throughput to the system. Standing wave ratio meters help determine how much of the RF unit's output power actually gets transmitted, and the optimum meter ratio is 1:1. A solid target of 1:3 or less is optimal. If the ratio is much higher, antenna adjustments are needed such as adjusting the mast height, ends of antenna height or distance of feedline run.

Figure 5 - SCTE Portable Radio and Power Supply

Assembling the radio should be practiced and familiarized by the station operator. A common setup will have: copper ground wire connected to the ground screw, HF 50 Ohm antenna feeder coax cable, power supply, handset connector cable and audio out like a small speaker or headphone connection. Make sure the surrounding area is clean and without obstruction of airflow. Also leave room for a pad of paper and pen to capture notes. Remember, the area may be lacking good light and air conditioning due to the situation that warrants the operator being at the station. If there is notice that wind-based storm could be the cause of the deployment do not deploy the equipment until after the storm has subsided. If the antenna (in particular) is exposed to the elements during storm conditions, it could be subject to serious damage or even blown away.

5.2. Antenna

The SHARES program has many frequencies allocated for use by authorized licensed participants. Versatile antenna systems should be secured to allow for transmission across many of the confidential frequencies. Some antennas are deployed and configured to transmit optimally on a particular frequency. That approach is good for the national net that will leverage one or two frequencies. However, to have the greatest flexibility and readiness to deal with the unknown of the deployment as well as propagation

conditions, it is recommended to select an antenna system that is either tunable or broadband in nature. Tunable antennas can match the maximum output power to the frequency the radio is set to transmit on.

A broadband antenna is very flexible. Usually this is a terminated folded dipole (TFD or T2FD). These are available in many lengths and while all will work, length determines the lowest effective frequency supported by the antenna. Typical TFD lengths can be 60, 90 or even 120 feet long. For NVIS communications the TFD should be mounted in an "inverted V" configuration with the center apex up about 5-10m with the ends sloping downward to a height of about 2-3m from the ground. Precise height values and shape are not critical, often deploying the emergency station will present unique installation opportunities and antenna manufacturers will also offer installation recommendations. The antenna is connected to the transceiver using coaxial cable. A high-quality double-shielded 50 Ohm cable such as RG-214 or LMR-400 is recommended. Up-front planning is essential to ensuring readiness in time of true need.

Figure 6 - SCTE Portable Broadband Antenna and Mast Kit

One item to note, HF communications is subject to propagation condition changes that are impacted by time of day, sun activity, time of year and station location. Propagation is probably the biggest wildcard in the art of emergency HF communications. SHARES leverage many HF channels to help offer various opportunity to find a frequency where propagation is less of an interference factor. As mentioned above, online solar index resources such as the United States Weather Prediction Center

(https://www.swpc.noaa.gov/products/planetary-k-index) offers reports of geomagnetic conditions that can impact station experience. Another notable resource to have programmed in the radio is the United States National Institute of Standards 5MHz 10MHz 15MHz and 20MHz channels. This service will help give station operators a sense of how that range of frequencies are performing at that time. The government service announces the time along with other important information such as Atlantic high seas warnings at 8 and 9 minutes after the hour, and a Pacific high seas warning is broadcast at 10 minutes after the hour.

6. Station Powering Considerations

As mentioned in the introduction, power is essential to the services we provide. Looking at the current state of the utility grid here in the US, as our importance of service to our subscribers increases so shall the level of preparedness be adjusted to match that expectation. The good news is that deepening our power resiliency for our last lines of communication does not need to drain capital budgets. Looking at flexible, renewable, portable, power stations to support an "off grid" mentality to our last line of communications is commercially viable at the time of this publication. A search for 2000kWh with peak power output of 4000W would provide a good reference point to power the station for several days with moderate use. Some of the newer portable power stations can incorporate multi-modes of charging to include solar, traditional AC and even DC from automobiles.

By way of reference, Cybersecurity and Infrastructure Security Agency (CISA) has released recommended guidelines for critical communication infrastructure. This can be found at: https://www.cisa.gov/sites/default/files/publications/Factsheet%20Resilient%20Power%20Best%20Practices.pdf. As suggested, there are various levels of power preparedness that are recommended when removing risks for vital communications. This breakdown provides different levels of backup to grid power that can be considered:

- **Level 1**– Least-cost best practices that provide a commercially reasonable chance of maintaining power for at least **three days/72 hours** under all-hazard conditions (for example, three days of fuel is stored onsite to maintain critical loads).
- **Level 2**– Provides a best-efforts approach to maintain power for at least **seven days** under all hazards.
- **Level 3**– Generally covers the most critical infrastructure where power should be sustained under all-hazard conditions for a minimum of **30 days**.
- **Level 4**– Typically limited to the most critical military/federal/National Essential Functions communications infrastructure where power should be sustained with no unplanned downtime under all hazards in **excess of 30 days**.

Ultimately it will be up to each of the broadband providers to determine what level of readiness each station should adhere to. History has demonstrated that having 72 hours of standby off grid power is a good model.

7. SHARES Station Setup

Pre-planning for this all-important asset is vital. Identifying who in the employee pool will be authorized (both at the company level and SHARES NDA level) and familiar with the equipment before deployment is critical. It is important to plan for both a fixed and portable station to help ready your organization for the disaster that warrants the activation of the SHARES network. Flexibility of the portable kit enables the operator to move the asset into and out of areas of need. Having a lightweight mast, antenna, hardened radio, and power source will provide the reliable equipment in time of need. Practicing setting

up and properly packing up the station during non-activation days will provide the operator skills necessary to deal with the stresses of getting vital information in and out of areas impacted by incident. A permanent station located at a mission critical building that already has plans for continuity can extend its value during disruptions. The roof can serve as a good permanent deployment location for an antenna and proper site survey is advised prior to selecting an antenna. Be sure to secure the station equipment as you would any valuable network asset.

Turning to the portable deployment needs, procedure wise, setting up the station in a logical manner can help ensure quick readiness to get on the air. First, identify available space for the antenna, RF unit and power supply. This will normally be limited to how much length coax feed line that connects the antenna to the RF unit is on hand. Experience has demonstrated that 100' is a practical length of a portable setup. Also, try to select an area where shade is readily available. Being present at the station for long hours and the added impact of the sun could result in additional bodily stress. Having a portable canopy is recommended.

The next steps involve setting up the station and all its components. After the location for the station has been determined, begin to setup the mast, necessary guy wires and antenna. Remember to connect the coax to the antenna before hoisting the antenna. Next connect the coax feedline to the RF unit. This will prevent accidental transmission by the RF unit without a load attached that may result in equipment damage. After the RF unit is connected to the antenna, connect the RF unit to the power supply. Finally, refer to the radio manufacturer's manual/recommendation for proper grounding technique as this will vary from radio to radio. However, typically there is a grounding screw on the back of the power supply that would allow for copper ground wire to be attached to a metal ground rod.

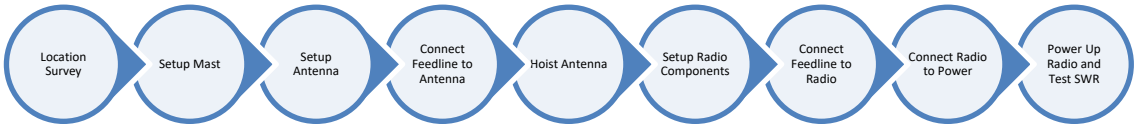


Figure 7 - Simplified Portable Station Deployment Steps

These basic steps should allow an operator to now transmit and receive signals on the SHARES frequencies. If performance is poor, conduct a simple SWR test and adjust the antenna as necessary. Actions could include raising/lowering the center mast or ends of the antenna or confirming a good dry fit of the feed line to the connections on the RF unit and/or antenna. Another opportunity to help improve performance is to run a simple metal cable on the ground under the antenna from one end to the other.

8. Digital Modes

In this section some common modes of digital communications or non-push-to-talk methods of message delivery are discussed.

8.1. HF Email

Email is such a valuable asset. It is hard to imagine conducting business without this method of communication. Disaster communications can leverage Winlink Global Radio Email®. Winlink is an all-volunteer project of the Amateur Radio Safety Foundation, Inc. (ARSFI), a non-profit public benefit corporation with no beneficial owners. Its original purpose (started in 1999) was to provide a very long-

range radio path for radio amateurs who did not have access to “land-line” communications needed to send and receive email messages. Subsequent uses have been oriented toward providing partial backup of email services for emergency communication service agencies during a local commercial communications outage or communications overload. Transfer speeds and available bandwidth do not allow for complete replacement of traditional Internet based email services but when essential messages need to be relayed reliably, Winlink can get the job done.

SHARES Winlink HF email system operates on approximately 80 dedicated channels which are not listed in the SHARES channel list or net list. The channels are known to the SHARES Winlink Express software and are updated electronically via Internet on a regular basis (preferred) or over the air (very slow). Winlink needs to be downloaded off the Internet to a viable laptop/desktop in advance, configured and tested. An authorized and registered separate call sign is required for the Winlink station. Like traditional Internet based email, Winlink requires a unique email address that is routable by both the Winlink platform as well as traditional Internet based email. When local Internet access is down, the email client will communicate via a HF modem connected to the computer and HF radio. Mail routing can leverage BOTH traditional SMTP and Winlink delivery methods. The Winlink client is configured to connect to a remote Winlink email gateway where the message can be properly relayed to either another operator on the Winlink network or Internet.

Figure 8 - SCTE HF Enabled Email Station

During a “worst case scenario” where traditional Internet/email is unavailable, station operators can leverage the radio-only Winlink mode. Messages are transferred to remote message servers (RMS) designated by the recipients as their message pickup stations. Each Winlink user can register up to three message pickup stations for redundancy. Modems specially designed for HF transmission (Pactor for example) are used to establish and route the traffic over the specific frequencies.

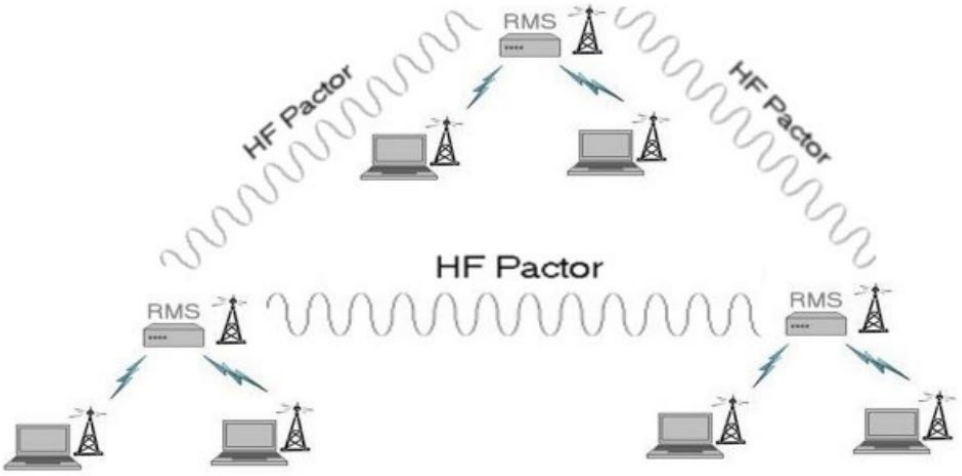


Figure 9 - Winlink HF Email Routing

Note, using this method of moving messages will be slower than what we are accustomed to with highspeed based Internet mail routing. Wherever possible the message should be plain text, simple and to the point as to avoid congesting the remote message servers.

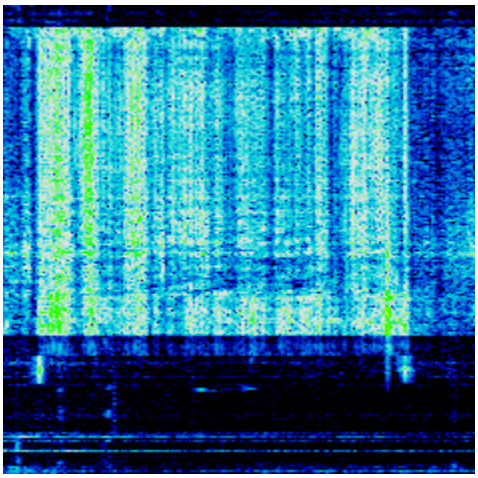
8.2. Automatic Link Establishment

Automatic Link Establishment (ALE) is an automated calling system that brings telephone-like functionality to HF communications. Instead of forcing operators to listen for calls, ALE uses digital signals to indicate that one station wishes to connect/call another. The receiving transceiver then alerts its operator (much like a ringing telephone), who can then accept the call. Once the call is accepted the two stations can use voice communications or switch to a digital mode, e.g. MT63, to transfer data, much like using a modem on a telephone call.

ALE is designated for use on the channels designated “SHARES ALE Net.” SHARES stations are encouraged to program these frequencies into their ALE radios. Sounding is a process to determine link quality analysis can be enabled to measure connectivity with other ALE stations. The ALE radio should not exercise its sounding function more than one-hour interval (ninety minutes is the preferred setting for equipment that supports it). ALE address codes and ALE frequencies are provided to licensed SHARES operators when completing their application to the program. The availability of ALE operations may be different from the hours of operation for voice nets. One of the strengths of using ALE is the radio will help optimize the experience/voice performance based on up to the minute propagation conditions.

8.3. MT63

During times of poor propagation, station operators and SHARES net controllers can leverage a broadcast technique to help get the important message through. SHARES has standardized on MT63 modulation/demodulation methods for this use case. MT63 is an orthogonal frequency division multiplexed (OFDM) digital data mode aimed for use in high noise environments. MT63 was developed by Paweł Jałocha. MT63 is designed for keyboard-to-keyboard (like text messaging) conversation modes on HF bands. MT63 distributes the encoding of each character over a long time period and over several tones. This code and symbol spreading implementation is key to its robustness under less than ideal conditions. The MT63 mode is very tolerant of mistuning, as most software will handle 120 Hz tuning offsets under normal conditions. MT63 uses 64 binary phase shift keying (BPSK) (2-PSK) channels placed in 500, 1000, and 2000Hz of bandwidth. There are 2 main modes of transmission: short interleaving and long interleaving. With short interleaving, MT63's robustness is somewhat compromised in exchange for lower latency (time to end of transmission). With long interleaving (what SHARES leverages), MT63 operates at its best robustness in exchange for a longer latency (about double the latency of short interleaving). Figure 10 represents the source broadcast visual and the computer converted text.



BT
UNCLAS
SUBJ: MT-63 Broadcast - 23 June 2021
1. Conditions continue to be very poor so please listen for stations attempting to c,k 3YJ?;(_GM3=▯(2N\`XfUi8nJUroI station does nop get back to a station please advise you have a "relay" and Net Control will get you to check that station into the net.
2. Please keep an eye on the weather; as noted this week there has been several severe storms that moved through the various parts of the country causing much property damage and the cost of approximately 13 lives. Please use caution if storms are or move into

Figure 10 - Visualization of MT63 Transmission

8.4. Encryption

The SHARES network does permit the use of encryption of messages. Traffic exchanged should be unclassified in nature. The use and access to the encryption software is restricted to SHARES members. Note, if a message is being encrypted, be sure that employees having the authority to permit such a message to the SHARES recipient is approved by senior company employees. Remember, the nature of SHARES is to exchange important information pertinent to the continuation of restoring conditions to business as usual.

9. Voice

The most common mode of communication for SHARES is voice. When operating the station, it is important to remember this service has a mission of safety and vital information exchange so providing the minimal amount of information that is required to get the message across is key. A typical exchange would be:

"This is SHARES CALL SIGN (spelled out phonetically) I repeat SHARES CALL SIGN (no phonetical spelling) in STATE of OPERATION (physical location) with the following priority traffic for the SHARES NET: insert message."

Listening stations will often acknowledge the message and request additional information such as how long you will remain on frequency, status of station power (estimated runtime) and wellbeing of people/operator at the station.

It is important to log any corresponding messages and call signs that are engaged with the exchange of information. Several hours could pass, conditions may change, or additional information could be received that impact an initial touch base with another SHARES operator and it is important to remain in contact with the NET for proper up to the minute information relay.

10. Local Communications

In the area where the HF skip zone is experienced (region under the NVIS propagation where there is very little to weak signal) an alternate supporting means of radio communication is required. This could be a man-pack designed for short distance HF communications (like in the military). This solution would be very tactical in nature supporting the actions of actual restoration processes. Key for this plan is not to depend on towers like in the case of cellular infrastructure as this infrastructure may not be present (why SHARES is being activated).

11. Conclusion

A key provision to having a solid plan for business continuity is to have the necessary tools for addressing the situation ahead of time and not trying to scramble and find the tools needed DURING an incident. SHARES station deployment needs to be planned for ahead of time, coordinated with team members on when to exercise (ideally monthly) the equipment and keep up to date with the SHARES coordinating office to ensure the latest frequencies/information is at hand when and where it is needed.

In summary, here are key components required to get on the SHARES network:

1. ALE ready HF radio

2. Broadband HF antenna capable of transmitting and receiving on frequencies between 3 MHz and 30 MHz
3. Antenna mast system
4. Proper space and elevation of 10 feet or higher to enable radio signal propagation
5. Necessary coax cabling to connect antenna to radio
6. Reliable power supply compatible with the selected HF radio
7. Backup power in the event primary power fails
8. Laptop computer with digital broadcast demodulation software such as MixW or Fldigi configured to receive MT63 transmissions

Finally, when evaluating the setup of radio stations, consider securing both a fixed and portable solution. The portable solution can be shipped/setup when and where needed. A logical location for a fixed station should be at a hardened facility such as a data center, network operations center (NOC) or other strategic critical facility designed to meet SCTE 226 Class A specifications. This 2021 Cable-Tec Expo paper along with SCTE 206 and SCTE 239 best practices will assist cable broadband providers optimize their readiness of major disasters.

Abbreviations

ALE	automatic link establishment
ARSFI	Amateur Radio Safety Foundation
BPSK	binary phase shift keying
CISA	Cybersecurity and Infrastructure Security Agency
EM	electromagnetic
DHS	Department of Homeland Security
FEMA	Federal Emergency Management Agency
HF	high frequency
kW	kilowatt
LED	light emitting diode
MHz	megahertz
NDA	non-disclosure agreement
NET	network
NOAA	National Oceanic and Atmospheric Administration
NOC	network operation center
NVIS	near vertical incidence sky wave
OFDM	orthogonal frequency division multiplexed
Ohm	ohms
RCS	regional coordinators for SHARES
RF	radio frequency
RMS	remote message server
SCTE	Society of Cable Telecommunications Engineers
SHARES	SHAred RESources
SMTP	simple mail transport protocol
SSB	single side band
SWR	standing wave ratio
TFD (T2FD)	terminated folded dipole
TP	toilet paper

US	United States
VHF	very high frequency

Bibliography & References

Signal Identification Guide https://www.sigidwiki.com/wiki/MT63

CISA Power Best Practices:
https://www.cisa.gov/sites/default/files/publications/Factsheet%20Resilient%20Power%20Best%20Practices.pdf

History of Cable by California Cable & Telecommunications Association:
https://calcable.org/learn/history-of-cable/

A collection of documents and forms related to the SHAred RESources (SHARES) High Frequency (HF) Radio Program: https://www.cisa.gov/publication/shares-documents

WinLink HF Email: https://winlink.org/

Guide for the Selection of Communication Equipment for Emergency First Responders:
https://www.ojp.gov/pdffiles1/nij/191160.pdf

ARRL Resources

Emergency Power for Radio Communications 2012
https://www.arrl.org/shop/Emergency-Power-for-Radio-Communications

Portable Operating for Amateur Radio 2108
http://www.arrl.org/shop/Portable-Operating-for-Amateur-Radio

Seven days of ionospheric conditions as observed by a global network of ionosondes
https://www.sws.bom.gov.au/HF_Systems/6/5

Tracking Round Trip Time Latency in the MSO Network

A Technical Paper prepared for SCTE by

Michael Overcash
Principal Engineer
Cox Communications
6305 Peachtree-Dunwoody Rd, Atlanta, GA 30328
404-269-6595
michael.overcash@cox.com

Alan Skinner
Principal Engineer
Cox Communications
6305 Peachtree-Dunwoody Rd, Atlanta, GA 30328
404-269-0845
alan.skinner@cox.com

Owen Parsons
Engineer
Cox Communications
6305 Peachtree-Dunwoody Rd, Atlanta, GA 30328
404-269-4998
owen.parsons@cox.com

Daniel Sciscoe
Network Engineer
Cox Communications
6305 Peachtree-Dunwoody Rd, Atlanta, GA 30328
daniel.sciscoe2@cox.com

Elizabeth Vitale
Network Engineer
Cox Communications
6305 Peachtree-Dunwoody Rd, Atlanta, GA 30328
elizabeth.vitale@cox.com

1. Introduction

Forget throughput – latency is the new standard of internet quality. Whether it's a glitchy videoconference or a "laggy game," it's increasingly important to know how latency is impacting customers, and how it interacts with the network components we control. In this paper we will describe work done with Raspberry Pi-based test points that measure roundtrip time, jitter and packet loss, using realistic UDP streams. We will also share some of the early data we've collected and discuss what we've learned so far.

2. Lag Overview and Project Motivation

Subscribers use "lag" to describe an aggregation of latency, jitter, and packet loss; a poorly performing service or application is described as "laggy". Historically, subscriber internet use was dominated by HTTP web browsing and streaming protocols that download a video segment at a time – neither of which is particularly sensitive to lag. Today, popular applications like real-time video conference and online gaming are extremely lag sensitive. In addition, the FCC SamKnows program now tracks latency in addition to speed.

CableLabs is addressing this need in the Low-Latency DOCSIS (LLD) program, and vendors are introducing various product features intended to improve lag. But how can operators know if these new products are effective? Lab testing can help of course, but there is no substitute for field loading and actual customer traffic patterns. Many new features over the years have shined in a lab and provided less-than-stellar results once deployed. Likewise, every new product comes with a slew of configuration parameters … are these parameters tuned correctly? Vendors typically provide recommended starting values, and often these values persist into perpetuity without being critically examined to ensure that they're optimal. We can do better, but only if we can measure the results each time we turn a knob.

Just as operators have monitored node utilization for years, operators need a platform to monitor and track lag over time throughout the network. Thus, the Cox LagSpy program was born.

2.1. LagSpy Goals

Cox has the following goals for the program:

- Use realistic UDP streams to measure latency, rather than ICMP pings.
- Ability to distribute test points widely throughout the Access Network.
- No special configuration of subscriber CPE equipment (e.g. no need for port forwards.)
- Upgradable with ability to add new features and test protocols over time.
- Low hardware cost.
- Configurable network utilization.
- Portable software

2.2. LagSpy Overview

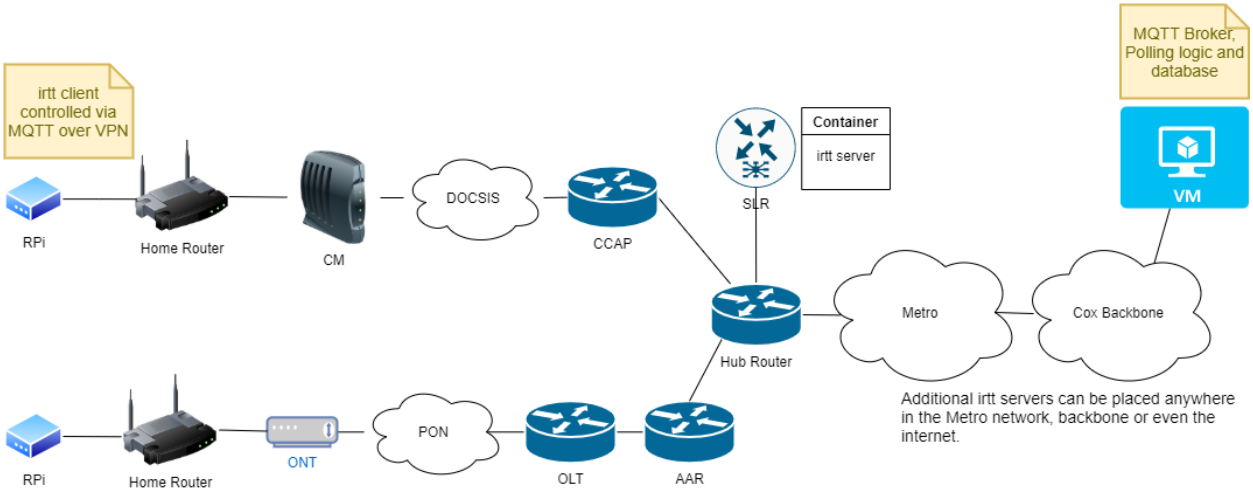


Figure 1 – LagSpy Proof of Concept Architectural Overview

In the proof of concept architecture, Raspberry Pi single board computers were mailed to employee volunteers. These Raspberry Pi devices are nicknamed “Lag-Pis”. Each Lag-Pi implements a network test client called IRTT (Isochronous Round-Trip Time). The Lag-Pis are controlled by a virtual machine on the internet known as the Poller, which also aggregates and stores the test results. The Poller instructs each Lag-Pi when to perform a test and specifies the IRTT server to test against.

The IRTT test involves a client and a server (similar to IPERF). The Lag-Pis implement the client role, and the servers are set up at interesting points in the Cox network. For the proof of concept trial, we are setting up IRTT servers at the Service Layer Router (SLR) connected to the Hub Router. The SLR was selected since neither the Access Router itself (CCAP or OLT) nor the hub router can act as an IRTT server. To minimize latency to the Lag-Pi, we wanted something as close to the access network as possible. We are also setting up IRTT servers at two of our Regional Data Centers (RDCs) for comparison to the SLR.

While our primary focus in the proof of concept is the DOCSIS Network, the architecture is transport agnostic and can run on any access network (DOCSIS, PON, Cellular Data, etc.)

2.3. Why IRTT?

IRTT is a widely available open-source package which generates a customizable UDP stream to measure route trip time and jitter, among other things. Here is an example of an IRTT test and its results:

```
irtt client -i 20ms -l 172 -d 30s --fill=rand --sfill=rand --hmac=0x<redacted> -q irtt-
telemetry.coxlab.net:22112
[Connecting] connecting to irtt-telemetry.coxlab.net:22112
[184.176.185.20:22112] [Connected] connection established
[184.176.185.20:22112] [WaitForPackets] waiting 352ms for final packets

                         Min     Mean   Median      Max  Stddev
                         ---     ----   ------      ---  ------
                 RTT  78.92ms  84.55ms  83.55ms  117.3ms  3.17ms
          send delay   -1.24s   -1.23s   -1.23s   -1.21s  2.25ms
       receive delay    1.31s    1.32s    1.32s    1.35s  2.22ms

       IPDV (jitter)   1.93µs   2.32ms   1.13ms  34.77ms  3.15ms
           send IPDV    110ns   1.89ms    925µs  19.11ms  2.41ms
        receive IPDV    754ns    740µs    274µs  34.42ms  2.31ms
```

```
    send call time   12.9µs     72µs             932µs  46.5µs
       timer error    100ns    129µs             827µs   107µs
 server proc. time   4.45µs   9.39µs             128µs  4.86µs

              duration: 30.3s (wait 352ms)
  packets sent/received: 1471/1471 (0.00% loss)
server packets received: 1471/1471 (0.00%/0.00% loss up/down)
    bytes sent/received: 253012/253012
      send/receive rate: 67.5 Kbps / 67.5 Kbps
          packet length: 172 bytes
            timer stats: 28/1499 (1.87%) missed, 0.64% error
```

In this example, the UDP payload size was set to 172 bytes, the inter-packet interval is 20 ms, and the test ran for 30 seconds. The total bandwidth was 67.5 Kbps, approximating a UDP audio stream. The round-trip time was on average 84.55 ms with an average jitter of 2.32 ms. These are the parameters Cox is currently using in the trial, but we are investigating other streams to model (see Section 3.3.2).

IRTT is launched in either client or server mode. The client generates the test traffic, which is reflected back by the server. The client compares the original packet to its reflected version to calculate the performance statistics. In our architecture, the Lag-Pi is acting as the client.

IRTT is extremely easy to deploy. On a Debian/Ubuntu Linux VM, you can install it simply by running `apt-get install irtt`.

2.3.1. UDP versus ICMP for Latency Testing

Many tools (e.g. the Ookla Speed Test) use an ICMP-based tool like ping for latency measurements. We selected a UDP-based tool as we believe this more accurately reports the subscriber experience.

UDP based measurements are more accurate because:

- Different QoS is generally applied to UDP versus ICMP.
- Internally, many devices implement ICMP as a control plane protocol and UDP as a data plane protocol. So for example, many devices employ hardware acceleration for TCP and UDP, but hardware acceleration is rare for ICMP.
- Real applications use TCP or UDP to transmit data. No common applications deliver user data using ICMP.
- Many devices rate limit ICMP handling for DDOS protection. IRTT sends a continuous stream of UDP data, but this is impossible in ICMP. A high ICMP packet rate is often treated as a ping flood DoS attack and will be blocked.

2.4. Why Raspberry Pi?

Raspberry Pis were selected for the pilot based on their low cost and small form factor. They can be cheaply mailed in bubble envelopes. We used the 2GB model of the Raspberry Pi 4B. Note that the networking is significantly improved from the Raspberry Pi 4 versus the 3. Specifically, the Raspberry Pi 4 can support gigabit ethernet speeds while earlier models were throttled by a slow USB 2.0 bus connecting the main SoC to the networking chip. This architecture limited the ethernet throughput to about 300 Mbps.

We will not use Raspberry Pis for large scale deployment. For mass deployment, we will incorporate a LagSpy client into Cox managed gateway devices.

2.5. Employee Pilot

We solicited 19 employee volunteers for the pilot from a DOCSIS working group and mailed pre-configured Lag-Pis to them. The volunteers simply plugged in the power adaptor and connected the Lag-Pis to an ethernet port on their LAN. The volunteers are distributed across multiple Cox markets.

Note that while the Raspberry Pi 4 supports Wi-Fi, we have disabled it and are relying exclusively on gigabit ethernet for LAN connectivity. Wi-Fi introduces a significant amount of latency and jitter, and the focus of this project is on the access network. However in a future project, we could easily measure Wi-Fi latency and jitter in this architecture simply by enabling the Wi-Fi on the Lag-Pi, and setting up an IRTT server on the access point.

3. LagSpy Technical Deep Dive

3.1. Software Architecture

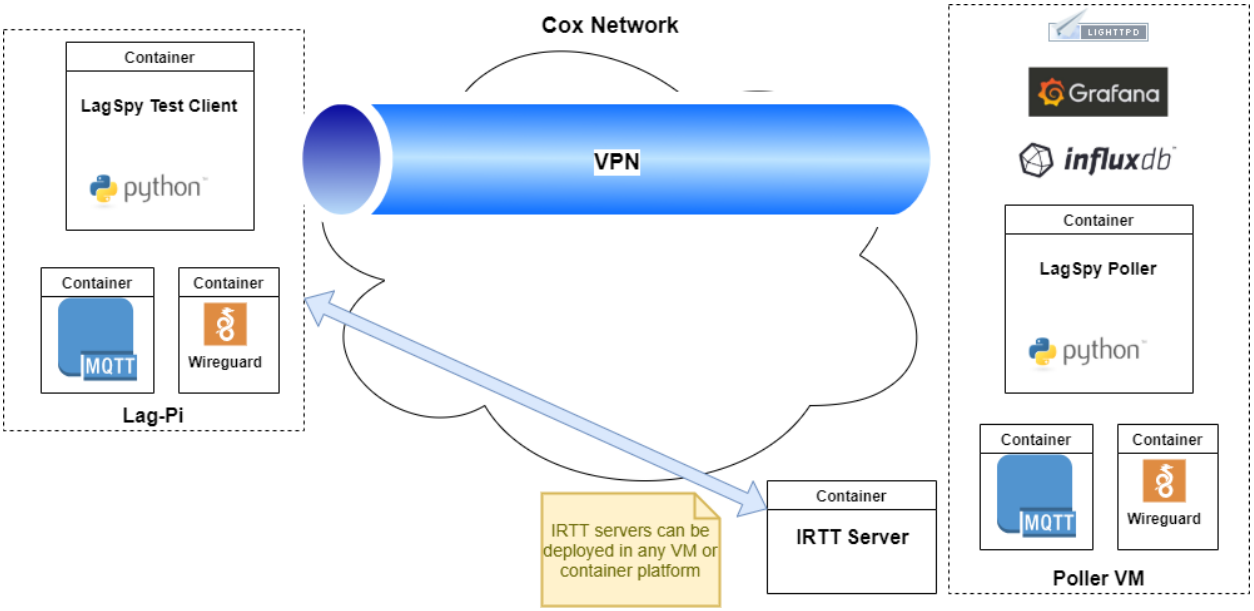


Figure 2 – LagSpy Software Stack

The principal components of the Lag-Pi are:

- The LagSpy Test Client, written in Python 3.8.
- Eclipse Mosquitto to implement an MQTT client.
- Wireguard to establish a VPN connection to the Poller for command and control.

The principal components of the Poller are:

- The Lagspy Poller, written in Python 3.8.
- Eclipse Mosquitto to implement an MQTT broker and localhost client.
- Wireguard for a VPN endpoint.
- InfluxDB to import and aggregate data from the Poller for visualization.
- Grafana for visualization of test results.

- Lighttpd (primarily used to upgrade the Lag-Pi.)

Note that MQTT and Lighttpd are only exposed over the VPN interface over the 10.13.0.0/16 subnet.

The IRTT Servers are implemented in containers, which can be deployed virtually anywhere as IRTT is a simple protocol that only exposes a single UDP port. Docker (and other container frameworks) can strictly limit the resources available to an applications, including memory, CPU, and network sockets, greatly reducing the risk to the platform.

3.1.1. Connectivity Driven by Client

The VPN tunnel is initiated from the subscriber side, meaning that no special configuration is required on the home gateway. The connection is initiated from the LAN side. The Wireguard `PersistentKeepalive` feature prevents the home gateway's NAT state from timing out due to inactivity.

Once the VPN tunnel is established, the Poller can initiate MQTT traffic at will to the Lag-Pi. No port forwarding is required.

Likewise, IRTT traffic is initiated by the Lag-Pi (client) and does not require special forwarding rules.

3.2. Command and Control

MQTT, a popular IOT control protocol, is used to manage the Lag-Pis. MQTT is a Publisher-Subscriber (Pub/Sub) protocol. Clients publish messages to named channels called "topics". Clients can subscribe to any topic of interest. MQTT creates a reliable mechanism to establish both 1:1 communication with individual Lag-Pis, and also to send messages to multiple Lag-Pis at once.

The MQTT Broker coordinates the message forwarding and is implemented on the Poller.

For LagSpy, all message payloads are in JSON format.

Table 1 – LagSpy MQTT Topics

Topic	Arguments	Direction	Description
connect/hello	n/a	Lag-Pi → Poller	Register with poller and keepalive
connect/enroll/<mac>	MAC address of Lag-Pi	Poller → Lag-Pi	Provide VPN credentials to Lag-Pi
connect/link_ok/<mac>	MAC address of Lag-Pi	Poller → Lag-Pi	Keepalive response
irtt/start/<mac>	MAC address of Lag-Pi	Poller → Lag-Pi	Start IRTT test
irtt/results	n/a	Lag-Pi → Poller	Results of IRTT test
iperf/start/<mac>	MAC address of Lag-Pi	Poller → Lag-Pi	Start IPERF3 test
iperf/results	n/a	Lag-Pi → Poller	Results of IPERF3 test

3.2.1. Wireguard Certificate Enrollment

The Wireguard VPN requires that each peer have a unique Private Key and a statically assigned IP address in the VPN subnet. As we did not want to manually provision each Lag-Pi with VPN credentials prior to shipment, we implemented a dynamic VPN credential mechanism over MQTT.

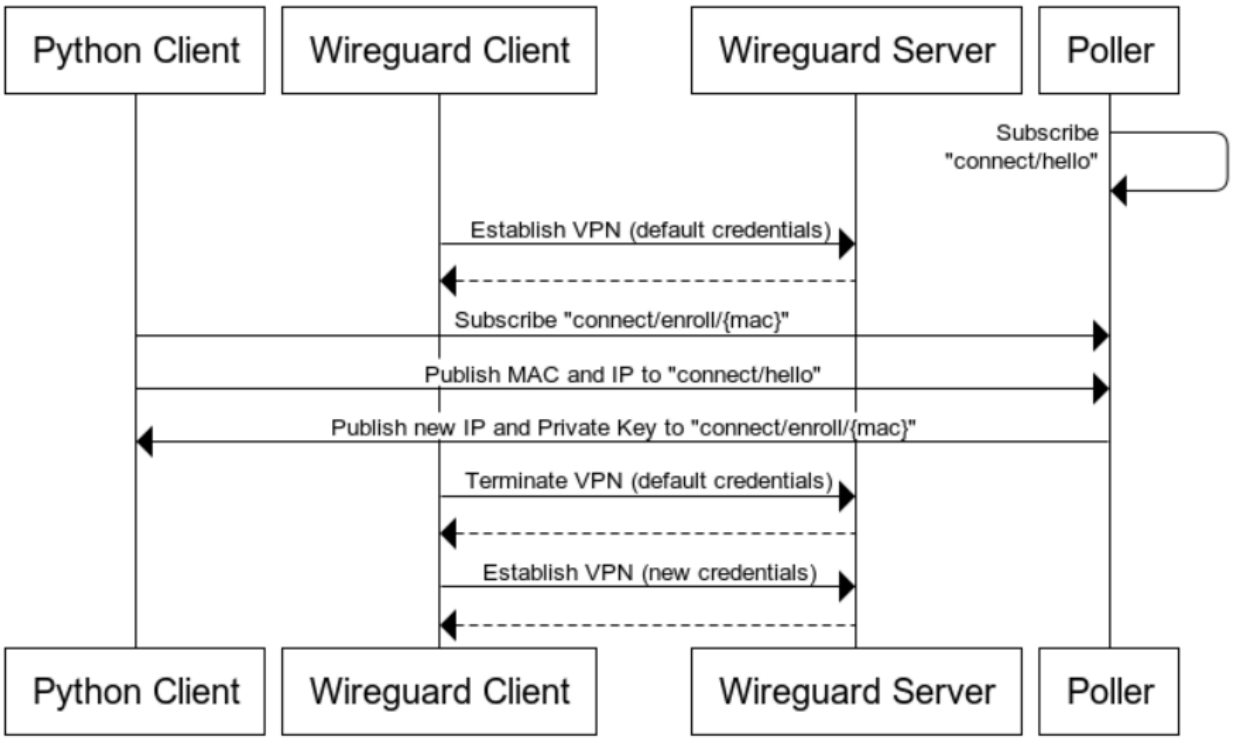


Figure 3 – Wireguard VPN Enrollment Process

A default set of credentials are used to establish the permanent credentials. Wireguard only allows one endpoint at a time to connect with the default private key. This enrollment process can gracefully handle a small amount of contention for the default credentials, but this algorithm will need to be revisited as we increase scale.

3.2.2. Network Failsafe Keepalive

If the Wireguard VPN fails, then the Lag-Pi becomes unmanageable. To mitigate this, each Lag-Pi runs a cron job that verifies the ability to download a small file from the Poller. If the file download fails, the script ensures the Wireguard container is running.

3.2.3. Debugging with Mosquitto

MQTT is a pub-sub protocol, meaning that any authenticated party can subscribe to a topic. Therefore a user on the Poller can subscribe to interesting topics for debugging purposes, and sniff the control messages sent over the VPN. This is very useful for troubleshooting.

In the example below, we can observe the “connect/hello” messages received from the Lag-Pis.

```
lagspy@poller-dt1-phx-0:~$ docker exec -it mosquitto mosquitto_sub -t
"connect/hello"

{"publicIpv6": "::", "publicIpv4": "68.109.32.146", "ip": "10.13.0.16", "mac":
"e4:5f:01:3b:18:23", "topic": "connect/hello", "message": "hello", "seqNum":
5393, "timestamp": "2021-07-19 15:27:49.829820"}

{"publicIpv6": "2600:8800:1a1:1a00::b71a", "publicIpv4": "", "ip":
"10.13.0.12", "mac": "e4:5f:01:21:a0:d4", "topic": "connect/hello", "message":
"hello", "seqNum": 2924, "timestamp": "2021-07-19 15:27:54.517621"}
```

3.3. IRTT Deep Dive

3.3.1. Limitations

IRTT attempts to break up the measured Round Trip Time (RTT) into "Send Delay" and "Receive Delay". However this decomposition is based on injecting a timestamp into the test packets, which requires precise time synchronization between the client and server. This doesn't seem to work in any practical environment that we've tested using Raspberry Pis or even Windows PCs. If the RTT is low enough, one of the components will be reported as a negative number. As we are fairly confident that the Cox network does not support time travel, we must conclude that this is due to clock skew between the client and server.

We have attempted to achieve better time sync by using GPS modules without success. We also tried to use NTP, where the NTP peer, the NTP server, and the NTP client were all on the same ethernet switch. This didn't work either.

We suspect the only way to get this to work is to use IEEE 1588 (Precision Time Protocol) with IRTT endpoints that use precision real-time clocks. This is out of scope for the LagSpy program, since our long-term plan is to use consumer grade internet gateways. Integrating LagSpy into a cable modem that implements DOCSIS Time Protocol could provide a way to get this precision.

IRTT sends symmetrical bidirectional test traffic. For example, if IRTT is configured to send 60 kbps upstream, then an identical 60 kbps stream is generated in the downstream direction as well. Unlike iperf3, it is not possible to perform unidirectional testing.

However, due to the nature of the DOCSIS protocol and the use of TDM in the downstream vs. TDMA in the upstream, the major contribution of latency and jitter is in the upstream. Downstream traffic of this nature will simply be forwarded along by the CMTS into the plentiful frames available to each modem's SFID (assuming the modem is below its QoS limit). This generally happens at or very near real time. In the upstream, however, the modem must request a timeslot in a contention-based interval first, then wait for the CMTS scheduler to allocate a timeslot, then wait for the MAP message to arrive, and finally send the data. This process is more sensitive to network congestion and has more possibility of variation due to limited contention intervals. Therefore, we make the assumption that the changes in latency and jitter over time are primarily due to the fluctuating conditions of the upstream.

3.3.2. IRTT Traffic Profiles

IRTT streams are highly customizable, so we can model a number of latency sensitive applications. Currently we are using a profile that simulates an audio stream, but we plan to add more profiles to our testing toward the end of 2021.

CableLabs has studied typical data streams from collaboration apps (see References). Their summary findings are below.

Table 2 – CableLabs Video Confrerencing Bandwidth Summary. Applications were anonymized by CableLabs.

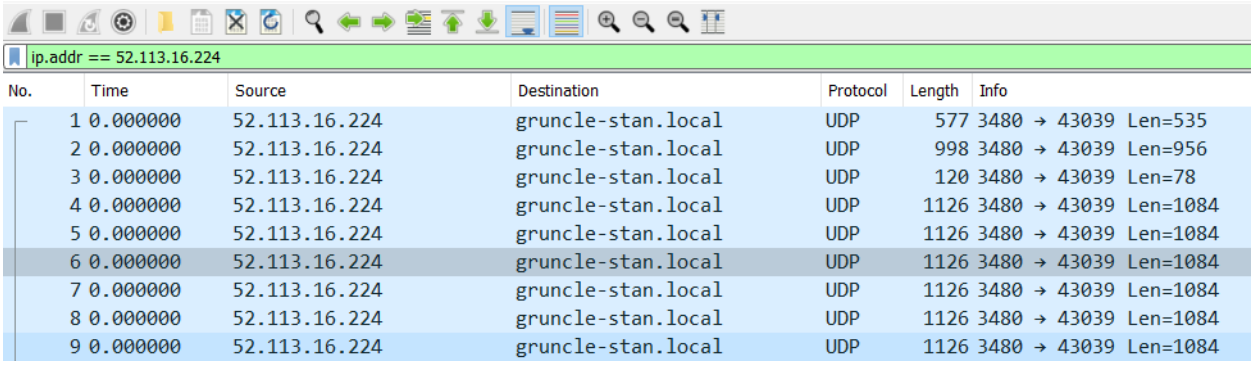
Application	Typical Upstream	Typical Downstream
Application A	< 500 kbps	Between 1 and 2 Mbps
Application B	200 kbps typical; one outlier at 2 Mbps	Approx. 1 Mbps
Application C	Approx. 350 kbps	Approx. 2 Mbps
Application D	Between 200 kbps and about 1.8 Mbps	Cluster at 500 kbps and cluster at 3 Mbps

While video conference traffic is highly asymmetrical, IRTT can only model symmetric streams. We recommend modeling the upstream data rate since most latency and jitter should be introduced in the upstream.

3.3.2.1. Characterizing Application Streams

Characterizing application streams is straightforward with a packet sniffer application such as Wireshark.

- Optional: place an ethernet switch with a mirror port between the device under test (DUT) and the LAN. These switches are widely available for under $50.
- Launch the app generating the traffic on the DUT.
- Start a Wireshark capture. For apps that run on a computer, you can run Wireshark on the same machine. For something like a game console, the optional ethernet switch must be used. Or, many prosumer/commercial home routers have a sniffer capability.
- Collect data for several minutes, then end the capture.
- Identify the traffic of interest and apply a display filter. The filter can be applied in the upstream, downstream, or bidirectional as desired. In this example:
    - Use `ip.addr==52.113.16.224` for bidirectional.
    - Use `ip.src==52.113.16.224` for downstream.
    - Use `ip.dst==52.113.16.224` for upstream.
    - Take care to filter based on the remote application server (public IP address) rather than the LAN device.



- Go to Statistics→IO Graphs in Wireshark. Set up Display Filters for upstream, downstream, and bidirectional. You can inspect both bitrate and packet rate by selecting the appropriate option for the Y Axis.

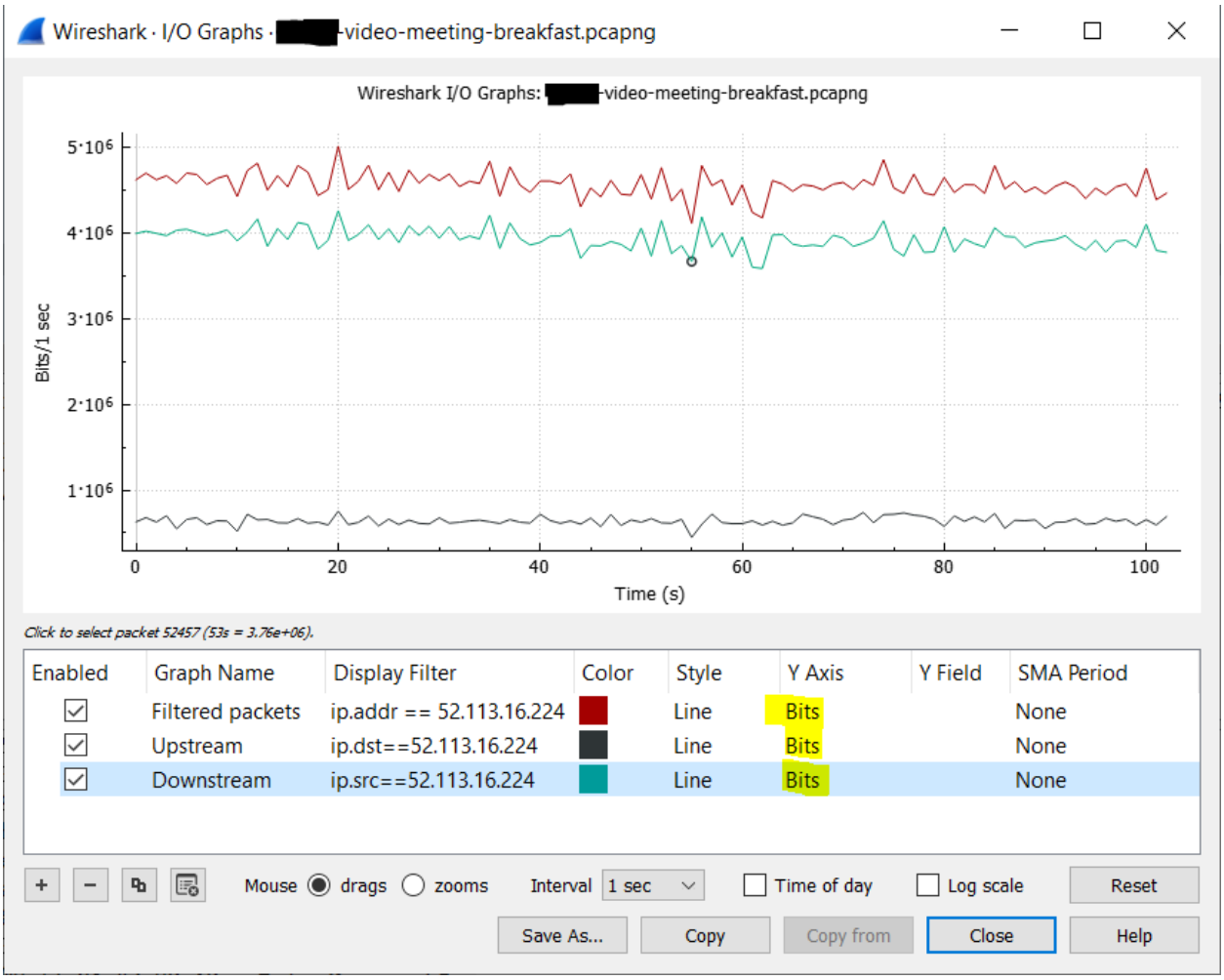


Figure 4 – Example video conference data rate. Note that "Bits" is selected on the Y Axis, and the Interval is set as 1 second. This is an asymmetric stream, with a typical upstream rate of 500 kbps and a typical downstream rate of 4 Mbps.

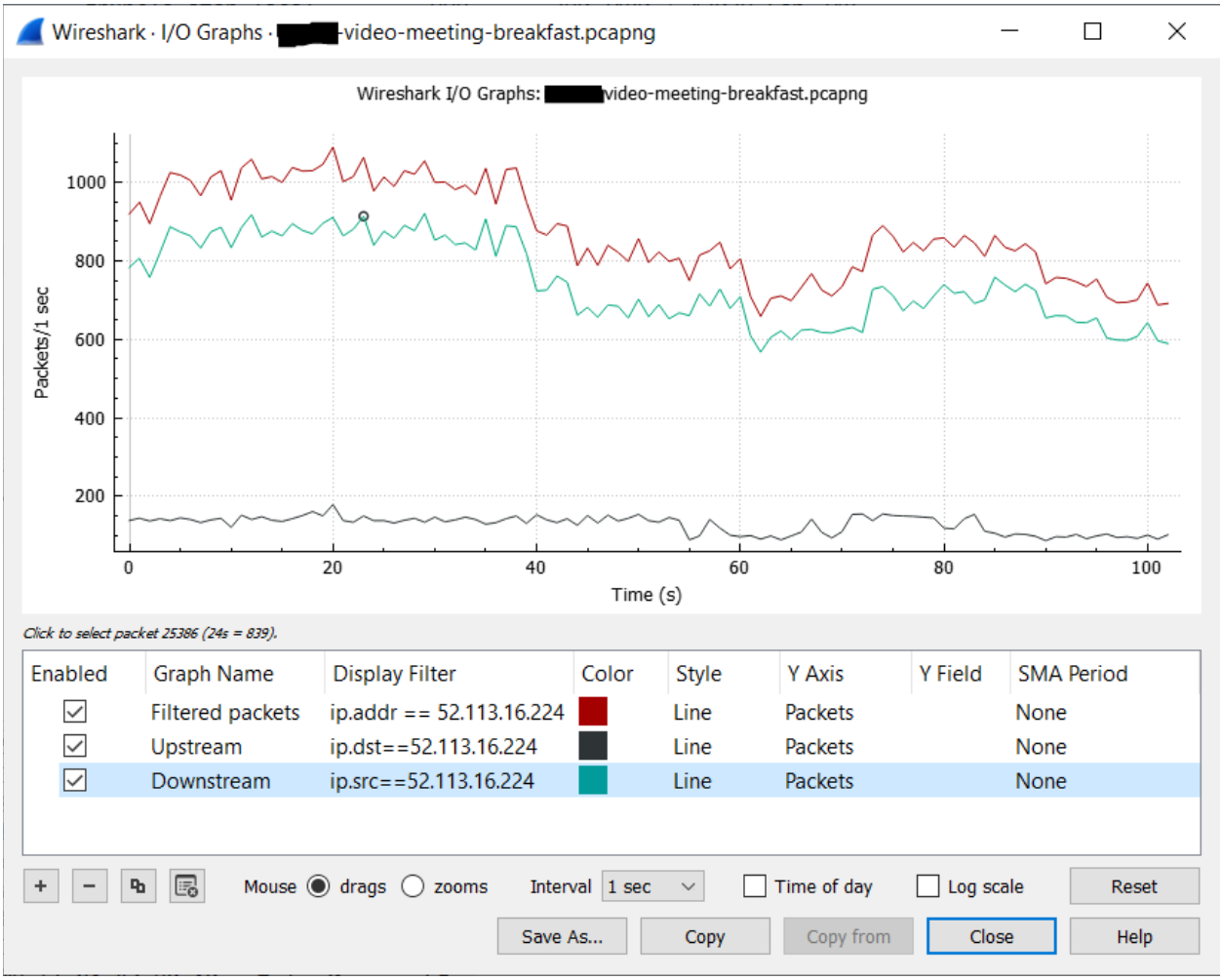
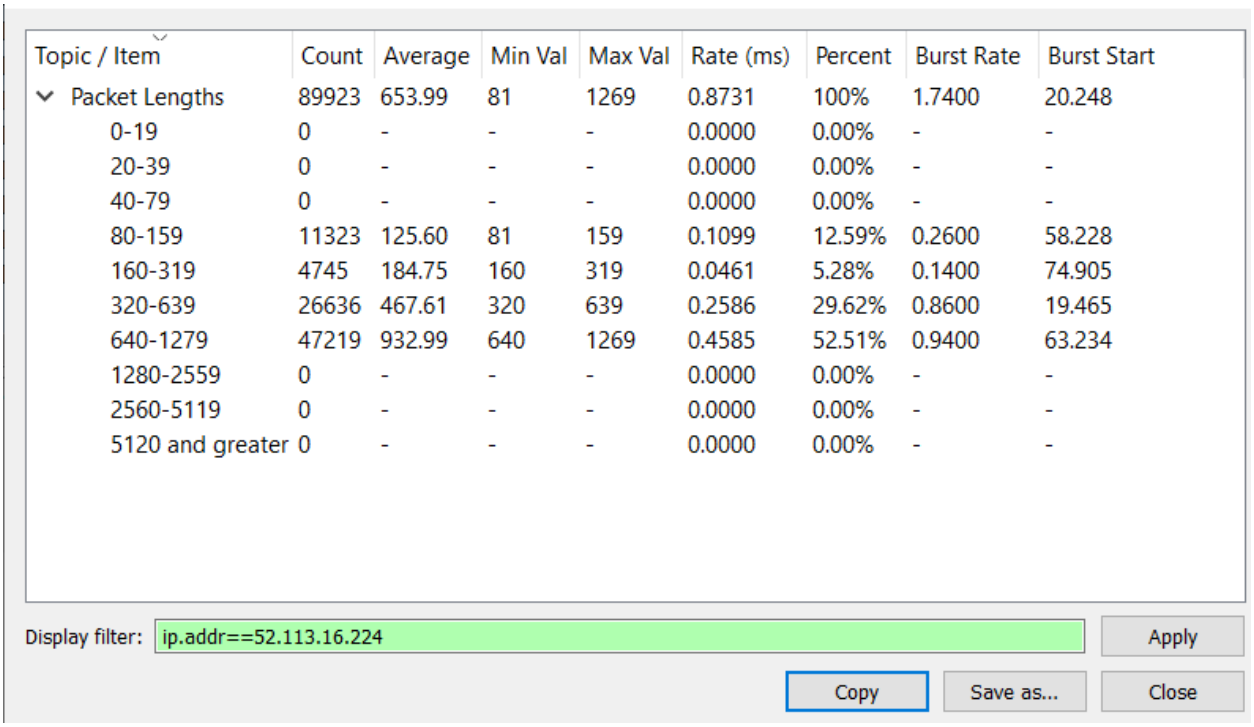


Figure 5 – Example video conference packet rate. Typical upstream rate is 100 pps, and typical downstream rate is 800 pps.

- Inspect the packet length in Wireshark using Statistics→Packet Lengths. Apply the desired Display Filter.



Topic / Item	Count	Average	Min Val	Max Val	Rate (ms)	Percent	Burst Rate	Burst Start
Packet Lengths	89923	653.99	81	1269	0.8731	100%	1.7400	20.248
0-19	0	-	-	-	0.0000	0.00%	-	-
20-39	0	-	-	-	0.0000	0.00%	-	-
40-79	0	-	-	-	0.0000	0.00%	-	-
80-159	11323	125.60	81	159	0.1099	12.59%	0.2600	58.228
160-319	4745	184.75	160	319	0.0461	5.28%	0.1400	74.905
320-639	26636	467.61	320	639	0.2586	29.62%	0.8600	19.465
640-1279	47219	932.99	640	1269	0.4585	52.51%	0.9400	63.234
1280-2559	0	-	-	-	0.0000	0.00%	-	-
2560-5119	0	-	-	-	0.0000	0.00%	-	-
5120 and greater	0	-	-	-	0.0000	0.00%	-	-

Figure 6 – Example video conference packet size. Typical packet size is 933 bytes. Since upstream and downstream histograms were very similar, only the bidirectional analysis is shown here.

From this analysis and focusing on the upstream, this application can be modeled in IRTT using the following parameters:

- Interval (`-i` parameter): 1 / 100 pps = 10 ms
- Length (`-l` parameter): 933 bytes – 28 bytes = 908 bytes. Note that we subtract the IP and UDP header size to calculate this parameter.

As a sanity check, IRTT reports a throughput of 733.5 Kbps, which is reasonably close to the observed value of approximately 500 Kbps.

3.4. Security Considerations

3.4.1. Lag-Pi (IRTT Client)

The Raspberry Pi is not a hardened hardware platform (e.g. it does not have secure boot or secure storage.) We are only using Raspberry Pis for a limited employee trial and will migrate to managed and secured hardware in the next phase.

The LagSpy application does not listen on any open ports. All command and control traffic is secured over a Wireguard VPN tunnel.

3.4.2. Poller

The Poller is listening on a Wireguard server port. All other services (MQTT, HTTP) are restricted to the VPN interface. In other respects, the same hardening considerations apply to the Poller as any standard server with an open port.

3.4.3. IRTT

The IRTT application is written in Golang and is not subject to the buffer overflow-based attacks that plague C/C++ applications.

By default, the IRTT server listens on UDP port 2112 and is discoverable using standard network scanning techniques (including Shodan.io). At a minimum, this allows an attacker to waste network resources by sending test traffic to the server, and could be the basis for a reflection attack.

To mitigate this risk, we recommend the following:

- Override the default port number
- If possible, disable IPv4 on the IRTT server and use IPv6 for IRTT testing. The IPv6 address space is sparse and more difficult to scan.
- Use the IRTT `-hmac` option to enable the HMAC feature. When enabled, the server will not establish a connection or otherwise respond to client traffic that doesn't use the same HMAC value. This means that the port will appear to be blocked by a firewall (filtered) in a standard nmap scan.

3.4.4. SLR

Cox utilizes routers at each hub site that are dedicated to hosting ancillary (non-data path) services. These routers are known as Services Layer Routers (SLR) and are deployed in pairs, directly connected to metro hub routers. The SLR is an appealing place to host an IRTT server because:

- The Docker implementation on the SLR imposes resource limits to mitigate against DoS and other resource-based attacks.
- The underlying router platform supports access controls that block traffic arriving on the IRTT address from impacting the other router functions.

3.5. Test Policy Configuration

Different Lag-Pis operate in different environments and some populations need to execute differents tests. Thus there is a need to implement a flexible policy framework to control how and when tests are executed on a given Lag-Pi.

Our current policy file subdivides testing into 5 different groups, as shown in Figure 7. Our two main group categories are testing groups and server groups. Testing groups specify tests for a member device to run and the server groups provide addresses of available servers.

For testing groups, each group has a list of devices with defined mac addresses. The devices with those mac addresses will run specific tests and write those results to a file in accordance with the permissions of the group. The irtt-testing group sets the mac address as "default", assigning all devices to the `irtt-testing` group automatically; this will require any connected devices to perform irtt tests. In contrast,

the `iperf3-testing` group requires that devices with mac addresses defined in the group, such as "ef:5f:01:3b:18:23", execute iperf3[1] tests in addition to irtt tests.

For server groups, each group contains a list of IP addresses running case-specific servers. These groups allow devices within the testing groups to obtain a list of Ips to run their tests against. For example, the `irtt-IPv4-server-Ips` group contains IP addresses running an IPv4 irtt server. As a result, when a device using IPv4 in the irtt testing group is looking for a server to run a test against, it would get the necessary IP from the `irtt-IPv4-server-Ips` group. Likewise, the `iperf3-server-Ips` group contains a list of Ips running an iperf3 server and the `irtt-IPv6-server-Ips` group contains Ips running an IPv6 irtt server.

```
groups:

        irtt-testing:
                group-name: irtt-testing
                permissions:
                        - run-irtt-tests
                        - write-irtt-results
                enabled: true
                devices: default

        iperf3-testing:
                group-name: iperf3-testing
                permissions:
                        - run-iperf3-tests
                        - write-iperf3-upstream
                        - write-iperf3-downstream
                enabled: true
                devices:
                        - e4:5f:01:3b:18:23
                        - e4:5f:01:3b:17:43

        irtt-IPv6-server-IPs:
                group-name: irtt-IPv6-server-IPs
                IPs:
                        - irtt-telemetry.coxlab.net
                enabled: true

        irtt-IPv4-server-IPs:
                group-name: irtt-IPv4-server-IPs
                IPs:

                enabled: true

        iperf3-server-IPs:
                group-name: iperf3-server-IPs
                IPs:
                        - 192.168.0.43
                enabled: true
```

Figure 7 – Policy file containing different testing and server groups.

[1] Iperf3 testing is not part of the core LagSpy functionality, but we are leveraging the LagSpy command and control framework to automate some iperf3 testing we are doing for product acceptance.

3.5.1. Server Autodiscovery

To avoid the need to manually assign each Lag-Pi to the closest server, we are implementing a simple autodiscovery algorithm to find the nearest server. Once per boot, the Lag-Pi will perform a traceroute to each server in the pool and will select the closest (by hops) for IRTT tests.

3.6. Data Visualization

To gather and present the metrics received from MQTT we have a three-part architecture. First, the metrics sent over the MQTT `irtt/results` topic activate a Python listener that decodes them, normalizes the data, then sends each measurement to an InfluxDB instance. Second, InfluxDB collects the data and applies any active retention policies, then makes the data available for consumers.

Lastly, we assembled a Grafana Dashboard to display a subset of the measurements returned by IRTT. Thanks to both Grafana and the InfluxDB query language (Flux), we're able to not only display the metrics as returned but also partition and display them however we wish. This includes building composite metrics, windowing, and aggregation across the whole metric population to make trends clearer or weed out noise that may be present in the data set.

Our current visualization tools plot results for individual Lag-Pi test points. As we increase scale and deploy more test points, we will need to develop tools to analyze the data in aggregate.

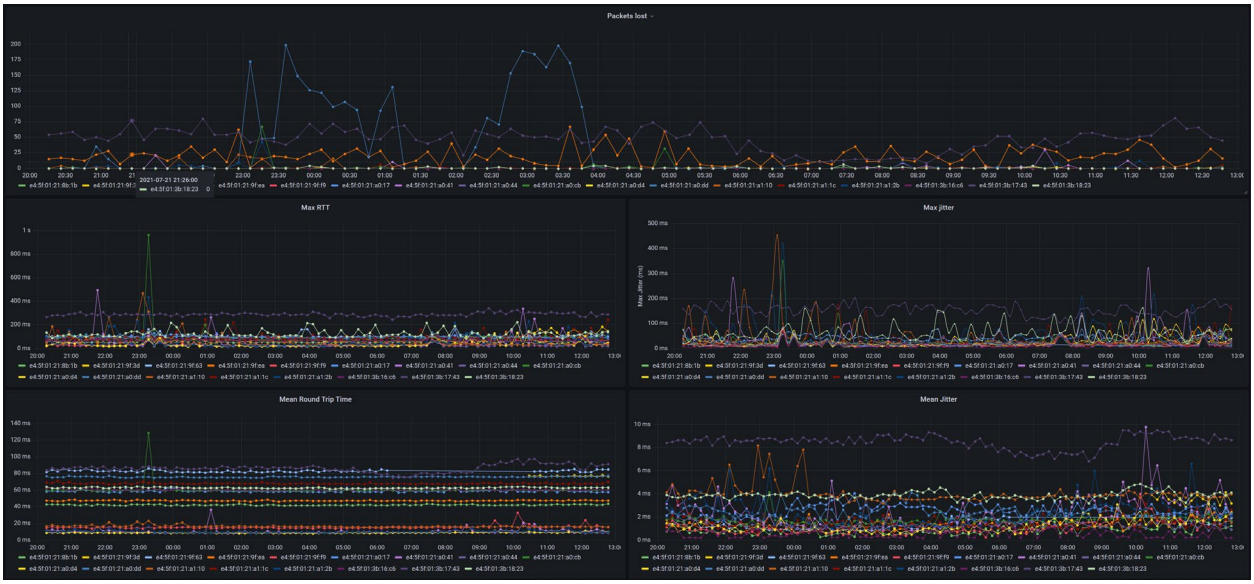


Figure 8 – Grafana Dashboard illustrating: Packet Loss, Max Round Trip Time, Max Jitter, Mean RTT, and Mean Jitter

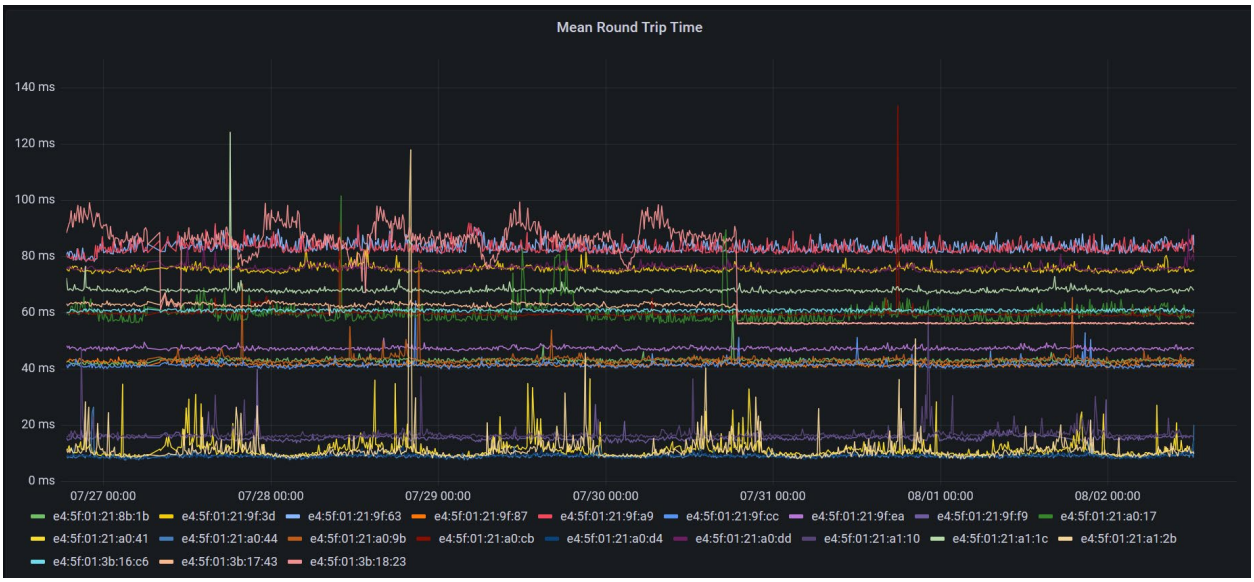


Figure 9 – Mean round trip time over several days of data. Note the cyclic increases for some devices (e.g. the yellow line near the bottom.)

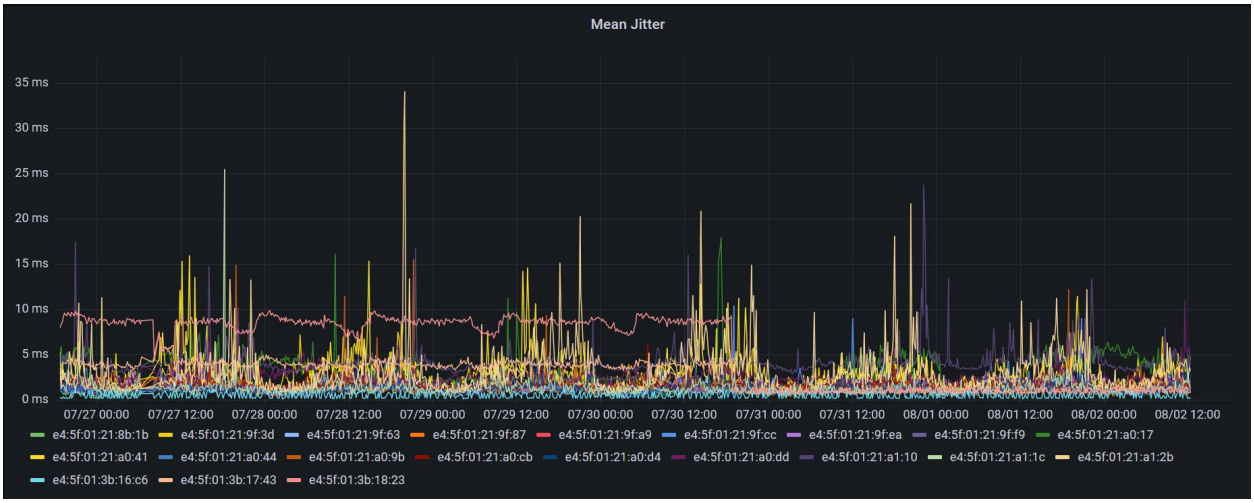


Figure 10 – Mean jitter over several days of data. Again cyclic behavior is evident for a few Lag-Pis.

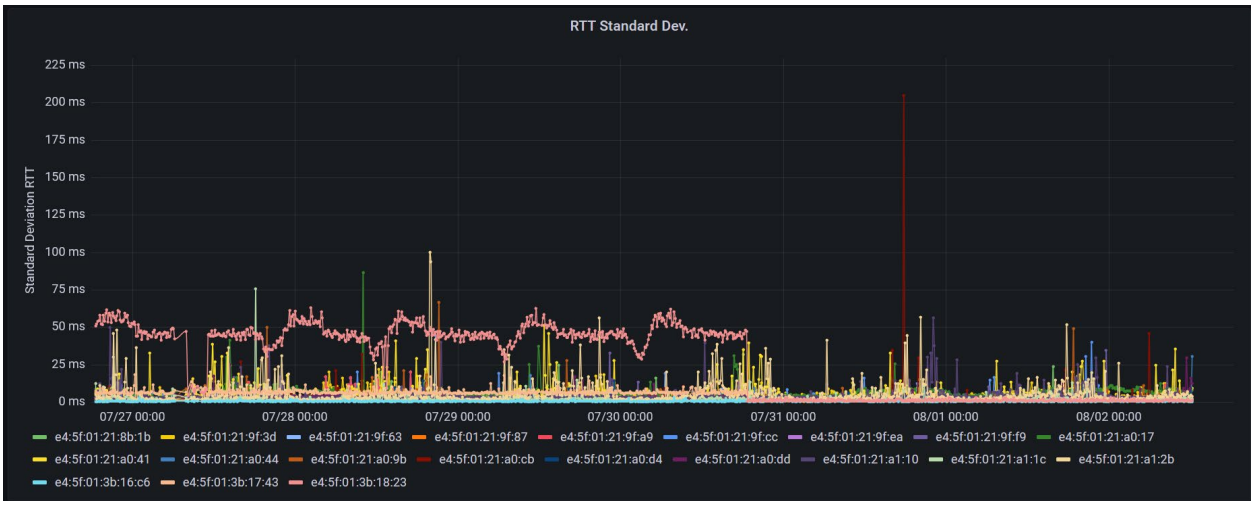


Figure 11 – Round trip time standard deviation across several days of data.

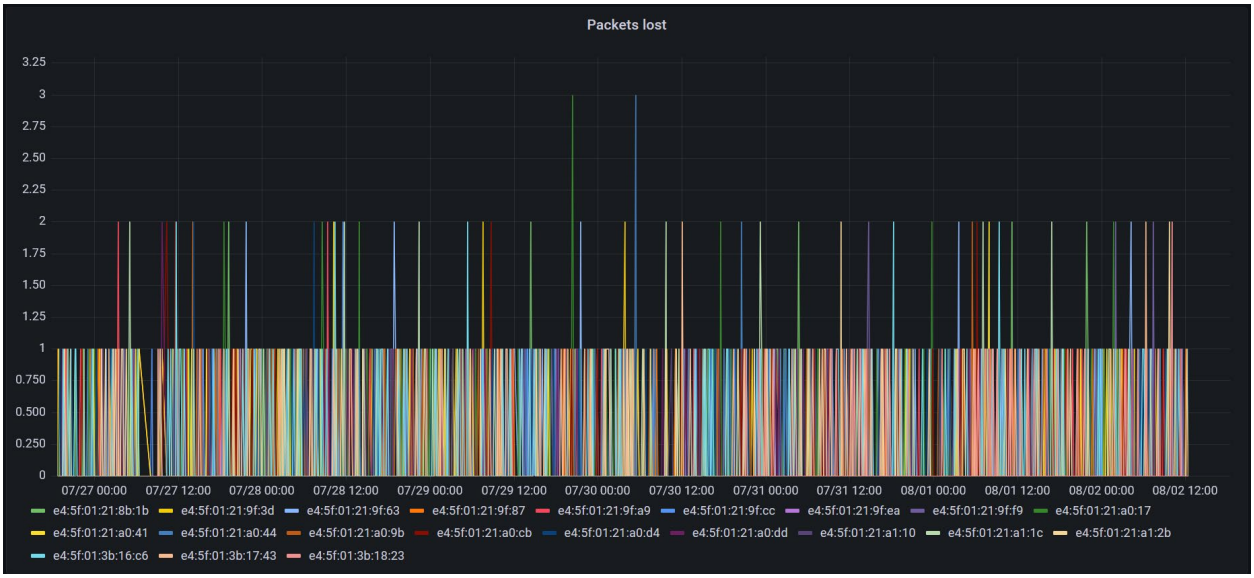


Figure 12 – UDP packet lost during test. Over this test interval, at most 3 packets were lost (out of 1500 sent during the test.)

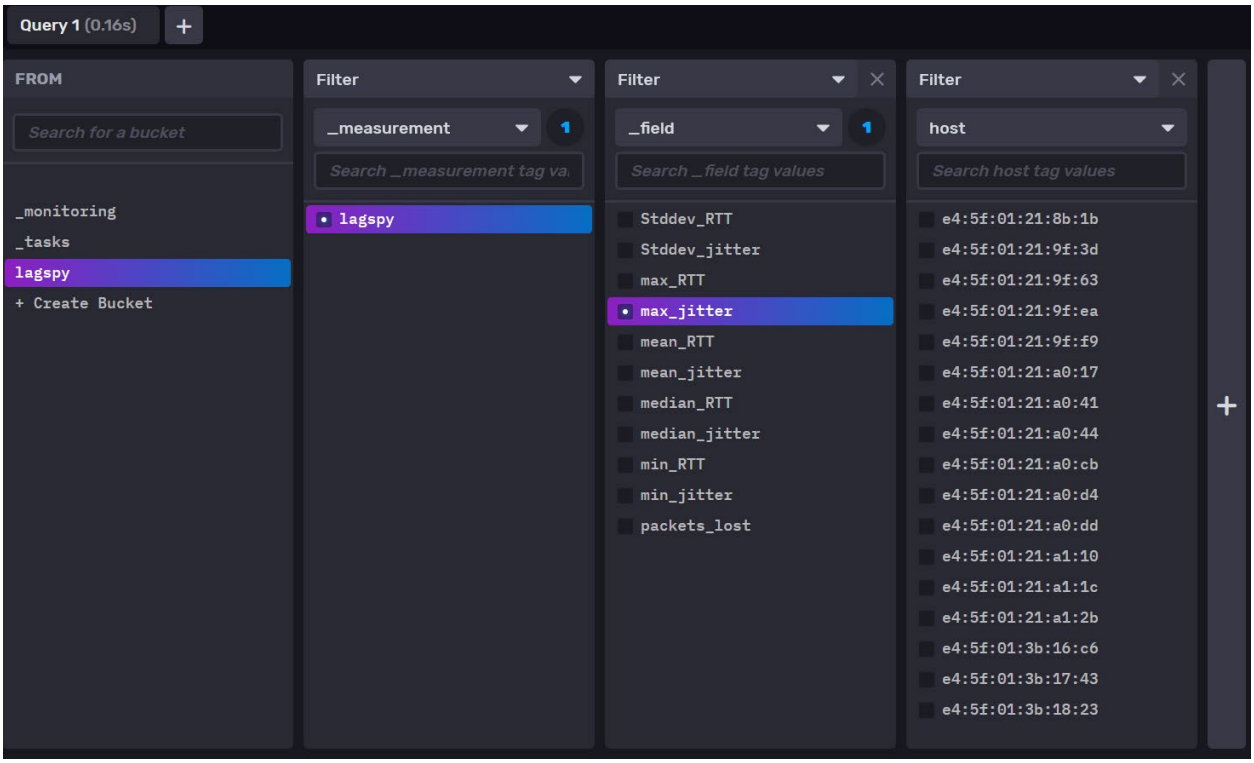


Figure 13 – InfluxDB Fields and Query Generator

3.7. Resource Usage

Application CPU and memory usage is minimal on the Lag-Pi. When idle, the CPU is 99.7% idle and there are 1.372 GB RAM free (on a 2 GB device). During an active IRTT test, 4.3% CPU is consumed by the IRTT process.

Table 3 – Application Resource Usage

Component	% Memory	Virtual Memory
Docker overhead (containerd, dockerd)	6.2%	3.4 MB
Python Test Client	0.6%	52 KB
IRTT Client	0.1%	879 KB

LagSpy is a lightweight application, especially on a platform that already has Docker running. This supports the long term goal of contributing LagSpy to a Linux-based gateway stack such as RDK-B or DD-WRT.

Note that Wireguard resource usage is primarily in kernel space and can't be easily measured.

4. Conclusion

This has been a project of discovery for our team, and we have learned the following lessons thus far:

- Approximately 50% of our Lag-Pi are deployed in an IPv4-only environment. Since Cox has enabled IPv6 for years, this was surprising to us. Many home routers still disable IPv6 by default.

- The Wireguard VPN implements an effective NAT keepalive using the `PersistentKeepalive` keyword. However this keepalive is disabled by default.
- There is a lot of variation in home network configuration, especially among our enthusiastic engineering volunteers. The sooner this functionality can be integrated into the home gateway, the better.
- The ability to remotely upgrade the Lag-Pi software has been critical even at this very early phase.
- Never underestimate the number of hurdles to jump through (e.g. security reviews) when setting up a server with a public IP address, no matter how trivial the service.
- There are still routers out there with outbound firewall policies.
- There is a software compatibility issue with NOOBS 3.5 and the Raspberry Pi 4B impacting about 10% of devices. Avoid this issue by installing Raspberry Pi OS directly onto the SD card.

Abbreviations

CCAP	Converged Cable Access Platform
CM	Cable modem
CMTS	Cable modem termination system
CSV	Comma Separate Value
DD-WRT	DresDeren-Wireless Router
DOCSIS	Data-Over-Cable System Interface Specification
DoS	Denial of service
DUT	Device under test
GB	Gigabyte
GPS	Global Positioning System
HMAC	Hash-based Message Authentication Code
IRTT	Isochronous Round-Trip Tester
IP	Internet Protocol
IPv6	Internet Protocol Version 6
IRTT	Isochronous Round-Trip Tester (open source application)
KB	Kilobyte
LAN	Local area network
LLD	Low Latency DOCSIS
Mbps	Megabit per second
MB	Megabyte
MQTT	Message Queuing Telemetry Transport
NOOBS	New out of the box software
NTP	Network Time Protocol
OS	Operating System
RAM	Random Access Memory
RDK	Reference Design Kit
RTT	Round Trip Time
SD card	Secure Digital card
SLR	Services Layer Router
SoC	System on a chip

UDP	User datagram protocol
USB	Universal serial bus
VM	Virtual Machine
VPN	Virtual Private Network
WAN	Wide area network

Bibliography & References

IRTT (Isochronous Round-Trip Tester) README.md, GitHub, https://github.com/heistp/irtt

Expanded Testing of Video Conferencing Bandwidth Usage Over 50/5 Mbps Broadband Service, CableLabs Inform[ed] blog, February 19, 2021, https://www.cablelabs.com/expanded-testing-of-video-conferencing-bandwidth-usage-over-50-5-mbps-broadband-service

Transitioning Advertising to IP Video

Technical Strategies for Migrating from QAM to ABR Video Advertising

A Technical Paper prepared for SCTE by

Jim Owens
Sr. Director, Video Advertising Solutions Product Management
CommScope
900 Chelmsford Street, Lowell, MA 08151
+1 (978) 614-3389
jim.owens@commscope.com

1. Introduction

Ever since the first cable networks launched in the 1970s, a major trend driving the programming of pay tv providers has been targeting. As more and more channels have launched to target specific interests, viewers have self-segmented. This, combined with the ability of cable networks to segment their networks into ad zones has enabled advertisers to target their ad dollars with increasing precision.

Today, the rise of Internet Protocol television (IPTV) and over-the-top (OTT) video has allowed content to be delivered to a growing array of connected devices. The transition to IPTV and the subsequent development of IP advertising technologies now enables service providers to take their advertising precision (and value) to a new level.

Many multiple system operators (MSOs) are betting on a transition from traditional quadrature amplitude modulation (QAM)-based cable TV to IP video based on adaptive bitrate (ABR) streaming using Internet-based approaches. To do so successfully, cable providers need a strategy that will enable them to gradually transition their legacy set-top box (STB) advertising platforms to the newer adaptive bitrate (ABR) dynamic ad insertion (DAI) systems without disrupting their existing, successful advertising businesses.

This paper provides an overview of how MSOs can prepare for the switch to IP video advertising while continuing to maximize the return on their legacy set-top box platforms.

2. Growth of IP Video

The use of IP and ABR video is increasingly prevalent. The IPTV market was valued at USD 72.24 billion in 2020 and is expected to reach USD 194.21 billion by 2026, at a compound annual growth rate (CAGR) of 17.89% over the forecast period 2021 - 2026.

Parks Research shows that 76 percent of consumers have an OTT service and 22 percent of broadband households have four or more services. As a result, some of the country's largest multichannel video programming distributors (MVPDs) have launched their own IP-based video services. In November 2016, AT&T introduced its DirectTV Now service (now AT&T TV), and Comcast followed suit in September 2017 with Xfinity Streaming service for existing Xfinity customers. In February 2019, Spectrum debuted its TV Essentials service.

One of the benefits of IP video is the ability to target addressable ads. Even today, the addressable ad market is well-established. According to USIM, currently, there are about 54 million MVPD (cable, telco or satellite) households that are linear addressable and 35 million that are ad supported video-on-demand addressable households.

However, the ability to target using traditional video delivery techniques to STBs is limited. Transitioning to ABR-enabled dynamic insertion offers more precise targeting and allows operators to charge more for the ads they deliver. Ultimately, the newer technology will provide benefits that the traditional set-top box can't offer. But as MVPDs build up the number of streaming subscribers to fully leverage this new technology, they need a strategy that allows for a graceful and gradual transition.

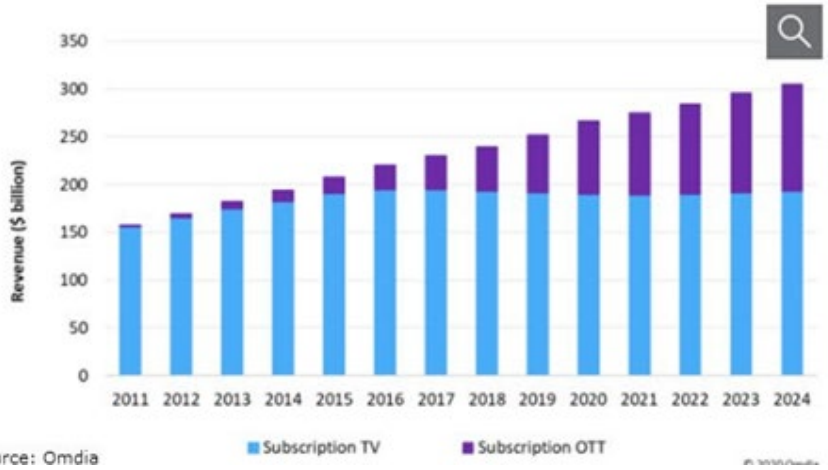


Figure 1 – Subscription TV and OTT Revenue Trajectory, 2011-2024

3. Legacy QAM STB Advertising vs. IP/ABR Advertising

While there are solutions for enabling limited addressable advertising capabilities in traditional STBs, this paper will focus primarily on network-based ad insertion techniques and their evolution as the industry migrates to IP video.

Today the majority of MVPD advertising dollars and subscription revenues come from the traditional set-top box platforms. Using MPEG-2 transport streams, ad servers and splicers, the platform inserts ads in pre-defined ad zones. The ad schedules, which are negotiated and bought, are locked 24 hours in advance. While the legacy STB platform still commands the bulk of the revenue for service providers, it is far from perfect.

For starters, ad zones are defined by geography, which is based on the service provider's access network topology and covers a relatively large area. Negotiating and selling the available ad space is also labor-intensive, involving advertising sales reps working with potential advertisers to fill all the available time slots. Because ad commitments must be locked in at least 24 hours in advance, service providers miss out on potential last-minute ad buys. Finally, the STB solution relies on a number of components that are hardware-based.

Conversely, IP-enabled advertising is a software-driven technology that offers more granularity and flexibility. It is enabled by adaptive bitrate technology that allows service providers to deliver multiple video quality levels within a video stream. Data within each stream is separated into fragments, each containing blocks of video encoded at different bit rates. The varying bit rates enable the media player to choose the appropriate bit rate based on the connection's performance level. The player relies on a manifest, which is essentially a list of available segments. At the simplest level, IP advertising is enabled by customizing this manifest for each stream and either adding or replacing existing segments to introduce ads that can be targeted down to individuals and sessions.

Beyond the technical performance advantages of ABR-enabled IP advertising, there are several significant business advantages. The IP-based technology works across any video device that connects to the internet. Therefore, service providers and their advertisers can reach viewers on whatever platform or

device they may be using. Because the technology is fully dynamic, ad decisions can be made in real-time, enabling advertisers to maximize their ad buy and helping service providers capture more last-minute sales. Ad sales can also more easily be automated using targeting criteria and real-time auctions.

The ability to target households, or even individual subscribers, with surgical precision is perhaps the most significant advantage of IP advertising. Addressability can be enabled via individual app and device-level data. Common targeting parameters include device, demographics, time of day/day of the week, content and category, location, etc. Advanced targeting can include content and purchase behavior, household income, product interest, and any other information available through the use of first and third-party data.

The advantage of this investment is obvious: providers can charge much more for addressable ads on IP—studies estimate up to 4-5 times more than for a regular ad. They can target down to the individual user and optimize ad delivery in real-time. This means higher ROI and the ability to capture new, hyper-targeted ad revenue streams.

The business case for making the transition to IP advertising is strong, but the market is still in the early stages. While IP offers better targeting and more dynamic sales models, the majority MVPD subscribers and ad revenue is associated with traditional STB advertising. Operators eventually need a unified solution that can address both markets.

The more immediate question for providers is how to make the switch in steps as they prepare their networks and organizations for the future while getting as much value as possible from their legacy STB platform. IP ad delivery utilizes entirely different technology, sales processes and ad decisioning. Implementing and adapting to these changes will take time, but there are steps that providers can take now and in the near future to prepare.

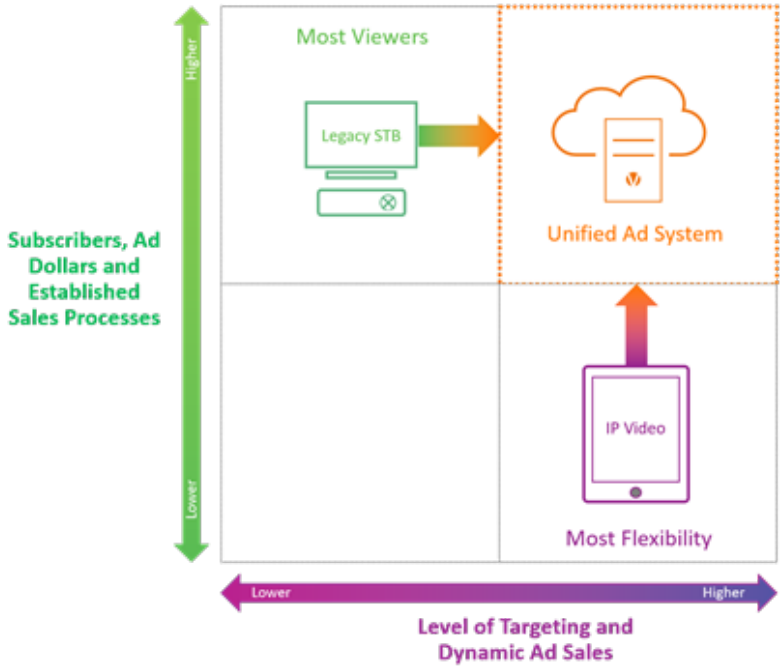


Figure 2 – Creating a Solution that Combines the Strengths of Traditional QAM and IP Advertising

4. Technology Overview

4.1. QAM Advertising

The QAM ad insertion stack consists of the ad insertion server and the splicer. Designed to perform frame-accurate operations on MPEG Transport Streams, the splicer can take an ad stream from the ad insertion server and replace the original bits in a live program stream. The ad insertion server is responsible for preparing and streaming the ads to the splicer, based on schedules published for each channel. The ad breaks are precisely signaled in band via the SCTE 35 standard. The splicer and ad insertion server use the SCTE-30 protocol to coordinate the timing of their activities.

Ad schedules are traditionally generated by a Traffic and Billing system, the function that resides in the ad sales operation and optimizes the placement of advertiser spots within the constraints of the ad buy.

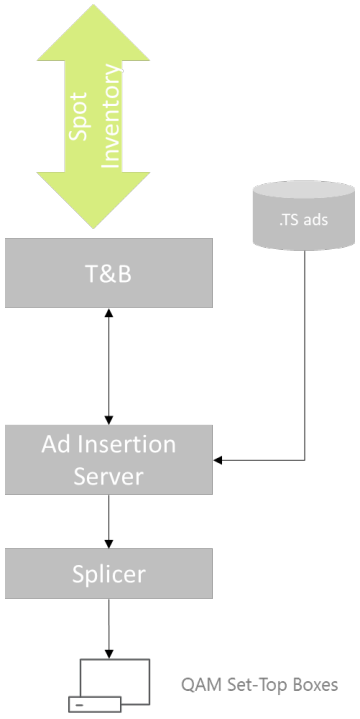


Figure 3 – QAM Advertising Stack

4.2. IP Advertising

There are two main approaches to inserting ads into ABR video, Client-Side Ad Insertion and Server-Side Ad Insertion. Both rely on the content being prepared so that segments align with ad insertion points and the appropriate markers identify the location of the ads. Client-Side Ad Insertions works by having the

client call to an ad server to obtain the correct ad. IP ad insertion is increasingly driven by Server-Side Ad Insertion, which will be the focus of this paper.

In IP video using ABR, the video segments are represented in a manifest, which lists the names of all of the segments that represent a live or on demand video stream. Each segment is also represented in multiple bitrates that can be selected based on the client's assessment of network performance and other factors.

In Server-Side Ad Insertion when an ad event is present, the Manifest Manipulator, using VAST or SCTE-130 standards, passes metadata from the client to an Ad Decision Service (ADS) and requests a decision on which ad to insert. The Ad Decision Service (sometimes just referred to as the Ad Server) identifies campaigns that are looking to target that particular client metadata and responds with one more targeted ads. The Manifest Manipulator then customizes the manifest with pointers to the new ad content segments and returns it to the client. Using this new manifest, the client then retrieves a personalized mix of content and advertisements (or other content).

Note that there can be multiple ADSs involved for a single piece of content that can be called depending on the ownership of the ad inventory. While the mechanics of that flow are outside of the scope of this paper, the Manifest Manipulator must be able to support multiple ADSs.

A separate process may be triggered to prepare the ads for delivery, particularly if they are retrieved from third parties dynamically. This may involve transcoding an ad from a mezzanine format into acceptable codec and bitrates or applying standard content profiles to an existing third-party ad.

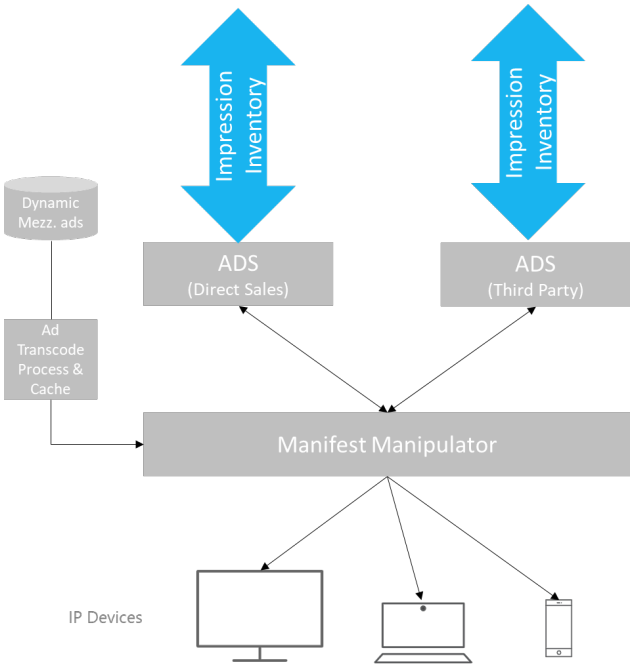


Figure 4 – IP Advertising Stack

4.3. Challenges of the Current Siloed Model

If treated independently, implementation of QAM and IP advertising models results in a siloed architecture, with lack of communication between the two and significant replication of resources. For example:

4.3.1. Separate Ad Decision Infrastructure

The separation of the ADSs for IP ad campaigns and a distinct Traffic and Billing system for scheduled ads limits the ability to generate revenue from the two systems:

- Difficulty in executing campaigns that address both IP devices and QAM STBs
- No opportunity to supplement local scheduled ads with dynamic impression-based ads
- Inability to sell unsold scheduled inventory with dynamic ads

4.3.2. Separate Ad Insertion Systems

The requirement to run and maintain two separate ad insertion systems duplicates operational and capex costs. Both a Server-Side Ad Insertion system based on Manifest Manipulation and a traditional QAM Ad Server and Splicer system need to be procured and maintained.

4.3.1. Parallel Ad and Entertainment Content Preparation

Just as the infrastructure to deliver the ads is replicated, duplicate systems exist to retrieve, process, and deliver entertainment content and ads. The cost of this delivery pipeline grows with the amount of content, the number of ads being targeted and the architecture of the network.

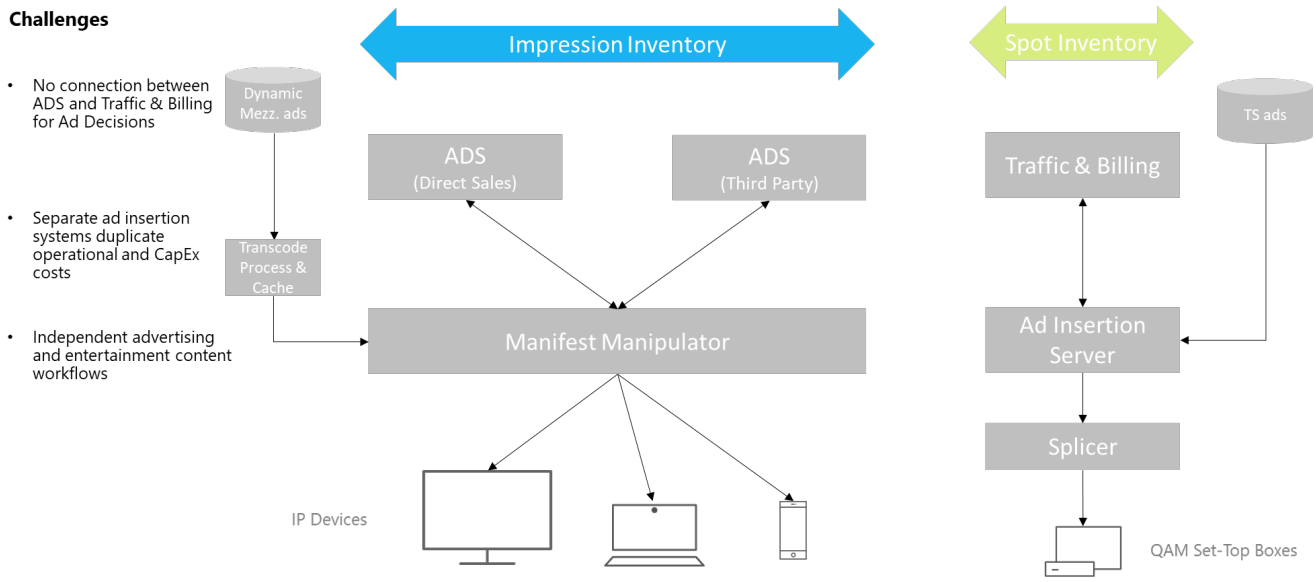


Figure 5 – Parralel IP and QAM Advertising Stacks

5. Business Goals for the Transition to IP

As operators make the transition to IP video and IP advertising, they must consider that traditional QAM video will continue to be delivered for some time and ensure that their existing advertising business is still efficient, profitable, and relevant. At the same time, they must invest in new solutions that leverage the capabilities of IP video, knowing that the number of viewers might remain relatively small for a period of time. This requires trade-offs and a concerted strategy that optimizes the existing business while positioning for the future.

Some of the business goals operators must consider when making this transition include:

5.1. Defend current QAM Advertising Business

As mentioned above, a key goal must be to defend the current advertising business while making the transition to IP video. This includes:

- Maintain existing business processes and ad revenue
    - Ideally the transition should allow for a period of time where existing ad sales, operations and technology can continue to run effectively on the existing QAM network without major changes.
- Extend zone-based ad sales to IP devices (IP Parity)
    - A common first step is to extend the current schedule-based ad sales to IP devices, often referred to as “IP Parity”. This means that ads purchased targeting QAM STBs by Ad Zone will appear on IP devices in that Ad Zone on the same services. This approach allows operators to establish an advertising system for IP devices without interrupting the existing business model. In this way, for a time the same ad sales process can address customers as they move to IP STBs and other IP devices.

5.2. Maximize Revenue

Once an IP advertising process has been established, operators can begin to not only take advantage of the new IP video capabilities but also optimize their legacy QAM advertising solutions to increase revenue. As will be outlined in more detail below, this involves leveraging the strengths of both platforms, beginning to unify some of technologies and enhancing the QAM advertising system to add some of the flexibility and targeting capabilities that are associated with IP advertising. These goals include:

- Monetize unsold ad inventory
    - Removing limitations on QAM ads sales that impose a 24-hour window. Operators can get an immediate revenue boost by making ads targeting QAM STBs more dynamic. Previously unsold inventory can be sold at the last minute, possibly using automated, programmatic techniques
- Enable a mix of scheduled and dynamic ads
    - At the same time, continue to support current ad sales and business processes, so that the system supports a mix of traditional schedule-based ads and dynamic ad placements
- Extend advertising to more channels
    - In many cases there are channels that are properly conditioned for ad insertion but due to low viewership or inadequate ratings data ads are not inserted by operators. The system

should be able to measure the viewing audiences on these channels and, if necessary, aggregate these viewers to make ad insertion on these channels profitable.

- Improve QAM advertising targeting capabilities
    - Leverage changes in the hybrid fiber coax (HFC) network architecture to enable finer-grained targeting of traditional QAM ads at the political precinct or neighborhood level.
- Enable campaigns across QAM and IP footprints
    - While transitioning from QAM to IP video, the solution should allow for operators to execute campaigns that address all device types (though not necessarily with the same targeting precision). This opens up opportunities for new types of campaigns and provides a model that is difficult for other players to offer.

5.3. Improve Operational Efficiency

Addressing a new IP video service that has different technical requirements and capabilities will require new technology and processes. Key to the success of this new advertising initiative will be the ability to optimize the different operational models to find efficiencies. Some of the critical areas include:

- Unify Ad Decisions
    - Eventually the solution should leverage a single "source of truth" for the subscriber data, targeting criteria and the campaigns that are available for execution. This system will incorporate both traditional schedules and dynamic campaigns and address all device types.
- Unify Entertainment/Ad Preparation
    - Great efficiencies can be found in standardizing on one video preparation workflow that can deliver entertainment content with advertising to QAM STBs and IP devices. Standardizing on ABR video formats as the source for both entertainment content and ads can reduce the costs of running two parallel video processing and preparation systems and prepare the operator for the all-IP video future.
- Unify Ad Insertion Systems
    - Similarly, relying on a single system for ad insertion/ad execution represents opportunities for opex and capex reductions versus purchasing and maintaining separate products.
- Transition to Virtualized Solutions
    - All of these migrations can also accelerate the move from hardware-based products to virtualized, software-based solutions that can reduce opex, capex and offer other advantages such as elastic scalability.

5.4. Enable New Business Models

Though not directly tied to the transition to IP video, operators are also looking to capitalize on new business models that leverage their unique knowledge of the end customer and position as owners of the delivery network.

- Insert ads on behalf of Content Partners
    - Cable operators are increasingly applying their ad insertion capabilities beyond the traditional inventory allocations on cable channels. Operators can insert targeted ads on national ad breaks on behalf of programmers and broadcasters on both traditional QAM STBs and IP devices inside and outside of the home. This opens up a new and very large part of the television advertising market.
- Enable Programmatic Sales Through Third Parties

- Operators can supplement their in-house ad sales with third party advertising networks to programmatically sell ad inventory that may previously have gone unsold.

6. Transition Strategy

6.1. Siloed Model

The siloed model represents serious challenges to achieving the business goals set out previously. However, there are some concrete steps operator can take to move their IP video advertising strategy forward.

6.1.1. IP Parity Through Processing of Existing Ad Zones Outputs

A common way to initiate IP advertising through IP parity in this model is to take the existing spliced channels and transcode and package them as separate IP video streams. In this model each ad zone becomes a separate version of the IP channel and is offered only to IP clients in that ad zone. As IP clients service by a CDN different techniques are employed to determine the location of individual clients and map them to existing ad zones. Support for out of home streaming may require additional sophistication. However, the goal of replicating Ad Zones on IP devices can be achieved without major changes to the existing systems.

Pros:

- Defend current QAM advertising business by extending scheduled ads to new IP subscribers
- Requires few changes to existing systems

Cons:

- Maintains dual processes for both QAM and IP video delivery
- Does not take advantage of the dynamic, targeted nature of IP clients
- Requires replicating video processing that may not be practical with large numbers of ad zones

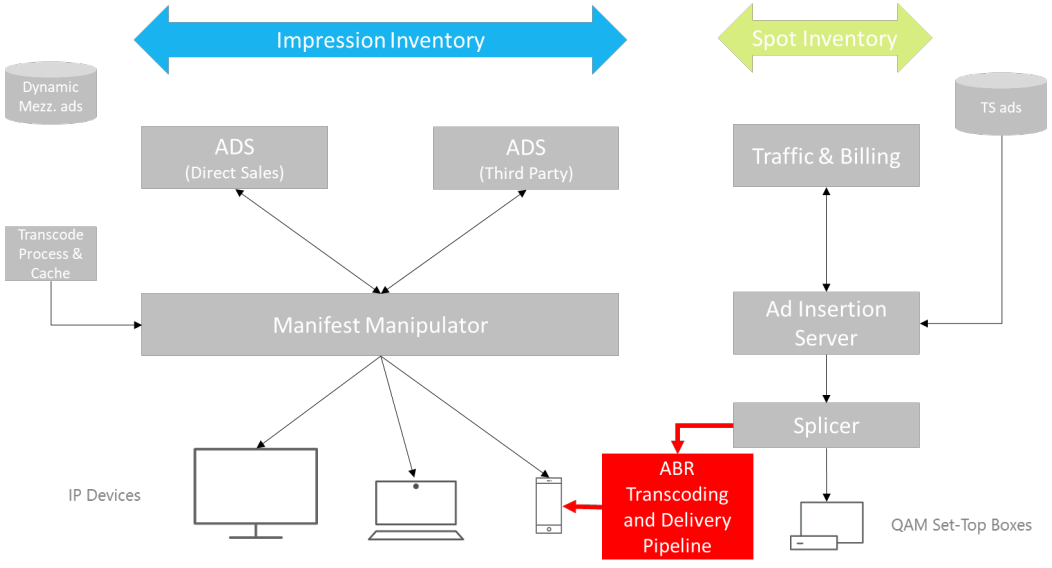


Figure 6 – Achieving IP Parity in a Siloed Model

6.2. Hybrid Model

The next step in the transition is a hybrid model, where both the IP and QAM advertising systems exist side-by-side but start to exhibit coordination and see reduction of redundant processes and components.

Some key steps in this stage are outlined below:

6.2.1. IP Parity Through Integrating Schedules into the ADS

A more sophisticated and efficient model of achieving IP Parity is by ingesting schedules into an ADS (or a component that can act as an ADS for schedule-based ads). The ADS processes the schedules and responds to queries from the Manifest Manipulator with the appropriate ad or ads. This approach has several advantages over the siloes model. First, it requires processing of only one version of each channel into the ABR format regardless of the number of ad zones. Second, it leverages a Manifest Manipulator for the ad insertion, establishing the infrastructure for future IP addressable ad insertion using standards-based protocols and processes. Third, it allows for the use of ABR content for IP ads, again leveraging a standard IP advertising process.

6.2.2. Add Dynamic ADS Integration to Existing Ad Insertion Servers

Just as a process can be established that delivers scheduled-based QAM ads to IP devices, the reverse can be enabled to allow the delivery of dynamic, impression-based ads to legacy ad insertion servers. By adding support for protocols such as VAST or SCTE 130 to the existing Ad Insertion Servers, the Ad Servers can identify ad avails that have not been sold and reach out to an ADS for a real-time (or near real-time) decision to fill the available spot. The ADS responds with ads that are part of campaigns targeting this inventory.

In conjunction with making the ad decision process more dynamic, the ability to collect real-time data from two-way STBs improves visibility and precision in ad measurement, which means that impression-based campaigns can be more dynamically interwoven with scheduled spots to better leverage inventory and increase ad revenue. It also expands the number of channels on which ads can be sold, due to precise viewership information for long-tail channels. This is accomplished by integrating with other elements of the video network, such as SDV servers that already aggregate channel tuning information.

This begins to accomplish the business goals of:

- Allowing unsold inventory on QAM STBs to be sold dynamically near the time of the ad placement
- Support a mix of schedule-based and dynamic impression-based ads on QAM STBs
- Begin to unify the IP and QAM decision systems
- More easily allow for the execution of campaigns across all device types
- Expand the number of channels that can support advertising

6.2.3. Add Support for IP Content to Existing QAM Ad Servers

Independently, or in conjunction with the integration of the Ad Insertion Server to the ADS, a further enhancement adds the ability of the QAM Ad Insertion Server to ingest ads in ABR format and stream them as MPEG-2 Transport Streams to a splicer. To do so, it may leverage the ad preparation process from the IP system, including transcoding and re-packaging.

This capability begins to achieve the business goal of unifying the ad preparation process and increases the universe of ad inventory available to the QAM ad insertion system.

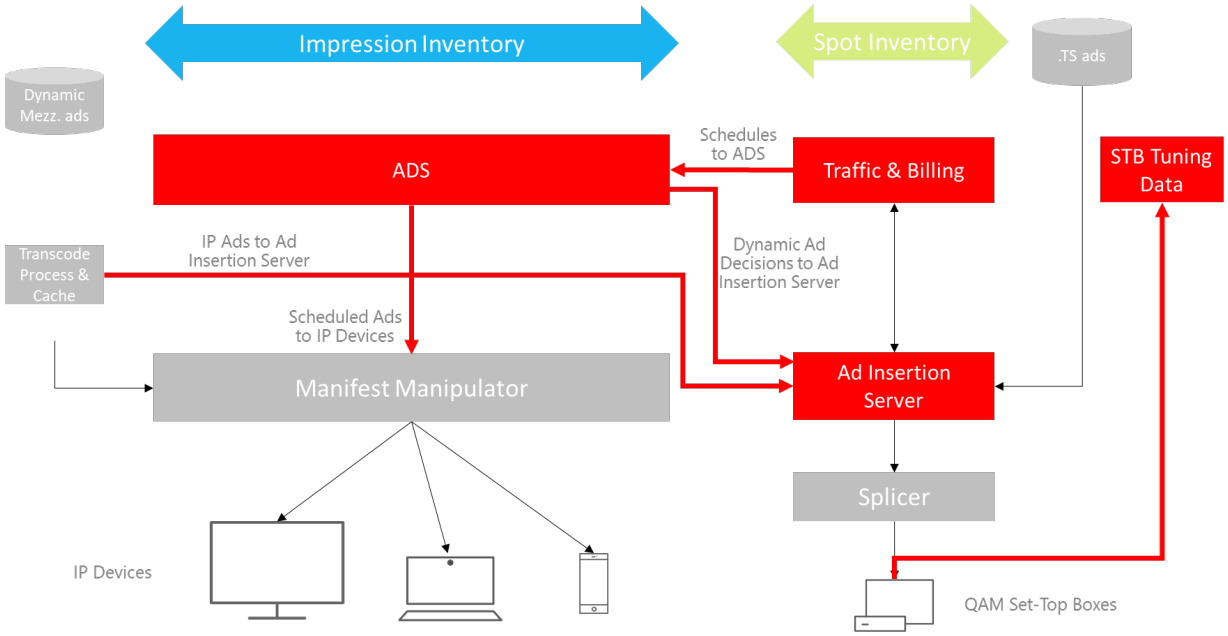


Figure 7 – Summary of Enhancements Possible in the Hybrid Model

6.3. Unified Model

The final model on the path to a full IP video delivery system is a unified model in which the advertising ecosystem has been consolidated into shared systems that support both IP and QAM devices yet fully support the capabilities of IP video advertising.

The characteristics of this model include:

6.3.1. Unified Ad Decisions Through the ADS

The ADS is the central point for all ad decisions, both IP and QAM. Dynamic ad decisions are used for ads delivered to both IP devices and QAM STBs. Scheduled ad sales may still be supported, but the system of record for these ad buys has shifted to the ADS. Note that multiple ADSs are still supported, as advertising inventory may be split between owners, and operators may call out to third party ADSs for decisions on ads sold by other parties in the ecosystem.

6.3.1. Unified Ad Insertion System

The unified model leverages the Manifest Manipulator for ALL ad insertions, using ABR video as the mezzanine format for all ad execution. Based on decisions from the ADS, the Manifest Manipulator generates both personalized manifests that establish a unicast stream to a specific device, but also "channel" manifests that can generate an ABR stream targeted by traditional Ad Zone or any other criteria desired. Similar to the way IP ad content support was added to the Ad Insertion Server in the hybrid

model, now the segments that comprise these ABR channels are converted to Transport Streams for delivery over the HFC network to QAM STBs. The result is a common ad execution platform based on virtualized software products, eliminating the need for traditional ad insertion servers and splicers.

6.3.1. Unified Ad and Content Preparation

Because all ad insertion is now done using ABR video as the mezzanine format, content preparation can be consolidated to a single workflow that uniformly processes and delivers video to QAM and IP devices. This step sets the business on a sustainable path of growth by further improving operational efficiency and adding flexibility to continue building the IP video network.

6.3.2. Enhanced Targeting of QAM Ad Insertion

Current Ad Zones offer powerful local targeting capabilities but are relatively broad compared to power of addressable IP advertising and are “hard wired”. Changes to the HFC architecture, such as switched digital video (SDV) (including the option of switching all QAM channels) and Distributed Access Architectures (DAA) represent opportunities to develop much more granular targeting for legacy STBs. For example, unique multicasts could be targeted to individual SDV service groups or DAA nodes, greatly enhancing the targeting capability of network-based ad insertion. The composition of the Ad Zone can also be flexible, dynamically assembled from the component service groups or nodes using software configurations.

The unification of content preparation and the efficiency with which channels variants with unique ads can be generated by the manifest manipulator make it relatively easy to generate the large number of targeted channels needed to take advantage of the greater number of smaller Ad Zones.

The unified model addresses the business goals of improving operational efficiency through:

- Unifying ad decisions
- Unifying ad insertion systems
- Unifying ad and entertainment content preparation
- Migrating all systems to virtualized solutions
- Taking advantage of enhanced targeting capabilities enabled by changes to the HFC architecture

Finally, the unified model establishes an advertising system that supports both IP and QAM devices, but is ready to transition to a full IP-only world without requiring any additional changes.

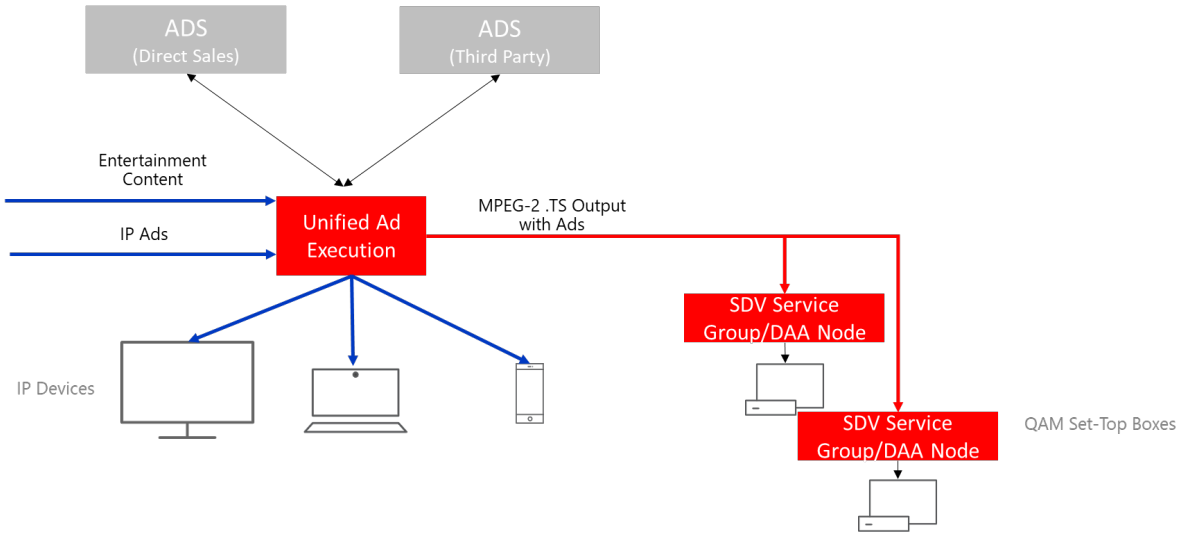


Figure 8 – Summary of Enhancements Possible in the Unified Model

7. Conclusion

As the chart below shows, different models address different business goals. The decision of how fast to proceed must be based on the return on investment, which will be unique for each operator based on factors such as the size of their subscriber base, sophistication of data collection and analytics, the maturity of their existing advertising business and the speed with which they intend to migrate to IP video.

	Transition Strategy						
	Silo	Hybrid			Unfied		All IP
	IP Parity - Xcode Ad Inserted Channels	IP Parity - ADS/Schedule Integration	Add Dynamic ADS Support to Legacy Ad Server	Add Support for IP Ads to Legacy Ad Server	Unified Ad Insertion Platform	Target QAM by SDV Service Groups/DAA Nodes	IP Advertising Platform
Business Goals							
Defend Value of Current Business					X		n/a
Extened QAM Advertising to IP Devices	X	X					
Maximize Revenue							
Improve Legacy Targeting Capabilities						X	n/a
Extend Advertising to More Channels			X		X		X
Leverage Unsold Advertisign			X		X		X
Enable Mix of Scheduled and Dynamic Ads							X
Enable Campaigns Across Both Footprints			X		X		n/a
Operational Efficiency							
Unify Decisions			X		X		X
Unify Content Prep				X	X		X
Unify Ad Insertion					X		X
Transition to Virtualized Solutions					X		X
Enable New Business Models							
Insert Ads on Behalf of Content Providers			X		X		X
Enable Programmatic Advertisng via Third Paries			X		X		X

Table 1 – Comparison of Business Stages in Addressing Business Goals

Abbreviations

ABR	adaptive bitrate
ADS	Ad Decision Service
CAGR	compound annual growth rate
DAA	Distributed Access Architecture
DAI	dynamic ad insertion
IPTV	Internet Protocol television
MSO	multiple system operator
MVPD	multichannel video programming distributor
OTT	over-the-top
QAM	quadrature amplitude modulation
SCTE	Society of Cable Telecommunications Engineers
SDV	switched digital video
STB	set-top box
VAST	Video Ad Serving Template

Bibliography & References

SCTE Advertising Specs; https://scte.org/standards

VAST specs; https://www.iab.com/guidelines/vast/

Turning On A Dime: The New Landscape of Adult Learning

A Technical Paper prepared for SCTE by

Shiloh McCoy
Supervisor, Technical & Safety Training
Charter Communications
4930 Energy Way, Reno NV
775-823-7730
Shiloh.mccoy@charter.com

Abbie O'Dell
Senior Director, Learning Services: Field Operations
Charter Communications
6399 S Fiddlers Green Cir, Greenwood Village CO
720-482-4205
Abbie.odell@charter.com

1. Abstract

This quantitative study considers the perceived effectiveness of virtual training events from the learner perspective. Data considered as part of the study include post-training participant evaluations gathered through the learning management system (LMS) of a large telecommunications operator, for both pre- and post-pandemic courses. Courses provided by eleven trainers from one geographic region were included as part of the study, and responses to evaluations were analyzed to identify whether a relationship exists between the trainers who received developmental training on adult learning theories and practices in the virtual classroom and participant perceptions of effectiveness. The findings indicate that participants have a more favorable perception of some elements of virtual training classes, when provided by instructors who have received additional training, and that instructors who have not received training receive lower scores in some categories. These findings indicate that adult learning theories and practices, if applied properly in the virtual classroom, create a learning experience that is just as effective as the traditional in-person classroom.

2. Introduction

Prior to the start of COVID-19 pandemic, the prevailing belief of many educators and business leaders was that in-person instructor-led training was far superior to virtual or online learning. Online learning experiences were considered a substandard or fallback option for workplace learning, and in many K-12 and post-secondary settings were often not considered at all. A report from the US Department of Education (2020) found that in the 2018 school year, 64.7 percent of post-secondary students were not attending any distance education or online courses, and only 16.6 percent were exclusively attending distance education courses. This paradigm shifted almost overnight, when the need for social distancing forced widespread migration to virtualized teaching and instruction for all learners, from K-12 and post-secondary to adult learning experiences in the workplace.

As a result of the shift, learning organizations have been challenged to re-evaluate how we transfer knowledge and specialized skills. Technologies such as Zoom, Microsoft Teams, and Cisco WebEx have replaced the traditional classroom environment in all levels of education. After a full year of distance learning, opinions vary widely on whether remote learning meets the needs of learners, or whether this learning modality should remain relegated to the role of a backup or emergency solution.

Modern theories of adult learning, known as andragogy, provide important insights into this question. Research indicates that effective and positive learning experiences can still occur in a virtualized delivery model. A 2010 meta-analysis of thousands of empirical studies by the US Department of Education found that students in online learning experiences performed slightly better than those in instructor-led events. Further inquiry is needed to gain additional insights specifically related to the pandemic and the experience of adult learners in the workplace, and could help expand the collective understanding of this important and timely topic.

3. Literature Review

The need for alignment to core andragogical theories in the online classroom is not a new concept. Malcolm Knowles, widely considered the founder of modern theories of adult learning, predicted in 1983 that the significant increase in electronic delivery of training would require adult learning professionals to learn how to “use the technology in congruence with principles of adult learning” (Blackwood & White, 1991).

3.1. Adult Learning Theories

The topic of how adults experience learning and transfer knowledge is a well-researched and robust area of study. Early works by Knowles (1975; 1980) outline the foundational principles that most subsequent studies have sought to prove or refine. The common characteristics of adult learners as understood within the profession (Merriam & Bierema, 2014; Abdullah, et al, 2008) are as follows:

- Self-Direction: Adult learners must be actively involved in the process of learning and have agency over the direction of the course, and in some cases, the selection of content.
- Experience: Adult learners bring their own life experience and knowledge to the learning environment, and effective facilitators of adult learning must recognize this and seek to connect the learners' experience to the topic being taught.
- Goal-Orientation: Adult learners have a goal in mind when attending learning experiences, whether that is to broaden their knowledge for a job or profession, or simply for personal development.
- Relevancy-Orientation: Adult learners must recognize the value or reason for the topic being taught, and tend to be more problem centered.

These principles have been further studied, refined and expanded upon since their introduction, but are commonly considered the core of adult learning practices, and apply in learning environments ranging from formal learning experiences such as post-secondary education to informal experiences found in social or community programs. Research on the use of these theories in the workplace learning environment is a fast-growing area of study (Caruso, 2018; Hendriks et al., 2018; Cookson, 2001; Grow, 1991), and the inquiries found in this study offer a meaningful addition to the literature.

3.2. Conversion to the Online Classroom

While the need for social distancing during the height of the pandemic required an immediate change to delivery methods, in many cases the content and learning outcomes were not in alignment with the modality. Davis and Arend (2013) clearly outline the different ways of learning and how a careful alignment between instructional methods and desired learning outcomes is critical to the success of any learning experience. Specifically in the context of online learning, Fein and Logan indicate that instructors must not directly transfer content originally designed for the in-person classroom without making the adjustments to the activities and program to better align with the learners' needs (2003). Further findings from Fein and Logan (2003) indicate that students rated courses more highly where the instructor made appropriate changes to adapt the course to the online environment.

Independent of the content itself, changes are required for the instructor in terms of methods and practices used for delivery. One of the notable challenges related conversion to online instruction is the complexity of the computer based classroom when compared to the traditional classroom. Instructors need to facilitate an environment where learners experience self-direction and construct meaning during the learning process, and must work to create meaningful discussion through improved listening skills and asking more facilitative questions to create dialogue (Davis & Arend, 2013; Fein & Logan, 2003). Realism of the learning environment is also key to success and learners benefit from the content being more project- and activity-based (Fein & Logan, 2003)

3.3. Andragogical Principles in Online Classrooms

Adult learning theory's application in virtual classroom environments is vast in scholarly literature. Arghode et al. (2017) and Deineha et al. (2020) have conducted systematic reviews on andragogical

learning principles that examine student's perceptions of excellent online instructional delivery. Findings suggest that instructor involvement is equally important and relevant when aligned with deep content (Arghode et al., 2017). Further evidence supports the theory that an online instructor can operate as a facilitator within the online student population by running activities and providing immediate feedback; this gives the students feel of control over their understanding and improved guidance throughout the course (Arghode et al., cites Yamagata-Lynch et al., 2015). Adult learning principles, when applied correctly, aid knowledge transfer by satisfying basic student needs. On the other hand, Arghode et al. (2017) and Deineha et al. (2020) also noted conflicts between students' perceptions when andragogical principles are applied for online classrooms in the same way they would be applied for in-person classes; resulting in negative student perceptions. Furthermore, the collaboration between teacher-to-student or student-to-student showed not to be perceived as necessary, stating that students were likely to favor individual work in an online setting (Arghode et al., 2017; Deineha et al., 2020). The literature analysis on adult learning theory in virtual classroom delivery shows a theme that adult learning practices for online environments require specific curating, similar to in-person adult learning. Likewise, a perceived gap suggests that little study has been done specifically for adult learners in a technical field under a large telecommunications operator.

The purpose of this quantitative study is to investigate the inclusion of foundational adult learning theories and practices in the virtual classroom for technicians of a large telecommunications operator, and analyze whether use of these methods in remote learning experience proves just as effective as in-person learning. To do so, we pose the following questions:

- Does the perceived effectiveness of virtual training change based on the use of adult learning principles/methods by the trainer?
- Do trainers who have received additional training themselves on adult learning in the virtual space receive better scores from participant on level 1 evaluations?

4. Research Methods

As detailed in the literature review, sources exist relative to the topics of adult learning theories and practices and the synchronous online environment, but limited research exists that specifically links the training of instructors back to the evaluation scores of learners. To that end, a quantitative method was selected for this study, in order to consider the existing data set related to learner reactions that is already captured in the Learning Management System (LMS).

4.1. Participants

The population considered in the study is comprised of employees within the field operations business unit of a large telecommunications operator, with the representative sample drawn from employees who completed any courses, both ILT and VILT, between July 1st 2019 and July 1st 2021.

The nonprobability sampling method was used to select one geographic region to consider as part of this study. This method was selected largely due to availability of data and because the particular region represents the phenomenon being studied (Creswell, 2003; Creswell & Guetterman, 2019). The region selected has a staff of eleven technical and safety instructors, some of whom received additional training and coaching on core adult learning theory and instructional methods in the virtual classroom, and others who did not.

4.2. Data Collection

The concept of levels of evaluation for learning experiences was first outlined by Kirkpatrick (1959, 1976, 1996) and is still a common method used to quantify impact of training. Kirkpatrick's model is comprised of four levels, typically explained as reaction (level 1), learning (level 2), behavior (level 3), and results (level 4). For purposes of this research, we are considering the level 1 evaluations for the selected group of learners. After completion of training courses, data were collected using a survey link sent via email to participants. Participants are not required to complete level 1 evaluations, and do not receive any follow up reminders if the evaluation is not completed. While the learner identities are captured automatically within the LMS, that information was removed from the data set for purposes of this study to maintain full anonymity of learners.

The evaluations contain a combination of question types, including traditional and modified Likert scale, Yes/No, multiple choice, and open-ended responses. For purposes of this study, the open-ended responses are not considered, due to the notable challenge with qualitative methodology and large numbers of responses which require significant time for accurate and meaningful coding and analysis to identify themes (Creswell, 2003; Creswell & Guetterman, 2019).

Question across the three evaluations ranged from the topic of pace ("The pace of the learning was…" with learners given modified Likert scale options) to topics of applicability ("Select the statement that best reflects your ability to apply the information"). In addition, one evaluation included questions that were specific to the performance of the instructor ("Did the facilitator of the course…Create an environment where I could ask questions and provide input?") answered using a traditional Likert scale.

4.3. Data Analysis

The data were exported into a spreadsheet from the Cornerstone™ LMS. Information that could be used to identify individual respondents (e.g. employee number) was removed from the data set. The trainers were each assigned a number (e.g. Trainer 1) which was used throughout the rest of the analysis. As noted above, the open-ended text responses are not considered within the scope of this study, and were removed from the data set. All Likert scale responses were coded using numeric values (e.g. 5 - Strongly agree, 4 – Agree, 3 - Neither agree nor disagree, 2 – Disagree, 1 - Strongly disagree) and Yes/No responses were coded as Yes=2 and No=1. A modified Likert scale was used for multiple choice questions with four possible answers, where 4 represents the positive response and 1 represents the negative response.

The database also required cleanup activities to remove responses not relevant to this study, such as questions asking for the name of the instructor. In addition, all non-answers/non-responses were removed from the data set.

The data were organized into four discrete groups, representing the time frame of pre- and post-pandemic, and further into groups representing those trainers who received additional training themselves on incorporating adult learning theories and practices into the virtual classroom, and those who did not. The groups were labeled Group 1, Group 2, Group 3 and Group 4 to enable simple comparative analyses of the data. See Table 2 for detail on group designations.

Table 1 – Group Designations

Group	Description	Start Date	End Date
1	Pre-COVID ILT	July 1, 2019	March 1, 2020
2	VILT, Prior to additional training being offered	March 1, 2020	May 1, 2020
3	VILT, Trainers who received additional training	May 1, 2020	October 1, 2020
4	VILT, Trainers who did not receive additional training	May 1, 2020	October 1, 2020

Trainers from the selected sample were assigned to groups based on specific criteria. All trainers from the selected sample were assigned to both groups 1 and 2, since all trainers delivered ILT prior to the pandemic, and all trainers delivered VILT sessions without receiving additional training during the early months of the pandemic. The trainers were then assigned either to group 3 or 4, based on whether they received developmental training in the following months. See Table 3 for detail on group assignments.

Table 2 – Trainer Group Designations

Code	Group 1	Group 2	Group 3	Group 4
Trainer 1	x	x		x
Trainer 2	x	x	x	
Trainer 3	x	x		x
Trainer 4	x	x		x
Trainer 5	x	x		x
Trainer 6	x	x	x	
Trainer 7	x	x		x
Trainer 8	x	x	x	
Trainer 9	x	x	x	
Trainer 10	x	x	x	
Trainer 11	x	x		x

After separating the data into four groups, the Likert scale questions were separated from the other types of scored questions, to enable numeric analysis using the same 1-5 scale for Likert, and the 1-4 and 1-2 scale for multiple choice and yes/no, respectively. Lastly, in the final data analysis all numeric values were converted to a percentage (e.g., for Likert scale the maximum number of points available is 5, so the final score for each category was divided into 5, while the multiple choice maximum score was 4, so the final score for those categories was divided into 4, and so on). This method enabled a more mathematically accurate comparison of participant response patterns.

As noted earlier, participation in level 1 evaluations after course completion is voluntary, so the results represent only those learners who chose to provide feedback. A total of 98 unique participants provided responses for courses attended during the considered time. Gathered responses were separated into the group categories identified in Table 2. In analysis of the four groups, no responses were received for any VILT courses provided during the identified time period for group 2. VILT courses were offered during this time period by the identified instructors, but no participants responded to level 1 evaluations for the VILT courses. As a result, group 2 data was not used as a comparative for the purpose of this study. Data were able to be considered related to group 1, identified as the pre-pandemic traditional ILT classroom. This dataset is used for the control in order to have a meaningful comparative for groups 3 and 4.

During the time period considered, data was collected using three different level 1 evaluations. To combine these data sets, all individual questions from the three level 1evaulations were reviewed and assigned to one of three possible categories, and one of ten possible subcategories. This categorization aligns with the foundational adult learning principles related to the questions, and enables us to consider the unique participant responses independently of the specific evaluation used, and thus identify trends in the responses. The three categories questions were mapped to are learner needs, learning technology, and adult learning theories and practices. Some subcategories may align with multiple higher-level categories, to enable deeper analysis of topics as aligned to adult learning theories. See Table 3 for detailed category and subcategory mapping.

Table 3 – Evaluation Questions and Categories

Categories	Subcategories
Adult Learning Theories & Practices	Instructor Overall Satisfaction
Adult Learning Theories & Practices	Classroom Management
Adult Learning Theories & Practices	Learning Objectives
Adult Learning Theories & Practices	Instructor Communication
Adult Learning Theories & Practices	Instructor Knowledge
Adult Learning Theories & Practices	Learning Pace
Learner Needs	Content Related to Job
Learner Needs	Learning Pace
Learner Needs	Overall Satisfaction/Recommendation
Learner Needs	Instructor Interaction
Learner Needs	Instructor Overall Satisfaction
Learner Needs	Learning Objectives
Learning Technology	Learning Technology

After mapping all questions to these ten subcategories, we identified significant challenges with categories of instructor interaction, instructor knowledge, instructor overall satisfaction, learning pace, and overall satisfaction/recommendation. Earlier level 1 evaluation models containing questions with these topics used a Likert scale, affording participants more range of possible answers; while the later version of the evaluations changed this concept to a yes/no question. The change in question and answer criteria caused significant variances between the two data sets, and we determined that the best and most accurate way to address this was to remove these categories from consideration within the scope of this research to avoid any potential inaccuracies in findings. Additionally, the subcategories of classroom management and instructor communication were not able to be mapped to the newer questions in the later level 1 studies, so these were also removed in order to avoid errors. The three remaining subcategories are considered within this study: content related to job, learning objectives, and learning technology.

5. Results

As noted in the literature review, adult learners tend to be practical and have a need for their learning experience to be relevant to their life or work (Merriam & Bierema, 2014). The subcategory of "content related to job" contains all questions from the three different level 1 evaluations that were determined to align with this concept. See Table 4 for all included questions for this subcategory. Participant mean

responses in group 1, the pre-pandemic traditional ILT group, were 86.53 percent positive, while group 3 and 4 responses were 89.29 percent positive and 87.50 percent positive respectively.

Table 4 – Content Relevance to Job Category

Question Text
I had sufficient opportunities during the training to apply and practice the concepts and skills presented.
This session has increased my ability to perform my current job.
Training activities reflected real world, on-the-job situations.
I plan to use this information in my current job.
The material covered is relevant to my job.
I would recommend this training program to a colleague in a similar position.
Select the statement that best describes examples and activities in the course.

The second subcategory analyzed was learning objectives, aligning with the adult learner's need for clarity on the goals and intended objectives of their learning experience (Merriam & Bierema, 2014). The questions included in this category address both the participant perception of their understanding of the objectives (e.g., "I had sufficient opportunities during the training to apply and practice the concepts and skills presented") and the participant perception of the facilitator's ability to clearly explain the objectives (e.g. "The facilitator clearly explained the program objectives"). See Table 5 for all included questions in this subcategory.

Table 5 – Learning Objectives Category

Question Text
The learning objectives were clear.
The facilitator clearly explained the program objectives.
I had sufficient opportunities during the training to apply and practice the concepts and skills presented.
Select the statement that best reflects your ability to apply information.
The facilitator clearly explained the program objectives.

For group 1, participant mean response was 87.20 percent positive, compared to 78.57 percent positive and 75.00 percent positive for groups 3 and 4 respectively.

The final subcategory of questions considered within the scope of the research was learning technology, which included questions designed to gather participant perceptions of whether the use of technology supported or detracted from their learning experience. See Table 6 for included questions in this subcategory.

Table 6 – Learning Technology Category

Question Text
Did the learning environment and/or technology support your learning?
Select the statement that best describes the use of technology in this course (if applicable).

The mean participant response for group 1 in the learning technology subcategory was 84.78 percent positive, compared to 82.14 percent and 98.75 percent for groups 3 and 4 respectively.

5.1. Discussion

While we anticipated that the mean evaluation scores for ILT (group 1) compared to VILT (groups 3 and 4) courses would show a higher positive response for ILT; that was only found to be the case for the subcategory of learning objectives, with the ILT group having a mean score of 87.20 percent compared to the much lower scores in groups 3 and 4 (78.57 and 75.00). This may indicate that instructors are more accustomed to clearly identifying the desired outcomes or objectives of the course in the traditional classroom compared to the VILT environment. Another interpretation could be that the participants scored this category higher for ILT because they were better able to understand and identify the objectives of the course, related to greater opportunities for hands-on practice of skills. Group 4, the trainers who received less developmental training on adult learning theories in the VILT environment, were lower than their peers who received the additional training, implying that they were less proficient at recognizing and including the importance of reinforcing learning objectives within this type of environment.

The category of content relevance to job was marginally higher for group 3, the group of trainers delivering VILT courses who received additional training on adult learning theories and practices. This finding implies that even in an online environment where learners are afforded limited opportunities for practice of the skills learned in the course, an instructor who uses techniques to engage the learners and construct meaning is able to foster an environment where learners are able to identify how they can use the information back on the job. An interesting pattern that also emerged in this category was that those trainers receiving this additional training scored higher in this category for VILT courses, even when compared to traditional ILT courses. Additional developmental training for instructors on adult learning theories and practices had an overall positive effect on learner perceptions, regardless of the type of classroom environment.

The final category, learning technology, provided surprising and important insight into learners' perceptions of the role of technology in the ILT environment compared to VILT. The ILT course perceptions related to technology use represent participant impressions do not relate to the use of an online/virtual teaching medium (e.g. Webex), but rather are connected to the use of technology labs or other materials found in the ILT setting. As such, this data point is less effective when used in comparison to the learner perceptions recorded for groups 3 and 4, in the exclusively VILT settings. What is unexpected about this data is the significantly higher (>10 percent) favorable scoring for group 4, the trainers who did not receive additional coaching on adult learning practices prior to delivering these courses. The trainers in group 3 who received additional developmental training on adult learning in the VILT environment also received coaching on use of tools within the online training platform, such as breakout rooms, polling, and other features. It is likely that these trainers subsequently utilized these tools during their sessions; and the participants' lower positive score may indicate that the use of these tools distracted them from the learning. The findings may also indicate that group 4 instructors relied more heavily on their instructional techniques rather than the online tools, again creating less distraction for the learner.

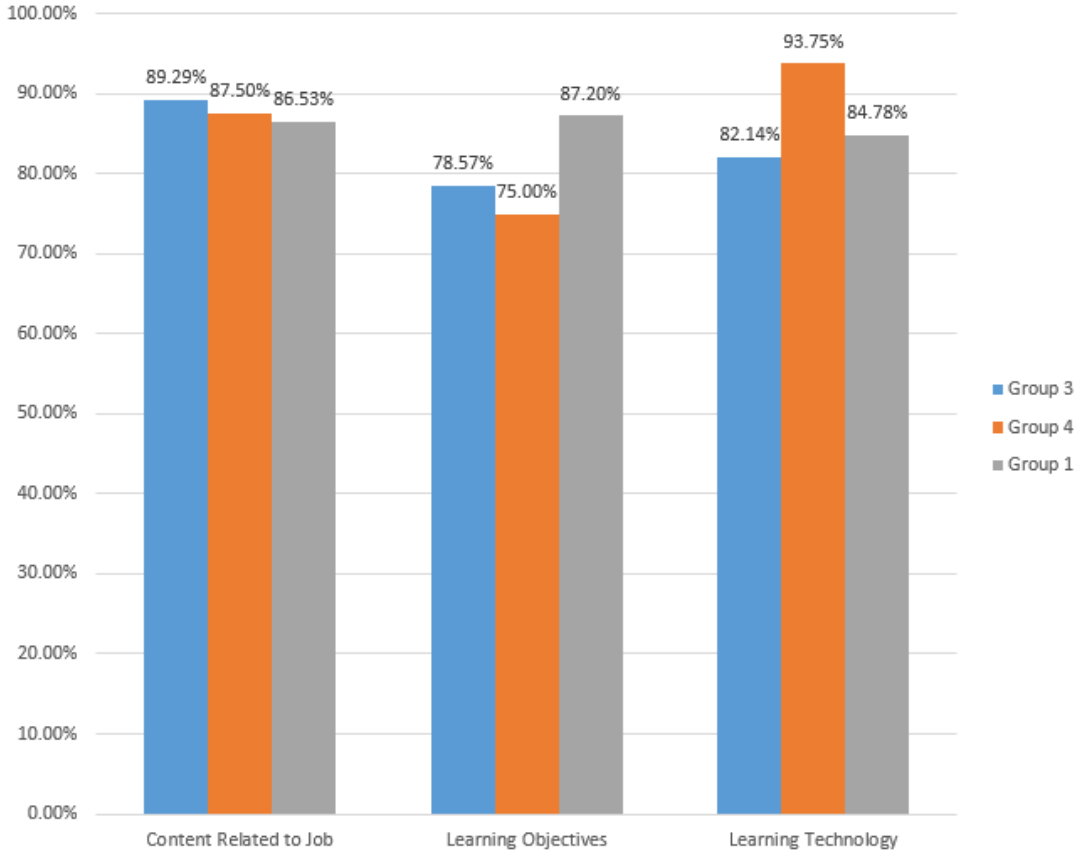


Figure 1 – Category and Group Comparison

5.2. Study Limitations

While this study provides a fascinating starting point for inquiry into this topic, some limitations exist related to the data available. The first consideration is related to the sample, and whether it accurately represents the population considered. Specifically, since level 1 evaluations are not a requirement of learners but are considered optional, we are only gathering the opinions and perceptions of those learners who felt strongly enough to take the time to respond. As a result, the responses may be skewed either towards a more positive angle (learners who felt strongly that the course went well and wanted to provide this feedback) or a more negative angle (learners who had a poor experience and want to share this). Greater accuracy could be obtained by sampling a set of the larger population and requiring learners to respond within the sample, to ensure that positive, negative, and neutral responses are all gathered and considered.

Another limitation of the study is related to the technology and the function of the LMS used to gather the data. Learners are not issued the level 1 evaluation invitation until the trainer providing the training marks the course roster as complete. As a result, if there are any issues with the roster close-out process, the learner may not even be given an opportunity to respond.

We were also challenged by the change of the level 1 evaluation questions during the time period considered. While the overarching topics were generally similar, as evidenced by the ability to categorize the questions for scoring analysis, it is certainly a challenge to have the required level of clarity on the data when there are significant differences in the data set. Specifically, several of the categories changed

from a 5-point Likert scale to a 2-point yes/no scale between iterations of the report, which rendered meaningful analysis within these categories impossible. Future studies of this type would be enhanced through the use of a consistent set of questions and question categories for the entire time period considered; which was not a possibility for this particular inquiry.

Lastly, the subjective open-ended responses that participants provided were not analyzed in this study. To appropriately code qualitative data, researchers must perform extensive review and analysis to ensure accuracy and objectivity (Creswell, 2003). Future studies of this type would benefit from a mixed-methods approach in order to gather both the quantitative responses that can be more easily analyzed, alongside open-ended qualitative responses from learners. Greater triangulation by considering perspectives from different sources beyond just the learners (e.g. trainers, operational leaders) could also strengthen the data and findings.

6. Conclusion

The findings of this research show a pattern of positive learner perceptions of virtual courses after the trainer has received additional training on virtual facilitation techniques and adult learning theory. Concepts from the literature related to application of adult learning principles in general, and specific guidance for the online classroom were supported in the findings. Although future study is needed to further explore this topic, the present study has enhanced the understanding of the relationship between application of adult learning principles in the online learning space, and positive participant experiences.

Abbreviations

LMS	Learning management system
ILT	Instructor-led training, used to indicate in-person/classroom training
VILT	Virtual instructor-led training, indicates online training
K-12	Kindergarten through twelfth grade school setting

Bibliography & References

Abdullah, M., Koren, S., Muniapan, B., Parasuraman, B., & Rathakrishnan, B. (2008). Adult Participation in Self-Directed Learning Programs. *International Education Studies, 1*(3), 66–72. https://doi.org/10.5539/ies.v1n3p66

Arghode, V., Brieger, E. W., & McLean, G. N. (2017). Adult learning theories: Implications for online instruction. *European Journal of Training and Development, 41*(7), 593-609. https://doi.org/10.1108/EJTD-02-2017-0014

US Department of Education. (2020). The NCES Fast Facts Tool – Distance Learning. National Center for Education Statistics (NCES) Home Page, a part of the U.S. Department of Education. https://nces.ed.gov/fastfacts/

US Department of Education. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Washington, DC: Author. Retrieved from www.ed.gov/about/offices/list/opepd/ppss/reports.htm

Blackwood, C. C. & White, B. A. (1991). Technology for teaching and learning improvement. In M. W. Galbraith (Ed), *Facilitating adult learning: A transactional process* (pp. 135-162). Krieger Publishing Company.

Caruso, S. (2018). Toward Understanding the Role of Web 2.0 Technology in Self-Directed Learning and Job Performance. *Contemporary Issues in Education Research*, *11*(3), 89–98. https://doi.org/10.19030/cier.v11i3.10180

Cookson, P.W. (2001). The online professional seminar: E-learning may aid professional development but there's no virtual miracle in sight. Education Week. https://www.edweek.org/ew/articles/2001/09/19/03cookson.h21.html

Creswell, J. W. (2003). *Research Design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage Publications, Inc.

Creswell, J. W. & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Pearson Education, Inc.

Davis, J. R., and Arend, B. D. (2013). *Facilitating seven ways of learning: A resource for more purposeful, effective, and enjoyable college teaching.* Sterling, VA: Stylus Publishing.

Deineha, I., Hromozdova, L., & Kovach, V. (2020). Realities of practical andragogy in the condition of the COVID-19 pandemic: migration pedagogy in Ukraine. ScienceRise: *Pedagogical Education, 5*(38). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3742203

Fein, A. D. & Logan, M. C. (2003). Preparing instructors for online instruction. *New Directions for Adult and Continuing Education, 100,* pp 45-55.

Grow, G. O. (1991). Teaching learners to be self-directed. *Adult Education Quarterly, 41*(3), 125–149. https://doi.org/10.1177/0001848191041003001

Hendriks, S., Sung, S., & Poell, R. (2018). Learning Paths of Customer-Facing Professionals in the Digital Age. *Journal of Workplace Learning, 30*(5), 377–392. https://doi.org/10.1108/JWL-01-2018-0023

Inverso, D., Kobrin, J., & Hashmi, S. (2017). Leveraging Technology in Adult Education. *Journal of Research and Practice for Adult Literacy, Secondary, and Basic Education*, *6*(2), 55–58.

Jelfs, A., & Richardson, T. E. (2012). The use of digital technologies across the adult life span in distance education. *British Journal of Educational Technology Vol 44* (2) 2013 338–351. https://doi.org10.1111/j.1467-8535.2012.01308.x

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society of Training Directors, 13*, 3–9.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), Training and development handbook: A guide to human resource development (2nd ed., pp. 301–319). New York: McGraw-Hill.

Kirkpatrick, D. L. (1996). Invited reaction: reaction to Holton article. Human Resource Development *Quarterly*, *7*, 23–25.

Knowles, M. S. (1975). Self-directed Learning. New York: Association Press.

Knowles, M. S. (1980). The Modern Practice of Adult Education: From Pedagogy to Andragogy. (2nd ed.). New York:Cambridge Books.

Merriam, S., & Bierema, L. (2014). Adult learning : linking theory and practice (First edition.). San Francisco, CA: Jossey-Bass, a Wiley brand.

Universal Aggregation For Service Convergence: Residential, Mobility & Business

A Technical Paper prepared for SCTE by

Michael Ting Wang
Network Architect III
Shaw Communications Inc.
2728 Hopewell Place NE, Calgary, AB,
Canada T1Y 7J7
+1 (403) 303-4054
Michael.Wang@sjrb.ca

1. Introduction

Network operators must modernize their networks to support the growing demand for faster internet connections, mobile backhaul and fronthaul, and business services. Universal aggregation is a converged networking approach that uses a shared coherent Dense Wavelength Division Multiplexing (DWDM) core with Internet Protocol (IP) routing and switching platforms.

As is the case for other Multiple-System Operators (MSOs), Shaw must remain competitive while striving to meet the intense challenges of delivering a better user experience during the COVID-19 pandemic. The pandemic has put increasing pressure on many operators and universal aggregation could prove to be a valuable tool for their current and post COVID-19 challenges. It provides the right speeds, footprint, aggregation, and automation capabilities with the ability to integrate Internet Protocol/Multi-Protocol Label Switching (IP/MPLS) with transport infrastructure.

This paper outlines Shaw's work to provide universal aggregation under its modular, coherent DWDM platform while providing greater capacity and significant network cost reduction. The legacy approach to deploying multiple interconnected and complex platforms for each type of service—such as residential, mobility, and business—often results in increased operational cost and difficulty troubleshooting.

For operators, universal aggregation creates a smarter, simpler, and more agile network. Shaw's universal aggregation network supports 100 Gigabits Ethernet (100GbE), Optical Transport Unit 4 (OTU4), 10 Gigabits Ethernet (10GbE), and Optical Transport Unit 2/2e (OTU2/2e) over 100Gbps & 200Gbps wavelengths. It uses a Zero Touch Provisioning (ZTP) system that simplifies operations and allows Shaw engineers to automate most of the deployment tasks. Currently, its maximum bandwidth capacity is 1 Terabit Per Second (Tbits/s) in a single rack unit at 224 watts. The paper will present a comprehensive overview of the implementation process and important considerations for the industry to replicate Shaw's success.

2. Drivers for Convergence

Operators' economic environment has changed significantly over the last two years.

Firstly, there has been a dramatic increase in traffic due to the explosion of residential data/video/voice, mobile xHaul, and business services. It has significantly reduced the spare capacity in the deployed networks and making it necessary to overbuild and introduce new transport technologies to satisfy projected traffic demand. Secondly, most operators are experiencing a need to reduce their operational costs, particularly in building and managing complex multilayer, multivendor networks. To simultaneously satisfy these two trends, operators need to integrate residential, mobility, and business traffic under a single transport network while also providing a simple and automated way to manage that network.

In network infrastructure, where in the past it may have been necessary to use separate equipment for optical transport (layer 1), Ethernet switching (layer 2), and IP/MPLS routing (layer 3), there is an evolution toward a model where all three tiers are integrated. The trends show a convergence of DWDM with Ethernet/client Optical Transport network (OTN)-Synchronous Optical Networking (SONET) as the initial step, and the integration of DWDM/Ethernet-OTN-SONET with IP/MPLS in the future. Separately, there is also a trend toward integrating three types of network services onto common infrastructure: residential, mobility, and business. Furthermore, network-layer convergence will integrate photonic with Ethernet/client OTN/client SONET and IP/MPLS.

Tables 1 and 2 demonstrate the concepts of traffic-type convergence and network-layer convergence, respectively.

Table 1 – Traffic Convergence

	Traffic Categories
Traffic Convergence	Residential
	Mobility
	Business

Table 2 – Network-Layer Convergence

	Network Layers
Network-Layer Convergence	Tier 1 - Optical Photonic
	Tier 2 - Carrier Ethernet
	Tier 3 - IP/MPLS (Future)

2.1. Convergence of Traffic Types

Since March 2020, cable networks have seen a 30.8% growth in downstream traffic and a 54.8% growth in upstream traffic (**ref. [9]**), and the enormous growth in the three service types is exacerbating network silos for operators. The convergence drivers for different traffic types will be discussed in detail below.

2.1.1. Residential Traffic

Over the years, the MSO landscape has proven itself to be dynamic. Residential triple-play once dominated revenue opportunities. Although cable operators are turning their attention to new revenue-generating opportunities in wireless and business services, residential is still a crucial revenue source. Now, operators are focused on aggressively reducing the operational costs of delivering residential services while continuing to meet customer demands for faster speeds.

Residential broadband services are generally carried over a separate DWDM/MPLS network from other types of services, managed by distinct network monitoring systems and, most importantly, live in their own priority structure. They have their own service tier, service-level agreement (SLA), latency requirements, and so on. The residential offerings act as an independent silo with the potential for significant operational inefficiencies.

If operators continue to operate their residential broadband products as separate silos they are at risk of failing to achieve the operating margins necessary to remain competitive due to a duplication of efforts and lack of focus. This could allow telcos that *can* offer multiple services on a converged platform to take their market share.

2.1.2. Mobility Traffic

After significant growth and success in video, data, and land line voice services, wireless is the next frontier for cable. Quarter over quarter, the US cable mobile virtual network operator (MVNO) business continues to see mobile subscriber growth (**ref. [1]**). As of 2021, less than four years after the launch of the first MVNO by Comcast, three US cable MVNOs combined have amassed millions of customers. Comcast and

Charter MVNOs utilize Verizon as the mobile network operator (MNO). Recently, Cox Communication also demonstrated an interest in starting an MVNO. The momentum is there, and executive leadership at cable companies has consistently shown strong support for and interest in growing the wireless business.

Comcast and Charter have been signaling for some time that they intend to build Citizens Broadband Radio Service (CBRS) based mobile networks in their existing cable footprints in an effort to reduce the compensation given to Verizon and other MVNO partners for use of their networks **(ref. [2])**. Their MVNO operations were intended as a way to build a subscriber base and brand in advance of owning their own wireless networks, despite incurring consistent Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) losses for an extended period of time. Cox—which had entered the wireless space a decade ago, only to exit after disappointing results—has signaled its intention to re-enter the wireless market through the purchase of a significant number of CBRS licenses across its cable footprint. The largest cable operators already have a dense network of millions of Wi-Fi hotspots that can easily be turned into 5G small cells to create a mobile network.

In contrast to the US, virtually all of Canada's largest cable and telco operators offer residential wireline, business wireline, and mobile services on their own infrastructure (**ref. [1]**). Rogers, Canada's largest cable and mobile operator, has been offering mobile services since 1985, with Vidéotron launching its wireless services in 2010, and Shaw acquiring Wind Mobile in 2015. The Canadian market faces strong competition from Canada's large incumbent telco operators, which have invested heavily in fiber to the home, connecting more than 60% of their broadband homes directly to fiber, and leveraging a robust Radio Access Network (RAN) sharing agreement to minimize their infrastructure costs.

Today's cable operators are tomorrow's mobile operators **(ref. [3])**, and behind every efficient wireless network, there must be a converged wireline network. The industry has reached the consensus that optical DWDM transport's capability of transmitting a large number of information streams simultaneously over a single optical fibre makes it the best possible solution for the most demanding xHaul needs. Operating separate residential and mobile optical cores is prohibitively costly for operators, while converging optical networks into a single, common core will significantly cut costs and improve network agility. A converged network will manage peak traffic load more effectively, improving both speed and performance.

2.1.3. Business Traffic

Operators of all sizes have found growing opportunities to take part in the lucrative and ever-evolving business services space. While the battle for the wireless dollar has gotten most of the ink in the trade press recently, the traditional telcos and cable MSOs are waging a war on another, less publicized front: the small and medium-sized business market. The telcos are moving aggressively into the residential video services market, where cable operators currently enjoy a high penetration rate. To compensate for the loss of residential video customers, operators have chosen to pursue perennially underserved small- and medium-sized businesses. And they are winning over these customers with Carrier Ethernet.

Operators typically deliver business services over disparate architectures. A high-security application, for example, runs over its own DWDM wavelength. Schools, financial institutions, and telephony that require fast switching (50-millisecond) protected services would be typically served via a SONET network. And low-priced Ethernet services run over shared bandwidth tunnels using Layer 2 or 3 architectures. In addition, business customers may favour symmetrical services, while residential customers are well served by DOCSIS, which is currently asymmetrical, typically offering 5:1–10:1 ratios in the downstream/upstream bandwidth. With business customers driving cable operators to better utilize existing

fibre plant and invest in new FTTP deployments, the trend is to offer multiple broadband pathways over existing fibre using DWDM.

2.2. Convergence of Network Layers

The layer-1 photonic network, the layer-2 carrier Ethernet network, and the layer-3 IP/MPLS network are designed, built, and expanded independently, and are often maintained by teams working in complete silos. This results in each layer being over-provisioned to cope with uncertain network demands (**ref. [4]**) and the networks being overprotected due to overlapping and redundant resiliency schemes on each layer.

This also results in additional Capital Expenditure (CAPEX) and Operational Expenditure (OPEX). In terms of CAPEX, transponders, switch ports, and router ports—the most expensive parts of the core of the network—become layered and the siloed infrastructure relies on large volumes of line cards for traffic hand-off between networking layers. Additionally, OPEX is increased because multiple teams are required to handle the provisioning and maintenance of a single service. Inefficiencies can also be seen in other areas, including a loss of time because of the need to deploy multiple processes and teams, and high complexity due to multiple independent network management systems (NMS) associated with each network layer, which results in high costs and poor network resource utilization and leads to poor monetization.

Multiple trends are driving operators to integrate DWDM optical transport with Ethernet/client OTN-SONET technology, primarily in metro and more recently in long haul networks. These trends include the demand for more bandwidth and greater agility with key applications, including cloud connect services, fixed broadband aggregation, mobile xHaul, and SONET migration.

2.3. Six Mandatory Attributes of The Unified Transport Platform

Operators have been asking for simpler, more cost-efficient converged network architectures that will enable them to concentrate on innovating revenue-generation services (**ref. [5]**). Based on Shaw's modeling, the converged agile optical framework must have following six attributes:

1) High bandwidth densification per rack space

For mobility operators, network densification means adding more cell sites to increase the amount of available capacity, which demands high bandwidth densification per rack space of the unified transport platform that contains Mobility xHaul, with a minimum threshold of 800Gbits/s per 1 Rack Unit (RU).

2) Building-block-approach scalability

Scalability is the ability of a system to expand without major modifications to its architecture. The concept implies the ability of a network system to accommodate a sudden increase in traffic volume gracefully and rapidly. After meticulous modeling by Shaw, it was found that the unified transport platform must be able to handle the unpredictable growth pattern of residential, mobility, and business traffic with building-block-approach scalability.

3) Full-set client interface support

The unified transport platform is required to support residential, mobility, and business services. Such a platform must be able to provide all required standard client interfaces, including Ethernet, OTN, and SONET.

4) Full-featured General Communication Channel (GCC0)

The unified transport platform has to integrate business services. Because of the power and rack space restrictions of business customer premises, installing an out-of-band management switch is usually not a viable option, or at least a very expensive one. Shaw's assessment concludes that in-band full-featured

GCC0 is a vital attribute of the converged agile optical framework. For a detailed explanation of GCC0, please go to Section 3.4.

5) Zero Touch Provisioning (ZTP)

Because the unified transport platform converges all service types, deployments are almost always multi-sited. Traditional turn-up based on manual initial configuration requires highly trained field personnel who are often difficult to find, and therefore multi-site deployment without ZTP tends to be slow, costly, and highly error prone. However, Shaw's analysis indicates that ZTP is a mandatory attribute for the unified transport platform.

6) Flex-grid

The concept of Flex-grid is related to the first attribute—high level bandwidth densification per rack space. To achieve high bandwidth densification, operators have to look beyond 100Gbps using a fixed grid. Only Flex-grid DWDM systems support 400G+ with high spectral efficiency. While fixed grid DWDM systems can still support "fat channels", it is at the expense of significantly lower spectral efficiency. Spectral efficiency is an important measure of how effectively or efficiently a fiber network transmits information.

3. Classic Optical Network

Classic Optical Networks are based on Legacy DWDM technology, which are hostile environments for service convergence. It is exceedingly difficult to implement service convergence in a classic optical network, if not entirely impossible. Legacy DWDM networks have the following defining features:

- It uses a chassis-based platform with low level bandwidth densification per rack space.
- Its scaling is inefficient.
- It is limited to SONET and 10GbE/40GbE client interfaces.
- It does not support line-side GCC0.
- It does not support ZTP.
- It uses a Fixed Grid (50 – 100GHz).

All the above characteristics present severe restrictions for network convergence. These characteristics are discussed in detail in the following sections.

3.1. Chassis-Based Platform and Low-Level Bandwidth Densification

All classic DWDM equipment manufacturers have chosen a chassis-based platform. In such a platform, each DWDM component is devised as a card to be installed in an empty slot of the chassis. For example, Arrayed Waveguide Gratings (AWG); Wavelength Selective Switches (WSS); Erbium-Doped Fiber Amplifier (EDFA); and most transponders/muxponders are typically constructed as individual cards. Figure 1 shows a typical 14 RU monolithic chassis.

A transponder/muxponder card is commonly installed vertically but still roughly takes about 1 RU. A typical legacy transponder/muxponder only has the bandwidth density of up to 100Gbps per 1RU. 100Gbps is significantly less than the minimum requirement of 800Gbit/s per 1 RU for a converged platform.

The transponders/muxponders work in pairs in a point-to-point system. At a specific source, the destinations for residential service, mobility service, or business service are usually different. Therefore, in a classic system this would require three pairs of transponder/muxponder cards at the source and destination, resulting in 3RU per site.

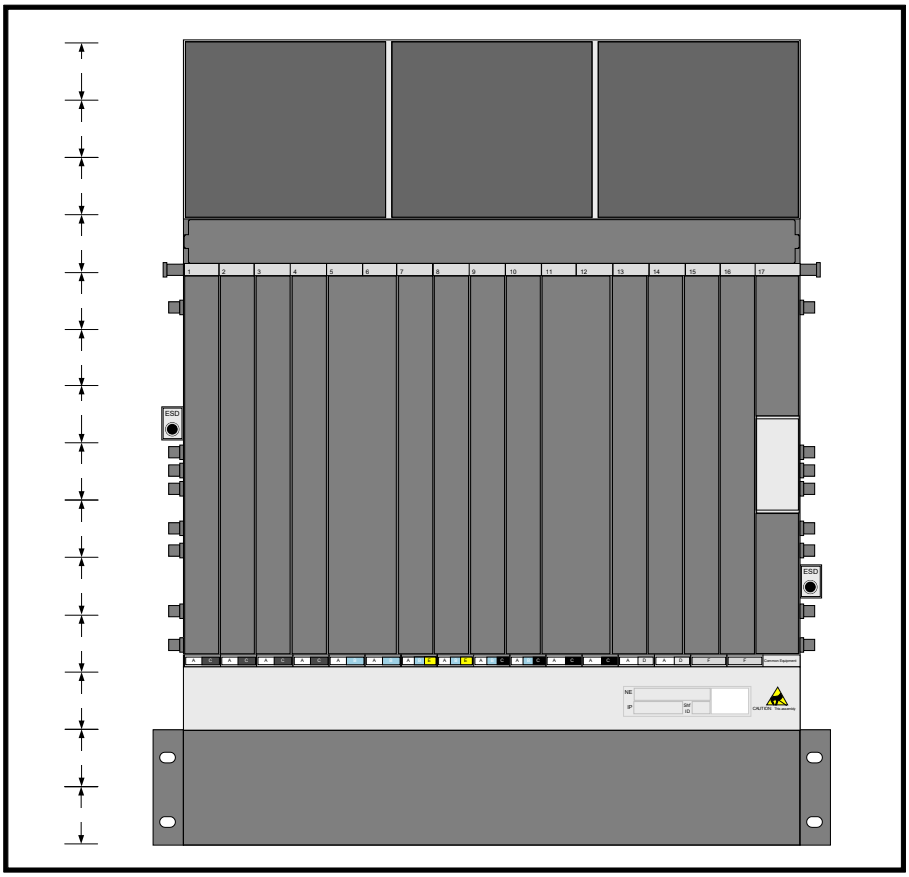


Figure 1 – Chassis Based Platform

3.2. Inefficient Scaling

The legacy monolithic-chassis DWDM systems are able to handle future traffic volumes on day one, but also require a large up-front investment for capacity that may not be needed for several years. This is particularly true for chassis equipped with specialized hardware such as terabit scale switching fabrics, as the fabric is part of the initial installation, even though transponders may be added over time. As for operators, space and power are also always at a premium, and classic optical networks require large space and power budgets at the beginning of deployment. The issue of inefficient scaling is particularly acute for remote mobility sites and business customer premises.

In order to achieve service convergence for mobility and business, MSOs need to examine modular DWDM systems with smaller footprints. For example, 1RU pizza-box blades that offer the advantage of efficient scaling for all operators.

3.3. Client Interface Limitations

Classic DWDM systems usually only support SONET client interfaces and low-rate Ethernet client interfaces. For a legacy transponder/muxponder, only three client formats are typically supported, including SONET OC192, 1GbE, and 10GbE, with only a few legacy DWDM vendors offering 40GbE and 100GbE

client interface support. While these three client interfaces are often sufficient for residential traffic, they fall short of the requirements for mobility and business traffic, which both require 100GbE and OTU4. In addition, many business customers also need OTU2/2e.

Because of the limited client formats, MSO operators are forced to move mobility and business traffic away from legacy optical networks and onto IP/MPLS networks. The traffic migration to IP/MPLS is considered suboptimal due to DWDM offering the most efficient use of optical fibre throughput capacity, as well as lower price points. As far as the client ports are concerned, classic optical network focuses on residential customers, while overlooking mobility and business customers. The client interface format limitation is a serious roadblock to service convergence.

Table 3 and Table 4 list the key parameters for OTN Frames and SONET OC192, for reference.

Table 3 – OTN Frames

OTUk	Bit Rate (Gbps)	Payload Rate (Gbps)	Payload Types
OTU4	111.809973	104.355975	100GbE
OTU2e	11.095730	10.356012	10GbE LAN, 10GFC (TTT)
OTU2	10.709255	9.995277	10GbE WAN, 10GbE LAN (GFP-F), STM-64/STS-192

Table 4 – SONET OC192

Acronym	Bit Rate (Gbps)	DS0	DS1	DS3
OC-192	9.9953	129,024	5,376	192

3.4. Client-Side GCC0 Only

General Communication Channel 0 (GCC0) is used for overhead communication between network nodes. GCC0 has operations, administration, and management (OAM) functions such as performance monitoring, fault detection, signaling and maintenance commands in support of protection switching, fault sectionalization, service-level reporting, and control plane communications. GCC0 has two bytes within OTN overhead. GCC0 is terminated at every 3R (re-shaping, re-timing, re-amplification) point and used to carry management information and GMPLS signaling protocol.

Legacy DWDM transponders/muxponders only support client-side GCC0. Optical Transport Network (OTN) standards define the GCC0; it is an in-band channel used to carry management information between transponder/muxponder pairs. GCC0 is a critical method to manage an optical device for a business customer site. Because of power limitations and rack space restrictions, it is very difficult to install a layer-2 management switch for out-of-band management in a business customer site. Figure 2 represents a schematic diagram on client-side GCC0.

Although client-side GCC0 is a useful feature, it is considered as only a partial implementation of GCC0. In client-side GCC0, the network interface device (NID) at the far-end site encapsulates management information into the GCC0 bytes of each OTU2/OTU2e frame. The NID sends out the OTU2/OTU2e frames from its network-side port, which has a grey (black and white) wideband pluggable, most likely at 1310nm. When the near-end transponder/muxponder receives the frames on its client-side port from the far-end NID, it decapsulates the GCC0 bytes from the OTU2/OTU2e frames and recovers management information from the GCC0 bytes.

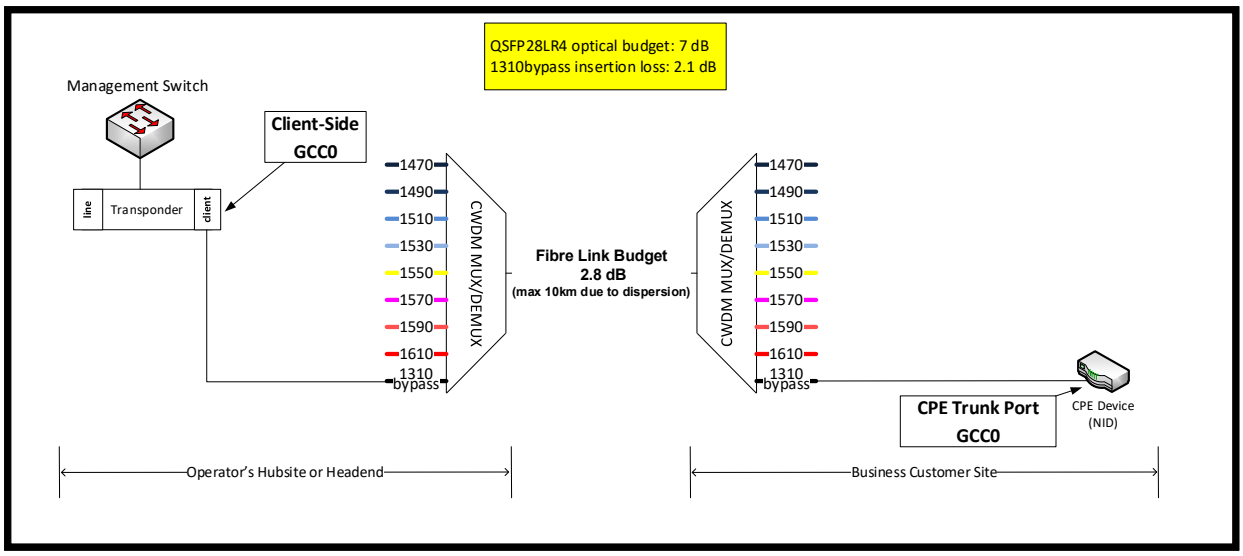


Figure 2 – Client-Side GCC0

While a management switch is necessary for the near-end site, it can be skipped for the far-end site when client-side GCC0 is supported on the near-end transponder/muxponder. It should be noted that the limiting factor for client-side GCC0 is the distance between the near-end and far-end. If there are only dark fibres between the near-end and far-end, the maximum distance for client-side GCC0 is approximately 10km, depending on the fibre rating. If CWDM is inserted between the near-end and far-end, the maximum distance is reduced to 5km. Though client-side GCC0's maximum distance is adequate for residential service and mobility service, it is insufficient for most cases of business service.

In summary, due to the aforementioned factors, legacy DWDM devices' lack of line-side GCC0 support considerably hampers the potential for service convergence.

3.5. Manual Turn-up Initial Configuration

As the name suggests, the goal of Zero Touch Provisioning (ZTP) is to install a networking appliance without the need for local configuration by a trained individual, making it possible for a new or replacement device to be sent to a site, physically installed, and powered up by a locally present employee without technical skills. The ZTP feature carries out the configuration and connection to the management system.

Legacy photonic-layer devices do not support ZTP, and therefore extensive manual configuration is required to turn up a new device in a classic optical network. Manual configuration is laborious, prone to errors, costly, and time-consuming. In this scenario, an individual with basic configuration skills and a laptop has to go onsite and configure the device for basic operation before its configuration can be completed remotely using the central management system **(ref. [6])**. Alternatively, the device can first be sent to a central location where it is staged before being sent to its final location for installation. This is also costly as it requires shipping the device twice, which could potentially mean passing through customs twice. And there is the risk of accidentally shipping an appliance with an IP-address destined for a different site.

For residential service, ZTP is helpful but it is not always needed. However, for mobility service at a remote site and business service at a customer premise, ZTP is critically important. With ZTP, a device can be shipped directly from any warehouse to the remote site or customer premise and installed as soon as it

arrives and be up and running within minutes of installation. This dramatically reduces the lead time, time spent on an installation, and number of configuration errors, which are significant benefits for mobility and business services.

Because legacy DWDM's chassis does not support ZTP, operators must look into next generation optical devices for service convergence. ZTP is becoming more widely supported as next generation optical equipment vendors realize their equipment can now be installed by local techs who may not be trained on provisioning and configuration.

3.6. Fixed-Grid Photonics

Older generation optical networks are based on 100 Ghz or 50 Ghz spaced photonic systems. These gridded networks can offer forty-eight or ninety-six fixed grid optical channels within the total 4800 Ghz C-Band Spectrum. This fixed grid spacing is based on the International Telecommunication Union (ITU) standard and has been the norm for most photonic systems for over 20 years. These systems use passive, ITU grid filters to provide wavelength ingress/egress. Second and third generation optical transponders running at 35 GBaud typically require 37.5 Ghz of optical spectrum, which fits perfectly into these ITU gridded filters. This version of the DWDM network has served the industry well for years.

As next generation transponders are moving towards 400GbE, they bring with them the need for larger per channel spectrum. This larger channel spectrum requirement exceeds that of what is available on these ITU gridded filters. The Fixed-Grid paradigm cannot support them, making this method obsolete.

4. Universal Aggregation

Universal Aggregation is a converged networking approach that enables the aggregation of traffic from SONET, Ethernet, and OTN services using shared fibre and the same optical DWDM platform. Under the Universal Aggregation paradigm, the DWDM systems are considered as both core technology and access technology. Behind the SONET, IP/Ethernet, and OTN services, the traffic sources include residential, wireless, and commercial.

In a traditional network, it is no wonder that performance and reliability are increasingly valued as operators must configure, shape, and optimize each service type individually on separate platforms, with each platform having its own vendor-provided technical support. Sometimes, this means dealing with multiple platforms with the same vendor. However, in the majority of cases, operators must deal with several different platforms from separate vendors. Managing multiple interconnected hardware deployments across diverse vendor environments significantly increases operational cost and makes troubleshooting a very complex task. Traditional separated network architectures are simply unable to scale quickly enough, negatively impacting time-to-market for new service revenues.

In traditional network architectures, each of the three network layers has its own network silo. There is growing evidence against siloed networks, contributing to its inherent weakness as part of the network architecture. Disadvantages of network silos include:

- An inability to better utilize or share the silo's resources.
- The extended period of time needed to deploy, manage, and upgrade in the siloed environment.
- Rising operating costs, given each silo's unique processes.
- The redundant building of each silo's own protection systems to meet availability requirements.

Recent developments in coherent optical technologies provide greater opportunities for operators to move toward comprehensive network convergence. In network converged systems, residential, mobility and business services coexist over the same fibres, forwarding tables, data planes, servers, etc. However, getting to this point will take time and careful planning. To help support the shift towards full network convergence, operators need a robust platform framework.

Universal Aggregation breaks up the aforementioned silos by converging all into a single, unified platform. This means aggregating mobile, residential, and enterprise traffic vertically, while at the same time aggregating SONET, Ethernet, and OTN services horizontally. Universal Aggregation enables a simple, compact, scalable, and efficiently converged infrastructure with coherent optic transport, saving operators from having to build different networks for different services. By supporting all services on a unified platform, Universal Aggregation reduces operational cost considerably and expands competitiveness substantially.

Table 5 – Universal Aggregation

	Traffic-Type Convergence	Network-Layer Convergence
Universal Aggregation	Residential	Tier 1 - Optical Photonic
	Mobility	Tier 2 – Carrier Ethernet
	Business	Tier 3 - IP/MPLS (Future)

Shaw has successfully implemented Universal Aggregation on a blade-centric architecture. It is a converged framework spanning access, transport, packet, and optical, and combining fast, simple deployment with a wide range of options for right-sizing capacity. Residential, mobility and commercial services are aggregated into one unified platform supporting SONET, Ethernet, and OTN. Shaw's implementation of Universal Aggregation has the following technical characteristics:

- High Bandwidth Density per Rack Space
- Efficient Scalability
- Full-Set Client Support - SONET, Ethernet, and OTN
- Line-side GCC0
- Zero Touch Provisioning
- Flexible Grid

Shaw's successful experience indicates that the most efficient unified technical platform of Universal Aggregation is a modular, coherent DWDM system with Ethernet/OTN/SONET client interfaces. In the section below, we will discuss the blade-centric platform in detail.

4.1. Overview of Blade-Centric Platform

As mentioned previously in Section 3, over-engineered, restrictive, chassis-based systems are not adaptable in the converged environment. A typical chassis is shown on the left side of Figure 3, and a typical modular blade is shown on the right side. The typical DWDM chassis is of 14RU, while the typical modular pizza-box is only of 1RU.

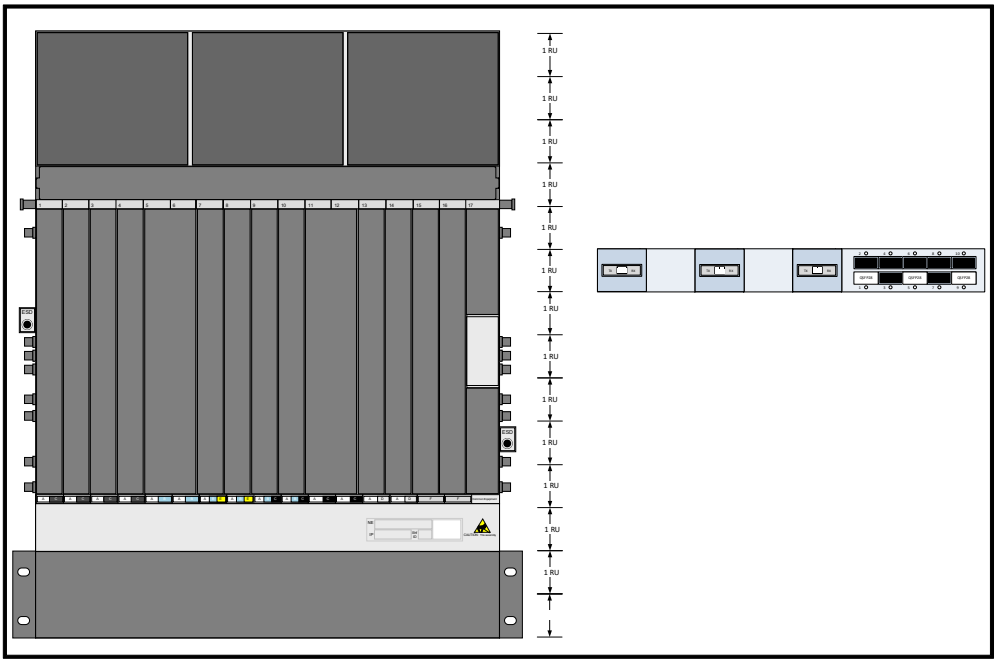


Figure 3 – Chassis vs Blade

The modular product line divides DWDM functions into three areas: ROADM blades, Splitter/Coupler blades, and Transponder blades, as shown in Figure 4. The ROADM-on-a-blade provides wavelength selective switching and amplification. While Figure 5 only shows two ROADM blades, our platform currently supports a maximum of eight ROADM blades (one main blade and up to seven tributary blades) that can be interconnected as a single Network Element (NE). A Splitter/Coupler blade has the main function of channel add/drop and replaces legacy static filters in CDC (Colorless, Directionless, Contentionless) configurations. Splitter/coupler blades are interconnected with the ROADM blades as one NE, and the Transponder blade is a transceiver with very high port density on both network and client sides, which in the future could be part of a single NE as well.

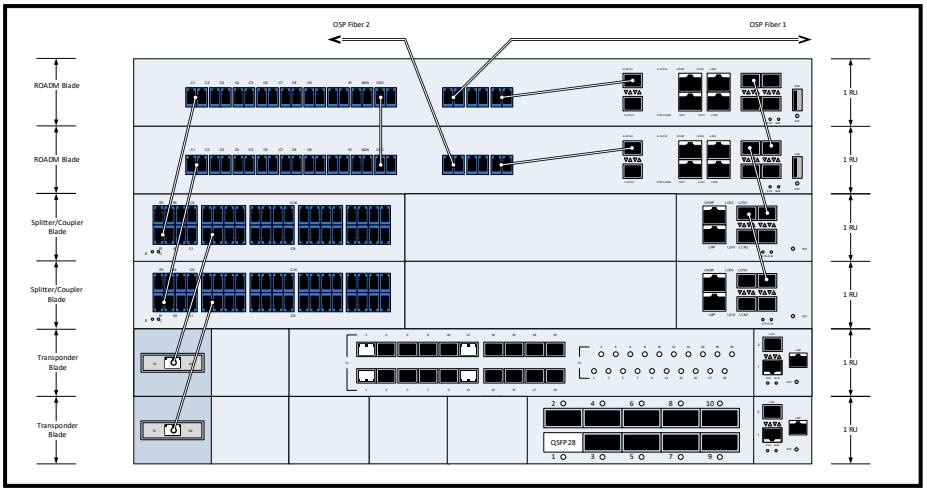


Figure 4 – ROADM Blades, Splitter/Coupler Blades, and Transponder Blades

This isolation approach allows vendors to rapidly implement and deploy new features in each secluded functional area when and as much as needed. Because of the modularity, vendors tend to be much more responsive to feature requests from operators.

4.2. High Network Densification

The blade-based platform is a simpler, more agile approach to creating an optical network for Universal Aggregation of residential, mobility, and business traffic. The blade paradigm is the only feasible platform for universal aggregation at this point in time due to the exceptionally high bandwidth density per 1RU. Figure 5 shows a typical implementation of Universal Aggregation on a modular platform. The transponder blade at the bottom of Figure 5 aggregates residential, mobility, and business into one 1RU pizza-box at the source. To achieve the same goal, the typical chassis-based platform would have to use three transponder cards, which means approximately three times of power consumption and rack space.

The transponder blade's line port bandwidth capacity can either be 100Gbits/s or 200Gbits/s, depending on the modulation on the line port, which ultimately depends on the fibre span loss and OSNR (Optical Signal to Noise Ratio) of the system. If the distance between the source and destination is shorter than 230km, it is likely that 200Gbits/s would suffice. If the distance is longer than 230km, however, capacity would have to be 100Gbits/s.

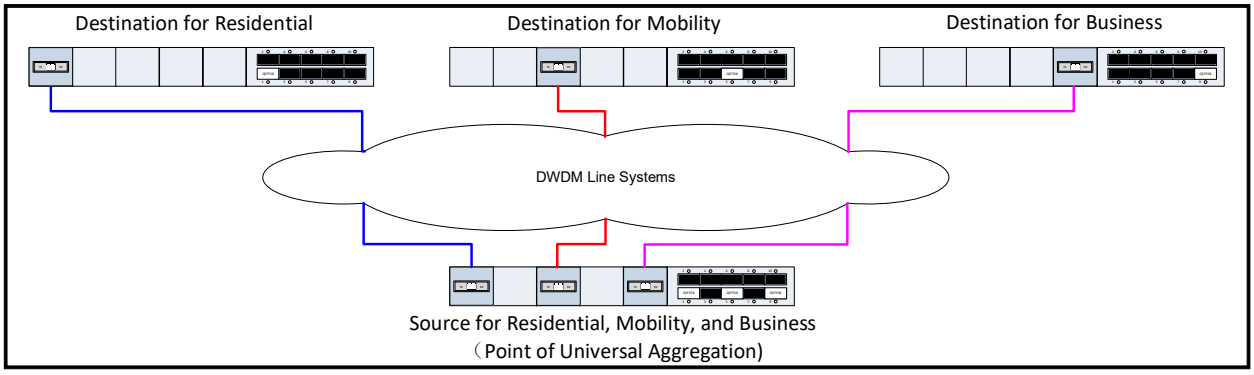


Figure 5 – Universal Aggregation on A Modular Platform

Figure 6 shows a typical transponder module with QSFP28 client interfaces. It has five line ports, and each line port has a maximum bandwidth capacity of 200Gbits/s. The total maximum capacity for the 1RU module is 5x200Gbits/s, which is 1Tbits/s.

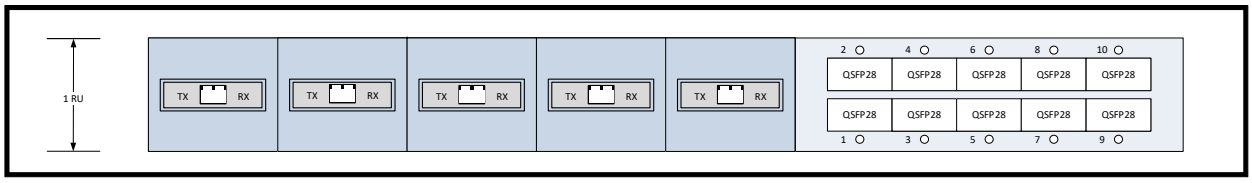


Figure 6 – Maximum Bandwidth Capacity of a Transponder Blade

While 1Tbits/s bandwidth capacity per pizza-box is considered extraordinary for a 1RU rack space, the power consumption is about 224 watts, which is on par with shelf-based transponder cards. In the blade-centric paradigm, the transponder blades consume a larger percentage of total power as compared to ROADM blades and splitter/coupler blades; this fact enables vendors to focus on independent improvements in power for transponder blades without the constraints of traditional converged shelves.

Most operators should be able to achieve an overall 30% power savings per 100Gbps bandwidth compared to legacy systems.

4.3. Efficient Scalability with Building-block-like Infrastructure

Traditional chassis-based platforms represent a significant initial capital outlay and the chassis backplane is inherently inflexible. On the contrary, a blade-based modular system requires a smaller, granular initial investment while allowing for a pay-as-you-grow approach. With single rack unit sized blades, the system is also space efficient and flexible.

The paradigm offers the advantage of efficient scaling for all operators. Scaling efficiently is one of the key network requirements for operators in the era of Universal Aggregation. The building-block approach to hardware allows for a low initial spend for year-one deployments with the ability to grow incrementally as traffic increases and more capacity is required. Many converged, monolithic-chassis DWDM systems, by contrast, are able to handle future traffic volumes on day one, but also require a large up-front payment for that capacity even when the capacity may not be needed for several years. This is particularly true for chassis equipped with specialized hardware such as terabit scale switching fabrics, as the fabric is part of the initial installation, even though transponders may be added over time.

4.4. Full Set Client Support - SONET, Ethernet, and OTN

Each service type demands a unique set of client interfaces from DWDM transponders. For residential landline telephony, SONET OC192 is needed. For residential Internet, both 10GbE and 100GbE have to be supported to communicated with CMTS. For mobility xHaul, 100GbE and OTU4 are required. Business customers request the widest range of client interfaces: SONET OC192, 10GbE, 100GbE, OTU4, OTU2/2e.

Table 6 summarizes the client interface formats required by each service type.

Table 6 – Service Type/Client Interfaces

Service Type	Client Interface Formats
Residential	SONET OC192,10GbE, 100GbE
Mobility	100GbE, OTU4.
Business	SONET OC192, 10GbE, 100GbE, OTU4, OTU2/2e

Chassis-based legacy platform only supports a limited subset of the client interfaces listed in the above table, ruling out monolithic chassis as a platform for Universal Aggregation. However, the blade-centric platform supports all client interfaces in Table 6. As listed in Table 7, Module A supports SONET OC192, 10GbE, and OTU2/2e, while Model B supports 100GbE and OTU4. Two models of modular blades are also shown in Figure 7, below.

Table 7 – Model/Client Interfaces

Model Type	Client Interface Formats
Model A	SONET OC192,10GbE, OTU2/2e
Model B	100GbE, OTU4.

Comparing Tables 6 and 7 confirms that together, Models A and B will provide full set client interfaces for residential, mobility, and business services, and that modular blades are ideal for Universal Aggregation of these three types of traffic.

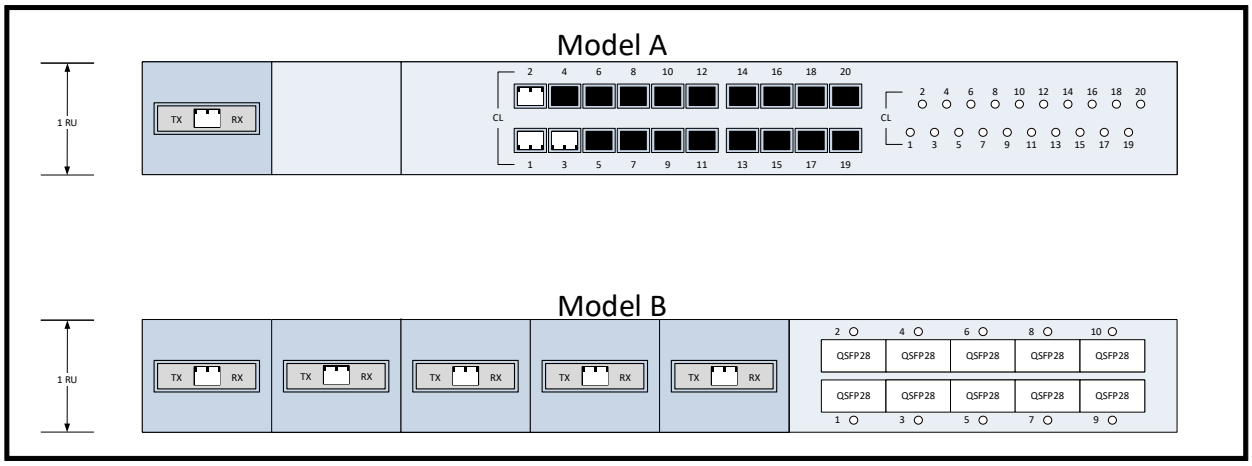


Figure 7 – Model A and Model B

4.5. Line-Side GCC0

As discussed in Section 3.3, legacy DWDM systems only support client-side GCC0. Blade-based next-generation DWDM systems support both client-side GCC0 and line-side GCC0.

For line-side GCC0, the transponder/muxponder in the far-end site encapsulates management information into the GCC0 bytes of each OTU4 frame. The OTU4 frame then goes through the electrical-to-optical conversion and is sent out of the line port of the far-end transponder/muxponder on a specific DWDM wavelength. Upon receiving each OTN frame by the near-end transponder/muxponder, the GCC0 bytes are extracted and the management information is retrieved. In using line-side GCC0, there is no need for a management switch in the far-end site, although a management switch is still necessary for the near-end site. The maximum distance for line-side GCC0 is about 80km depending on the fibre rating, even with CWDM inserted. This distance is adequate for most business services.

Figure 8 below is the schematic diagram of line-side GCC0.

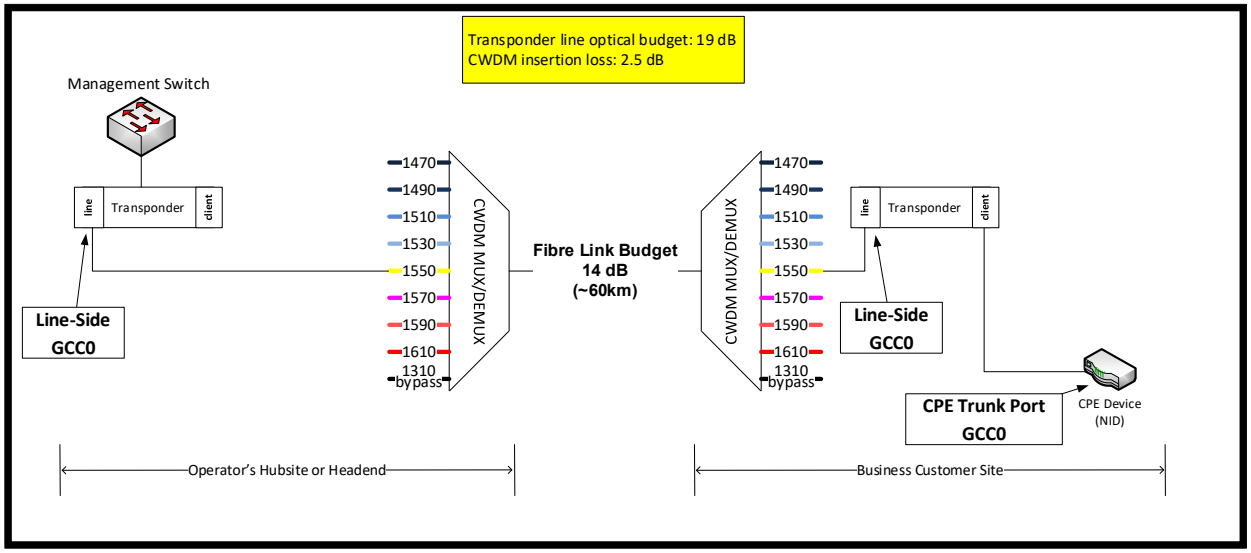


Figure 8 – Line-Side GCC0

4.6. **Zero Touch Provisioning (ZTP)**

The blade-centric next-generation platform also supports ZTP. Shaw has set up several pre-staging centres in various regions. Before shipping a blade to the site, Shaw engineers will pre-provision the blade and save the configurations to the USB key attached to the blade. The blade will then be shipped with the USB key to local technicians. When the local technicians receive the blade and USB key, all that is required is a plug-and-play. Figure 9 shows a typical set-up of a pre-staging lab.

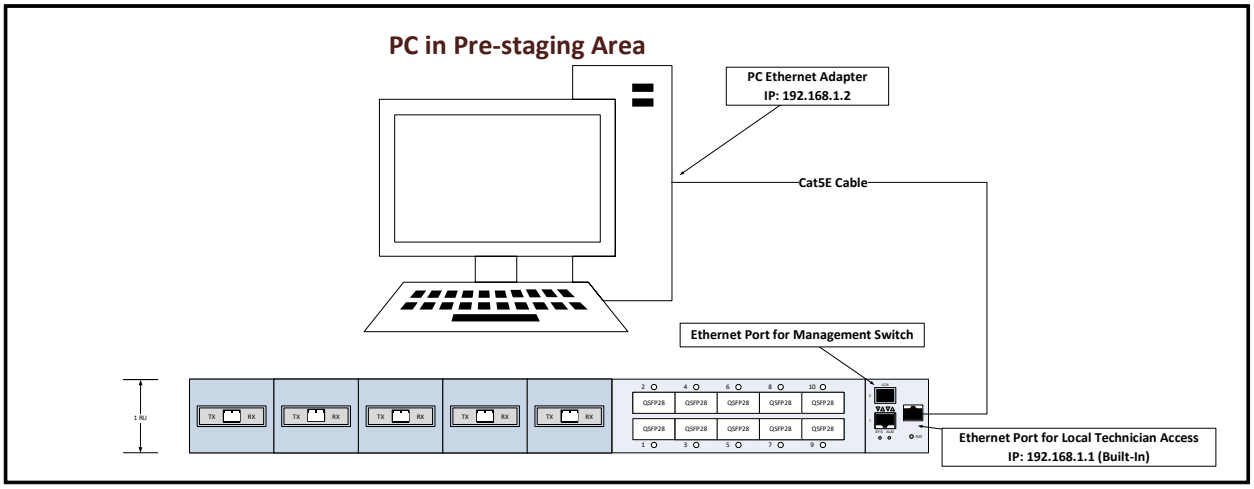


Figure 9 – Zero Touch Provisioning

4.7. **Flexible Grid**

Flexible Grid, also written as Flex-Grid, is in contrast to Fixed-Grid which was discussed previously in section 3.6. Flex-Grid combines two concepts together: finer wavelength granularity and the ability to join adjacent wavelength slots together to form arbitrary sized channels. From hardware perspective, Liquid Crystal on Silicon (LCoS) based Wavelength Selective Switch (WSS) makes a ROADM Flex-Grid.

Universal Aggregation requires 400Gbps transponder/muxponder because residential, mobility, and business each normally requires 100G client bandwidth. The blade-centric next-generation modular platform provides a 400G transponder blade with Flex-Grid. With the 400G blade, operators will be able to transport more than twice the information within the same spectrum on Fixed-Grid. Operators can actually further increase the spectral efficiency by 25% with the use of Flex-Grid by allowing the subcarriers in DWDM superchannels to be squeezed more closely together. While Flex-Grid drastically improves spectral efficiency, the modular 400G transponder pizza-box does not require Flex-Grid as absolute pre-requisite, it will also work on existing legacy Fixed-Grid for backward compatibility.

The superchannels enable operational scaling by allowing operators to turn up optical capacity in larger increments with the same effort. Going forward, next-generation optical transport networks will need to make the most use of the flexibility of advanced coherent modulation technologies. More advanced flexible coherent modulations will support a wide range of modulations tailored to specific applications of residential, mobility, and business. Flexible DWDM grids will enable more efficient and flexible use of optical spectrum to maximize capacity.

5. Conclusion

This paper has outlined the benefits of Universal Aggregation on a modular platform, which is a promising paradigm for Cable operators.

Under the paradigm of Universal Aggregation, it is unnecessary for operators to build separate networks for different traffic types of residential, mobility and business. By aggregating all three types of service on a unified platform, Universal Aggregation reduces operational cost considerably and expands competitiveness substantially.

In traditional systems, each network layer has their own network silo. In the Universal Aggregation model, these silos are all broken down. The consistent unified platform across optical, Ethernet, and IP/MPLS will smash down all network silo walls.

Shaw's Universal Aggregation utilizes a blade-centric platform with the following features:

- It has astoundingly high bandwidth density per 1RU with a 90% increase in available system bandwidth comparing with legacy systems.
- It uses a building block approach to scalability, which only requires a small initial spend while allowing for continuous growth through pay-as-you-go.
- It supports 100GbE, OTU4, 10GbE, and OTU2/2e over 100Gbps and 200Gbps wavelengths.
- Its line-side GCC0 offers in-band management for remote mobility sites and business customer sites.
- It implements a Zero Touch Provisioning (ZTP) system that simplifies operations and allows engineers to automate most of the deployment tasks.
- It supports Flex-Grid which increase the spectral efficiency by 25% for 400G blades.

Abbreviations

3R	Re-shaping, Re-timing, Re-amplification
10GbE	10 Gigabits Ethernet
100GbE	100 Gigabits Ethernet
CAPEX	Capital Expenditure
CBRS	Citizens Broadband Radio Service
CPE	Customer Premise Equipment
DSP	Digital Signal Processor
DWDM	Dense Wavelength-Division Multiplexing
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortization
EDFA	Erbium-Doped Fiber Amplifier
GCC	General Communication Channel
GMPLS	Generalized Multi-Protocol Label Switching
IP	Internet Protocol
IP/MPLS	Internet Protocol/ Multi-Protocol Label Switching
ITU	International Telecommunication Union
LCoS	Liquid Crystal on Silicon
MSO	Multiple-System Operators
MVNO	Mobile Virtual Network Operator
NE	Network Element

NID	Network Interface Device
NMS	Network Management System
OAM	Operations, Administration and Management
OPEX	Operational Expenditure
OSNR	Optical Signal to Noise Ratio
OTN	Optical Transport Network
OTU2/2e	Optical Transport Unit 2/2e
OTU4	Optical Transport Unit 4
RAN	Radio Access Network
ROADM	Re-configurable Optical
RU	Rack Unit
SONET	Synchronous Optical Networking
Tbits/s	Terabit Per Second
UA	Universal Aggregation
WSS	Wavelength Selective Switch
ZTP	Zero Touch Provisioning

Bibliography & References

[1] *Jennifer Andréoli-Fang, John T. Chapman*, "*Cable and Mobile Convergence - A Vision from the Cable Communities Around the World", SCTE•ISBE, 2020.*

[2] *Jeff Heynen, "Cable's Fixed-Mobile Convergence Future", www.delloro.com, 2020*

[3] *Jeff Baumgartner, "Cable's evolutionary path leads to mobile, convergence", www.lightreading.com, 2020.*

[4] *Fernando X. Villarruel, "Framework for Convergence of Services on The MSO Network - Using the Principles of Network Slicing", SCTE•ISBE, 2020.*

[5] *Kevin Bourg, Sergey Ten, Peter Wigley* "*An Overview Of Optical Architectures Necessary To Achieve 5G's Key Performance Indicators, SCTE•ISBE, 2020.*

[6] *Timothy Maenpaa, "Addressing Unrelenting Growth In Backbone Fiber Systems Using Next Generation Photonics And Automation", SCTE•ISBE, 2020.*

[7] *Chad Andrews, Steve Canepa, Bob Fox, Marisa Viveros, "The end of communications services as we know them - How 5G and edge computing will help define who wins in the booming digital economy", IBM Institute for Business value, 2021*

[8] *Koby Reshef, "Paving the Path to 400G Migration", www.pipelinepub.com, 2021.*

[9] *National Cable & Telecommunications Association, "The Asymmetric Nature of Internet Traffic", www.ncta.com, 2021.*

Acknowledgements

I would like to express my thanks and gratitude to Damian Poltz (SVP, Wireline Technology & Strategy) and Felipe Arroyo (Manager, Optical Networks), who provided the opportunity to undertake this project, as well as their guidance and input throughout. I would also like to express my gratitude to Lili Ti (Advisor, Communications) for her help with editing this paper.

Unleash the Power of Cloud Computing for CMTS

A Technical Paper prepared for SCTE by

John Chapman
CTO Broadband Technologies, Cisco Fellow
Cisco Systems Inc.
170 W Tasman Dr, San Joe, CA 92677
408-526-7651
jchapman@cisco.com

Tong Liu, Ph.D
Principal Engineer
Cisco Systems Inc.
300 Beaver Brook Road, BOXBOROUGH, MA 01719
978-936-1217
tonliu@cisco.com

1. Introduction

Not so long ago, the CMTS was all about the hardware, the chassis, the cable line cards with embedded CPU, memory and storage. It had long development cycles and expensive/lengthy processes for deployment and upgrade. The virtualization of the CMTS removed the need of the specialized hardware by providing the CMTS services in software running on generic servers. By doing so, it has greatly improved the efficiency of the CMTS hardware infrastructure and started the cloud native transformation of the CMTS software.

With the virtualization laying out the foundation, moving the CMTS to the cloud is a natural step forward to scale the CMTS beyond the on-premises physical servers. Cloud provides on-demand compute, memory and storage resources with high scalability, availability and security. It offers users the economies of scale and better cost efficiency with a pay-as-you-go cost model. It reduces the application development time and simplifies the deployment/upgrade processes by managing the underneath infrastructure and platform on behalf of the users.

CMTS cloudification is not just a rehosting effort, it is an optimization process that requires proper service decoupling and workload placement. The cloud computing services can be “resource-aware” or “serverless”. The resource-aware service, such as Amazon Elastic Compute Cloud (EC2), allows users to control the resource allocations and autoscaling at a relatively coarse granularity.

The serverless service, such as AWS Lambda or Amazon DynamoDB, fully manages the resources on behalf of the users with built-in fine-grained auto scaling capabilities. The geographical location of the cloud platform also matters as it impacts the network latency and the data transport cost. The distributed and inhomogeneous cloud environment presents an interesting and challenging problem for the cloudification of the CMTS.

In this paper, we explore the cloudification solutions for the CMTS. Following the introduction, we first overview the vCMTS structure in Section 2 and the public cloud infrastructure in Section 3. We then explain the cloudification motivations in Section 4 and describe the different architecture choices in Section 5. We examine the different cloud hosting options in Section 6 and explore the CMTS service placement strategies in Section 7. In section 8, we show an example for deploying the CMTS services on AWS. In Section 9, we explore how to take advantage of the serverless platform by converting the stateful DOCSIS processes into stateless functions. In Section 10, we discuss how to start the cloudification with a proof-of-concept (PoC) effort. In Section 11, we study how to deliver CMTS as a service by using CI/CD and the cloud powered multi-tenant architecture. Finally in Section 12, we conclude the paper.

To simplify the text and precisely describe the intended cloud entities, we use the Amazon Web Services (AWS) terminologies in this paper and provide links to AWS original definitions in the Reference Section. Similar terminologies and definitions can be easily found on other public cloud providers’ platforms such as Microsoft Azure, and Google Cloud platform,

2. vCMTS Architecture Overview

The vCMTS virtualizes the physical CCAP core in the DOCSIS distributed access architecture (DAA) [1] as shown in Figure 1. It is typically installed in a headend or hub that is connected to both the converged interconnect network (CIN) and the service provider (SP)’s core network. Function wise, the vCMTS contains the following service domains with distinctive timing requirements:

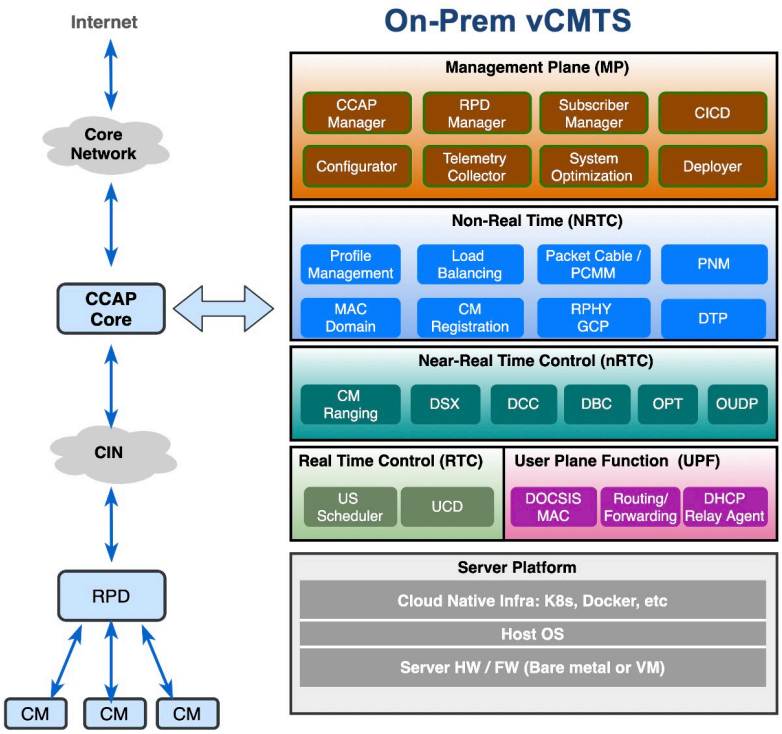


Figure 1 – vCMTS in Cable Distributed Access Architecture

Management Plane (MP)

MP contains the non-real-time (latency > 1sec) management functions to control/optimize the CMTS network elements, for deployment, configuration, health monitoring and carrying out any corrective actions if needed.

Non-Real Time Control (NRTC)

Non-real time (latency >1sec) control, NRTC, refers to the CMTS control functions that have a timing budget of 1second or above, such as CM registration, load balancing, profile management, and proactive network management etc.

Near-Real Time Control (nRTC)

Near-real time (10ms – 1sec) control, nRTC, refers to the CMTS control functions that have a timing budget of a few hundreds of milliseconds, such as DOCSIS ranging, DSX and DBC etc.

Real Time Control (RTC)

Real-time (<10ms) control, RTC, refers to the CMTS functions that have a timing budget of single-digit milliseconds, such as DOCSIS upstream scheduling and functions, UCD update transaction and MAP replications.

User Plane Functions (UPF)

> The term of UPF is borrowed from mobile for describing the data plane functions that require high-throughput and low latency, such as DOCSIS MAC frame processing, packet routing/forwarding, and DOCSIS downstream scheduling.

The latency values for RTC, nRTC, and NRTC are adapted from the Open Radio Access Network (ORAN) Alliance [2].

The vCMTS today is hosted by a dedicated server or a server cluster on premises that contains the compute, memory, storage and networking resources, and certain server/platform applications, such as Docker, Kubernetes, Radis, Kafka etc., to run the containerized CMTS microservices.

3. Public Cloud Infrastructure

Public cloud delivers virtualized computing services on-demand globally through internet and/or direct connections. Today the major public cloud platforms have practically covered all locations that have a critical mass of Internet users.

Depending on the distance relative to the end users, the public cloud infrastructure can be either in the Region or at the Edge using the AWS terminologies, defined as below.

Region

> A Region is a physical location around the world that contains multiple data centers within a geographic area, such as US East (Northern Virginia), US West (Northern California) etc. Each Region consists of multiple, isolated, and physically separate Availability Zones (AZs).
>
> Each AZ has one or more discrete data centers with redundant power, networking, and connectivity in a Region. Within the Region, an AZ is physically separated by a meaningful distance, many kilometers, from any other AZ, although all are within 100 km (60 miles) of each other.

Edge

> The Edge infrastructure provides the computing services as close to the endpoints as necessary, deployed as the cloud-managed hardware and software in locations outside the Region and even onto the customer owned devices themselves. Edge is typically used by latency sensitive or data intensive applications which would benefit from the local processing for avoiding the latency or network cost for shipping the workload to the Region.
>
> AWS edge offers multiple edge computing solutions including AWS Outposts for on-premises, AWS Local Zones for metro areas, and AWS Wavelength for the 5G Edge.

Figure 2 shows the AWS Region map and edge networks in North America [3]. Similar coverage can also be found on other public cloud platforms. For a vCMTS, both the Region and the Edge may be used for cloudification.

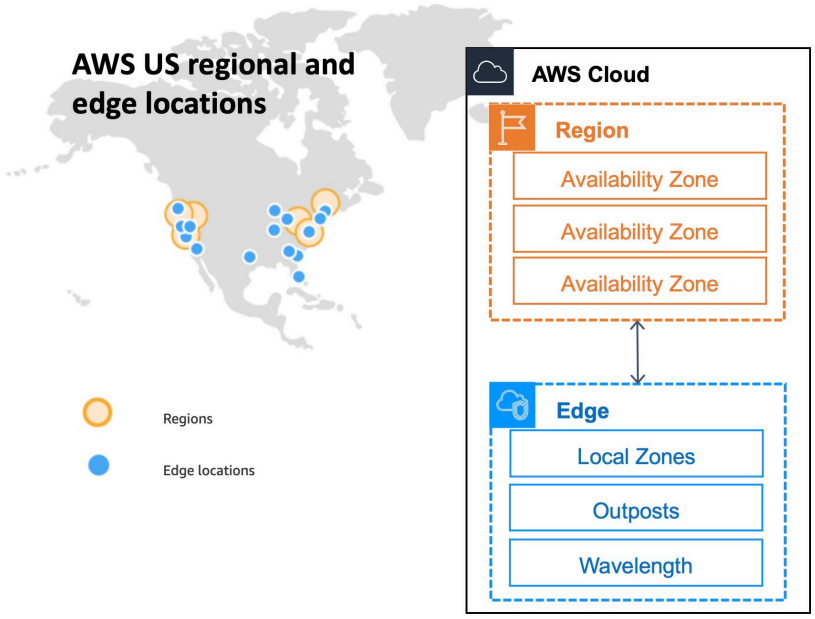


Figure 2 – AWS Infrastructure and Region Map in North America (2021)

4. Moving the vCMTS to Cloud

Just like virtualization removes the need for the specialized hardware, cloudification aims to remove the need for owning the physical servers, instead use the cloud compute services with a consumption-based cost model.

Moving to the cloud benefits both the CMTS vendors and the operators. For the CMTS vendors, cloudification simplifies the infrastructure/platform development, so they could have more time/resource innovating on their unique market differentiators. For the CMTS operators, cloudification promises global coverage, built-in high availability and rapid scaling capabilities. With on-demand cloud computing, operators have less risk for over-provisioning or under-provisioning, no need to worry about upfront costs and on-going server maintenance (provided by the cloud provider), or patching and software upgrades (provided by vendor).

The chosen public infrastructure / platforms should best suit the CMTS workloads in terms of proximity, performance and cost efficiency. Hybrid and multi-cloud strategy may be used to adapt to specific CMTS deployment needs.

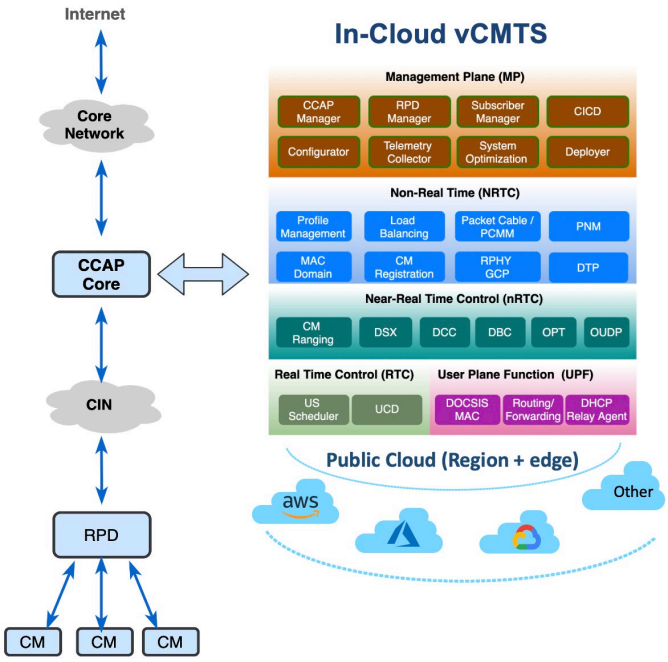


Figure 3 – vCMTS Hosted in Public Cloud

5. Cloudification Archtecture Choices

Migration to the cloud is hardly a one-step move. It involves multiple iterations to achieves the optimum efficiency. Figure 4 shows the typical architectural choices that may be used along this journey.

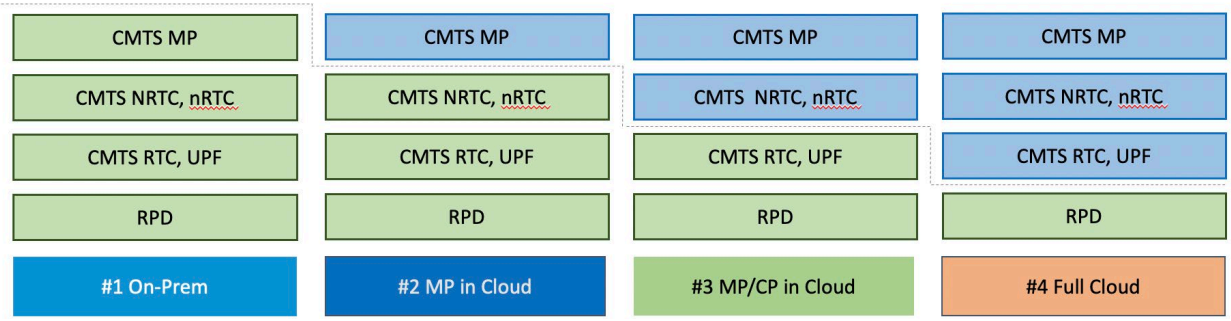


Figure 4 – CMTS Cloudification Choices

#1 On-Premises

This is the vCMTS with all the CMTS service domains hosted on premises and no cloud involved. This represents the majority of vCMTS deployments today.

#2 MP in Cloud

By simply moving the MP to the cloud, this could be a phase-one approach for cloudification. It also accomadates the FMA DAA where DOCSIS is built in the Nodes attached to the plant.

#3 MP and CP in Cloud

This is a hybrid development/deployment environment involving both cloud computing and the private on-premise compute platform. Specifically, the MP and the DOCSIS control plane (CP), including the NRTC and nRTC functions, are in the cloud, while the UPF and the RTC functions remain on-premise. This option requires that the DOCSIS control and user planes are separated with proper interface definitions.

#4 Full Cloud

The CCAP core is fully moved to the cloud. More details can be found in Section 8 where a full cloud deployment example is presented using both the Region and the Edge platforms. A Region-only full cloud setup can be found in Section 10 to test drive the cloud computing platform. This scenario applies if the cloud provider is providing Internet transit services or if the vCMTS is paired with a test harness that includes virtual DAA nodes and virtual CMs.

6. vCMTS Hosting Options

The cloud environment offers a variety of hosting options for the vCMTS services, as shown Figure 5. This includes the options of physical servers on premises, a "resource-aware" virtual server platform, or a "serverless" platform.

The vCMTS today typically uses the physical servers on premises, owned and managed by the service provider. It scales statically by the number of physical servers, with each server as a pooling unit for all resource types including compute, memory, storage and network interfaces. Once installed, the vCMTS capacity is fixed regardless the actual workloads. Provision is typically needed based on the workload in the worst case scenario, resulting in wasted resources during idle time.

The on-premises hosting option does have its advantages for handling the latency sensitive and data intensive workloads, such as UPF and RTC. Because of this, all major cloud platforms have the on-premises extensions, either as physical servers, such as AWS Outposts, or software agents/packages running on the user-owned servers.

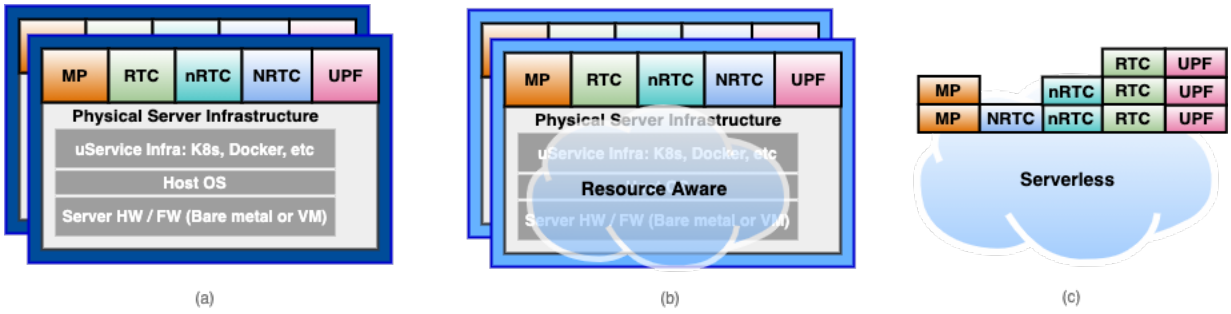


Figure 5 – vCMTS Server Platform Options

The “resource-aware” platform is the traditional cloud computing platform provided by virtual machines (VMs) overlaid on physical hardware, such as Amazon EC2 [4]. On this platform, users can configure the VM instances and organize clusters without the need to own the underneath physical server. In fact, they only rent a fraction of the cloud infrastructure to run the workload and achieve the economies of scale of sharing the infrastructure with others. Since users can control the resource type and running time, this platform is suitable for long-lasting or stateful workloads. However, there is the overhead to configure the VMs and the autoscaling is in the number of VMs, a coarse granularity comparing to the next serverless platform.

“Serverless” is a new cloud computing technology that keeps the infrastructure/platform completely hidden from the user [5]. It aims to simplify the application development/deployment processes, and enables a fine-grained, transaction-based consumption model. The serverless resources are automatically scaled on the user’s behalf as the workload changes. The serverless platform started with AWS Lambda, a function-as-a-service platform, and later expanded to include any resource-unaware backend services, such as Amazon DynamoDB for database, Amazon Simple Storage Service (Amazon S3) for binary storage, and AWS Fargate for running the containers.

The serverless platform is rapidly evolving but is comparatively less mature. For example, it is limited today in hosting stateful or latency sensitive applications, as the underneath compute resource is ephemeral and invisible to user applications. To take advantages of serverless platform, applications need to be stateless, event driven and able to tolerate the invocation delays.

7. CMTS Workload Placement Strategies

CMTS is a complex system, producing a wide variety of workloads that have diverse requirements on the infrastructure/platform, such as the distance/latency to the cable modems (CMs), processing speed, memory size, I/O performance, and CPU types etc. On the other hand, the cloud environment is distributed and inhomogeneous in nature. Having an optimum workload placement is critical to minimize cost, ensure the CMTS performance.

Figure 6 shows the mapping of the key workload requirements (y-axis) to the available cloud platform types (x-axis). It reveals the suitability for placing a CMTS workload onto a specific platform, as described below:

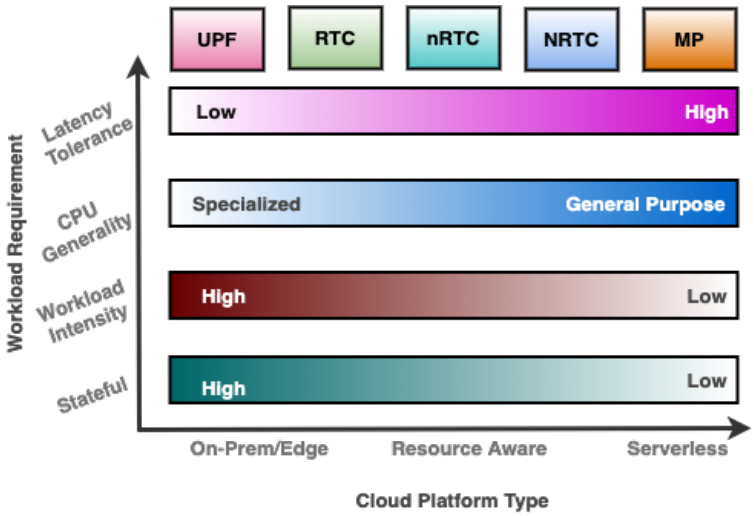


Figure 6 – CMTS Workload Placement Strategies

UPF and RTC

The UPF and RTC workloads are suitable for the on-premises/edge platform, as both types are latency sensitive and/or data intensive.

The UPF workloads are time sensitive and data intensive, up to millions packets per second per Service Group (SG). The RTC workloads are timing sensitive and compute intensive. The DOCSIS US scheduling decisions need to be made every millisecond with a processing time limited to the low hundreds of microseconds. Accurate DOCSIS timing is also required to align MAPs with the US physical layer. The RTC relies on the UPF for forwarding the real-time signaling traffic. The combined platform service access delay must not exceed the overall latency budget of low single digit of milliseconds.

nRTC

The nRTC workloads are the DOCSIS control protocols with a time budget of a few hundreds of milliseconds, including the signaling propagation delay and the processing delay. Examples of the nRTC workloads include ranging request handling, dynamic service flow addition/change/deletion (DSX), and dynamic bonding group change (DCC). The nRTC timing requirement is more relaxed than the UPF and RTC requirements, but tighter than the NRTC and MP workloads.

As an optimization strategy, the nRTC workloads can be placed in the Region to free up the Edge resources for UPF and RTC. Since the in-region serverless platform today have non-negligible invocation delay that may exceed the nRTC timing budget, the in-Region resource aware platform, such as Amazon EC2, may suitable for the nRTC workloads.

NRTC and MP

The NRTC workloads, such as CM registration, MP workloads, and as SG configuration, are typically infrequent and latency tolerant (minimum one second). Occasional usage spikes may occur, for example, during a planned upgrade/maintenance window or upon an unexpected power outage. In which case, the NRTC may experience a CM registration storm with potentially thousands of CMs trying to register at the same time.

These characteristics match well with the serverless platform, where the NRTC and MP would benefit from the built-in elasticity and the economics of the fine-grained cost model. During the normal operation time, the serverless platform may not incur any cost as the NRTC or MP workloads are very low and may well be within the free tier of the serverless services. During the rare event like the CM registration storm, the serverless platform can rapidly scale up to meet the demand and scale down once as storm fades away.

In summary, the CMTS workload placement is an optimization process to minimize cost within the boundary conditions for the CMTS to perform at a given scale.

8. Full Cloud CMTS Deployment Model

A full cloud CMTS may be deployed on both the edge and the Region for an optimum workload placement. Figure 7 shows a full cloud CMTS deployment example on the AWS platform, where the UPF and RTC functions are placed at the edge and the nRTC, NRTC, and MP functions are placed in the Region.

The AWS virtual private cloud (VPC) service is used to form a virtual network to logically isolate the edge and Region resources for the CMTS. The CMTS VPC is connected to the on-premises network via an AWS Local Gateway, which further connects to the CIN and the SP core network via a SP router (SPR). This setup allows the DOCSIS upstream and downstream data traffic to traverse the same path as in the vCMTS today. For the CMTS control and management traffic, a dedicated connection known as AWS Direct Connect, is used between the Region and the edge to reduce the transport latency and jitter.

Within the Region, multiple AZs are used to add geo-redundancies for the control and management plane. The nRTC is placed on the resource-aware platform meet to the near-real time latency requirement. With explicit configurations, auto-scaling can be enabled for the nRTC across multiple AZs. The NRTC and MP can be placed on the serverless platform, thanks to the extra timing budget, which has built-in auto-scaling capability and better cost efficiency with its fine-grained, transaction-based cost model.

To support the user plane and control/management plane separation, the following interfaces/protocols are used to facilitate the communication between the CMTS service endpoints across the Edge and the Region. Each CMTS endpoint is assigned with an IP address from its hosting subnet reachable by other endpoints through the border gateway protocol (BGP).

Punt Path:	For the US UPF to punt the upstream DOCSIS signaling to RTC, nRTC or NRTC via TCP or UDP, such as CM Bandwidth Requests to RTC, Ranging Requests to nRTC, and Registration Requests to NRTC.
Inject Path:	For the CMTS control plane to inject the DOCSIS signaling messages to the DS UPF via TCP or UDP, such as MAP and UCD messages from RTC, Ranging Response messages from nRTC, and Registration Responses from NRTC.
Rest API:	For MP to configure/manage the CMTS control plane and user plane services. The REST APIs can be carried over HTTP and leverage existing interface definitions based on the DOCSIS YANG data models.
Telemetry Streaming	For MP to monitor the CMTS control plane and user plane operations. The telemetry data can be carried in GRPC and using the existing DOCSIS model driven telemetry interface definitions.

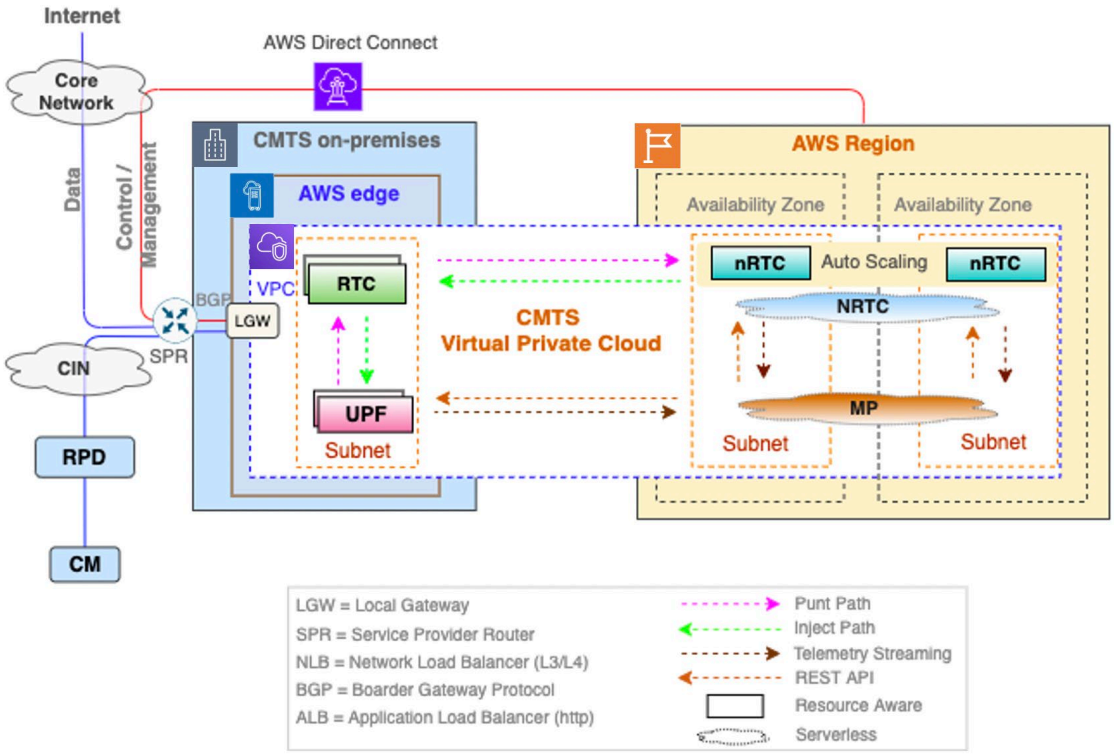


Figure 7 – Full Cloud CMTS Deployment Example on AWS

9. Serverless DOCSIS State Handling

Given the simplicity and economics appeal of the serverless platform, we are interested to explore how it can be used to support DOCSIS which is known for being deeply stateful. The DOCSIS control protocol is based on states of several DOCSIS entities, such as service groups (SGs), cable modems (CMs) and service flows (SFs). DOCSIS applications are multi-step procedures, such as profile management, load balancing and channel resource management for resiliency. The cloud serverless platform on the other hand has no affinity to the underlying compute infrastructure, and it is not possible to hold states across invocations. To provide serverless DOCSIS services, the code needs to be written in a stateless style.

Given the serverless technologies available today, the following two methods can be used for DOCSIS state handling,

- using the serverless data stores to persist the states externally, and
- using the serverless workflows to orchestrate the stateless functions.

This stateless approach to programming is also required to implement horizontal scalability which provide elasticity and thus better economics with resource-aware platforms.

9.1. Using the Serverless Data Stores

This is a key technique in the cloud native design that uses the external datastore to persist the states across function invocations. States are externalized as events passing between the functions through the data store while the functions themselves remain stateless. Examples of the datastores on a serverless

platform include Amazon DynamoDB, an NoSQL database with single-digit millisecond access time or microsecond access time with cache accelerations, and Amazon S3 for large binary object storge.

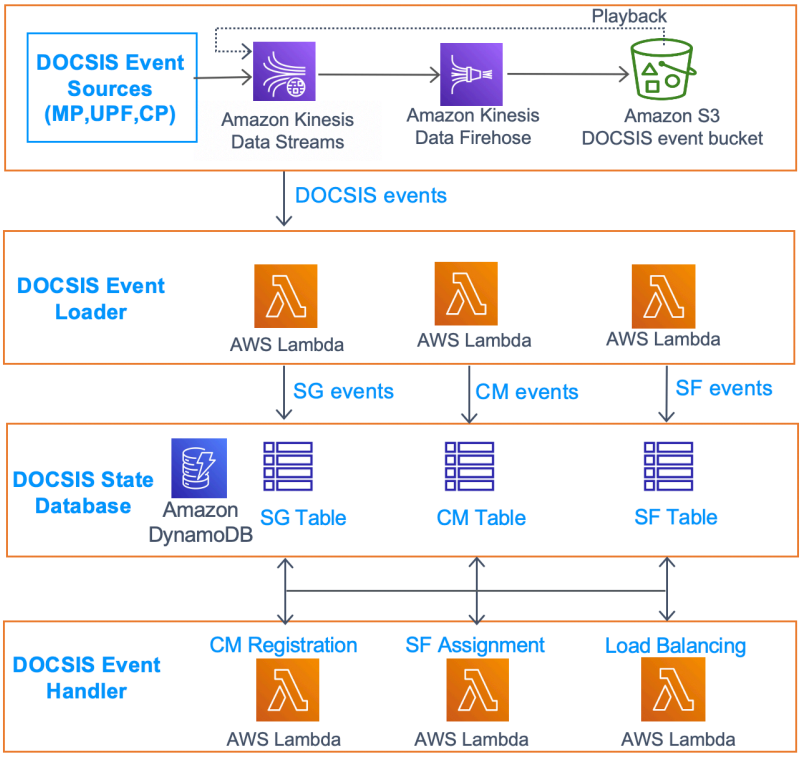


Figure 8 – DOCSIS State Handling Using Stateless Datastore

Figure 8 shows an example for using the serverless data stores to handle states of the non-RTC workloads. It uses an AWS serverless platform that contains AWS Lambda for serverless compute, Amazon DynamoDB for serverless database, Amazon Kinesis for serverless streaming, and Amazon S3 for serverless storage.

The design is based on an event driven architecture where the state information is passed as events between the stateless functions. The state changes of the DOCSIS SG, CM, and SF entities are represented as events from the management, user plane and control plane. These events are pushed into Amazon Kinesis, a data streaming service that keeps the event sequences in different shards or service groups.

These events are then pulled by the DOCSIS event loaders (implemented as AWS Lambda functions) and sorted into the corresponding Amazon DynamoDB tables. The loaded events will then trigger the DOCSIS Event Handler (also implemented as AWS Lambda functions) that watch the tables to process the state changes.

As an optimization, multiple handlers can be triggered for a single event. For example, a CM REG-REQ arrival event will trigger the CM Registration handler to process the REG-REQ, the Load Balancing handler to process the CM's load balancing group configuration, and the DS/US resiliency function to provide the channel list for the CM's transmit and receive channel sets.

In this example, there is also an Amazon S3 bucket to store the event history in longer term, which serves as the single source of truth that can be used to derive the DOCSIS states for debugging, fault recovery and AI/ML analytics.

In summary, by converting states into events, a stateful DOCSIS microservice can be disaggregated into a set of serverless functions. The fine-grained stateless functions then allow the serverless platform to quickly adapt the computational resource allocation to the shifting workload requirements, therefor achieving higher resource efficiency and lower cost comparing to the resource-aware platform.

9.2. Serverless Workflows

The DOCSIS application is typically a higher-level construct that contains various functions working in coordination with each other to accomplish the desired tasks. The execution order of the functions reflects the application states and the state transitions as the dependency or precedence of the execution of one function to another.

On the cloud serverless platform, there is a workflow service that can be used to record the function execution sequence, such that individual functions do not need to know about the application states. Additionally, since the workflow captures the prior knowledge of the execution order at runtime, it can be used for the platform level optimizations, for example, by warming up functions, pre-fetching data and instructions to achieve better performance.

Figure 9 is an example of using the AWS step function to implement DOCSIS profile management. AWS Step Functions is a serverless workflow service that can be integrated with other AWS services, such as AWS Lambda, Amazon DynamoDB, Amazon Kinesis, and Amazon SNS to accomplish a complicated procedure with many logical steps.

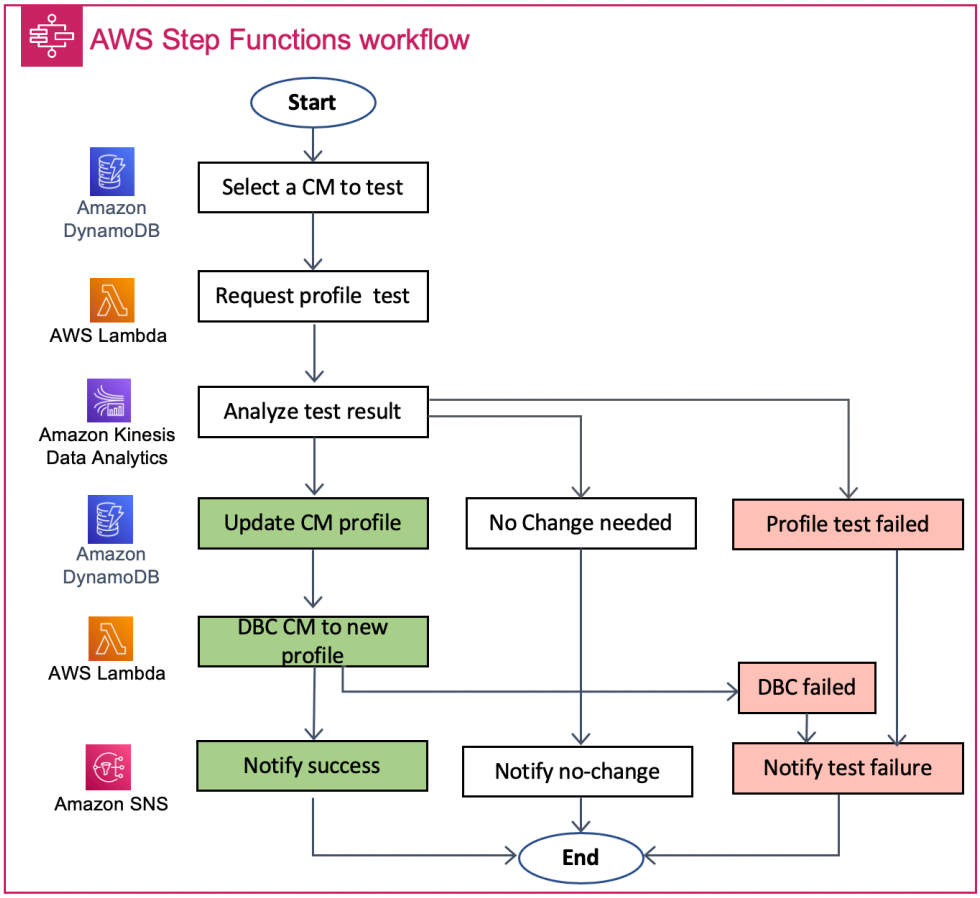


Figure 9 – DOCSIS Profile Managment Using Stateless Workflows

10. Getting Started

Moving to cloud is an interesting and challenging paradigm shift. It requires deep understanding of both the cloud services and the CMTS resource usage characteristics to maximize the cloud benefits. Since it may take years to mature to that level of understanding, cloudification is an iterative optimization process that requires learning while doing.

Given the on-prem CMTS has already been virtualized, a lift and shift strategy to the cloud is feasible to jump start the cloudification by moving a copy of an existing vCMTS and data to public cloud with minimal or no redesigning or modifications. This essentially is a test-driven development model to identify issues caused by the cloud hosting environment, and resource consumption hotspots for targeted redesign and optimizations.

This cloudification strategy motivated us to start a proof of concept (PoC) project using AWS as shown in Figure 10. In this setup, the vCMTS, the Cisco cnBR and the cnBR manager, are placed in an AWS Region that covers the on-premises lab location where the RPD and CMs are located. The cnBR in the Region is connected to the lab's router via an AWS Direct Connect. The PoC is used evaluate all vCMTS workloads in the cloud environment. Finding of this study will help refine the cloudification strategy and determine the architecture choice.

The in-region full cloud solution as used in this PoC can also be used as a convenient testing/development environment, which can be quickly created on-demand without the expansive lab expansions.

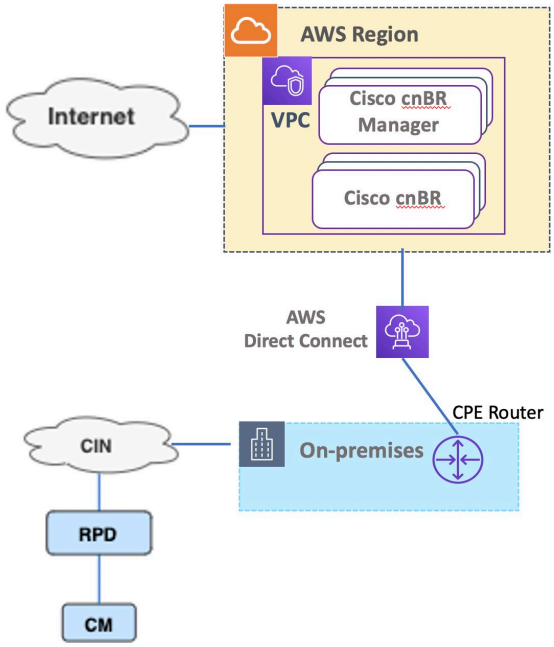


Figure 10 – Poof of Concept for CMTS Cloudification

11. CMTS-as-a-Service

With cloudification, the CMTS can be delivered as a service to the cable operators through the Internet or the direct connect services offered by the cloud providers. To make the CMTS-as-a-Service successful and sustainable, multi-tenancy is the key to bring down the CMTS development and maintenance cost.

In its basic definition, the CMTS multi-tenancy is an architecture in which a single CMTS entity serves multiple cable operators. Such entity can be the continuous integration and continuous delivery (CI/CD) pipeline that produces the production software and various cloud resources required to execute the CMTS functions.

11.1. CI/CD Simplification

CI/CD can be pictured as a pipeline, where new code is submitted on one end, tested over a series of stagers (code, build, test, release, deploy) and then published as operation-ready code. In the traditional single tenant environment, the complexity of the CI/CD pipeline goes linearly with the number of tenants. The CMTS vendor needs to maintain separate pipeline for each operator, slowing down the CI/CD life cycle, and wasting development /deployment resources.

With multi-tenancy, only one CI/CD pipeline is required as shown in. Figure 11, which can dramatically reduce the CI/CD complexity, improve quality and accelerate innovations, as all tenants would use the same code, the same operational infrastructure, and the sane deployment procedures.

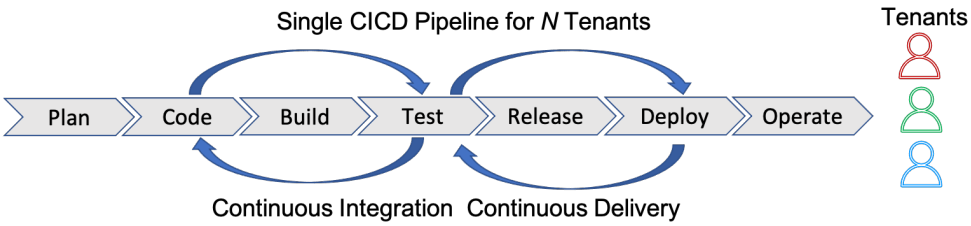


Figure 11 – CI/CD Simplification With Multi-tenancy

11.2. Multi-tenant CMTS Resource Sharing/Isolation

A multi-tenant CMTS achieves the cost and operational efficiency by sharing the same CMTS software and the cloud infrastructure among multiple operators, while keeping the tenants from accessing another tenant's resources. It boils down to a constrained optimization issue to properly partition the CMTS resources among the tenants either statically and/or dynamically.

There are a wide range of factors to consider for the resource partitioning across different tenants [6]. For example, a big operator who has large and constant workloads may be suitable for dedicated resources, a small operator on the other hand may find it more feasible to use the multi-tenant pooled resources. The workload characteristics also impact the partition strategy. For example, it's more economical to use the pooled resources for the DSX transactions, as they are only triggered occasionally. The goal of the multi-tenant CMTS is to efficiently serve everyone and automatically scale up and down based on current requirements demanded by the tenants.

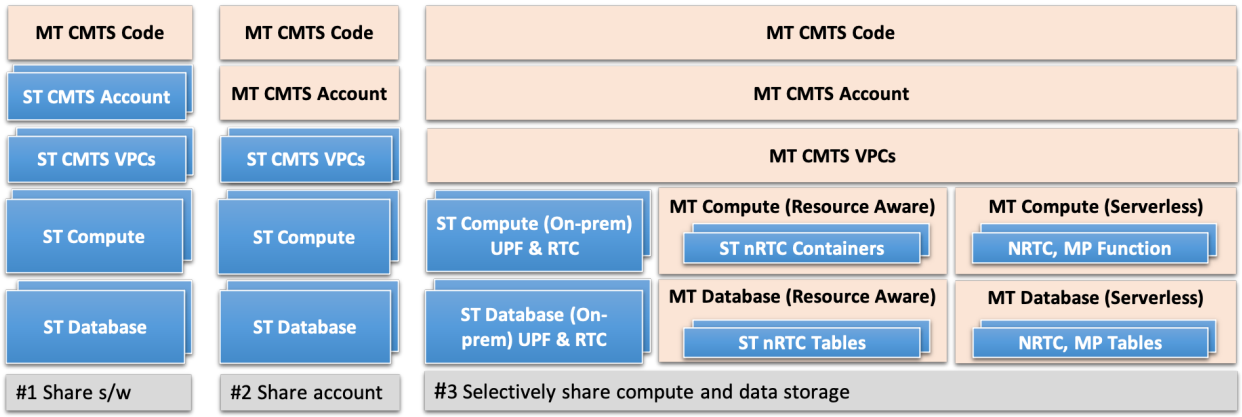


Figure 12 – Multi-tenant CMTS Resource Sharing and Isolation Options

Figure 12 shows three different resource partition options of the tenant stack including CMTS software, CMTS account, the VPC network construct, and the compute/data storage resources. These options may be applied in any combination to fit various tenant requirements.

The first option is the most isolated approach, where tenants only share the multi-tenant (MT) CMTS software. A single-tenant (ST) account is assigned to each tenant to keep all the moving parts of the tenant stack in complete isolation.

The second option pushes the isolation to the next level with all tenants running the same code in the CMTS provider's account. One or more unique ST CMTS VPCs are assigned to each tenant to enforce the isolation with the network boundaries. VPC based isolation avoids the certain account provisioning

restrictions and gives the CMTS provider more centralized control across all tenants' deployment and operation. However, this option also has the same scaling issue as the first option, as the number of VPC grows, the management agility and resource efficiency decrease.

The third option has much more granular resource partitions among the tenants. It offers the best efficiency, agility and cost benefits among the three options. However, it requires a much more comprehensive tenant isolation strategy. Since operators are only naturally separated by their headend/hub locations, the on-premises UPF and RTC functions are completely isolated in compute infrastructure and data stores. For the CMTS services hosted in the Region, tenants are isolated by containers on the resource aware compute platform, and by functions on the serverless compute platform.

For the data stores, tenants can be isolated by tables or even table entries via the tenant identifier. Multiple cloud utilities are available to simplify tenant isolations. For example, using AWS Identity and Access Management (IAM) to implement run-time, policy-based isolations, using Amazon Elastic Kubernetes Service (Amazon EKS) namespace to isolate the ST container groups, and using the Amazon DynamoDB partition key to isolate ST tables.

12. Conclusions

The decision for moving the CMTS to the cloud is obvious: you pay for what you use, scale when you need, and spend less time managing the infrastructure/platform, and end up with more time to innovate for your business.

Cloud is different from the on-premises server environment with its distributed Region and edge locations and various resource-aware and serverless compute/storage platforms. Moving the CMTS to the cloud cannot be a simple rehosting process. Instead, it requires iterative optimizations to make most of the cloud environment.

With cloudification, the CMTS can be delivered as a service, enabled by CI/CD and the cloud based multi-tenant architecture. This opens new business opportunities for both CMTS vendors and cable operators improving business agility, productivity and cost efficiencies.

With the virtualization laying out the foundation, the cloudification of the CMTS can start with a lift-and-shift followed by a test-driven development model for targeted redesign and optimizations.

It is the time to unleash the power of the cloud computing for the CMTS.

Abbreviations

API	Application programming interface
AWS	Amazon Web Services
BGP	Boarder Gateway Protocol
CM	Cable modem
CCAP	Converged cable access platform
CIN	Converged interconnect network
CI/CD	Continuous integration, continuous delivery
CMTS	Cable modem termination system
CP	Control Plane
DS	Downstream
DSX	Dynamic service addition/change/deletion
FMA	Flexible MAC architecture
GCP	Generic control protocol
gRPC	Google RPC
MAP	DOCSIS bandwidth request
MP	Management Plane
MT	Multi-tenant
nRTC	Near-real time control plane
NRTC	Non-real time control plane
RPD	remote PHY device
RTC	real-time control plane
SG	Service Group
SF	Service Flow
SP	Service Provider
ST	Single tenant
UPF	User plan function
vCMTS	Virtualized CMTS
YANG	yet another next gen (data modeling language)

Bibliography & References

[1] https://www.o-ran.org/
[2] “CM-SP-R-PHY-I14-200323: DOCSIS Remote PHY Specification”, CableLabs, 2020
[3] https://aws.amazon.com/about-aws/global-infrastructure/regions_az/
[4] https://aws.amazon.com/ec2/
[5] https://aws.amazon.com/serverless/
[6] https://aws.amazon.com/partners/programs/saas-factory/tenant-isolation/

Up Your Uptime With Automation

A Technical Paper prepared for SCTE by

Nancy McGuire
Executive Director – Reliability Engineering, Operational Support
Comcast Cable
1800 Arch street, Philadelphia, PA 10103
(215) 286-2290
Nancy_McGuire@cable.comcast.com

Kathy Fox
Vice President – Product Management, XOC
Comcast Cable
609 Odin Rd, Coudersport, PA 16915
(856) 912-7850
Kathy_Fox@cable.comcast.com

1. Introduction

This is being written in the summer of 2021, a year-plus since the global pandemic that sent many of us home to work, and created an entire category of 2020 memories – "my/our COVID project" – ranging from sourdough bread to bathroom remodels to fitness journeys. If there was such a thing as an unofficial, job-related COVID project, automation would assuredly qualify! From plant maintenance to residential product support, circumstances arose over the past year that catalyzed us to align processes and achieve automation.

Maybe this feels familiar to you: Volumes upon volumes of tickets. Late night/early morning pages and dispatches. On-call team member burnout. All are symptoms of large, complex systems, built and patched and rebuilt over decades, and probably showing multiple signs of wear. You, of course, "just" want to ensure you know what is going on, where, and when.

But how do you sort what is truly an impacting event that requires immediate attention, from the what can be a daunting level of "informational noise"? How do you get the right eyes on the right problem at the right time?

The answer that worked for us was a duet of process alignment and automation. In this paper, we will review some of the automation opportunities we identified, how we identified them, and the impact they had on our teams and business operations functions. We will review two different approaches to address these opportunities – informational noise reduction, and optimization by correlation -- in two quite different areas of our operation: first, by our field plant maintenance technicians, supported by our XOCs (eXcellence in Operation Centers), and second, in our national Residential Product Support team.

2. Getting Started – Process Alignment

In consolidating and post-consolidating environments, a necessary course of work is to find "sameness," for purposes of scale. For our national and field teams, we have been focusing on sameness across all functions. However, overall "sameness" only delivers that initial goal of industry consolidation and geographic clustering -- scale. As important as scale, and the next step after achieving it, is internal synchronicity. We define that here as "the best of best practices, applied at scale," across the consolidated organization. Internal synchronicity optimizes how things are done, and it does so at scale. It's why we're on a quest, partly characterized in this paper, to identify best practices/processes/ solutions and implement them everywhere.

As we identified areas of differences, an initial area of activity we believed could serve large benefits to customers and employees was identifying and implementing best practices around node outage treatments. When we dispatch plant maintenance technicians for node outages during the day, we are pulling them off other work that is in their queue. When we dispatch plant maintenance technicians after hours, we incur overtime, and when there is enough on-call activity, those technicians can burn out. It is therefore important that we dispatch them only to nodes where they can take action to fix problems. It is about getting the right technician to the right place at the right time. It is worth noting, plant maintenance technicians are the more expensive technicians, with the pricier bucket trucks, as compared to fulfillment technicians. And we all know that when dispatched at night, technicians deserve for it to be a legitimate issue.

We needed to determine what steps we could take to reduce the number of times plant maintenance trucks get dispatched, only to find that the node was out because the power was out, or some other event the technician was not able to repair – the classic "NTF" (no trouble found). We have three different divisions and as we looked at node outage practices, we found different processes for each of the divisions.

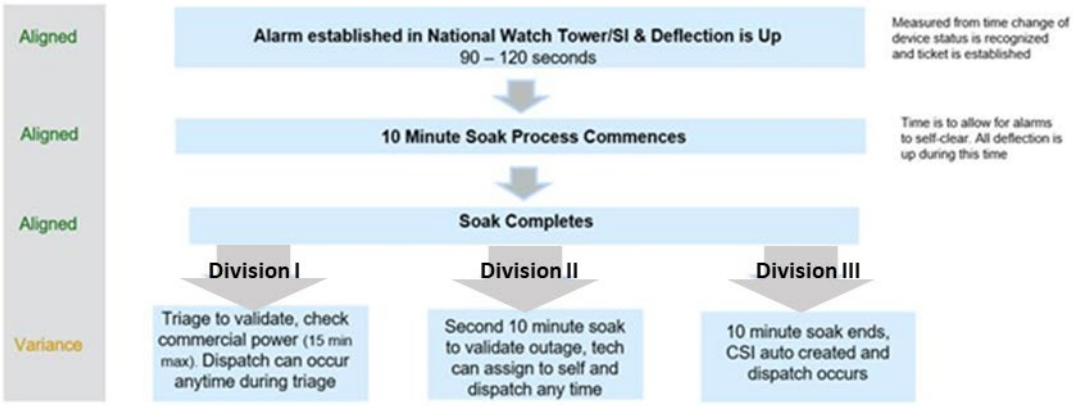


Figure 1- Traditional Division "Node Soak" Process

As we see in Figure 1, above, there was alignment in three areas, but variance existed once a node completed the initial 10-minute "soak," a term we use to indicate the period of time during which node components self-clear any alarms. It is important to note that nodes go into soak when a certain percentage of modems served by that node go through a registration state change (to offline).

The division teams came together to review their processes and resulting data, comparing apples to apples, to see who had the most efficient plant maintenance dispatch records, based on total dispatches, fix codes, and NTFs. Perhaps not surprisingly, folks can have a good bit of passion for their individual tools and/or processes. So, after a notable amount of review of data and discussion of the data, the node outage process utilized by Division I was identified as the best solution for the customer, the employees, and the company. We referred to the activity as "outage pre-verify," because the XOC was verifying that the outage exists before dispatching staff. Once we had a process name and agreement, the team then focused on developing a thorough implementation plan which included detailed training and communication plans; a review of the LOQs (Lines of Questioning) used by the Division I team during the triage activities; determining and preparing for expected HFC (Hybrid Fiber-Coax) desk task increases because of the triage activity; and expected truck roll decreases related to plant maintenance. Once we had developed this detailed plan, we launched it on a region-by-region basis across divisions 2 and 3. In some areas, we launched more quickly than originally planned, because it was such a good story, from a plant maintenance perspective.

3. The Results of Process Alignment

How has the Outage Pre-Verify work been received by the two other division field teams? Charles Detwiler, a Comcast Maintenance Supervisor in the Pittsburgh area summed it up this way:

"Outage Pre-Verify has reduced overtime and man hours drastically. HFC techs are putting information in tickets, so we now have a better understanding of what we are going to be troubleshooting and there is also much more trust that this is a legitimate outage. We just had a regional call and reviewed the significant truck roll reductions since rolling out Outage Pre-Verify. Lastly, our techs are less stressed because we are rolling fewer trucks at night, so we are getting more rest."

Steven Musembi, VP (Vice President) of XOC for Comcast's Central Division credits the introduction of Outage Pre-verify as *"a direct contributor to node performance in the Central division. OPV (Outage Pre-Verify) freed up labor via transaction reduction to help find and address plant issues, improving the customer experience for all of our customers."*

Issues that are detected via our tools and rule sets, which become tasks for our HFC technicians, are referred to as Dashboard Tasks. Whenever there are node issues that hit our thresholds, that creates a dashboard task. The numbers speak volumes. Figures 2 and 3, below, show the total number of dashboard tasks and the number that were cancelled as a result of XOC outage pre-verify work. The overall result was reduction of all dashboard tasks by nearly a third (29.32% and 28.5%.) Internally, we were also able to break these out by daytime and after-hours Plant Maintenance dispatches, to show the overtime savings. Again, these reductions happened and continue to happen because the HFC desk technicians are following through on the same set of OPV (Outage Pre-Verify) LOQs (Lines of Questioning) to properly triage a node outage.

OPV Details - Division 2 Fiscal June 2021			
Region	**Total Dashboard Tasks**	**# Cancelled Dashboard Tasks**	**% Canceled Tasks**
1	14765	3847	26.05%
2	11679	3620	31.00%
3	8105	1624	20.04%
4	7807	3326	42.60%
Totals	**42356**	**12417**	**29.32%**

Figure 2- Division 2 Dashboard Tasks canceled based on triage

OPV Details - Division 3 Fiscal June 2021			
Region	**Total Dashboard Tasks**	**# Cancelled Dashboard Tasks**	**% Canceled Tasks**
1	7487	1974	26.30%
2	5582	1962	35.15%
3	3843	978	25.45%
4	4241	1193	28.13%
5	2134	532	24.93%
Totals	**23287**	**6639**	**28.51%**

Figure 3 – Division 3 Dashboard Tasks Canceled based on triage

Daytime Savings – Division II

Region	# of Dash Tasks between 8 am and 5 pm	# Canceled WT Tasks between 8 am and 5 pm	% of Canceled WT Tasks Relative to the # of Dash Tasks (8 am to 5 pm)	Total Tme Savings in Hours (Tasks 8am-5pm *1.5 hours)
1	5,649	1,451	25.69%	2,176.5
2	4,610	1,357	29.44%	2,035.5
3	3,850	820	21.30%	1,230
4	3,387	1,408	41.57%	2,112
Total	17,496	5,036	28.78%	7,554

Figure 4 – Breakdown of One Division's Plant Maintenance Cancelations – Daytime hours

Daytime Savings – Division III

Region	# of Dash Tasks between 8 am and 5 pm	# Canceled WT Tasks between 8 am and 5 pm	% of Canceled WT Tasks Relative to the # of Dash Tasks (8 am to 5 pm)	Total Tme Savings in Hours (Tasks 8am-5pm *1.5 hours)
1	2,924	852	29.14%	1,278
2	2,343	888	37.90%	1,332
3	1,672	507	30.32%	760.5
4	1,661	563	33.90%	844.5
5	911	239	26.23%	358.5
Total	9,511	3,049	32.06%	4,573.5

Figure 5 – Breakdown of Another Division's Plant Maintenance Cancelations – Daytime hours

4. Additional "Sameness" Opportunities

The Division I team also shared its operating practice for MSOs (Multi-Soaking Outages). A multi-soaking outage is the term used to describe a node that drops into soak, which as a reminder occurs when a certain percentage of modems associated with that node experience a registration state change, indicating that "something is up" at the node. What's different in a multi-soak outage is that the node self-clears before the 10-minute soak time elapses, but alarms and re-soaks itself again, in short order. The division's process was to triage and dispatch help to nodes that dropped into soak, and then cleared, four or more times in any rolling 24-hour period. Because we were going to be reducing plant maintenance truck rolls with our Outage Pre-Verify work, it made sense to also implement this same practice across all three divisions. We knew that implementing the MSO functionality would drive additional HFC technician tasks as well as additional plant maintenance-related truck rolls. However, the HFC leadership teams were gaining efficiencies by reducing tasks in other areas, and the Plant Maintenance teams were very capable of taking on this work, given the significant reductions in truck rolls as a direct result of the Outage Pre-verify work.

Figure 6 shows that our plant maintenance teams are, on average, finding and correcting node multi-soak issues more than half (57.5%) of the time. Again, these are intermittent outages that we are correcting before they become a larger outage – a big lift for the customer experience.

Division II	Region	Tasks	% Region
	1	137	54.80%
	2	102	72.30%
	3	160	57.60%
	4	77	72.00%
Division Total		**476**	**61.30%**
Division III	**Region**		
	1	177	48.10%
	2	136	58.10%
	3	145	63.60%
	4	132	54.30%
Division Total		590	57.50%

Figure 6 – Multi-Soaking Outage (MSO) Repair Results

With the solid progress being made for MSO with the setting of 4 times in a rolling 24 hours, we are poised for the divisions to move to 3 times within 24 hours, to drive even more reliability into the infrastructure – which automatically means fewer outages for our customers! One division has already moved to 3 times in 24 hours, while our two other divisions preparing to make the move as well.

5. Align the Process, Then Automate – One Way

Once the teams aligned on the process, we needed to lighten the load for our HFC technicians, who had taken on a great deal of additional triage activities. And now that we were all on the same process, we could build consistent automation. All those canceled jobs were canceled because of the excellent work performed by these HFC technicians – and we did not add headcount to our HFC desks to support the outage pre-verify or the MSO work. Together, the division and national team developed a user story to capture a description of a software feature from an end-user perspective. This allowed us to to reduce "eyes on glass" time for the HFC desk technicians. Considerations for the user story included:

- Utilizing automation and other tool inputs to validate node outage events and reduce pre-verify questions, using already-available data for:
    - Verification of device registration state
    - Confirmation of no upstream activity – flatline node
    - Validation of power supply status
    - Determining "storm mode" status
    - Verification of headend or CMTS (Cable Modem Termination Server) alarm
- Saving an estimated 3 minutes per pre-verify task (on average, a pre-verify task takes 7 minutes prior to the automation), by reducing the number of manual actions needed for task completion
- Saving an estimated 30K+ hours per year at the XOCs, on eliminated workflows.

This automation work is in progress for implementation in 2021. Future automation enhancements will include checks of Xfinity Home devices served by a node, to determine connectivity type (backup / cellular), to further the accuracy of automated checks related to commercial power availability.

6. How Covid Impacted Automation Activities

When you are supporting the flagship products for a major cable/broadband company, it is imperative that they work at or better than the expected 99.999% of the time. This certainly made 2020 a year to remember! The pandemic required teams and companies to shift perspectives, cultures and make some really challenging decisions. Employees were having difficulty with the lack of in-person social engagement, hallway conversations and the biggest issue, the blurring boundary between work/life hours.

7. Necessity: The Mother of…*Automation!*

We had to find those solutions fast! But first we had to identify what could be automated. We looked at redundant tasks such as pulling logs from devices, ticket entry, ticket assignment and triage. Previously, our offshore team did the ticket creation, while another offshore team did the initial triage. Outsourced resources also played a major part in "air traffic control" - assigning tickets to the correct teams for mitigation, as well as escalating to on-call resources when needed. Fortunately, our teams were slightly ahead of the curve here and had been doing the data analysis prior to this occurring. This event just fast-tracked the groundwork we had already laid out.

Having a head start in finding and implementing automation drastically reduced our fears about team burnout and not being able to support our products at the same level. While there were obvious candidates, we were still faced with the challenge of *how* we find automation opportunities.

The undisputed answer: data analysis and best practices.

Specifically, we built reports to understand what our biggest drivers were, and what could we do to streamline our workflows. We also looked for "low hanging fruit," or, the quick wins. We engaged with our engineering teams to dive into what was important and what was lower priority that could wait until business hours. We worked to create a portal so that our internal customers could create tickets on demand and escalate if needed. We worked on auto-assignment, auto paging and correlation. We developed "soak" periods for alarms and alerts, to give electronics a grace period to reset themselves. Not only did we work to make our processes and systems more efficient and user-friendly, but we also worked to make them smarter.

We also needed to understand who our internal customers were and what needs they had. Any solution we developed needed to be user-friendly and convenient. What existing tools could we share with them to help them do their jobs, and potentially prevent the escalation in the first place? What opportunities are "hiding in plain sight," for anomaly detection and/or reductions in MTTM (mean time to mitigate)?

8. How We Automated It

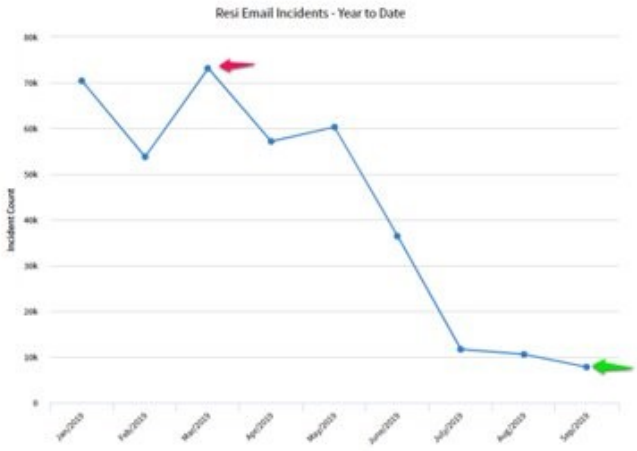
In March of 2019, we started taking a hard look at the volume coming into the Production support team. We broke it down by the primary service, looking at each individually. We ran ticket reports to understand volume, trends, and commonalities. Our highest volume related to an email platform. In March of 2019, there were over 74,000 tickets created for this platform alone. Once we did the analysis, we discovered that a majority were auto-resolving, when the main event resolved. In other words, we

were generating tickets for events where they were not the primary issue, but rather an impacted service. It was informational noise.

One of the other findings was that, the outsourced team was largely serving as a pass-through for tickets. An alarm or alert would trigger and be assigned to the Tier 1 queue. The Tier 1 team would respond by reassigning the ticket to a Tier 2 team. Ahem: If no investigation or triage is happening, then why is a ticket entering any queue?

These top-two findings guided our actions and next steps. Our first action was to correlate events with primary drivers. As an example, if one VM (Virtual Machine) went into an alarm start and there were 50 affected applications on that VM (Virtual Machine), we will generate 51 tickets. In the updated noise reduction effort, we now have 1 ticket that has 50 events appended. This reduces volume while still tracking impact. It significantly reduced overall volume and allowed us to find and trend the truly impactful events before they became major incidents.

Based on our analysis of the data, we were able to identify and define this correlation solution, which we presented to our engineering partners. We wrote out exactly what problem we wanted to solve, and what our desired outcome would be. Once we had these defined, we were able to figure out the actions needed to get from the problem statement to the solution. We then presented this to our engineering teams to kick off our efforts and garner buy-in. During this collaboration process, we also developed a deeper understanding of the platforms we support, strengthened our relationship with the engineering partners and most importantly, provided results. This also allowed the engineering teams to focus on the stability of the product rather than sort through a high volume of informational noise.



Created	Incident Count	Percentage of Incidents
Jan/2019	70,429	18.47%
Feb/2019	53,794	14.11%
Mar/2019	73,147	19.18%
Apr/2019	57,165	14.99%
May/2019	60,299	15.81%
June/2019	36,449	9.56%
July/2019	11,676	3.06%
Aug/2019	10,582	2.78%
Sep/2019	7,792	2.04%
Total	381,333	100%

Figure 7- volume reduction via correlation

The impact of this implementation also had other benefits. By limiting what alerted to the affected service, application, or infrastructure, the mean-time-to-mitigate was reduced and customer NPS (Net Promoter Score) saw a rise in positive feedback.

"I lead an engineering and operations team for enterprise-scale complex applications. Large events would often lead to a storm of alerts making root cause analysis slow. Nancy's team came in and implemented automatic correlation that minimized the noise to just the important alerts. This has significantly reduced our MTTM (mean time to mitigate), which improves our customer NPS."
– Gabriel Satterlee - Director of Engineering Operations

This process was the beta, or proof-of-concept (PoC), and was proven to work. Based on these results, we began to push correlation activities out to our other supported products and applications. When we trended the data, post implementation, we noted that there was significant reduction in both MTTT (mean-time-to-triage) and MTTM (mean-time-to-mitigate).

Part of the overall reduction effort involved reviewing the alerts we had in place. By reviewing these alerts, we were able to understand what was actionable, as opposed to a warning or for awareness. We developed a better understanding about how to adjust thresholds where it made

sense, and anticipate what related events could be correlated. We were effectively able to reduce the noise and focus on hidden problem areas via deep dive investigations.

Most of the automation efforts involved escalation reduction, by looking at three factors:

> *1- Are these escalations actionable or not? If not, we can adjust thresholds to ensure actionable next time or remove them entirely.*
>
> *2- Reduction of duplicate escalations/tickets: Do we need 5 tickets for the same issue? Typically, that answer is no. Our approach was to merge similar alerts into the same ticket to reduce toil and escalate a single time for a single issue.*
>
> *3 – Using problem management to focus and deep-dive on the top "repeat offenders." This helped to prioritize Engineering/Developer teams around specific issues to improve customer experience and reduce load on our teams.*
>
> *These three factors were key in both deprecating our offshore resources and improving quality of life (work/life balance) for our teams.*
>
> ***~Brian Seeley, Sr. Manager Production Support***

As this process unfolded, we realized that we were able to automate a lot of the manual work – meaning that our Tier 1 , serving as a catch-and-dispatch team, was not required to support the internal product team, because we were directly assigning to the responsible resolver teams. Those that were not directly assigned to an engineering team went directly to our Tier 2 Triage desk for a more technical review. In some instances, we were even able to move resolution steps all the way to Tier 1 customer Care and, eventually, a customer-facing application, so that they could "self serve" for issue resolution, without having to call us. More on this later.

We did still have the challenge of ticket entry. We lacked a means for our field teams or other business partners to report an issue to us, beyond email. The process involved emailing our Tier 1 team, which would acknowledge the request, create the ticket and assign it to our Tier 2 team for initial review, while pulling of necessary log files. After Tier 2 completed its work, it would be assigned to an incident management team or engineering team, depending on the triage steps. After the Incident Management or Engineering teams completed their piece, the ticket went back to Tier 1, to follow up with the issue reporter. The number of exchanges between assignment groups caused a lag in final root cause or mitigation of the issue.

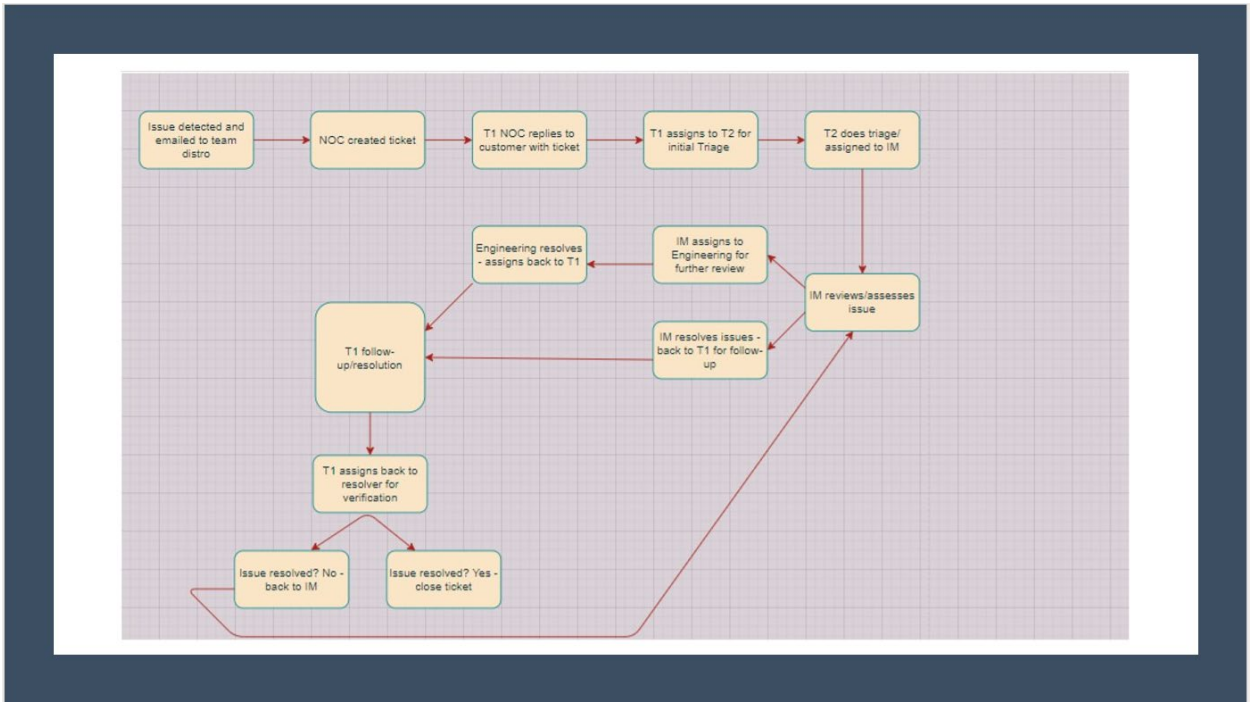


Figure 8- Pre-portal Workflow

Working with our ticketing tools team, we were able to provide requirements for a ticket portal that anyone with the organization could access. This allowed users to directly submit their issues and receive a ticket number at time of entry. We also allowed a user reporting an issue to escalate to the on-call resource, which is especially important after hours or on weekends. ***Note, we did include a disclaimer to "use with discretion," as it will generate a page to the on-call resource's phone.*

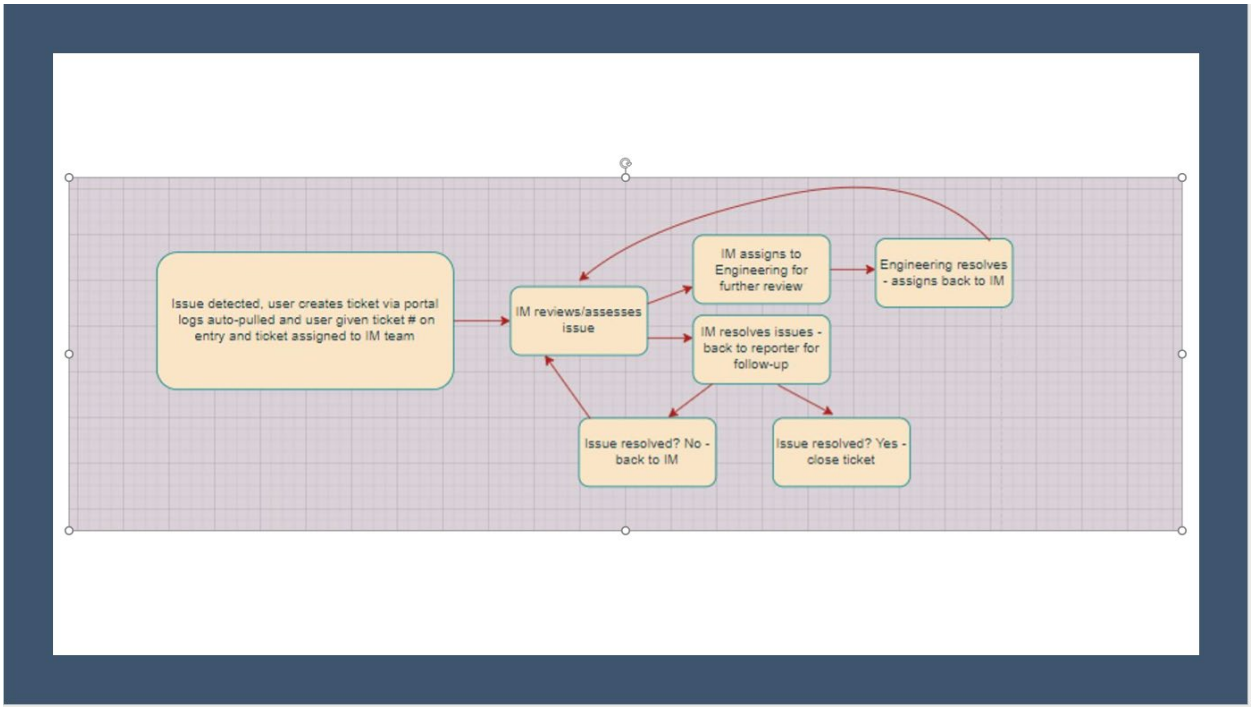
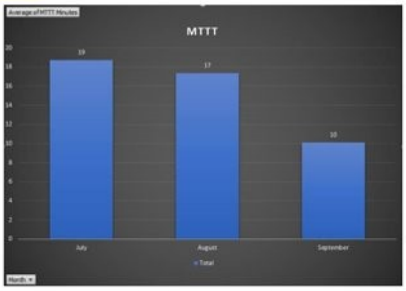


Figure 9- Post-portal Workflow

The portal allowed us to simplify and streamline the ticket process, lessen the number of reassignments, and reduce the MTTM. In addition, the portal and auto-assignment work, we were able to automate pulling device logs when needed (typically a single-customer reported issue). By automating the log pulls, we were able to demonstrate a fair reduction in MTTT and MTTM. We also saved upwards of 2,000 hours (about 2.5 months) of manual labor, and automated most of the Tier 2 work, increasing the opportunity for cost savings.

Mean Time to Triage:
We see an approximate 8-minute decrease for the month of September when comparing to July & August.



Mean Time to Mitigate:
We see an approximate 30-minute decrease for the month of September when comparing to July & August.

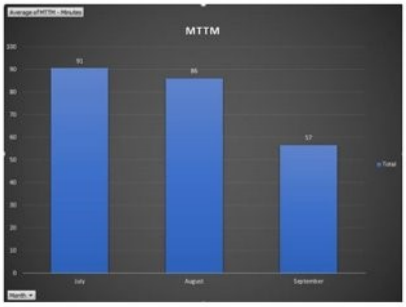


Figure 10 - Mean Time Graphs

Previously, we mentioned tool-sharing and being able to get fixes out to Tier 1 care agents, and eventually directly into our consumers' hands, via an app. When we subsequently analyzed volume, we started seeing trends. What we noticed was that two API (Application Program Interface) functions accounted for most actions taken to fix a customer's escalated issue. This required them to call Tier 1 support, which then escalated to the advanced team, which created a ticket for us to email to the Tier 2 team. That is a long way around, when the task is to quickly escalate!

We worked with our development engineering team to safeguard the API and make it accessible to our Production Support team. We validated the functionality of the tool into which these APIs were integrated. Once we signed off on it, we opened it up to the advanced care team. This prevented escalations to us and our engineering partners, but still created an escalation to the advanced care team. It saved mitigation time for the customer -- and we knew we could do better. Comcast uses ITGs (Interactive Troubleshooting Guides) for our Tier 1 agents, so we worked to get the correct LOQ (line of questioning) into the ITG, with a wrapper around the APIs so that Tier 1 agents could execute and resolve the customer's issue at point-of-contact, by clicking a link. This readily reduced the mean time to resolve as well as escalations – but what if we could help the customer and prevent the call altogether? By leveraging our existing technology and available customer-facing applications, we did just that. We made these functions available directly to the customer in the Xfinity App.

9. Conclusion

A careful review and analysis of your network, systems and applications helps identify areas of sameness, best practices, and resultant opportunities for process alignment and automation. When these opportunities are missed, it can slow the process of automation, potentially causing re-work. Categorizing your findings helps define your proof of concept and identify the low hanging fruit (the quick wins). We found it helpful to write out a problem statement, including an anticipated solution or expected result. From there, determinations can be made about what tools to leverage, to get from problem to solution, and to start defining the "How" of automation opportunities.

In reviewing the automation journey, we categorized opportunities into three "buckets." The first was the "low-hanging fruit" / quick wins, which in our case was the ticket auto assignment. The second was event correlation, which involved a comparatively medium level of effort. We had the foundation already available but needed to ensure that we had proper mappings and service flows in place before implementation. The third was the more difficult, longer-range automation plan of getting resolutions pushed out customer-facing tools, like our customer care app, Xfinity Assistant. This was a multi-phased approach that took a lot of coordination with various teams.

What drove our decisions was the result of extensive, targeted data analysis, which identified the redundant task-type work. If automation is on your to-do list, start by looking for repetition – what tasks are team members are being asked to perform over and over? This also helps to prioritize your automated efforts. By analyzing our email platform, for instance, we were able to determine that there were two main functions, or API calls, that were leveraged to resolve in the range of 90% of customer reported issues. This was a separate opportunity, of the windfall variety, that was identified while we were reviewing options for event correlation.

One consistent theme that emerged throughout all our automation efforts was that it must be measurable. We have the data to baseline where we started, and can continue to measure results to prove (and occasionally disprove) the value of the work being done. Are we seeing the benefits of this effort, as we anticipated? If yes, proceed. If no, re-evaluate: Either we are missing a key component, or this really does not have the ROI (return on investment) we originally envisioned. In which case, it may make sense to deprioritize this automation task and move to something that will have greater impact.

Keep in mind: Data can come from a lot of sources. It can also be a combination of different source types that can give you what you need to measure performance and results accurately. Sometimes, you'll find that you have this great idea, but no way to prove it because the data does not exist. That does not mean you throw away your idea! Reframe the question. Work to identify, define, and implement the right logging to build your data. It is likely a gap that needs to be addressed.

Last not least: Stay agile. Flexibility and adaptability should be the forefront of any effort. Remember to identify the big wins as well as the low-hanging fruit. Ensure that your data tells the story and paints your vision, which will help you to garner the support of your teams for buy-in. Because ultimately, talking about automation is great, and we can sing its praises -- but without stakeholder and customer buy-in, it is not going anywhere. It is extremely difficult to enact change. Most employees, team members and even family members are resistant to changing the way they do their daily tasks. By understanding other teams' processes and procedures, you can highlight the benefits they will see by implementing the

suggested methods. You will also find that they have critical processes that you may want to leverage into your playbook – which brings us back to pollinating best practices throughout the organization.

Abbreviations

API	Application Programming Interface
HFC	Hybrid Fiber Coax
ITG	Interactive Troubleshooting Guide
LOQ	Line of Questioning
MSO	Multi-Soaking Outage
MTTD	Mean-Time-To-Detect
MTTM	Mean-Time-To-Mitigate
MTTT	Mean-Time-To-Triage
NTF	No Trouble Found
OPV	Outage Pre-Verify
VM	Virtual Machine
XOC	eXcellence in Operation Centers

Upstream OFDMA Anomaly Detection and Triaging

A Technical Paper prepared for SCTE by

Jay Zhu
Senior Engineer
CableLabs
j.zhu@cablelabs.com

Karthik Sundaresan
Distinguished Technologist
CableLabs
k.sundaresan@cablelabs.com

CableLabs
858 Coal Creek Circle, Louisville, CO,80027
3036619100

1. Introduction

Upstream Orthogonal Frequency Division Multiple Access (OFDMA) technology in DOCSIS 3.1 is starting to be rolled out in the field. Operators are beginning to test Upstream OFDMA channels in the lab and in the field and are discovering various intricacies in getting the upstream OFDMA to work robustly. Lower frequencies in the upstream spectrum can be noisy and making use of those portions of the spectrum tougher. Upstream RxMER looks very different than the Downstream RxMER, due to the noise funneling characteristics on the HFC plant, the additive nature of noise has a large impact at the CMTS upstream receiver.

As operators roll out OFDMA technology, they are starting to collect data on the performance of these OFDMA channels. This includes the US RxMER data, IUC usage hours, profile definitions etc. As a cable industry we are just starting to comprehend the OFDMA channel performance. Analyzing the US RxMER data and the IUC data is a powerful tool in understanding the performance of each of the node segments and the individual modems. This paper will discuss methods on how to analyze the upstream network data. It will discuss algorithms on how to logically extract the outlier modems and node segments. This paper will discuss methods for anomaly detection, historical behavior analysis, pattern recognition, classification and condition evaluation in the access network data. Combining the analysis of data along, with network topology and device location, it is possible to create a general view of the plant condition and isolate problem sources. The paper will implement methods on how to assign a health score to modems and network segments in an effort to triage which are the top priority nodes that operators need to work on. All this will enable operators to reduce upstream OFDMA troubleshooting and problem resolution time, reducing operational costs and enhancing network reliability.

2. Upstream Background

DOCSIS® 3.1 is now largely deployed in the field. This has primarily focused on a very successful roll out of the Downstream Orthogonal Frequency Division Multiplexing (OFDM) technology. Operators now are beginning to test Upstream Orthogonal Frequency Division Multiple Access (OFDMA) and are now deep into understanding the various intricacies in getting the US OFDMA to perform robustly. The first step is for an operator to get a good stable initial configuration of the OFDMA channel (location channel parameters, IUC definitions etc.). Once an operator can get CMs operating reliably on the OFDMA upstream channel, the operator can then start thinking about how to improve the reliability and efficiency of that upstream channel.

In the upstream direction, the cable system may have a 5-42 MHz, 5-65 MHz (Europe), 5-85 MHz, or 5-204 MHz pass bands. While a DOCSIS 3.1 CM supports a minimum of two independently configurable OFDMA upstream channels with each occupying a spectrum of up to 95 MHz, the challenge has been to find appropriate space in the spectrum to locate these channels. Operators who are running a 5-42 MHz plant are trialing out OFDMA in the space available after the spectrum used by 3 or 4 SC-QAM channels. Operators with mid split (5-85) plants usually have up to 10 SC-QAM channels and are making space for an OFDMA channel by using some of the spectrum just below 85 MHz and turning off a few SC-QAM channels. Operators in Europe with a 5-65 plant typically have 3 or 4 SC-QAM channels are using the remaining space an OFDMA channel, either the OFDMA channels go from 20-45 MHz or 45-65 MHz, depending on where the SC-QAM channels are.

The OFDMA upstream multicarrier system is composed of either 25 kHz or 50 kHz wide subcarriers. For a 95 MHz channel, this equals 3800 25 kHz spaced subcarriers or 1900 50 kHz spaced subcarriers. DOCSIS 3.1 Upstream transmission uses OFDMA frames. Each OFDMA frame is comprised of a configurable number of symbols (K = 6 to 36). Several transmitters may share the same OFDMA frame

by transmitting on allocated subcarriers of the OFDMA frame. The upstream spectrum is divided into groups of subcarriers called minislots. Minislots have dedicated subcarriers, all with the same modulation order ('bit loading'). OFDMA minislots are 400 kHz wide and have either 8 (50 kHz) or 16 (25kHz) subcarriers. The modulation order of a minislot, as well as the pilot pattern used may change between different transmission bursts and are determined by the profile definition.

2.1. Upstream Network Data

CMTS and CM support features and capabilities that ca be leveraged to enable measurement and reporting of network conditions. These Proactive Network Maintenance (PNM) features deliver metrics which operators can use to identify undesired impacts such as plant equipment and cable faults, interference from other systems and ingress noise. With this information operators can make modifications necessary to improve conditions and monitor network trends to detect when network improvements are needed.

The OFDMA technology comes with a few different PNM measurements: Upstream Capture for Active and Quiet Probe, Upstream Triggered Spectrum Analysis, Upstream FEC Statistics, Upstream RxMER Per Subcarrier, Upstream Equalizer Coefficients, Upstream Impulse Noise Statistics, Upstream Histogram, Upstream Channel Power. These eight features are detailed in the [PHYv3.1] specification. The purpose of these upstream PNM functions is to analyze the upstream in various ways. These goals include measuring the plant response, understanding the underlying noise floor, having a wideband spectrum analyzer function on the CMTS, gathering statistics of burst/impulse noise occurring in a selected band, understand the linear response of the upstream cable plant, monitoring upstream link quality via FEC and related statistics, understanding the nonlinear effects in the channel such as amplifier compression and laser clipping, providing an estimate of the total received power in a channel and becoming aware of the upstream receive modulation error ratio (RxMER) for each subcarrier.

2.1.1. US RxMER

A CMTS uses upstream probes for ranging-related functions such as determining transmit pre-equalizer coefficients. A CMTS also uses the upstream probe to take an RxMER (received modulation error ratio) measurement. The CMTS grants probe opportunities to a CM in a P-MAP message with the "MER" bit set. When the CMTS receives the probe transmissions from the CM corresponding to such a grant, it performs the RxMER measurement and uses the results in its decision making. It also populates the corresponding MIB object or can upload a RxMER per subcarrier file via TFTP, for the operator's information. Some CMTS implementations also measured the RxMER in an alternate fashion, they are using the actual data transmission bursts from a CM to do a measurement.

2.1.2. Other US PNM data

The other Upstream PNM features (Upstream Capture for Active and Quiet Probe, Upstream Triggered Spectrum Analysis, Upstream FEC Statistics, Upstream Equalizer Coefficients, Upstream Impulse Noise Statistics, Upstream Histogram, Upstream Channel Power) are in various stages of maturity on different CMTS platforms.

As of the writing of this paper, we have had access to a lot of Upstream RxMER data, but have not been able to gather a meaningful set of samples from the field for any of the other PNM data types. In the future, as we get access to more data samples of those types from CMTS in the field, we can start analyzing those (for a future paper). For this paper, we focus on the US RxMER data that we have and start building methods and tools and ways to visualize and analyze this data set and figure out what kinds of upstream evaluation we can perform.

3. Upstream Observations

3.1. Introduction to Upstream RxMER Data

The upstream RxMER data is a PNM metric measured by the CMTS to report receive modulation error ratio on a per subcarrier basis. Similar to the downstream RxMER data, the upstream RxMER data can be helpful in identifying upstream impairments over frequency as well as performance fluctuations over time. The following is an upstream RxMER capture measured from 23.9 MHz to 51.4 MHz. Typically, noise floor is higher at lower frequencies. Although Pre-Equalization can help compensate the tilt, we still observe a slight inclination in the data from the lower frequencies to the higher frequencies.

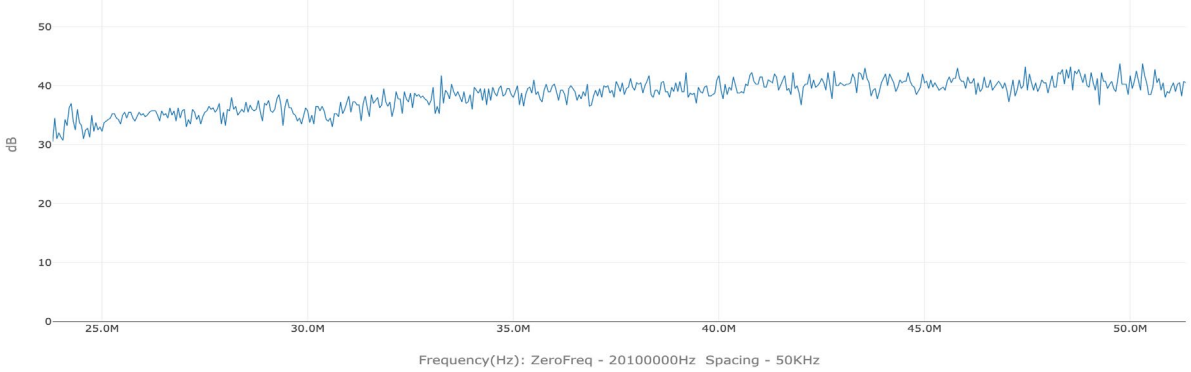


Figure 1 – An US RxMER sample on a OFDMA channel from Single CM

When more samples are captured over time and are displayed in one plot, we can observe the changes of RxMER values over time. By capturing more upstream RxMER samples and analyzing the over-time features of the data, there is opportunity that intermittent impairments and field events can be observed and categorized. In this paper, the discussion is focused an upstream RxMER dataset that was collected from multiple OFDMA interfaces during a 2-week timeframe.

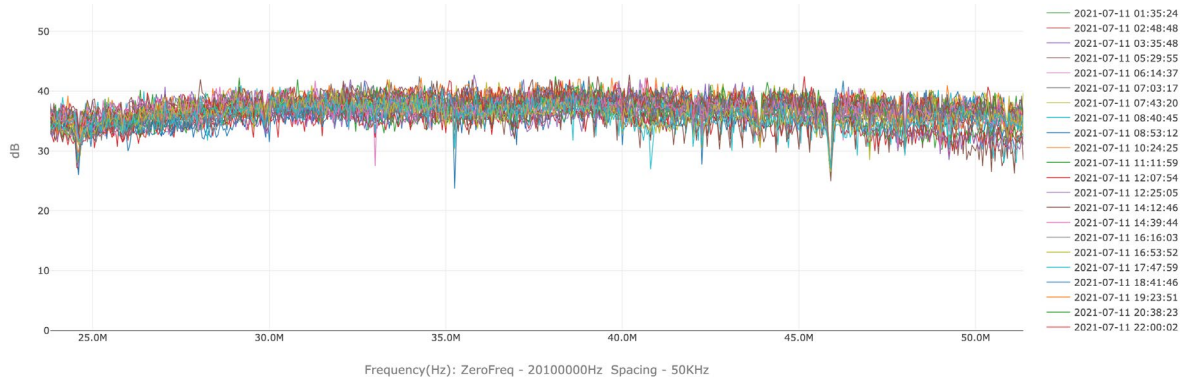


Figure 2 – Multiple US RxMER samples, OFDMA channel, Single CM (lab)

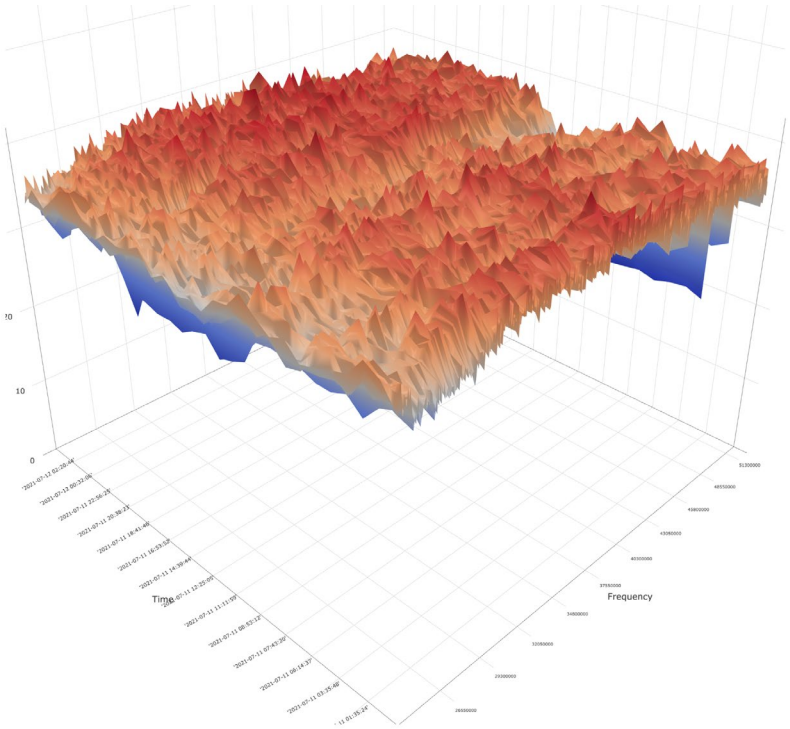


Figure 3 – 3D view of US RxMER samples, time versus frequency

3.2. Variations and Impairments

The following figure shows upstream OFDMA RxMER samples collected from a CMTS in the field. A lot of variation in RxMER values is observed in this visualization. The RxMER values of unused/inactive subcarriers are reported by the CMTS as 0xFF (63.75 dB), which are the spikes seen in the graph. The maximum difference between the RxMER values on a same subcarrier can be as much as 20 dB, which indicates that the condition of the upstream OFDMA channel is constantly changing. Ingress noise can also be observed on captured samples.

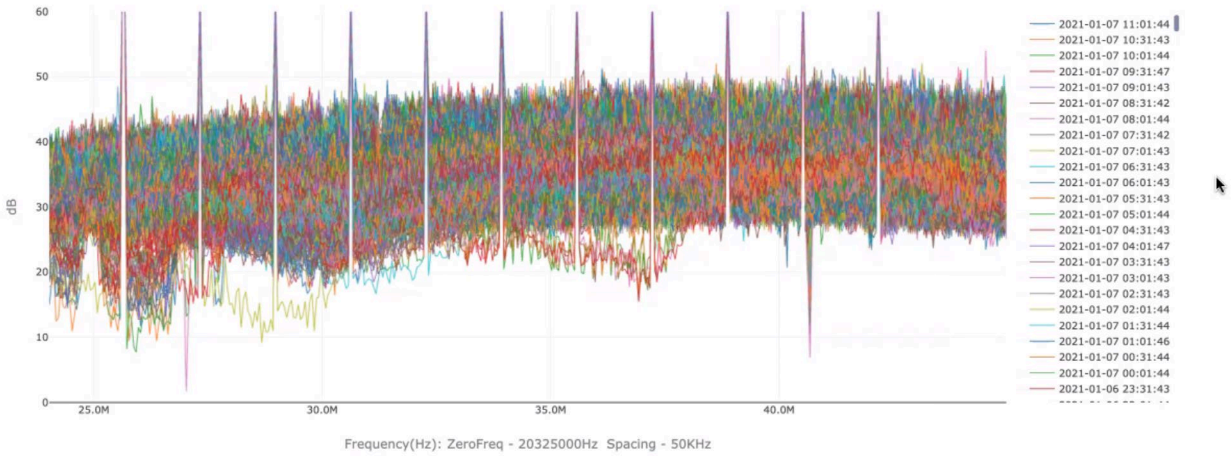


Figure 4 – Multiple US RxMER samples, OFDMA channel, Single CM (field)

In the RxMER captures collected from a different CM, we observe less variation in RxMER values. We also observe standing waves across the spectrum.

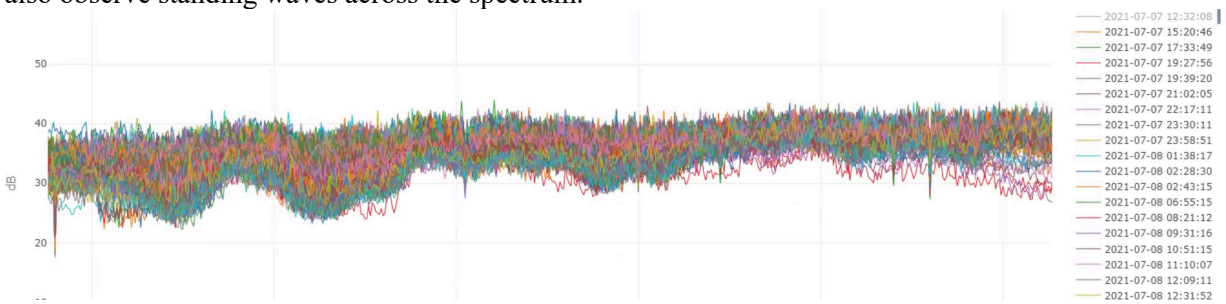
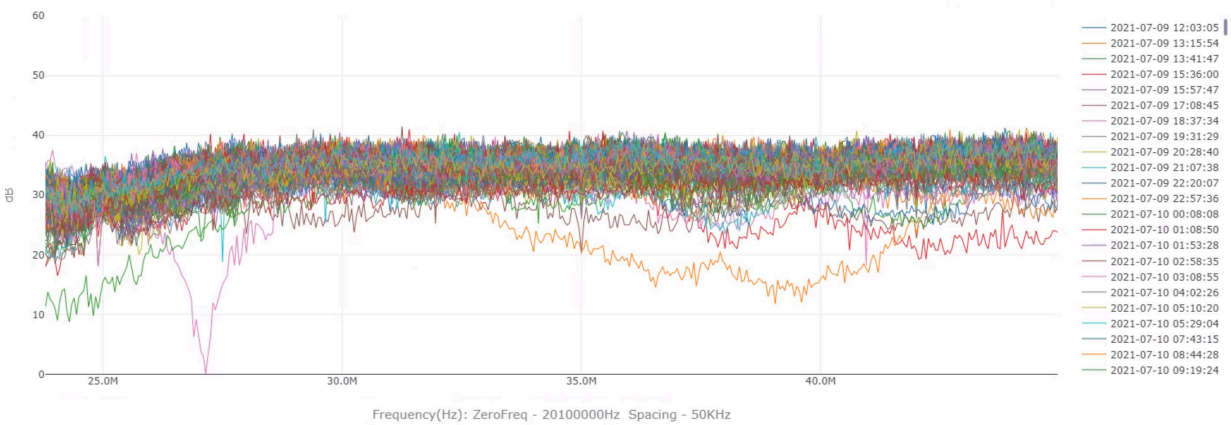
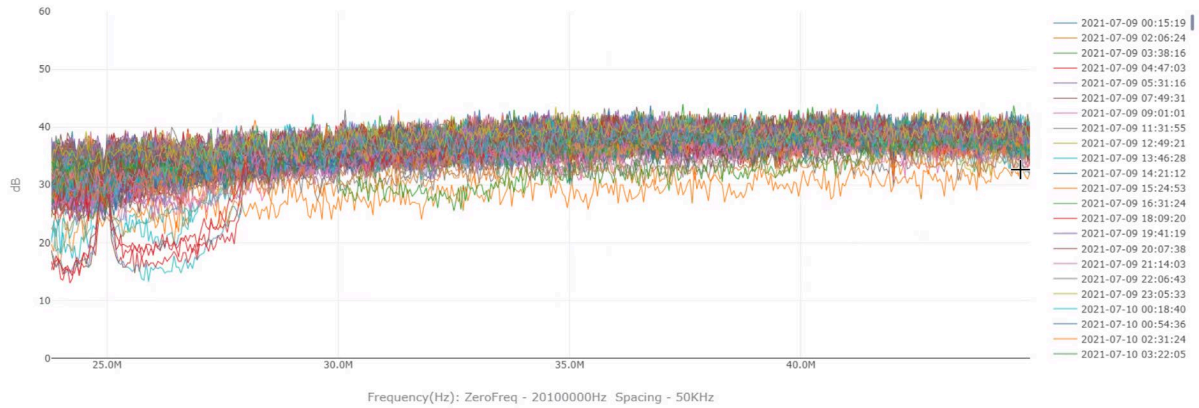


Figure 5 – Multiple US RxMER samples, OFDMA channel, Single CM (field)

The following RxMER captures show relatively tight value distribution but have many outliers throughout the capturing period.





Based on the observations of such variations, we develop statistical analysis methods to further extract information from the upstream RxMER data, which are discussed in the following section.

4. Statistical Analysis

When assuming that the RxMER values captured from each subcarrier are following normal distribution while the impairments are absent, it is helpful to use percentiles, variance, skewness, and kurtosis calculated from the RxMER values over time to summarize the distributions and provide insights into the behavior of the impairments. Each of the statistical calculations is discussed in the following sections. We also discuss initial use cases of the statistical metrics, observations of RxMER time series on different frequency ranges, and why the data suggests that PMA is necessary for the robustness of the upstream OFDMA deployment.

4.1. Percentiles

In order to analyze the upstream OFDMA RxMER values captured over time for each CM, we calculate different percentile values from each of the subcarrier's historical RxMER values. We calculate 1%, 2%, 5%, 10%, 15%, 20%, and 50% values from all the RxMER values captured on each subcarrier and include the minimum and maximum values to illustrate the range and distribution densities. By calculating the delta values between the percentile traces, one can automatically identify frequency ranges that may need attention as well as intermittent impairments on the data captured over time, as highlighted in the following figures.

For example, when we calculate the deltas between the minimum value, 1%, 5%, 15%, 50%, and the maximum value, and observe large differences between 1% and 5% traces while seeing smaller delta values between the other percentile traces, such as the condition indicated in the third highlight in the following figures, an intermittent impairment can be identified. On the other hand, if the percentile values are evenly distributed but the averaged percentile delta values are high, such as the condition shown in the first highlight, a persistent impairment can be identified.

Figure 6 – RxMER over time with persistent impairments and intermittent impairments

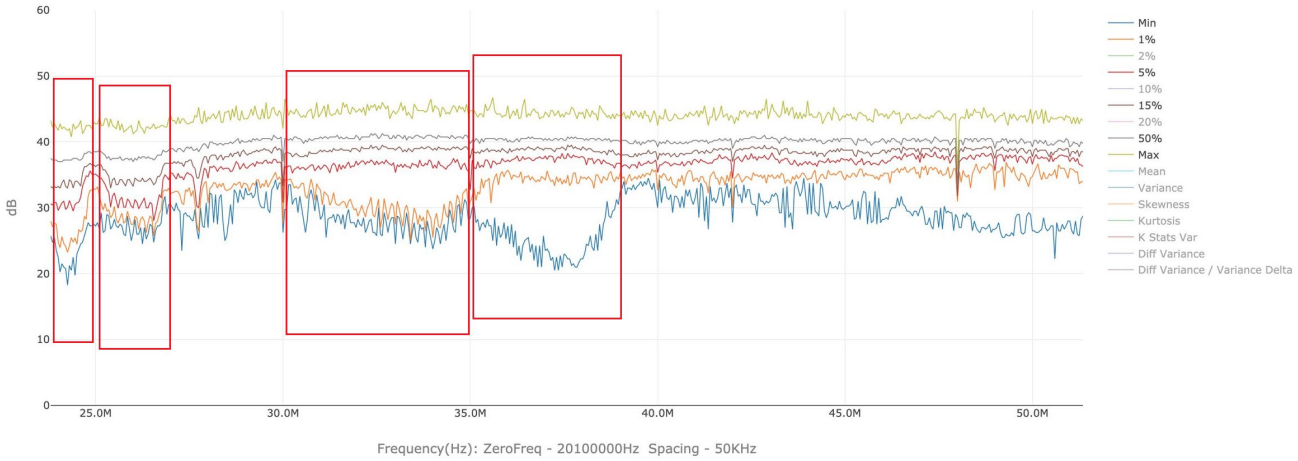


Figure 7 – Percentiles and min, max values calculated for each subcarrier

4.2. Variance over Time

It is well known that variance is a statistical measurement of the spread between numbers in a dataset. In the application we developed for this research, other than calculating the percentiles we calculate the variances of the RxMER values captured from each subcarrier and visualize the variance calculation results over frequency.

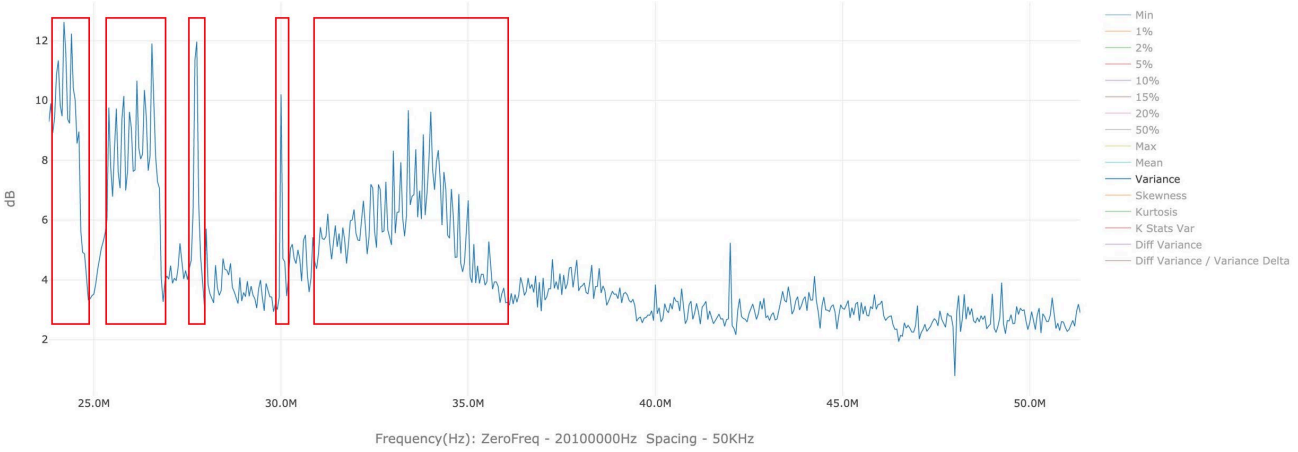


Figure 8 – Variance values calculated for each subcarrier

Relatively high variance can indicate that the RxMER values captured from the subcarriers are unstable over time. However, variance by itself would not provide sufficient information to differentiate persistent impairments and intermittent impairments/events, as both can cause instability of subcarriers' RxMER values which leads to high variance values. By defining thresholds for the variance, it can be used as the first step of selecting unstable subcarriers based on their upstream RxMER values captured over time. And in order to extract sufficient information for categorizing the impairments observed over time at

high-level, such as identifying persistent impairments versus intermittent impairments, we introduce skewness and kurtosis calculations of the RxMER values from each subcarrier.

4.3. Skewness over Time

Skewness is a measure of the asymmetry of the probability distribution. The value of skewness can be positive, zero, or negative. Positive skewness indicates that there is more weight in the right tail of the distribution, whereas negative skewness indicates the opposite.

Figure 9 – A CM with persistent impairments and intermittent impairments

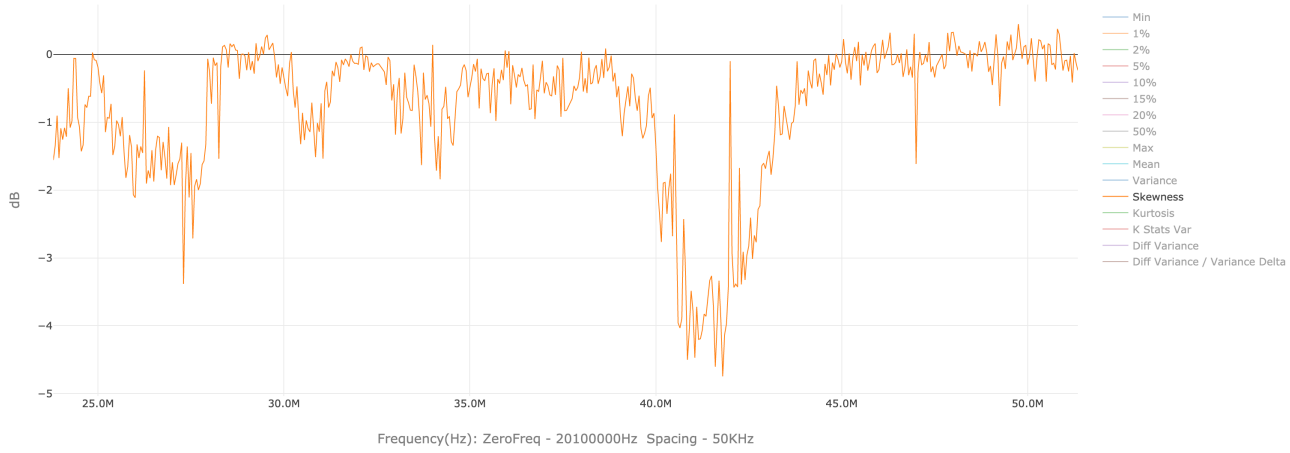


Figure 10 – Skewness calculated for each subcarrier

Considering that the distribution of a subcarrier's RxMER values over time should be approximately symmetrical when no impairment is present or only persistent impairments present, large absolute skewness values (especially when the skewness is negative) can be used to identify subcarriers affected by intermittent impairments/events.

Combining skewness with variance, it can be inferred that when both of the variance value and the absolute skewness value are large, the subcarrier is primarily being affected by intermittent impairments;

and when the variance value is large and the skewness value is close to zero, the subcarrier is primarily being affected by persistent impairments.

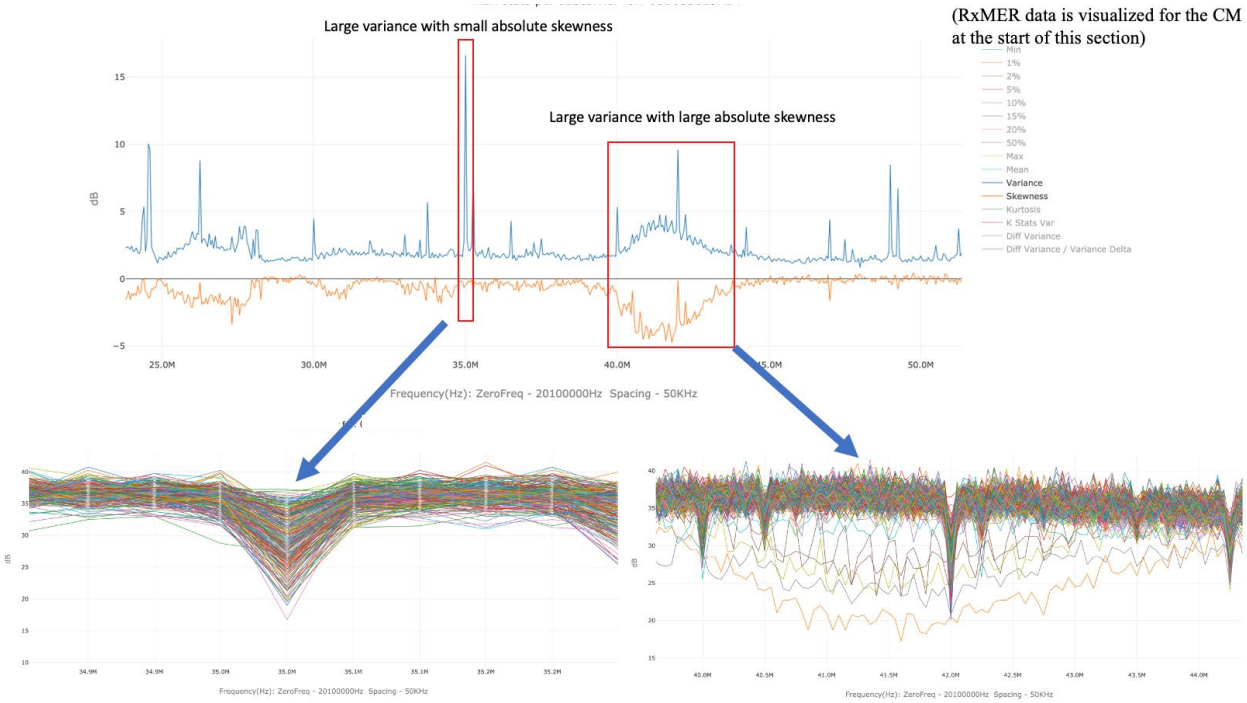


Figure 11 – Persistent and Intermittent Issues identified by variance & skewness

The bottom of Figure 11 shows that a persistent issue presents and is identified by using the combination of variance and skewness at 35 MHz of the spectrum. The variance of this subcarrier's RxMER values is large, however, the skewness is close to zero based on the calculation and the observation that the RxMER values are evenly distributed under the impairment.

Another impairment between 40 MHz and 44 MHz is identified as an intermittent issue since the variance values and absolute skewness values are large. The behavior of both of the identified impairments can be further confirmed when we visually check from the bottom of the 3-dimensional graph (Figure 12) generated from the upstream RxMER samples captured over time.

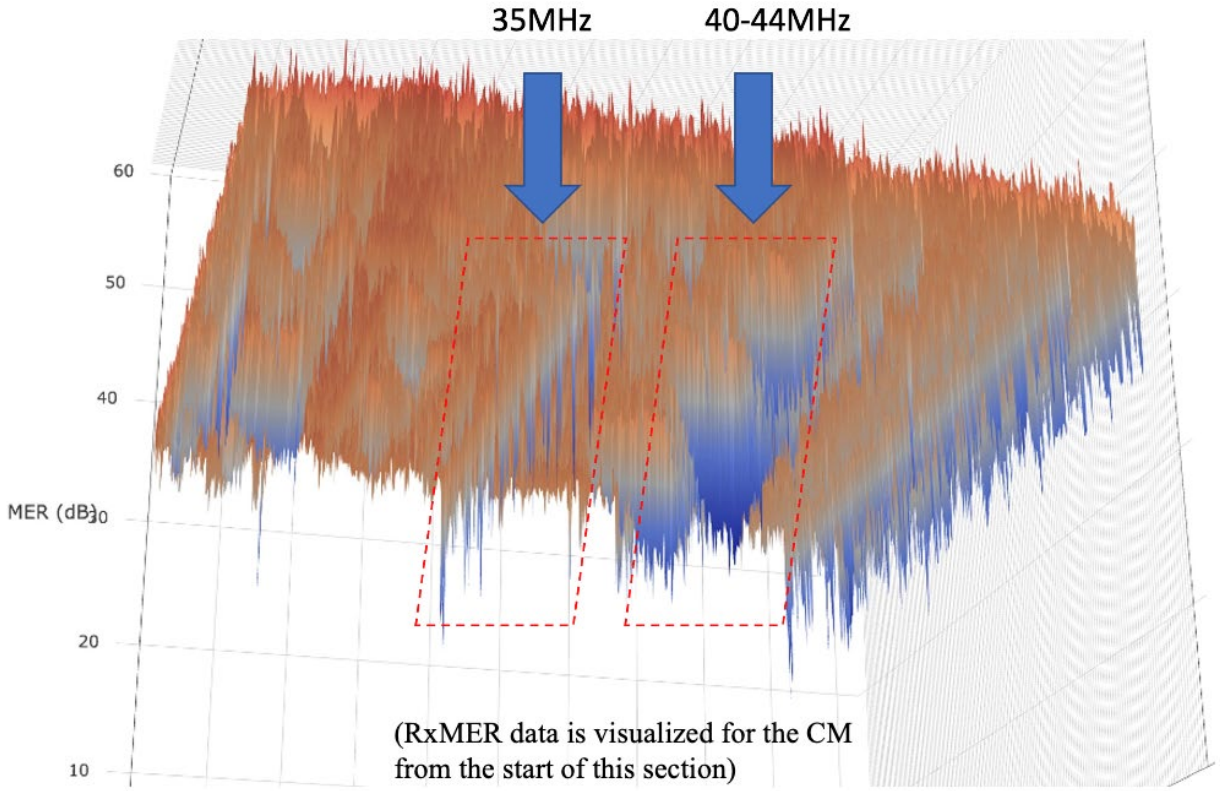


Figure 12 –3D graph of the RxMER samples captured over time

At 35 MHz, the impairment is persistent. It constantly affects the RxMER values and is almost captured by every RxMER sample. On the other hand, the impairment between 40 MHz and 44 MHz can be considered intermittent since it only presents occasionally in the 300 RxMER samples visualized in the 3D graph.

4.4. Kurtosis over Time

Kurtosis is a measure of the "tailedness" of the probability distribution. Higher kurtosis corresponds to greater extremity of deviations, which can be correlated with variance and skewness to further confirm if an identified impairment/event captured by the RxMER data is persistent or intermittent.
In our application, we calculate excess kurtosis (the value is 0 when the distribution is normal) of RxMER values over time for each subcarrier. From the observations, the combination of large absolute skewness values and large kurtosis values emphasizes the intermittent behavior of an impairment. This can be a promising technique to process a large number of RxMER samples captured from each CM over time, remove the noise, and filter out information that can be used for potential automatic anomaly detection in OFDMA, as demonstrated in Figure 13.

(RxMER data is visualized for the CM at the start of this section)

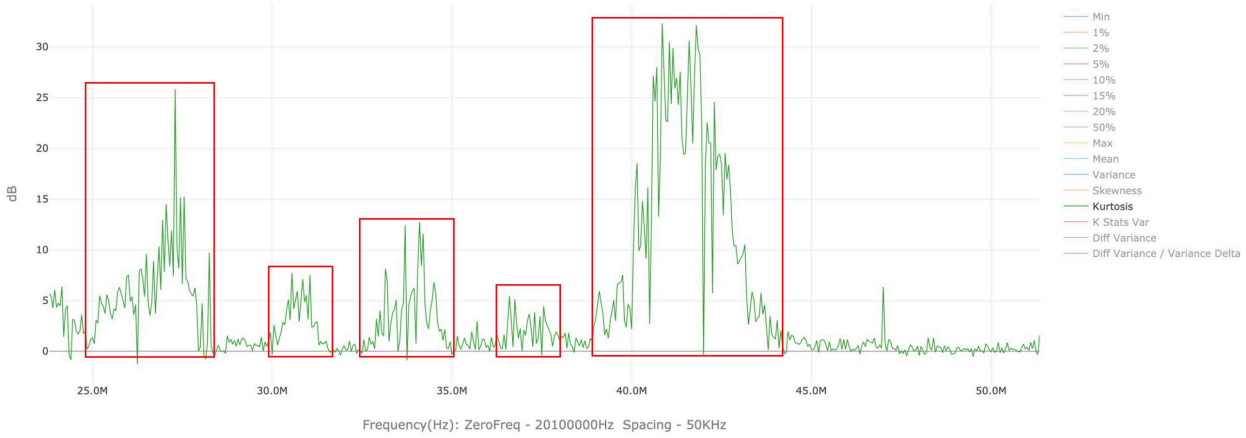


Figure 13 – Intermittent issues identified by kurtosis

4.5. RxMER Time Series on Different Frequency Ranges

Another way to analyze RxMER data is to view the variations of small frequency range (practically a multiple of the subcarrier size). Here we choose a unit of 1 MHz, which for the 50 kHz spacing is 20 subcarriers. The idea is to take an average of the RxMER for those twenty subcarriers and then plot that average value as it changes over time. We do this for every 1 MHz of the spectrum, and then certain patterns become more apparent. In order to observe RxMER value changes over time sequentially, we calculate the average RxMER values of each 1 MHz frequency within the channel and visualize the averaged RxMER values with the sample capture timestamps (as shown in Figure 14).This visualization method provides insights into the unstable nature of the upstream and helps us identify issues that can be potentially identified by algorithms we may develop.

Figure 14 – Averaged RxMER time series for one CM

One of the patterns (see Figure 15) we observed is periodical changes across the whole channel over the course of the day, and these varying RxMER patterns repeat every day. The variation is significant, anywhere from 5 to 10 dB.

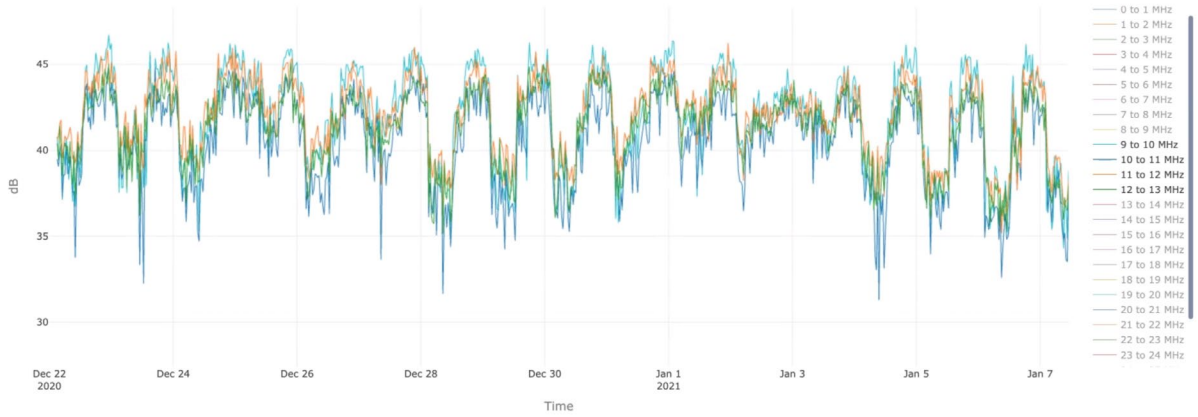


Figure 15 – Recurring daily variation in RxMER

In another example, (see Figure 16), an intermittent issue happened on July 5th can be identified between 11 MHz and 15 MHz of the OFDM channel (34.9 MHz – 38.9 MHz), where there is a sudden drop in the average RxMER values by about 18 dB.

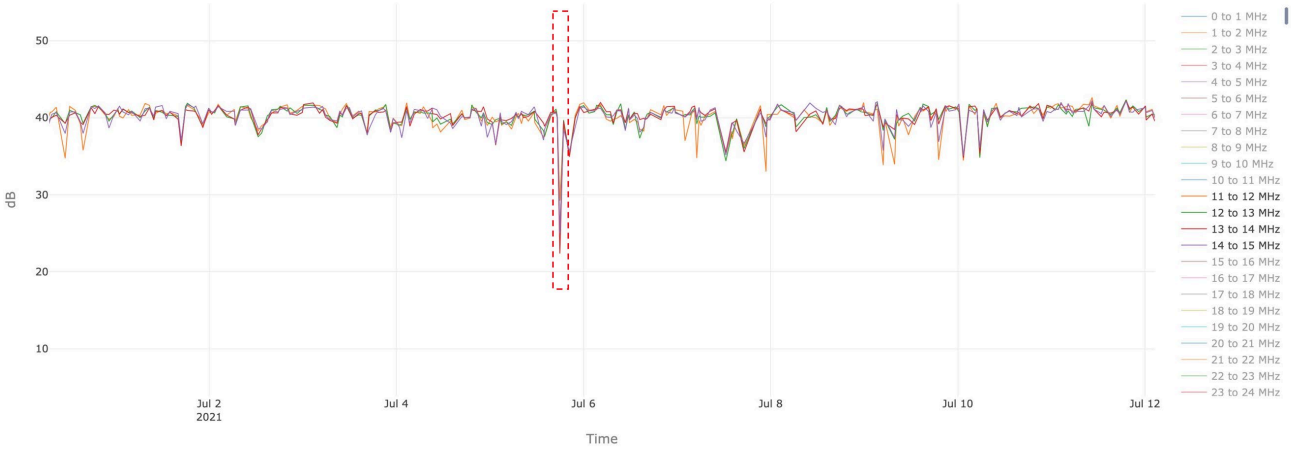


Figure 16 – Intermittent issue observed on RxMER time series

We can then target the problematic RxMER sample using the timestamp, as shown in the top half of Figure 17. Comparing the RxMER sample that has issues (the yellow trace) with a clean sample captured on the same day (the red trace), it indicates that an intermittent issue happened and significantly affected the result of the RxMER measurement.

In addition, because of noise funneling in the upstream, such an issue may also be captured by the measurement of upstream RxMER on a different CM around the same time, as shown in the bottom half of Figure 17.

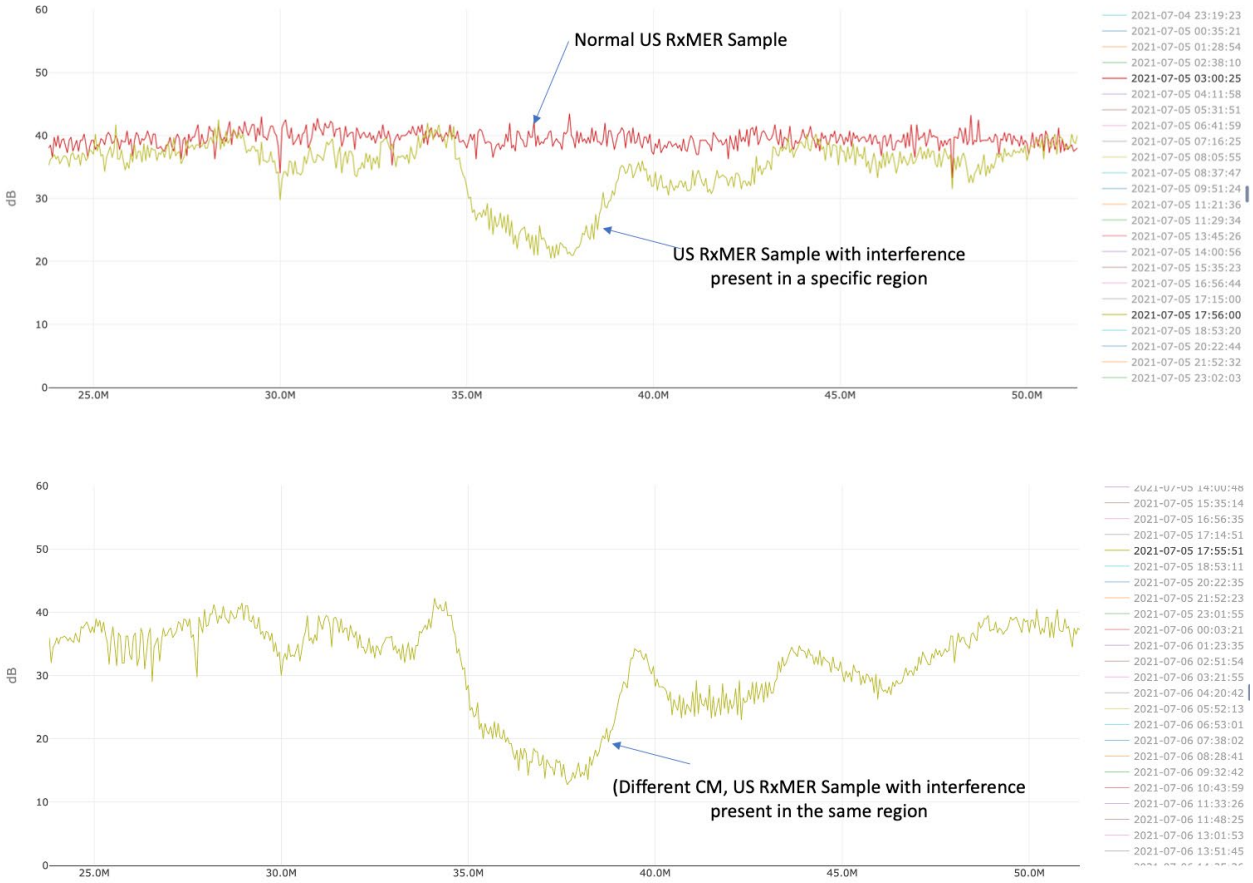


Figure 17 – A problematic frequency region on multiple CMs

To provide another example of such intermittent issues being highlighted by RxMER time series, as shown in the top of Figure 18, a significant event is identified using the data collected from a different CM where all subcarriers had a 30 dB drop in their RxMER values. This outlier sample is as shown in the bottom half of Figure 17. The cause of this outlier sample is under investigation, this could be a CMTS measurement bug or it could be the nature of the plant noise manifesting itself as a low RxMER.

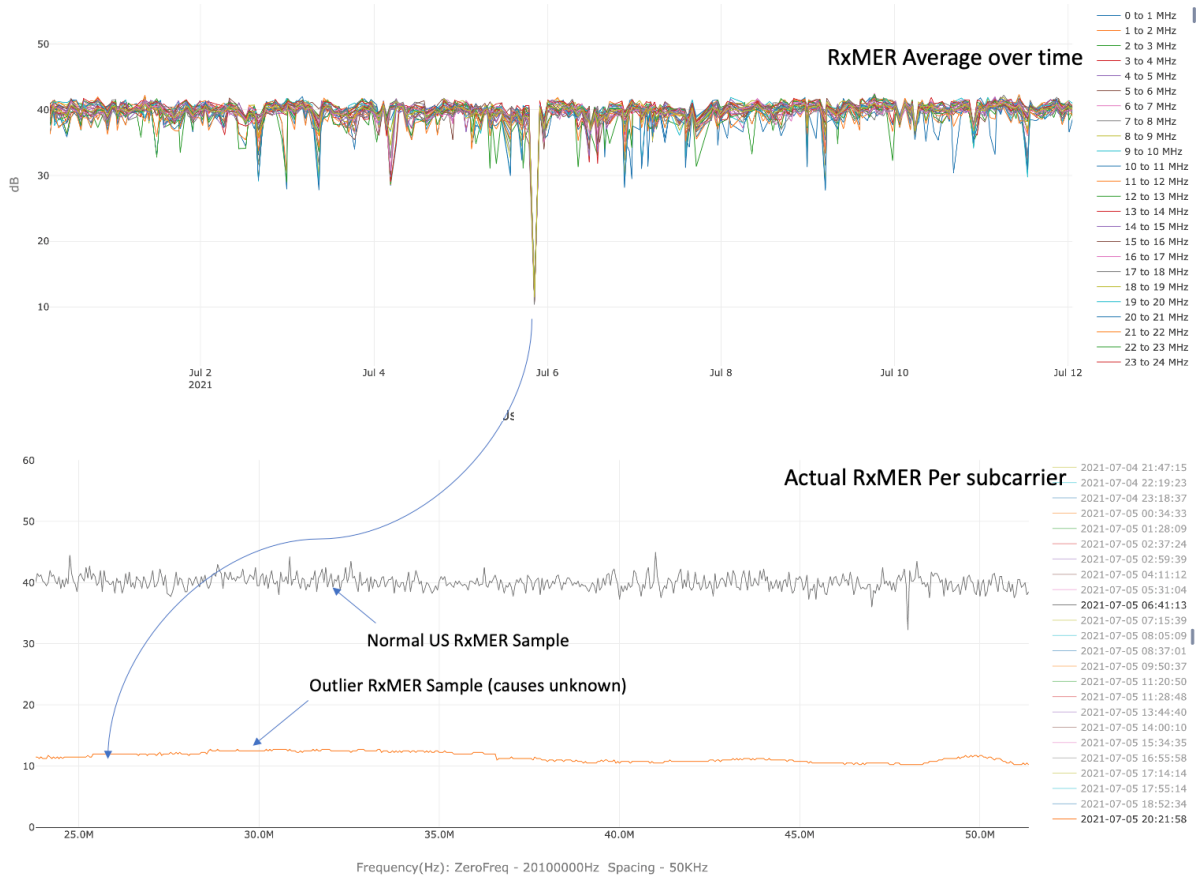


Figure 18 – A different intermittent issue observed on RxMER time series

In another example, one of the observations we made were sudden drops in average RxMER across the whole channel. For example, in Figure 19, we can see on Dec 27th the RxMER for the CM drops down by 10 dB, and then a few days later goes higher and fluctuates between two or three levels.

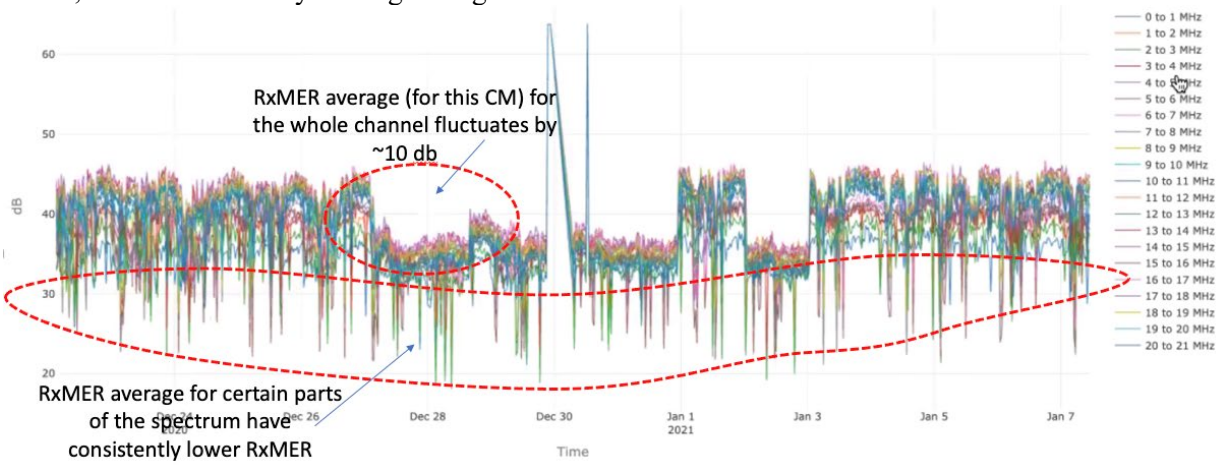


Figure 19 – RxMER Variation across time.

By using the average RxMER time series data, intermittent issues (both unknown issues and repetitive issues) can be identified with low computational cost. For the goals of proactive network maintenance (PNM), such identified issues can be characterized and categorized for the operators to be aware of and potentially reduce maintenance time and cost. In addition, when the root causes of the issues are located and isolated, the features of these issues can be fed into advanced models such as machine learning based classifiers for automatic anomaly detection and root cause inference.

4.6. Clustering Analysis on Statistical Measures

In order to analyze the features extracted from the upstream RxMER data collected over time using variance, absolute skewness, and kurtosis, and research how the impairments can be automatically isolated and categorized, we first calculate the statistics for all of the subcarriers of CMs on multiple OFDMA channels, and then generate a 3-dimensional graph to show if the statistics of the subcarriers' RxMER data are naturally clustered in 3D space.

As shown in Figure 20, there are two main clusters in the 3D graph, one of which has relatively large variance values and small absolute skewness and kurtosis values, whereas another cluster has small values in variance but relatively large values in both absolute skewness and kurtosis.

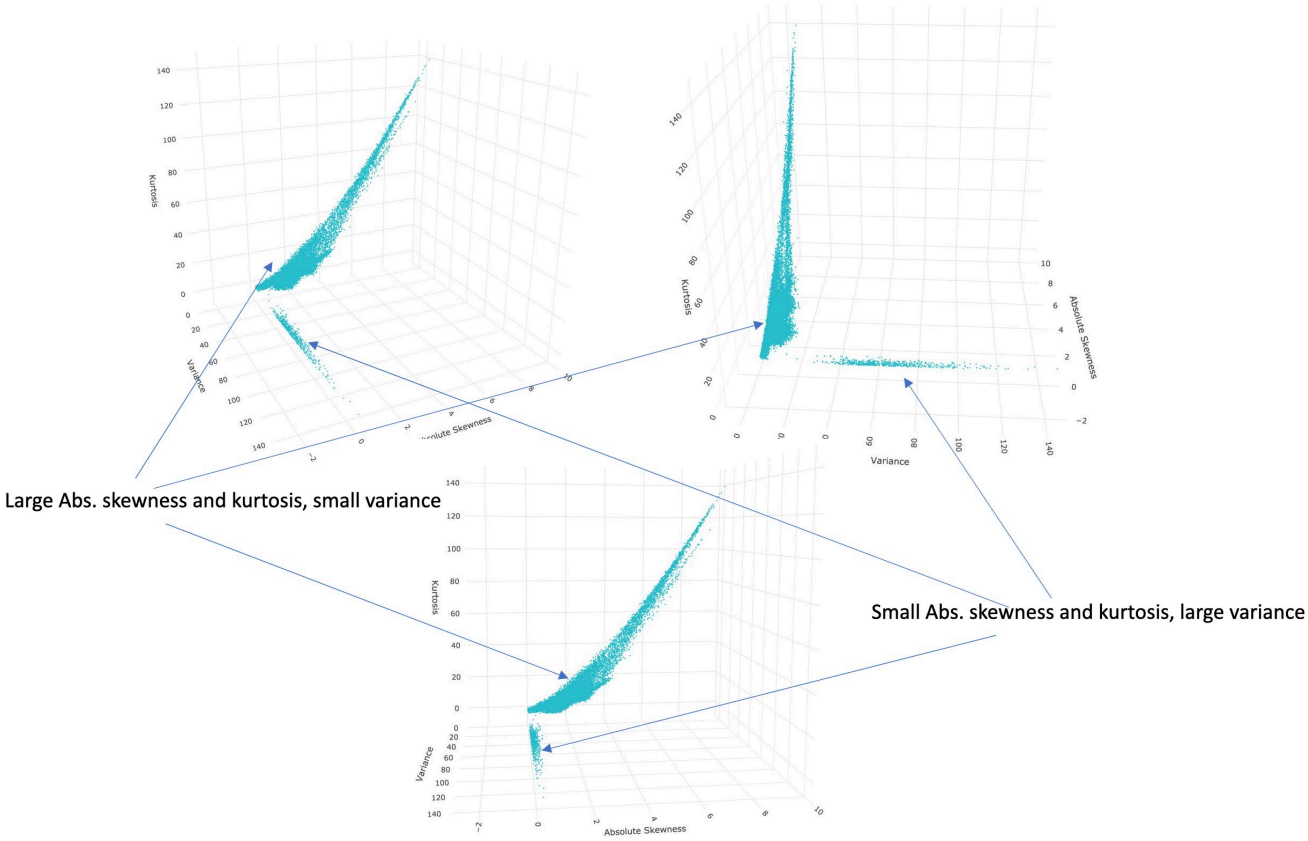


Figure 20 – 3D graph of variance, absolute skewness and kurtosis values

It can also be inferred that the cluster that shows significance in absolute skewness and kurtosis suggests positive relationship between the two statistics. In other words, large kurtosis and absolute skewness both can be used as criteria to identify intermittent impairments on the OFDMA channel, and they potentially reinforce each other.

In addition, this analysis helps identify thresholds that can be applied to the statistics in creating a threshold-based method to automatically detect subcarriers under impairment. Also, centroids can be calculated for these clusters to categorize the impairment types.

More research is ongoing around understanding the impairment types in upstream OFDMA channels. Calculating the statistics of upstream RxMER values over time can be a promising feature extraction method for machine learning models. For example, a 1-dimensional convolutional neural network can use the normalized variance, skewness, and kurtosis traces over frequency as input taken from 3 different channels. A regression model (see [ANOMALY DETECTOR ICPHM 2020]) can be developed to automatically localize and classify impairments observed by RxMER data in upstream OFDMA.

4.7. Time Clustering – PMA Use Case

By analyzing the upstream RxMER data collected from the field, an important note can be made that the condition of the OFDMA channel can vary significantly from moment to moment, which can lead to frequent IUC shifting if the CMTS is enabled to automatically adjust upstream IUCs for CMs based on FEC error rates and RxMER measurement results. To automatically create a set of IUCs that can provide optimal operational robustness and channel capacity, upstream PMA can be adopted. The upstream PMA algorithm uses time clustering to group all upstream RxMER samples collected from all CMs over time in order to automatically design optimal IUCs, which has been proven to be beneficial in field trials.

There are 2 algorithms that are implemented in upstream PMA which we call algorithm 2b and 2c, see [US PMA SCTE 2020]. Algorithm 2b is more aggressive in capacity improvements compared to algorithm 2c, whereas algorithm 2c designs more robust IUCs. To provide examples, we calculate and visualize the time each CM spends on each IUC to provide insights on how the CMs may utilize the IUCs based on simulated CMTS criteria. As shown in Figure 21, IUCs created by algorithm 2b are used by the CMs evenly in time, as the IUCs are designed to prioritize channel capacity gains. In Figure 22, CMs use IUC 5 and IUC 6 primarily, as the bit loadings of the IUCs are lower to ensure robustness and fewer IUC shifts.

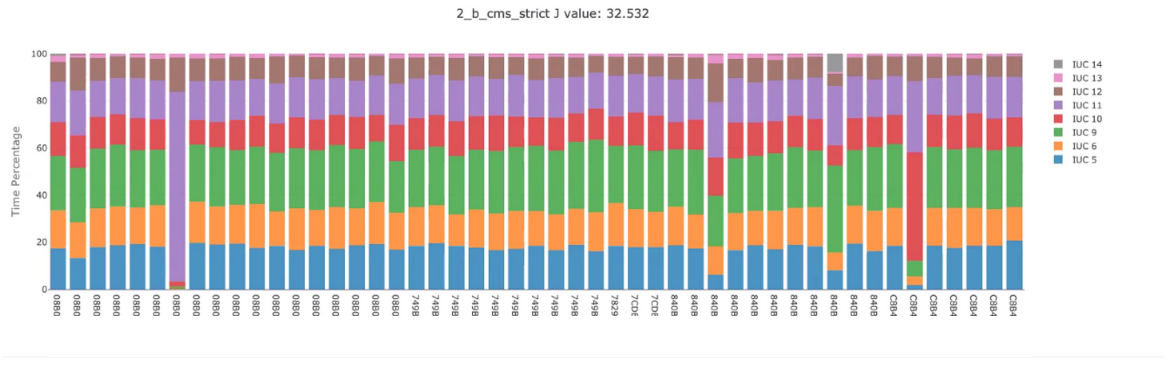


Figure 21 – IUC usage time distribution per CM (method 2B)

Figure 22 – IUC usage time distribution per CM (method 2C)

5. Data Analysis

We have observed some interesting patterns in the OFDMA channel data collected by the CMTS on modems in the field. These include "outage" patterns on a particular modem on a particular OFDMA channel, or invalid RxMER values for a CM across the channel etc. To categorize the issues in data reporting, we label them as "measurement discontinuities".

5.1. US RxMER Measurement Discontinuities

The first pattern that we see is that some modems have no RxMER entries for large parts of the data collection timeframe. As an example, in the figure below, during a data collection period of two weeks, we see that for a period of four days the modem does not respond and so the CMTS has no RxMER samples for that modem. This may be caused by but true outage or an issue at the cable modem or it could also be that the user just turned the cable modem off. Some data analysis would need to be done to correlate the outage with say data traffic on the modem and tease out the root cause of this outage. A second issue that we observe is when the CMTS reports value of 0xFF for all the subcarriers off that modem. Based on some discussions with the operator on the status we believe that this condition is when the CM is in partial service on that OFDMA channel but is maintaining data connectivity with the CMTS on the other channels.

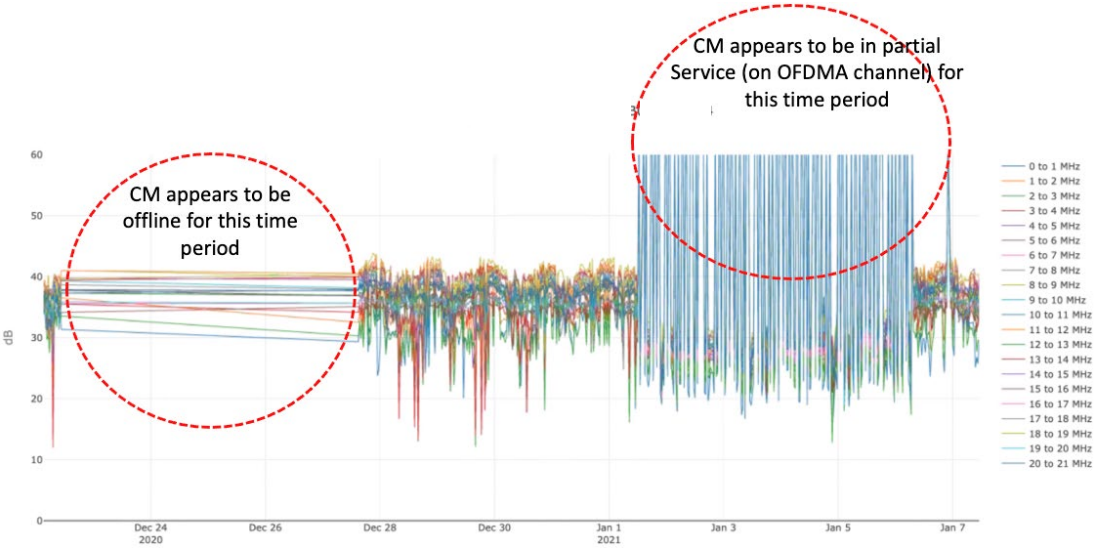


Figure 23 – One CM offline or in partial service

5.1.1. CM Outage Issues

Based on examples we saw in the previous section we know that multiple modems have outages at different points in time. This leads us to the next step of analysis by aggregating all such outage data from modems on the same node. The idea is to observe each of the individual CM outages on the node and try to correlate them over time to identify any meaningful events on the network.

Here we aggregated all the offline outage events (i.e., no RxMER samples) from the modems and are plotting them as a bar graph over time. Given that the data collection frequency was every hour we consider each gap in the data set as one event for that one modem. We calculate the number of such outage events and plot them on the graph below as an aggregate of all modems in the service group. The graph below shows that between Dec 24th and 28th there were many hour-intervals where up to 13 to15 modems in the service group were unable to respond to the RxMER request and work offline during that interval

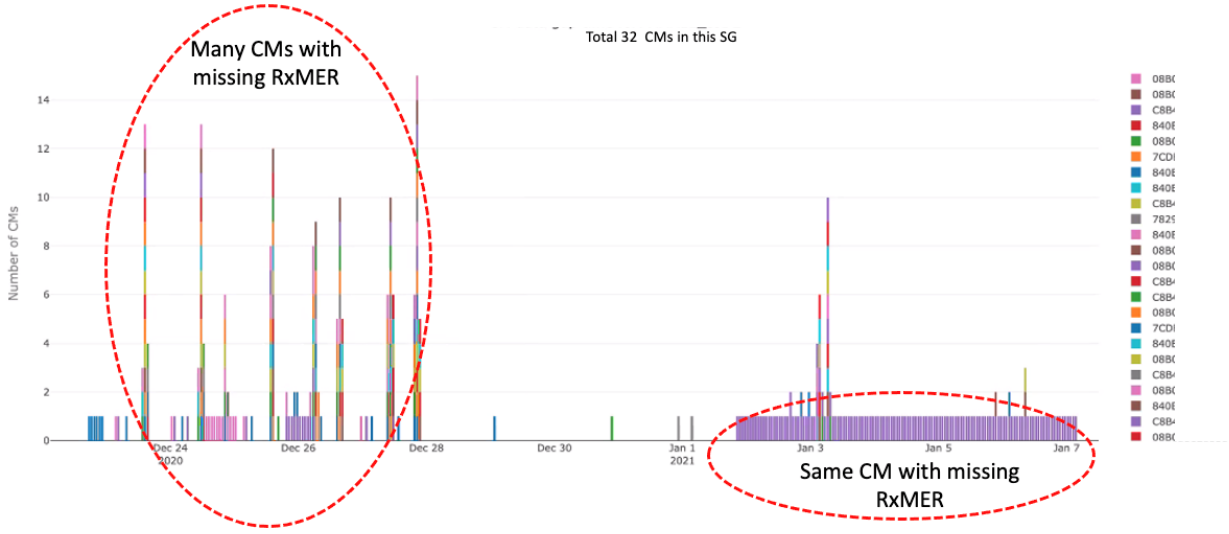


Figure 24 – Num of CMs Missing Data Samples across SG

We also looked at the amount of time each modem was missing data and plotted a histogram of the outages. The graph below shows the number of events which lasted from ~50 minutes to ~3 hours.

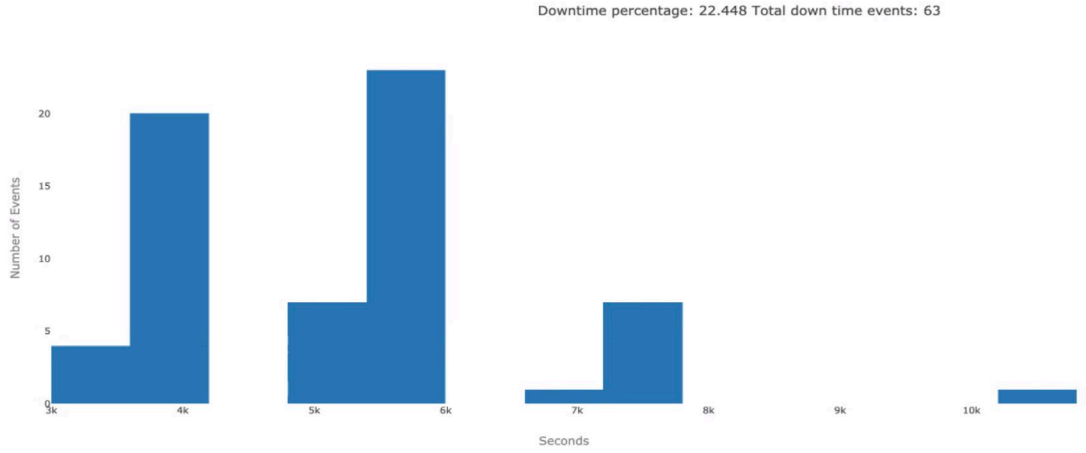


Figure 25 – Characterizing outage times

5.1.2. CM Invalid RxMER

The second type of measurement discontinuity are the events when the CMTS returns all 0xFF values (i.e., invalid samples) for US RxMER from the modems. Any collection interval which has a US RxMER file but with the data set to all 0xFFs counts as one event for that one modem. We calculate the number of such outage events and plot them on the graph below as an aggregate of all modems in the service group. The graph below shows that between Jan 1st and Jan 7th, there were many hour-intervals a few different modems (different colors for the actual CM MAC address), up to 2 at a time, were unable to respond to the CMTS request appropriately to get valid RxMER measurements. In particular two CMs, in green and gray looked to be struggling during that week.

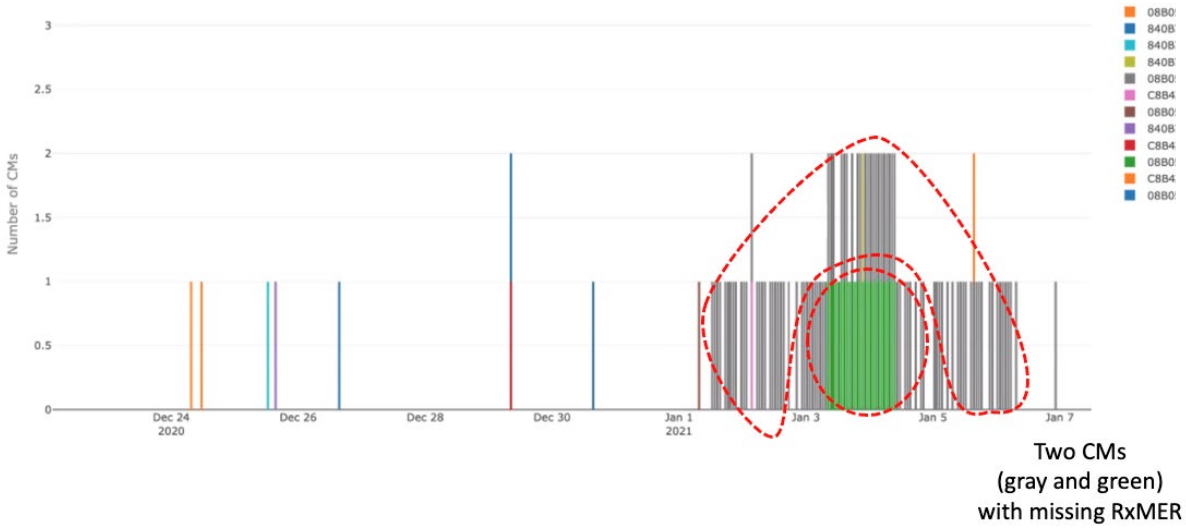


Figure 26 – Num of CMs offline across the whole SG (FF)

5.2. Health Score

Operators have identified a need to develop a set of metrics from the network which reflect how they affect the customer's network connection (speed, reliability, latency). Operators want to identify which factors have an impact on the service. Any network health metric needs to be ultimately linked to the customer impact.

A health score allows an operator to discern which alarms/events from the network are relevant. An operator would like to use a health score to prioritize problems and help triage problems and identify which areas/problems need truck-rolls.

One can think of various components to the health score, including downstream metrics upstream metrics node level metrics etc. Here we are just focused on the upstream component of an overall health score. We look at the individual CM level score and add a service group or node level score. At this point these upstream scores themselves could consist of multiple components though for this paper we are focused only on the RxMER data.

5.2.1. CM Score Upstream

Each of the parameters that we discussed above such as variance skewness, kurtosis, percentile, could all have different thresholds or areas of interest where we could flag the particular modem. One can also look at FEC statistics or profile/IUC statistics, and combined them into a CM score

5.2.2. Node Score Upstream

One can also aggregate all the measured metrics from all the modems on a given node and come up with an aggregate node score. The benefit of this view is for an operator to prioritize how a particular node is doing and how healthy is a node as compared to other nodes on the plant. This gives and operator a prioritized list of nodes which they can then work through the top issues.

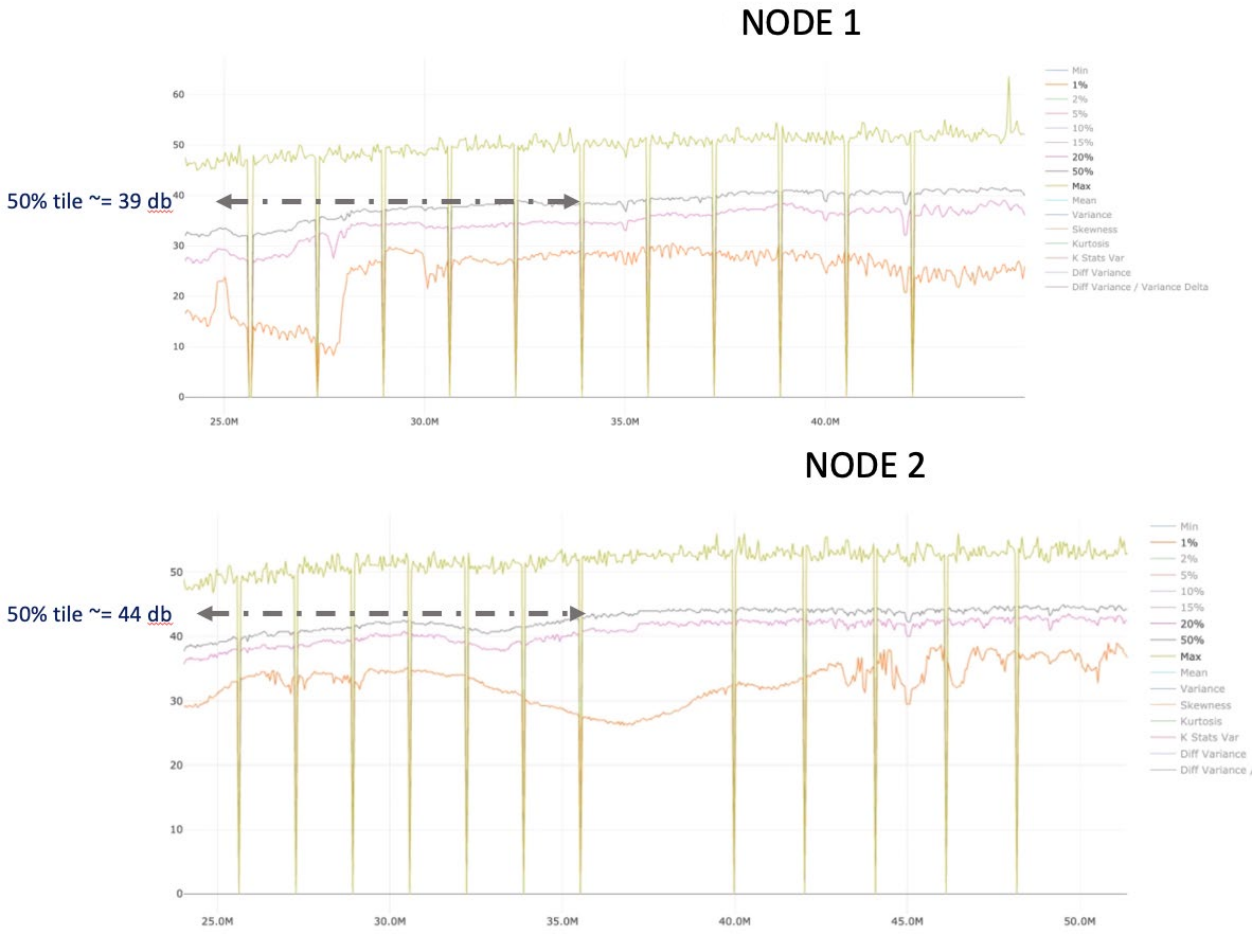


Figure 27 – All CMs All RxMER samples, Percentile, Node 1 vs Node 2

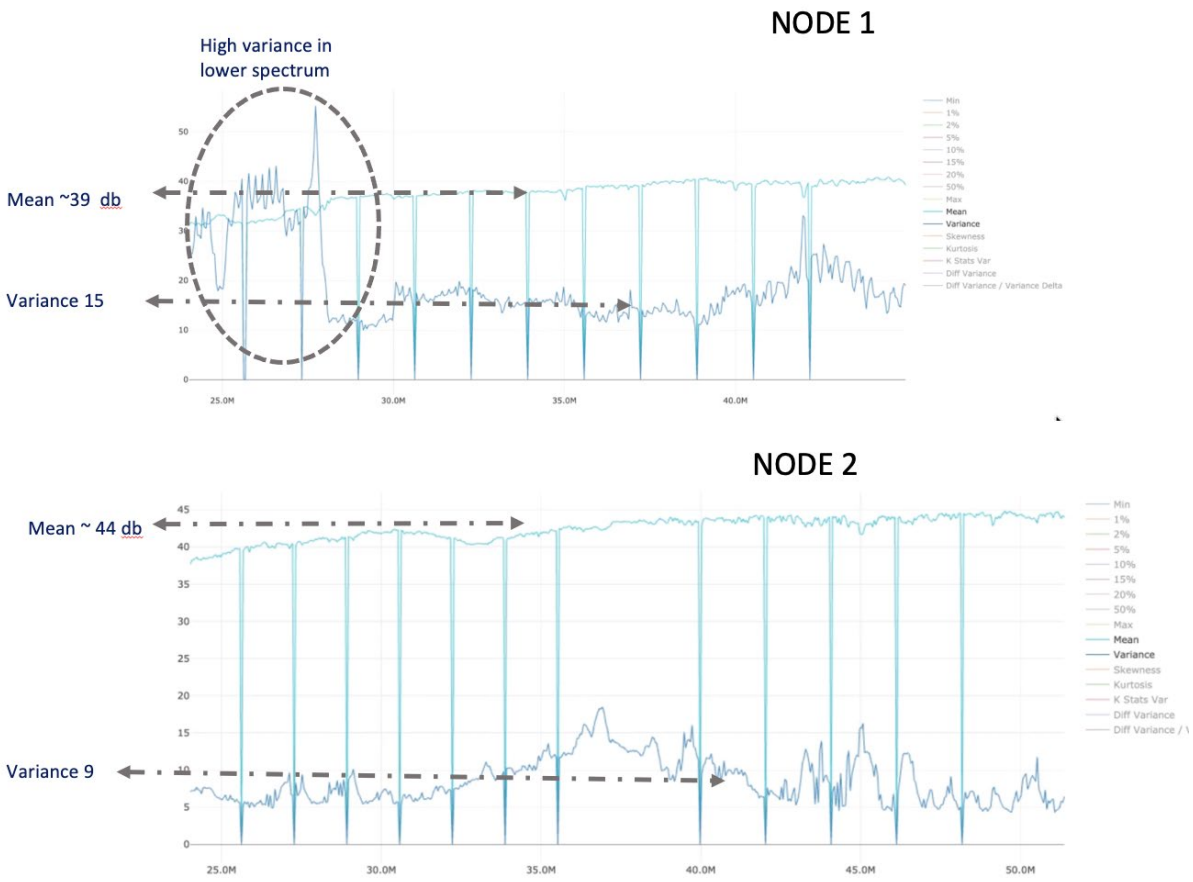


Figure 28 – All CMs All RxMER samples, Variance, Node 1 vs Node 2

5.2.3. US RxMER Data Analytics Application

CableLabs has developed an upstream RxMER data analysis application which can be used for a variety of functions [C3 CableLabs]. As shown in this paper one can visualize the RxMER of a cable modem, all the RxMER samples of a CM, all the RxMER samples of a node. There is also a table view off a service group where you can CB statistics off each individual cable modem. An operator can obtain the average variance, skewness, kurtosis in each part (frequency ranges) of the channel. An operator can also sort the CMs based on those values.

This is a very powerful tool in identifying the "outlier" CMs in the node/segment and then use that information for further analysis/ triaging of truck rolls etc.

CM MAC	Spacing	First Active Sub-carrier	First Active Sub-carrier Frequency	Interface Number	SG Name	Average Variance (lower spectrum)	Average Variance (higher spectrum)	Average Skewness (lower spectrum)	Average Skewness (higher spectrum)	Average Kurtosis (lower spectrum)	Average Kurtosis (higher spectrum)	Downtime Percentage	Gap Time Percentage	Show Data
08B0	50	74	24025000	34259153	AM	182.14	161.22	0.92	0.93	-0.38	-0.3	22.448	33.036	Show Data
08B0	50	74	24025000	34259153	AM	46.21	32.62	1.81	1.79	6.68	7.21	6.633	4.719	Show Data
7CDB	50	74	24025000	51003593	AM	29.22	25.73	1.54	1.43	9.03	6.4	4.081	2.679	Show Data
08B0	50	74	24025000	34259153	AM	27.42	24.05	-1.56	-2.21	5.34	8.64	0	0	Show Data
7CDB	50	74	24025000	51003593	AM	26.93	35.09	-0.45	-0.44	-0.54	-1.08	0	0	Show Data
94917	50	74	24025000	51003593	AM	26.02	24.24	0.38	0.29	6.28	3.55	1.148	0.256	Show Data
840B7	50	74	24025000	34259153	AM	25.03	18.81	-0.33	-1.03	2.16	5.93	0.255	4.591	Show Data
7CDB	50	74	24025000	51003593	AM	23.55	29.39	-0.45	-0.37	-0.25	-1.12	0	0	Show Data
840B7	50	74	24025000	51003593	AM	23.34	22.06	0.07	0.17	4.94	2.01	0.765	0.255	Show Data
840B7	50	74	24025000	34259153	AM	22.81	25.35	-1.17	-2.3	3.71	9.84	0	2.041	Show Data

Showing 1 to 10 of 113 entries

Previous 1 2 3 4 5 ... 12 Next

CM MAC	Spacing	First Active Sub-carrier	First Active Sub-carrier Frequency	Interface Number	SG Name	Average Variance (lower spectrum)	Average Variance (higher spectrum)	Average Skewness (lower spectrum)	Average Skewness (higher spectrum)	Average Kurtosis (lower spectrum)	Average Kurtosis (higher spectrum)	Downtime Percentage	Gap Time Percentage	Show Data
08B0	50	74	24025000	84570313	AM	4.15	2.37	-0.64	0.8	4.39	14.56	0	0	Show Data
7CDB	50	74	24025000	84570313	AM	7.03	4.42	-0.88	-0.65	2.7	3.16	0	0	Show Data
08B05	50	74	24025000	84570313	AM	7.04	4.84	-0.82	-0.87	2.67	4.17	0	0	Show Data
1CB0	50	74	24025000	84570313	AM	7.07	4.64	-0.85	-0.85	2.7	4.13	0	0	Show Data
840B7	50	74	24025000	84570313	AM	7.36	5.25	-0.92	-1.08	2.92	4.88	0	0	Show Data
08B05	50	74	24025000	84553937	AM	7.53	3.2	-0.39	0.8	0.54	2.92	0	0	Show Data
C8B42	50	74	24025000	84570313	AM	7.62	5.57	-0.97	-1.21	3.1	5.47	0	4.592	Show Data
9491	50	74	24025000	84570313	AM	7.74	5.34	-1	-1.19	3.18	5.32	0	0	Show Data
9491	50	74	24025000	84570313	AM	7.75	5.54	-1	-1.25	3.18	5.69	0	0	Show Data
9491	50	74	24025000	84570313	AM	7.92	5.72	-0.99	-1.34	3.04	5.88	0	0	Show Data

Figure 29 – All CMs All RxMER samples, Variance, Node 1 versus Node 2

5.3. Future Work

There are future topics we will continue to research on in order to develop advanced methods for automatic upstream OFDMA anomaly detection for cable operators to gain more visibility into the upstream OFDMA performance, potentially guide truck rolls and reduce network maintenance cost. These topics include defining anomaly categories in the upstream, anomaly detection methods, and correlations of different measurement data.

5.3.1. Defining Anomaly Categories in the Upstream

We observe many common impairments in the upstream RxMER such as the ingress-like noise observed between 15 and 27 MHz of the OFDMA channel as shown in Figure 30. However, the sources of the impairments mostly remain unknown today. Apart from this, the types and features of the upstream impairments vary from interface to interface, node to node, and area to area. In order to categorize the observed anomalies in the upstream, one may define the labels of the impairments in a hierarchical way. For example, the two parent categories could be persistent impairment and intermittent impairment. Under each of the parent categories, there could be sub-categories such as ingress, echo, or loose connector etc.

Defining impairment categories in the upstream OFDMA can provide critical insight into understanding performance affecting issues and how they need to be prioritized. It also guides the development of feature extraction methods and anomaly detection methods.

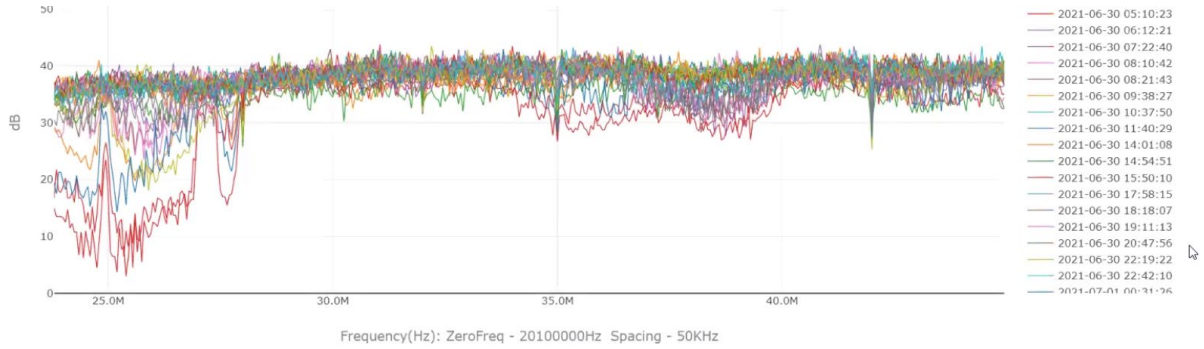


Figure 30 – Impairments at lower frequencies

5.3.2. Anomaly Detection Methods

Several different types of anomaly detection methods can be applied to the upstream RxMER dataset. Threshold based algorithms can be used to label impaired subcarriers and differentiate intermittent issues and persistent issues; 1-dimensional convolutional neural networks can be used to detect patterns over time (RxMER time series) or over frequency; 1-dimensional convolutional neural networks that has multiple input channels can be used to take aggregated data as input, for example, variance, skewness, and kurtosis of each subcarrier's RxMER values over time to perform pattern recognition of the anomalies; 2-dimensional convolutional neural networks can be used to directly consume RxMER sample captures over time (the 2 dimensions are frequency and time) and identify events with their frequencies and timestamps, similar to recognizing objects on pictures; Recurrent neural networks can be used to recognize patterns on RxMER time series. All of these candidate methods depend on well-defined categories/patterns and labeled datasets.

5.3.3. Correlation of Different Measurement Data

While the upstream OFDMA RxMER data already provides much information for performance monitoring, correlating it with other measurement data could have enhanced benefits. Upstream Pre-Equalization data can provide insights into how the CMTS configures the CMs to compensate impairments or other channel attributes; upstream triggered spectrum captures can provide detailed views of the upstream transmission status; upstream FEC error rates can be correlated with RxMER captures over time to analyze how much a certain impairment/event is affecting the robustness of the communication. All of these can be interesting research topics in upstream data analytics.

6. Conclusion

As OFDMA becomes more and more widely deployed, the upstream RxMER measurement capability is becoming available and mature on the CMTSs and MSOs have started to leverage this capability to gain visibility into the status and performance of OFDMA. By visualizing the upstream RxMER captures and producing different views of data, information such as upstream impairments and RxMER variations can be captured to support and drive research of upstream data analytics. In this paper, we discussed how statistical methods can be used to analyze RxMER values captured over time from each subcarrier and extract useful information from noisy upstream datasets. Intermittent issues and persistent issues can be isolated and differentiated by calculating variance, skewness and kurtosis on the upstream RxMER data captured over time. And by plotting the statistics in 3-dimensional space, we discover clusters that can potentially be leveraged to support automatic impairment detection/categorization. Analyzing the RxMER data captured over time by creating average RxMER plots for different frequency regions can provide views from another aspect of the data. Combining the RxMER over time data with identified CM reporting issues, MSOs can gain more visibility into the history of OFDMA channels' performance on individual CM basis as well as at node level. As of the goals of future research of upstream OFDMA data analytics, it is possible that more methods for feature extraction can be developed to simplify and aggregate measurement information. And intelligent models can be built to localize and classify upstream OFDMA impairments accurately and efficiently.

Abbreviations

3D	three-dimensional
bps	bits per second
CM	cable modem
CMTS	cable modem termination system
dB	decibel
DOCSIS	Data-Over-Cable Service Interface Specifications
FEC	forward error correction
HFC	hybrid fiber-coax
Hz	hertz
IUC	interval usage code
ISBE	International Society of Broadband Experts
kHz	kilohertz
MHz	megahertz
MSO	multiple-system operator
OFDM	orthogonal frequency division
OFDMA	orthogonal frequency division multiple access
PMA	profile management application
P-MAP	probe map
PNM	proactive network maintenance
RxMER	receive modulation error ratio
SCTE	Society of Cable Telecommunications Engineers
SC-QAM	single channel quadrature amplitude modulation
TFTP	trivial file transfer protocol

Bibliography & References

[DOCSIS PHYv3.1] DOCSIS 3.1 Physical Layer Specification, CM-SP-PHYv3.1-I18-210125, January 25, 2021, Cable Television Laboratories, Inc.

[DOCSIS MULPIv3.1] DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.1-I21-201020, October 20, 2020, Cable Television Laboratories, Inc.

[US PMA SCTE 2020] K. Sundaresan, J. Zhu, and J. P. Fernandes, "*Field Experiences with US OFDMA and Using US Profile Management", SCTE 2020*

[ANOMALY DETECTOR ICPHM 2020] J. Zhu, K. Sundaresan and J. Rupe, *"Proactive Network Maintenance using Fast, Accurate Anomaly Localization and Classification on 1-D Data Series", 2020 IEEE International Conference on Prognostics and Health Management (ICPHM), 2020, pp. 1-11, doi: 10.1109/ICPHM49022.2020.9187045.*

[C3 CableLabs] CableLabs Common code community, https://community.cablelabs.com/wiki/display/C3

Using AI in Network Planning and Operations Forecasting

Petar Djukic
Director AI & Analytics
Ciena Canada
Ottawa ON, Canada
pdjukic@ciena.com

Maryam Amiri
Lead AI Engineer
Ciena Canada
Ottawa ON, Canada
maamiri@ciena.com

1. Introduction

There are two main sets of costs in network planning that can be reduced through automation and artificial intelligence (AI):

- First, the process today is highly manual. It takes up the time of personnel who could be better utilized if the process was highly automated. Think of network planners with intrinsic knowledge of their network who spend most of their time updating Excel spreadsheets by hand. Their valuable experience could be used much better to maintain the network and to ensure that the customers are happy. ***Automation*** is a tool that can help reduce this set of inefficiencies.

- Second, network planning depends on the ability to forecast network traffic demands. Today this is typically done in an ad hoc fashion where an initial guess of the future network traffic demands is adjusted by essentially a gut feeling from multiple stakeholders. The problem with this approach is that it may result in many inaccuracies in the forecast. Overestimates in the forecast result in too much bought or deployed equipment and then underutilization of network resources. Underestimates in the forecast result in too little deployed equipment, lowered quality of service (QoS) observed by the customer and delayed service deployments, which must wait for equipment to be installed. ***Artificial intelligence*** (AI) is a tool that can reduce this set of inefficiencies.

We note that the two problems feed on each other and prevent either from being solved effectively. The first problem, which is one of automation is more obvious to network operators. As a network equipment vendor, we often hear of the customer woes caused by complicated planning spreadsheets. Through planning network deployments, we also observe first-hand the impact of inaccurate traffic forecasts. The two problems feed on each other as follows: the logic in one direction is “since we have to manually forecast network traffic there is no point automating the rest of the process as there will always be manual steps anyway” and in the other direction “since we manually do planning, there is no point automating the forecasting”. It is also quite possible that network operators are not fully aware of the latest forecasting tools, some of which we talk about in this paper and which can break this cycle of circular logic.

Our hypothesis is that increased automation and the use of artificial intelligence can reduce planning costs, while increasing service provider’s velocity and agility. The focus of this paper is on artificial intelligence and its automation, which should be a medium-term target for the industry. Network planning automation is a near-term topic which many vendors are addressing with their software solutions. There are many vendors advertising their ability to import spreadsheets and incorporate them into automation software, so we will not talk about this topic anymore.

Our main focus is on how AI can be used to automate the process. The goal of using AI technologies is automate as much of the network planning process as possible. Figure 1 shows how this could be done.

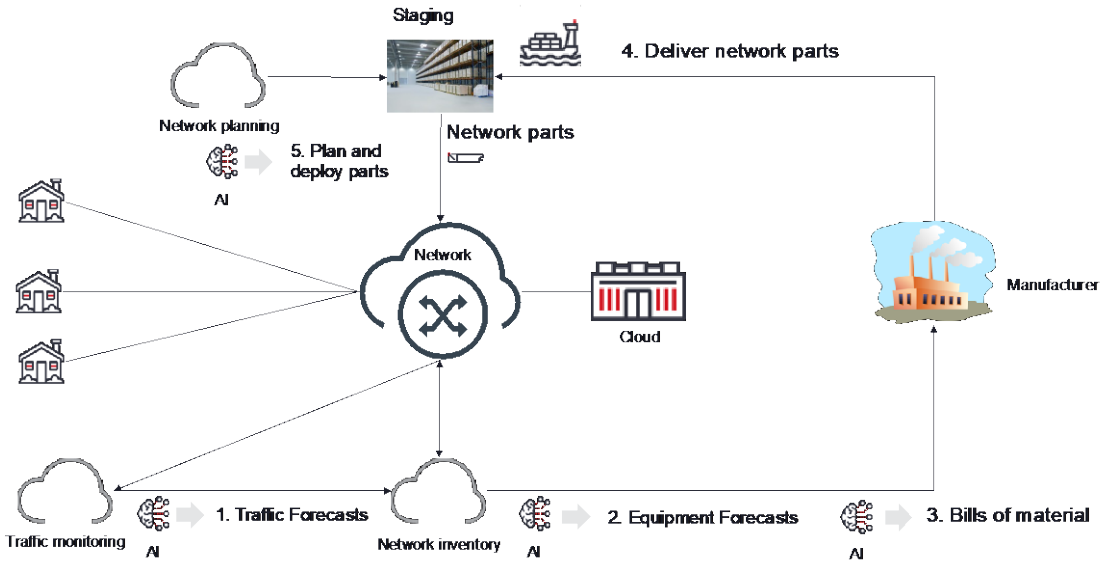


Figure 1 A Vision for a self-planning network

In Figure 1, the network connects the users to the cloud, where most of today's services reside. The traffic monitoring module collects network measurements, such as packet counters and network conditions (e.g., on fibers, or with IPFIX) and uses AI to create traffic forecasts (1). The traffic forecasts are passed to the network inventory module, which uses AI to create equipment forecasts – how much equipment to buy (2). The equipment forecast is passed on to another AI module, which optimizes costs and delivery times for equipment and creates bills of materials, which are sent to manufacturers (3). The manufacturer delivers network equipment to a staging area (4). Meanwhile, the network planning module uses AI to plan network deployments and dispatches technicians to install equipment from the staging area. Ideally, the only manual process is to install equipment and other parts of the process are fully automated with software and AI.

The rest of the paper is organized as follows. We start with a short description of how AI is implemented using deep neural networks (DNNs), following a description of a software architecture which is required to incorporate DNNs into network planning processes. Then we talk about forecasting techniques for network measurements, and we show some performance results using DNNs to forecast network traffic.

Throughout the paper we cite Wikipedia and AI blogs for various AI concepts. While this may appear to not be the most scientifically sound, we found these articles easy to follow and they always link to the more complete computer science papers for the keen reader. There is much DNN jargon used in the paper. We introduce DNN-specific terms in quotes "" to emphasize their jargon origins.

2. Foundations of AI technologies

AI technologies are based on the use of deep neural networks (DNN) for machine learning (ML). Machine learning is a computer science concept in which functional blocks are created by showing the computer examples of correct outputs from inputs, instead of explicitly writing functions that instruct the computer on how to produce outputs from inputs in structured programming (Wikipedia, n.d.). Instead of coding the algorithm, a generic machine learning algorithm is "trained" with examples of what the correct outputs are for given inputs. In recent years, DNN technology has elevated machine learning cognition to the level of human capability. For example, it is now possible to train a DNN-based machine learning algorithm to read a paragraph of text and answer questions about it more accurately than humans, or to categorize x-ray images better than radiologists (Zhang, et al.).

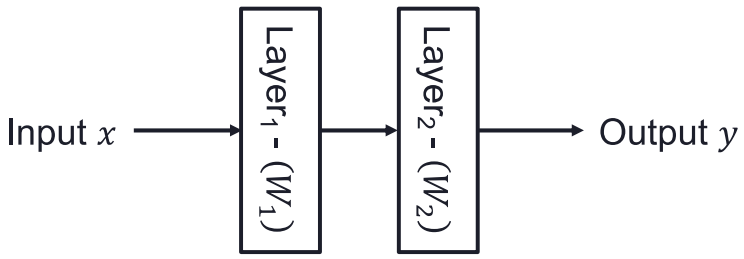
DNN technology is based on basic linear algebra components – matrix multiplication and addition and basic calculus –derivatives. Most of the knowledge required for DNNs has been around for hundreds of years since the time of Carl Friedrich Gauss (Wikipedia, n.d.) and Issac Newton (Wikipedia, n.d.). What is new at this time is that the advances in parallel computing have made it possible to deal effectively with large matrices and train very large DNNs. The most common computing platform are the Graphic Processing Units (GPUs) (Wikipedia, n.d.), which can be used even beyond DNNs, and they are now being complemented with Tensor Processing Units (TPUs) (Wikipedia, n.d.), which are specialized computing units for DNNs. AI accelerators such as TPUs are now found almost anywhere from being embedded in laptops, cellphones, and dedicated data center servers.

DNN-based AI technologies are bringing two main advantages to machine learning. First, as we mentioned DNN-based techniques are starting to perform tasks better than humans. Even though these tasks are limited in nature, it is still impressive to see this, and it opens the question of what other tasks DNN-based ML do better than humans. This is one of the questions we want to answer in this paper. Second, unlike other ML techniques DNN-based ML algorithms are highly automatable. This means that the role of humans in ML is now transferring from traditional manually intensive tasks such as feature engineering, feature selection and tuning of algorithms performed by data scientist, to more traditional software building, which automates these tasks. We explain these concepts shortly. An astute reader will notice that automation of ML is as important as the automation of the network operations to produce a self-planning network. The difference is that in ML this problem is well on its way to being completely solved.

For completeness and interest of the reader we now give a simplified overview of how DNNs make predictions and how they are trained. This is followed by an overview of the different aspects of DNN ML automation.

2.1. How DNNs make predictions

A DNN is a set of algebraic equations that describe how outputs are determined from inputs using matrix operations. A graphical version of the DNN representation is shown Figure 2a, which shows the most basic type of building block for DNNs, known as "dense blocks". The 2-layer DNN shown in the figure is shallow. A typical may have dozens of layers.



a) Graphical Description of a 2-layer neural network

$$y = \max(0, W_2 \times \max(0, W_1 \times x + b_1) + b_2)$$

b) Mathematical Description of the 2-layer neural network

Figure 2 An example 2-layer DNN used for forecasting

The network in shown Figure 2a evaluates an equation involving linear algebra shown in Figure 2b. In this example equation, $W_1 \times x$ is a matrix multiplication (Wikipedia, n.d.) and the max function ensures that the result of all operations is positive. Terms b_1 and b_2 are called bias for the layer. The input to the network is x, while the output is y, so the equation describes the functional block performed by the DNN. The output y is also called a prediction and in the case of forecasting, x is the historical time-series we are trying to predict while y is the future value we need to know for planning purposes. The input x is a mathematical vector whose components are called "features". Each feature is a separate input variable contributing to the output of the DNN. In the case of forecasting, "features" are past samples of the observed network measurements.

The simple set of algebraic transformations in this example is very powerful as it can be shown mathematically that a neural network with enough layers – depth – can approximate any function. For this reason, DNNs are known as "universal approximators" (Hanin & Sellke, 2018).

A pictorial description of a DNN shown in Figure 2a can be translated into the above equation in Figure 2b by an AI engineer and then made into a software program that performs the set of algebraic equations. In practice, the software piece is simple to write using libraries such as TensorFlow (Abadi, et al., 2015) and PyTorch (Paszke, et al., 2019). The DNN can also be exported into the Open Neural Network Exchange (ONNX) (Open Neural Network Exchange, n.d.), which describes the equations and can be loaded into many DNN software frameworks.

Going even further than the simple daisy chain we used, more complex structures could be built by using feedback – Long short-term memory (LSTM) networks and sparse matrices – convolutional neural networks (CNNs). LSTM networks are thought to be good for forecasting as they can model sequential nature of time-series where past values are related to future values. As it turns out, LSTM is not particularly good for network time-series as we show later in the paper. A more promising area for network time-series forecasting is the recent development of matrices that perform inverse fast Fourier transform – IFFT, matrices that perform the inverse

discrete wavelet transform – IDWT and matrices that perform polynomial functions. We talk about these later in the paper.

2.2. How DNNs learn

So far, we described the prediction or inference part of the DNN use. If we know the weights of the DNN (e.g., matrices W_1 and W_2 in Figure 2) then upon receiving the inputs, the set of matrix calculations described by the DNN is performed to determine the outputs. The outputs of the DNN are called the "predictions". This process of making prediction is sometimes called "inference". Weights are determined during a process of training.

For example, if we have the function

$$y = 2x^2 + 3x,$$

we can generate a dataset of training samples shown in Table 1 in the two left-most columns. With the dataset we can use a training function provided in open-source software such as TensorFlow (Abadi, et al., 2015) to determine a set of matrices W_1 and W_2 that result in the best approximation of the function. This function is called "fit" as it fits the weights of the DNN to the dataset during the training. The fit function minimizes the error of the predictions $\hat{y}$ for the dataset compared to actual values in the dataset y. The error can be measured with the Mean Absolute Percentage Error (MAPE) shown in the right-most column of Table 1.

Table 1 Example training dataset for $y = 2x^2 + 3x$

Input x	Actual output y	Predicted output $\hat{y}$	MAPE $\frac{\|\hat{y}-y\|}{y}$
1	5	4.5	10 %
2	14	15.6	11.5 %
3	27	25.4	5.9 %

This simple example may make it seem odd. Why is a DNN learning a known function when we could just be using the function itself? The true power of DNNs comes in when the function is not known but has been observed. In this case, the dataset is a set of observed values coming from the network and the underlying process that models it is not known. For example, the observed values could be packet counters or SNR measurements. In the forecasting case, the training procedure will determine a relationship between the past and future observed values. Once the DNN is trained it can be used for forecasting, by taking in past values and then giving out future values.

The great DNN research achievement in recent years has been to devise a training procedure that uses a set of inputs and outputs of a function to find the set of internal weights W to approximate the function. The training procedure uses "stochastic optimization" whose understanding essentially requires a PhD in mathematics or computer sciences. However, this understanding is not necessary to use DNNs as almost anyone who understands software development can write approximately 10 lines of code to create the DNN and train the function.

The training procedure is iterative and takes in a set of examples of inputs and outputs in batches. For each batch, the training procedure takes in the inputs and generates predictions using the current matrices. The predictions from the DNN are compared with the known outputs to find the error in the predictions (the "loss" function) and this error is used to adjust the weights. The adjustment is usually done using a gradient descent that takes a learning rate as an input and calculates the error of the predictions from the weights and number of predictions. The learning rate determines how quickly the descent happens and how closely to the best fit the training gets. The gradient of the whole DNN is calculated using "backpropagation" (Wikipedia, n.d.), which is an algorithm applied backwards through the DNN to differentiate it. Backpropagation is one example of automatic differentiation using the chain rule (Wikipedia, n.d.).

2.3. Tunning DNN Models

A DNN model is a trained DNN. A single DNN may have multiple models for different versions of the dataset, or different versions of the training algorithms. Each version may have the same structure (number of matrices, size of matrices, and flow through the matrices), but the weights may be different. In a parallel to software development, DNN models have different versions, which presumably improve with higher version numbers. Unlike software, a DNN model is not guaranteed to work well over time, as the inputs may have significant changes in their statistical properties. An example would be traffic demands changing if a new data center peering point is added to the network.

Tuning DNN models has historically been the task of data scientists. The tuning process involves four parts: feature engineering, feature selection, optimization of training hyper-parameters and selection of the DNN architecture. The reliance on data scientists for tuning of DNNs is now waning due to the introduction of automation, as we show in the next section.

Feature Engineering

Feature engineering is the process of modifying input variables to make them better during training and prediction. Historically, this was a very important part of machine learning and data scientists spent a lot of time on it. With the improvements in DNN technology, which during training learns the best representation of input variables, this process has become almost irrelevant. However, it is still important to scale the input and output variables to small range (typically between 0 and 1) to avoid numerical issues.

Feature Selection

Feature selection removes unimportant features. As the number of features increase so does the size of the DNN as each of the weight matrices needs to have the width of the vector it is multiplied with. A larger DNN size results in a longer time to make a prediction (and therefore to train the network). Most importantly, larger size means that the DNN training requires a larger dataset due to the "curse of dimensionality" (Wikipedia, n.d.). For example, if a DNN requires 300 samples per feature to be trained, adding 10 new features means that we need to have available another 3000 new training examples to get the equivalent performance. As the training data used for forecasting accrues historically, adding 10 new features to the inputs means that more measurements are required to train the DNN. Getting an extra 3000 new 15-minute

network samples for training may take as long as 4 weeks, so reducing training dataset size is a very important problem.

Feature selection is based on the premise that not all DNN inputs contribute equally to the output of a DNN. The basis of this premise is that during training, DNNs use the statistical correlation between input and output variables to deduce the best weights. Input variables uncorrelated to the outputs are not needed at the input to make predictions. Over the years, many feature selection approaches have been developed and this process can be automated through a search of required features (SciKit Learn, n.d.).

Selection of training hyper-parameters

Recall that during training input examples are grouped into *batches*, and that for each batch a gradient of the DNN is calculated using backpropagation and applied to the weights with a *learning rate*. The weights of the DNN are called "parameters" as they are determined during training, while the batch size and the learning rate are "hyper-parameters" as they are inputs to the training procedure.

Selecting the right batch size and learning rate are the easiest way to improve the performance of DNN during training. The process used for this is called hyper-parameter optimization or hyper-parameter tuning (Brownlee, n.d.) and is highly automatable and easily parallelized.

In the rest of the paper, we do not distinguish between training and hyper-parameter tuning. In practice, they are often combined into a single procedure.

Selection of a DNN architecture

The simple DNN example in Figure 2 uses two matrices of matrix weights W_1 and W_2, which are determined during the training process and tuned during the hyper-parameter search. However, the choice of the number and size of matrices may not be obvious for each data set.

At least two architectural parameters are unknown for even the simple example in Figure 2: the number of matrices and the size of each matrix. We used 2 matrices, which was easier to explain in this paper, but a typical DNN may have many more layers than that. The height of each matrix is also not readily known as only the width of each matrix is known (the height of the previous matrix). For more complex internal structures it gets complicated in terms of how the architecture is selected, and there is even an area of DNN techniques dealing with measuring the difference between architectures, called ablation (Wikipedia, n.d.).

To find the best architecture, in parallel to the hyper-parameter a network architecture search (NAS) (Wikipedia, n.d.) is also needed. The NAS space is much larger than the space of hyper-parameters, so this is much bigger problem to automate effectively. Typically, the search is done in unique directions, for example one direction could be a daisy chain of dense layers (shown in Figure 2a), while another direction may be a network with daisy chains of LTSM layers. In each direction the NAS is restricted to the number of layers and the height of each layer. Recently, this has been generalized in open-source software (OSS) by Google's Model Search (Google, n.d.).

3. Operationalizing DNNs with AI Software

We now pivot to perhaps a more interesting set of topics for network operators – how DNNs are operationalized with software. The main set of software available is free OSS. The AI software stack has evolved over the last 3 years in conjunction with the developments at the Cloud Native Computing Foundation (CNCF) (Cloud Native Computing Foundation: Building sustainable ecosystems for cloud native software, n.d.). The two main drivers are TensorFlow (Abadi, et al., 2015), which implements DNNs, and KubeFlow (Kubeflow: The Machine Learning Toolkit for Kubernetes, n.d.), which is a set of Kubernetes services used to create AI-specific distributed applications. Both OSS projects were initiated and are still strongly supported by large cloud providers (e.g., Google).

3.1. Microservices

The AI software stack is based on the concept of "microservices" (Wikipedia, n.d.), which are meshed with networking into distributed AI applications. This follows today's architectural patterns, where distributed applications based on microservices underpin today Software-as-a-Service (SaaS) software delivery model (Wikipedia, n.d.).

Microservices are implemented with Linux containers, which group Linux processes to have common permissions and resource limits. Communications with microservices are implemented with networking and commonly with a higher layer protocol such as the Hyper-text Transfer Protocol (HTTP) implementing a Representational State Transfer (REST) architectural style (Wikipedia, n.d.). The microservices are also called RESTful as they are assumed to not hold state between subsequent REST application programming interface (API) calls. Each micro-service exposes its API through Uniform Resource Locators (URLs) corresponding to each of its available functions. Microservices in the same distributed application communicate through networking, so they are not guaranteed to run on the same server, or even in the same datacenter.

Microservices can be combined into distributed applications, where there may be layers implementing different functionalities (e.g., the web interface layer and the database layer). To improve security inside the distributed applications, microservices-based distributed application use a "service mesh" (Wikipedia, n.d.) – an overlay topology interconnecting the microservices in a security conscientious way. As the RESTful communication approach does not provide stateful transfer of data, a messaging bus (service) may be employed to simplify communications between microservices in the same distributed application.

The distributed microservices application architecture is very common in cloud applications, and it has been driven by OSS delivered through the CNCF (Cloud Native Computing Foundation: Building sustainable ecosystems for cloud native software, n.d.). The key contribution to CNCF came from Google in the form of Kubernetes (Kubernetes: Production-Grade Container Orchestration, n.d.), which is container resource orchestrator. Most of the work of the CNCF revolves around building software required to make distributed applications managed by Kubernetes easy to make.

3.2. DNN models as microservices

A trained model is described with an ONNX file, which contains the description of the DNN layers and the weights of each layer. ONNX files can be read by all major DNN software libraries and the used to make predictions. An ONNX file does not make predictions, it only describes a network and its weights. When loaded into ONNX hosting software, the software provides a way to make predictions with the stored DNN model. Inside software, a DNN model is used with a function "predict", which takes as an input features x and outputs an estimate $\hat{y}$.

The functional DNN model fits with the RESTful approach of stateless services. The DNN model does not hold state between subsequent prediction calls. The model can be served on a unique URL with an HTTP post request carrying x and the serving microservice returning the prediction $\hat{y}$. This serving functionality is available in all major DNN software distributions. For example, TensorFlow Serving (Tensorflow: Serving Models, n.d.) and TorchServe (PyTorch: TORCHSERVE, n.d.), provide generic serving containers, which can load a DNN model from an ONNX file and then provide a service access point for the model's predict function. In a distributed application, a serving container becomes is a microservice serving multiple models, each with a unique URLs.

3.3. AI architecture

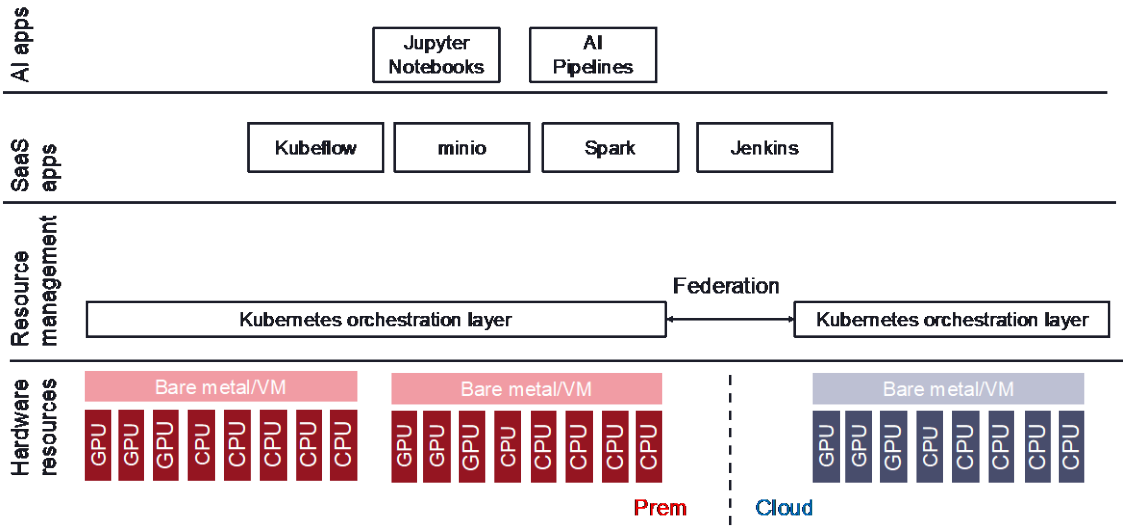


Figure 3 Example AI software stack

The AI software stack is shown in Figure 3 and contains several microservices-based distributed applications running over a distributed hardware architecture. All architectural components in the figure are OSS and are taken from CNCF and TensorFlow family of software. The main AI OSS component, KubeFlow, is spearheaded by major cloud companies, who use it in their cloud and are the basis of their automatic ML (AutoML) offerings. Only the knowledge of Kubernetes is required to setup and maintain the software. It does not take very long to install and set up the stack, or to apply software upgrades. In our opinion, building AI software without using the

Kubernetes and KubeFlow OSS ecosystem would be a strategic mistake, which may result in much unneeded development efforts and may likely result in a complete overhaul of AI software architecture a few years later.

3.4. Architectural Layers

The “hardware resources” are shown at the bottom of Figure 3. They may reside in a private cloud, on the premise close to networking equipment (the edge cloud), or in the public cloud. Hardware resources are made available through the Kubernetes API. We note that the hardware resources are not homogeneous or equally distributed. For example, there would be many more compute and storage resources in the cloud then on the premise (“infinitely” so). The Kubernetes resource manager uses the resources according to their cost and latency requirements of the distributed application. Specifically, one would expect to use edge or network resources first if they are available, to achieve best latency, and spill over other less latency constrained workloads into the cloud.

The “resource management” layer in Figure 3 is a federation of Kubernetes container orchestrators, which manages the hardware resources across multiple clouds. There is a Kubernetes instance associated with each separate set of hardware resources (e.g., cloud or edge). The Kubernetes orchestrator keeps track of available hardware resources and allocates the resources requested for each microservice and spins up the microservice on those resources. Separate instances of Kubernetes can be federated to allow for a seamless use of resources regardless of their location. Kubernetes also provides services useful for a microservices-based architecture: a domain name service (DNS) which maps service URLs to IP addresses, load balancing, automatic scaling based on measurements of service latency, monitoring of microservices health, and restarting them when necessary, and a global registry implemented as distributed key-value store. In terms of resource separation and security, Kubernetes provides namespaces which are logically separated groups of microservices.

The “SaaS apps” layer in Figure 3 is a set of distributed applications installed on top of the Kubernetes resource manager, which are required to build distributed AI. For illustration purposes, we show some of the more useful services for AI: MinIO (MinIO: Object Storage for the Era of the Hybrid Cloud, n.d.) can used to store datasets, by providing an S3 (Amazon S3: Object storage built to store and retrieve any amount of data from anywhere, n.d.) compatible object storage, with the ability to store objects (files) transparently on the local servers or in the private or public cloud; Spark (Apache Spark: Lightning-fast unified analytics engine, n.d.) is a distributed in-memory analytics engine capable of processing vast amounts of data, so it can use to process datasets; Jenkins (Jenkins: build great things at any scale, n.d.) is an integration and delivery automation software; and KubeFlow is an AI pipeline orchestrator and DNN model tracker. A Kubernetes cluster may have many other services co-existing with these, depending on how it is used.

The “AI apps” layer is at the top level of the AI software stack in Figure 3. AI apps are built using the KubeFlow SaaS layer. KubeFlow provides facilities for launching of AI specific containers hosting Jupyter Labs Notebooks (Jupyter: Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of

programming languages, n.d.) and for creation of AI pipelines, which are distributed applications used for training and validation DNN models. KubeFlow also contains services for tracking and delivery of DNN models.

3.5. AI Pipelines are DNN model factories

AI pipelines are dynamic distributed applications that train DNN models. Each model has its own pipeline, which includes data processing, training with NAS and hyper-parameter selection, and validation steps. Figure 4 shows an example AI pipeline. A pipeline is specified using KubeFlow's domain-specific language (DSL), which describes the containers to be used at each stage of the pipeline. In the example, there is a container for processing data, container to train a model and a container to test a model. When the pipeline is started, Kubeflow spins up the containers in the order specified and ensures that each part of the pipeline finishes before containers for the next part of the pipeline are spun up.

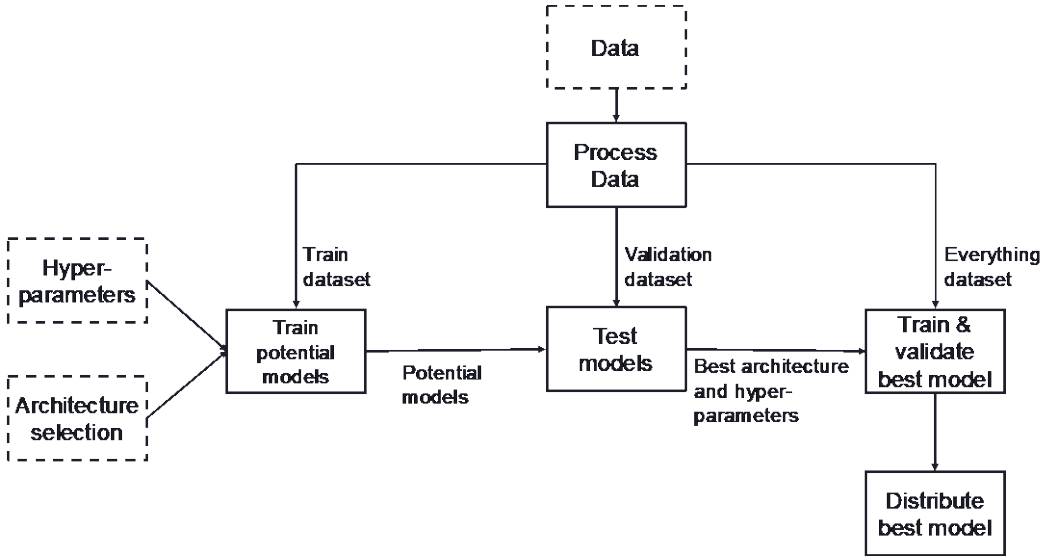


Figure 4 An Example AI pipeline

In the example, data is input to the "process data" step, which requires KubeFlow to provide the data to the process data container, which it also spun up for the pipeline. In practice, data may be in an S3 bucket stored locally or in the cloud. This data processing step may use SaaS services of Spark to process a large volume of data. Note that the data is split into a training and testing dataset to follow common machine learning methodology. The test dataset is provided to the training step, which uses NAS and hyper-parameter search to train many DNN models. This part of the pipeline uses Katib (Kubeflow, n.d.) to create many instances of the same training container with different inputs and runs them in parallel on the Kubernetes cluster. Potential models are evaluated using the validation step. The models are ranked based on the performance of the loss function (error) as shown in Table 1. The DNN model with the lowest error corresponds to the best architecture and hyper-parameters for the training dataset. These are passed to an instance of the training container, which trains the best model with all available

data. The model is validated using hold out data from the dataset. The final step is to distribute and serve the best model.

Figure 5 shows the relationship between AI pipelines and the regular network applications. On the left side of the figure, AI pipelines are a special type of distributed application, used specifically for DNN model search with NAS and hyper-parameter tunning. In our software stack, AI pipelines exist in their own Kubernetes namespace. Each pipeline implements a specific AI use case. For example, network traffic forecasting, which is the topic of this paper, is one use case, which is quite different from other use cases. We show two other use cases in Figure 5: an AI use case that trains a DNN for optical signal-to-noise SNR calculations; and an AI use case that trains a DNN model for network design and planning. The latter may be a DNN model implementing a reinforcement learning approach for path selection in the network. The output of an AI use case is a DNN model. As the DNN models are trained in a namespace different from where they run, then stored in a DNN model registry, which is accessible from all namespaces.

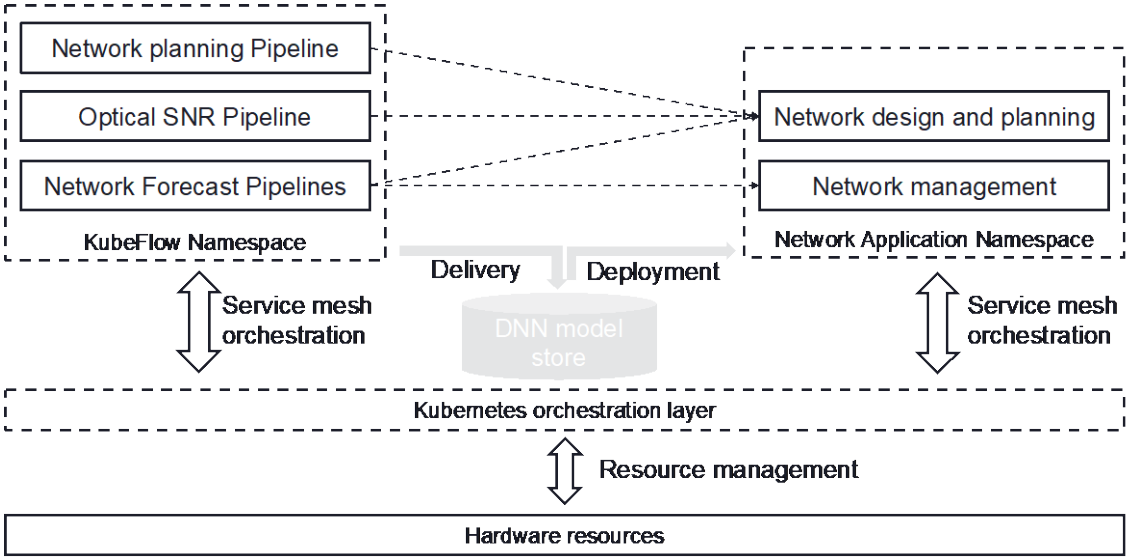


Figure 5 AI pipelines and other distribute network applications

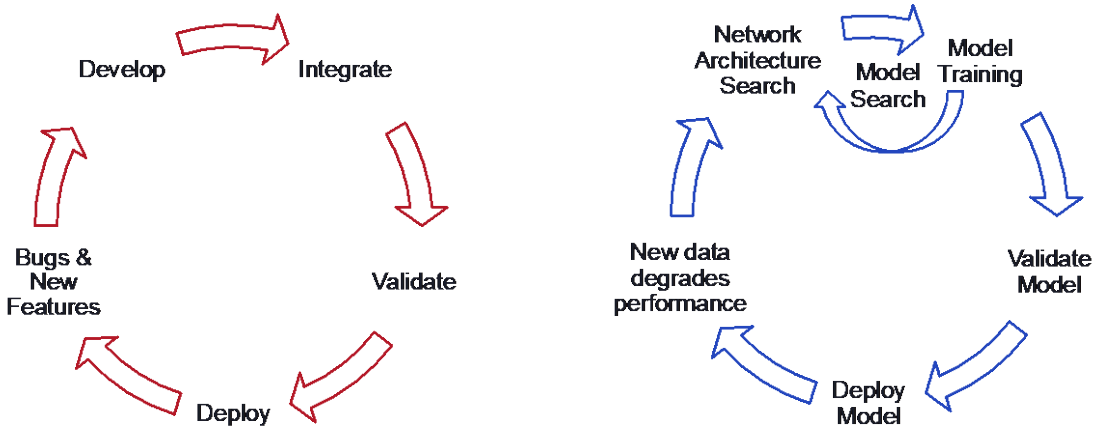
The right side of Figure 5 shows the network applications using the DNN models trained by the pipelines on the left side. The dashed arrows indicate where pipelines deliver models. Note that the network forecasting is many applications, as forecasting could be done at different scales for planning, or management. The model is delivered to the DNN model store and then it is deployed in the corresponding network application.

In the example, we assume that the network applications are hosted on the same Kubernetes cluster. This could be done in separate namespaces as shown in the figure to assure that the network applications are secure and have guaranteed resources. The figure shows several pipelines, the “network planning pipeline”, the “optical SNR pipeline” and the “network forecast pipeline”. Each pipeline produces a DNN model for a specific purpose and with a separate

dataset. To produce the model, pipelines may use the same pipeline template (e.g., template in Figure 4), or they may have different pipeline templates. For example, one of the pipelines may use transfer learning (Wikipedia, n.d.), while other pipelines may use reinforcement learning (Wikipedia, n.d.) to produce their DNN models.

3.6. The MLOps Cycle

The architectural relationships in Figure 5 create an opportunity for MLOps. MLOps (Google) are the equivalent of DevOps (Wikipedia, n.d.) for AI. DevOps is generally accepted way of creating cloud-based software application with a tighter integration of software development and delivery, than what has been done historically. Integral to DevOps is the concept of combined Continuous Integration (CI) and Continuous Delivery (CD). The process is typically done with an automation tool like Jenkins. The CI/CD process is shown in Figure 6a. As the code is being developed, it is automatically tested and integrated and then validated. Testing and integration are done in the development environment, while the validation is done in the testing environment. If the new software is compliant with validation tests it is transferred to the production environment. Even in the production environment the software can be tried out before full deployment. This can be done with Kubernetes, which is instructed to only deploy some portion of the newly created containers into production and load-share application between the new and old containers. Only if the new containers perform satisfactorily, the new containers replace all the containers.



a) DevOps Cycle with CI/CD b) The MLOps Cycle

Figure 6 DevOps vs. MLOps

DevOps are enabled by the ability to automate testing, integration and validation and then automatically deliver containers to the cloud. This ability is right now coming directly from the use of microservices and Kubernetes, combined with CI/CD automation. Microservices architecture allows incremental upgrades (one micro-service at a time). Kubernetes provides a platform to test microservices through namespaces. For example, with Kubernetes it is possible to easily create namespace where different versions of the same microservices are logically separated on the shared hardware and incremental testing of new features and bug fixes can be

done. Kubernetes also provides facilities to gradually upgrade a microservice in production by gradually replacing copies of old versions with new upgraded ones and to monitor the new versions and to roll them back if problems are noticed.

The MLOps cycle is shown in Figure 6a. This is the equivalent to CI/CD for DNN models, with some fundamental differences. The first difference is that the DevOps cycle is triggered by bugs in the code or new features, while the MLOps cycle is triggered by new data, which degrades the performance of the existing DNN model. The second difference is that the DevOps cycle starts by development of the code, while the MLOps cycle starts with a network architecture search and model training, which includes hyper-parameter optimization. Instead of coding a new DNN model, the new model is found automatically using new data. Validation and deployment parts of the cycle are essentially the same. The new model is packaged as a microservice and is updated in the service mesh just like any other microservice.

3.7. Automatic Machine Learning (AutoML)

The combination of AI pipelines with NAS and hyper-parameter search is known as automatic machine learning (AutoML). The search for the best combination of data processing, network architecture and hyper-parameters is completely automated and the DNN model factory in Figure 5 can produce DNN models without human input.

While this may sound magical as the data scientist expertise and experience is greatly reduced and removed from the ML process, we note that there are several points to AutoML that require some human input:

- First, the NAS and hype-parameter space may be quite large. This means that the search space could be so large that the time to create a new model may exceed operational needs. A data scientist or an AI Engineer would be involved to decide on the *reasonable* search space that may generate *good* DNN models.
- Second, model performance in production must be monitored and evaluated as data coming from the network changes over time. A sufficient change in the network data will trigger the MLOps cycle. However, a data scientist still needs to monitor the performance and its trends and make judgment calls when to trigger the MLOps cycle or debug DNN model performance for causes of degradation.
- Third, the role of AI models in the distributed network applications needs to be well understood. This requires a human contribution in understanding the domain of application and to design the system that automates downstream actions of network operators.

So, while data scientists will still be needed in the world of AutoML, their role in the process will change. Instead of each data scientists training a single model over a period of weeks or months, that data scientist will be able to train hundreds of models in a day and debug the select few for performance issues. Data scientists may also take on the role of a translator (Analytics translator: The new must-have role, n.d.) to bridge the gap between business needs and AI technological capabilities.

4. Forecasting the Network Demands with Artificial Intelligence

So far, we have talked about existing and soon-to-be-available AI architectures and components. This section talks specifically about how DNNs can be used for forecasting in the network.

Forecasting is the process of taking in historical values for a network measurement and producing estimates of future values. Network measurements are stored in databases as time-series, "series" is used because measurements are collected uniformly, so only a sequence of measurements is required and time can be calculated by a position in the series.

Several kinds of forecast are possible, depending on what the inputs and outputs are:

- Single-variate forecasts are for time-series containing a single measurement, for example packet counts from a single router link.
- Multi-variate forecasts are for time-series containing multiple measurements, for example packet counts from multiple links on the same router.
- Single-step forecasts produce the next value in time, so if packet counters are collected every 15 minutes, a forecast at 8:00AM will be for the packet counter at 8:15AM.
- Multi-step forecasts produce a sequence of future values, so if packet counters are collected every 15 minutes, a forecast at 8:00AM can be for the next day and produce 96 values (for every 15 minutes in a 24-hour period).

There are only 4 valid forecast types, given the two kinds of inputs and outputs.

The content of this section is best suited for single-variate forecast of single-step or multi-step kind. However, we also show results for multi-step forecasts.

4.1. Forecasting network traffic

Forecasting approaches are based on behavioural models of underlying processes that reflect in the data. For example, some generalities about human generated data are almost always true: human activity almost always increases, and human activity is seasonal. The growth in human activity is especially reflected in the growth of Internet traffic, which seems to only go up (Cisco, n.d.). There is also seasonality in the Internet traffic (Hen & Karlesson, 2019) reflected both in which direction the network traffic flows and in the volume of traffic. For example, the difference is pronounced during a single day with high business traffic during work hours and low business traffic during the evening and high entertainment traffic during the evenings, which are also temporarily different across the world. Generally, seasonality is also noticeable during different days of the week with weekends and holidays showing high consumption of entertainment and low business traffic.

4.2. Traditional forecasting approaches

A generally accepted forecasting approach is to decompose the time-series and extract the trend and seasonality components (Hyndman & Athanasopoulos). The reminder of the time-series includes noise, which needs to be smoothed out. Figure 7 shows the decomposition process. The input is historical time-series. The trend is extracted first, followed by the seasonality. The remainder of the time-series is smoothed. The trend can be used to make long-term forecast (in the order of months, quarters, and years), while the seasonality can be used to make forecasts for

repeating patterns in the time-series (daily, weekly). The smoothed remainder can be used to make short-term forecasts.

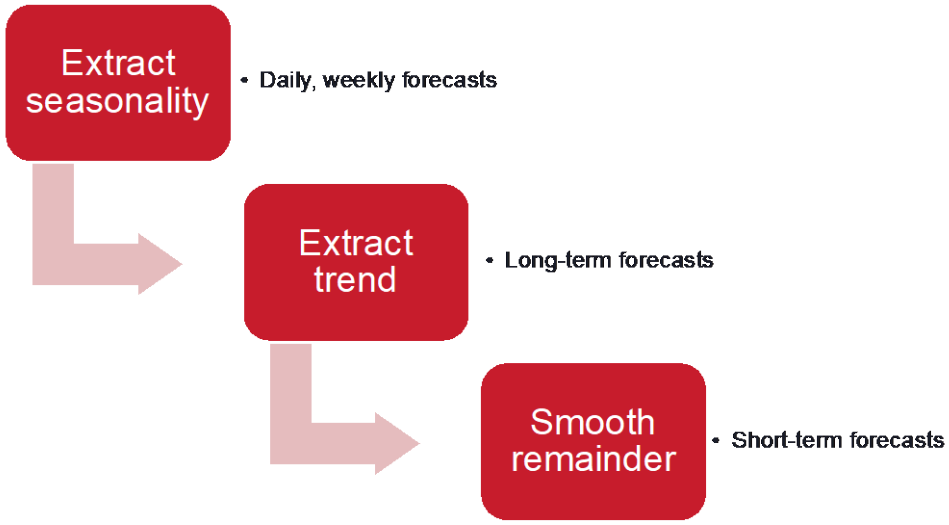


Figure 7 Forecasting time-series decomposition

Traditionally, time-series decomposition has been a highly manual process. The time-series structure is not known before analysis done by someone with a lot of forecasting expertise. The decomposition can be done automatically using one of the newer tools (Prophet: Forecasting at Scale), however a forecasting expert should still ensure that the automation has gone smoothly as the forecasting model in (Prophet: Forecasting at Scale) makes very specific assumptions about the time-series. The forecaster makes an informed decision for each of the time-series components trend, seasonal and smoothing. A typical forecasting expert works with a small dataset and uses judgment in deciding which forecast makes the most sense. For example, forecasting quarterly GDP over the last 50 years only has 200 points and a forecasting expert may use their knowledge of economics and information available outside of the time-series to decide if GDP will go up or down in the next quarter. Compared to network data, this is very small dataset – there are close to 200 data points in two days of 15-minute bins. So, some of the analysis done by traditional forecasting technique doesn't translate well into the networking domain.

In terms of time-series feature, an expert forecaster may use some of the following approaches:

- Seasonality can be estimated in many ways including taking averages at repeating time, e.g., by find an average traffic on Monday 9:00AM-9:15AM, by isolating data in this repeating period of time and then averaging, or with more complex approaches such as estimating the periodicity in the frequency domain.
- Trend is typically chosen as the best line that fits through the data. To make things simple to understand, a forecaster usually picks linear or exponential trend (Wikipedia, n.d.), but more complex methods such as piece-wise linear trends (Hyndman R. J.) are also possible.

- Smoothing is used to remove the noise from the time-series. A familiar from of smoothing may be the moving average (MA) smoother available on stock tracking platforms (Ermey, 2019). The moving average is also statistically the most likely estimate of the next value in the time-series. More sophisticated smoothers also exist, namely the auto-regressive moving average (ARIMA) family of estimators (Wikipedia, n.d.), which model temporal relationships between samples of a time-series. Some forecasters feel confident in using
  ARIMA forecasters for multi-step forecasts, however assumptions on the underlying data should be checked before getting overconfident with this method.

It is important to take stock of the state-of-art in existing forecasting approaches: they are highly dependent on matching the model in the forecasting algorithm to the time-series, because historically there wasn't that much data available for forecasting. Both reasons are why human intervention is required for traditional forecasting approaches.

As we have shown earlier, one of the advantages of using DNNs is that they can learn the best model for a time-series, given enough data. In the networking use cases, there is enough data. We now go over some of the approaches that can be used to automate forecasting with DNNs.

4.3. Forecasting with DNNs

Probably the most accepted DNN forecasting approach is to use recurrent neural networks (RNNs) (Wikipedia, n.d.). RNNs use recursion to pass data from the previous step to the current step as shown in Figure 8. In the figure the DNN block is identical throughout the network, meaning that recursion happens, and the last output is the function of all inputs. The RNN structure models time dependencies in the time-series. The block could be made from anything, but a popular structure long-term short memory (LSTM), which is numerically stable.

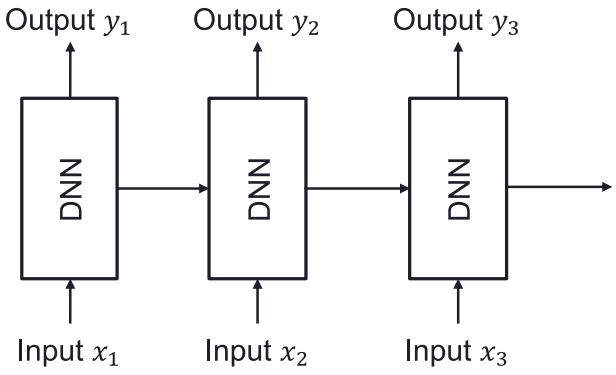


Figure 8 Recursive Neural Network

The assumption of the recursive relationships of time sample is reminiscent of that ARIMA forecasting models. As we already mentioned, this approach is valid for short term forecasts (smoothing) after trend and seasonality have been removed. The assumption of the recurring

relationships in the time-series isn't always true, or important for time-series forecasting. The more important assumptions are around trend and seasonality, which should be accounted for.

Recently, a residual network structure (Wikipedia, n.d.) was used to model this behaviour. In a residual network, it is possible to have skip connections, that allow addition or subtraction of various output blocks. The N-BEATS doubly residual network architecture (G, n.d.) for time-series forecasting is shown in Figure 9. Following from left to right, on the bottom, it can be seen that what is happening in the network is a trend estimate being determined first and then being subtracted from the input. Then the seasonality estimate is determined and subtracted from the "residual" of the input (input with trend subtracted). Finally, the residual of both of those is smoothed out. On the top of the network, the outputs of the three estimates are combined for the forecast.

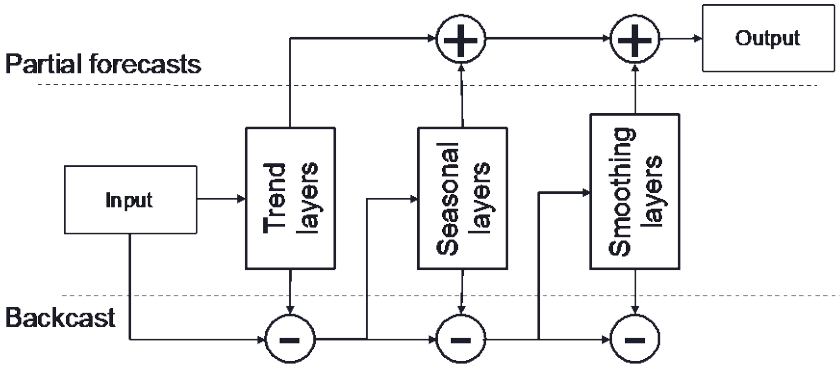


Figure 9 Time-series decomposition

As an example, we used the NBEATS (G, n.d.) network and compared it to the winner of the M4 forecasting competition (The M4 Competition Team, n.d.) dataset. The winner of the competition was a modified LSTM, which was used to win the competition and it required human intervention to perform well. The M4 forecasting competition dataset has over 100,000 financial and manufacturing time-series. Algorithms are compared against a simple baseline, which is used to ensure that there is something to learn in the dataset. The baseline algorithm uses the last seen value as the forecast of future values. It may be hard to believe, but this simple baseline is very hard to beat on real-life datasets.

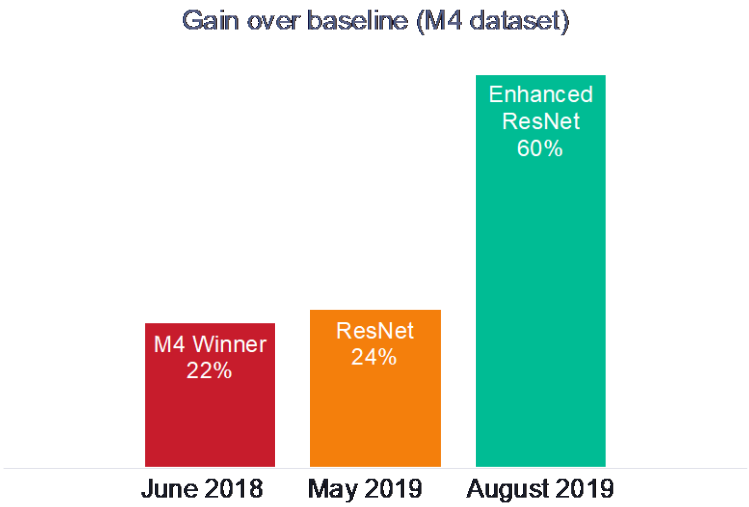


Figure 10 Performance comparison of DNN approaches (public dataset)

In the Figure 10, NBEATS is labeled as ResNet. The relative gain is measured on the error of one method over the other method. So, 22 % gain of the M4 Winner over the baseline means that the error (MAPE) of the winner is 22 % less than the winner. It should be observed that NBEATS improves upon the M4 Winner, but not by a lot. Based on our observations of the NBEATS performance, we developed a new ResNet based algorithm and it shows an improvement of 60% over the baseline (labeled Enhanced ResNet). As far as we can tell, the Enhanced ResNet is the world's best forecaster as it works better than the M4 competition winner and the next best forecaster.

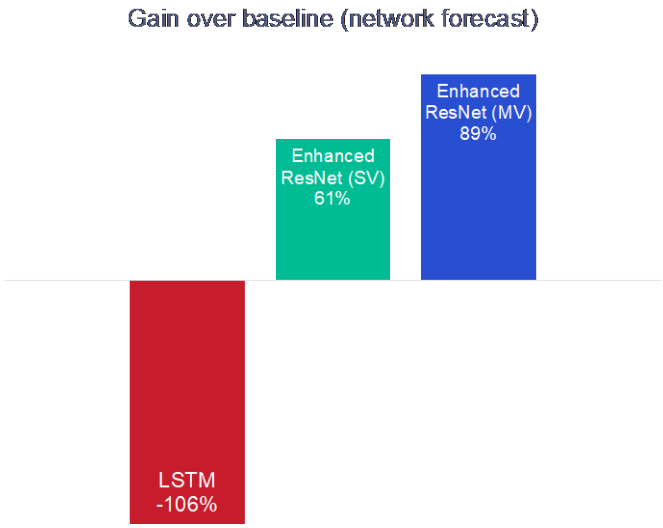


Figure 11 Performance comparison of DNN approaches (network dataset)

Real-world network datasets are quite different from the financial series in the M4 dataset. For the most part, financial time-series are smooth and without sudden changes. On the other hand, network time-series often have unexpected steps and spikes. We tested the LSTM and Enhanced ResNet algorithms on a network time-series in our possession. The data was collected from a large service provider. The results are shown in Figure 11. Note that LSTM does not work well on this dataset. It performs much worse than the baseline. We are confident this is because LSTM assumes too many dependencies on time, so it doesn't work when the time-series have abrupt changes.

On the other hand, the Enhanced ResNet algorithm (shown as "Enhanced ResNet (SV)") works much better than the baseline. Furthermore, the Enhanced ResNet algorithm can be extended to forecast from multi-variate inputs (shown as "Enhanced ResNet (MV)"). Network time-series have dependencies on other time-series, so it is advantageous to combine them during the forecasting process.

5. Summary

This paper has talked about forecasting in the context of network operations. We have made the argument that network operations, especially its planning functions, need to automate all parts of their processes and that this cannot be done without automated forecasting. We have shown how automated forecasting can be done with DNNs and AutoML. We have also used network time-series from an actual network to show the power of forecasting with DNNs.

We believe that DNNs and AutoML have the potential to revolutionize the network planning process and make it more accurate and less costly than today.

Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
CI/CD	Continuous Integration (CI) and Continuous Delivery (CD)
CLI	Command-line interface
CNN	Convolutional Neural Network
CNCF	Cloud Native Computing Foundation
CPU	Central Processing Unit
DNN	Deep neural network
DNS	Domain Name Service
DSL	Domain-specific language
GPU	Graphic Processing Units
IDWT	Inverse Discrete Wavelet Transform
IFFT	Inverse Fast Fourier Transform
IGP	Interior gateway protocols
IP	Internet Protocol
HTTP	Hyper-text Transfer Protocol
IPFIX	IP Flow Information Export
LSTM	Long short-term memory
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
NAS	Network Architecture Search
ONNX	Open Neural Network Exchange
OSS	Open-source software
PoP	Point of Presence
RCA	Root Cause Analysis
RNN	Recurrent Neural Network
SaaS	Software-as-a-Service
S3	Simple Storage Service
SP	Service Providers
TPU	Tensor Processing Unit
URL	Uniform Resource Locators

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . Jozefowicz, R. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* Retrieved from Software available from tensorflow.org: https://www.tensorflow.org/

Amazon S3: Object storage built to store and retrieve any amount of data from anywhere. (n.d.). Retrieved June 5, 2021, from https://aws.amazon.com/s3/

Analytics translator: The new must-have role. (n.d.). (Harvard Business Review) Retrieved June 9, 2021, from https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/analytics-translator

Apache Spark: Lightning-fast unified analytics engine. (n.d.). Retrieved June 5, 2021, from https://spark.apache.org/

AWS. (n.d.). *Amazon S3: Object storage built to retrieve any amount of data from anywhere.* Retrieved 07 21, 2021, from https://aws.amazon.com/s3/

AWS. (n.d.). *AWS Simple Monthly Calculator.* Retrieved from https://calculator.s3.amazonaws.com/index.html

Brownlee, J. (n.d.). *Hyperparameter Optimization With Random Search and Grid Search*. Retrieved June 2, 2021, from https://machinelearningmastery.com/hyperparameter-optimization-with-random-search-and-grid-search/

Cisco. (n.d.). *Cisco Annual Internet Report (2018–2023) White Paper*. Retrieved June 9, 2021, from https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

Cloud Native Computing Foundation: Building sustainable ecosystems for cloud native software. (n.d.). Retrieved June 5, 2021, from https://www.cncf.io/

Ermey, R. (2019, January 31). *The Magic Of Moving Averages*. Retrieved June 10, 2021, from https://finance.yahoo.com/news/magic-moving-averages-173300478.html

G, K. (n.d.). *N-BEATS: NEURAL BASIS EXPANSION ANALYSIS FOR INTERPRETABLE TIME SERIES FORECASTING*. Retrieved June 11, 2021, from https://kshavg.medium.com/n-beats-neural-basis-expansion-analysis-for-interpretable-time-series-forecasting-91e94c830393

Google. (n.d.). *Google Model Search*. Retrieved June 2, 2021, from https://github.com/google/model_search

Google. (n.d.). *MLOps: Continuous delivery and automation pipelines in machine learning.* Retrieved June 6, 2021, from https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

Hanin, B., & Sellke, M. (2018). *Approximating Continuous Functions by ReLU Nets of Minimal Width.* Retrieved from https://arxiv.org/abs/1710.11278

Hen, H., & Karlesson, N. (2019, July 10). *Identification of Seasonality in Internet Traffic to Support Control of Online Advertising.* Retrieved June 10, 2021, from https://research.yahoo.com/publications/9113/identification-seasonality-internet-traffic-support-control-online-advertising

Hyndman, R. J. (n.d.). *Piecewise linear trends.* Retrieved June 10, 2021, from https://robjhyndman.com/hyndsight/piecewise-linear-trends/

Hyndman, R. J., & Athanasopoulos, G. (n.d.). *Forecasting: Principles and Practice .* Retrieved June 10, 2021, from https://otexts.com/fpp2/

IETF. (n.d.). *Operations and Management Area Working Group (opsawg).* Retrieved 07 13, 2021, from https://datatracker.ietf.org/wg/opsawg/documents/

Jenkins: build great things at any scale. (n.d.). Retrieved from https://www.jenkins.io/

Jupyter: Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages. (n.d.). Retrieved June 6, 2021, from https://jupyter.org/

Kubeflow. (n.d.). *Introduction to Katib*. Retrieved from https://www.kubeflow.org/docs/components/katib/overview/

Kubeflow: The Machine Learning Toolkit for Kubernetes. (n.d.). Retrieved June 5, 2021, from https://www.kubeflow.org/

Kubernetes: Production-Grade Container Orchestration. (n.d.). Retrieved 06 05, 2021, from https://kubernetes.io/

MinIO: Object Storage for the Era of the Hybrid Cloud. (n.d.). Retrieved June 5, 2021, from https://min.io/

Network monitoring. (n.d.). Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Network_monitoring

Open Neural Network Exchange. (n.d.). Retrieved from https://onnx.ai/

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., . . . DeVito, Z. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems 32*, (pp. 8024--8035).

Prophet: Forecasting at Scale. (n.d.). Retrieved from https://facebook.github.io/prophet/

PyTorch: TORCHSERVE. (n.d.). Retrieved June 6, 2021, from https://pytorch.org/serve/

Quittek, J., Zseby, T., Claise, B., & Zander, S. (2004). *Requirements for IP Flow Information Export (IPFIX).* Retrieved from https://www.rfc-editor.org/info/rfc3917

Roughan, M. (n.d.). *Internet Traffic Matrices.* Retrieved 07 16, 2021, from https://roughan.info/project/traffic_matrix/

Santos, O. (2016). *Network Security with NetFlow and IPFIX: Big Data Analytics for Information Security.* Cisco Press.

SciKit Learn. (n.d.). *Feature Selection*. Retrieved from https://scikit-learn.org/stable/modules/feature_selection.html

Tensorflow: Serving Models. (n.d.). Retrieved June 6, 2021, from https://www.tensorflow.org/tfx/guide/serving

The M4 Competition Team. (n.d.). *M4 Competition: Updates*. Retrieved from https://forecasters.org/blog/2018/01/19/m4-competition/

Time-series compression algorithms, explained. (n.d.). Retrieved 07 22, 2021, from https://blog.timescale.com/blog/time-series-compression-algorithms-explained/

Wikipedia. (n.d.). *Ablation (artificial intelligence)*. Retrieved June 6, 2021, from https://en.wikipedia.org/wiki/Ablation_(artificial_intelligence)

Wikipedia. (n.d.). *Autoencoder.* Retrieved from https://en.wikipedia.org/wiki/Autoencoder

Wikipedia. (n.d.). *Automatic Differentiation*. Retrieved June 5, 2021, from https://en.wikipedia.org/wiki/Automatic_differentiation
Wikipedia. (n.d.). *Autoregressive integrated moving average*. Retrieved June 10, 2021, from https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average
Wikipedia. (n.d.). *Backpropagation*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Backpropagation
Wikipedia. (n.d.). *Benchmarking Methodology for Network Interconnect Devices*. Retrieved 07 09, 2021, from https://datatracker.ietf.org/doc/html/rfc2544
Wikipedia. (n.d.). *Carl Friedrich Gauss*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Carl_Friedrich_Gauss
Wikipedia. (n.d.). *Command-line interface*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Command-line_interface
Wikipedia. (n.d.). *Curse of Dimensionality*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Curse_of_dimensionality
Wikipedia. (n.d.). *Data compression ratio.* Retrieved 07 21, 2021, from https://en.wikipedia.org/wiki/Data_compression_ratio
Wikipedia. (n.d.). *DevOps*. Retrieved from https://en.wikipedia.org/wiki/DevOps
Wikipedia. (n.d.). *Exponential Growth*. Retrieved from https://en.wikipedia.org/wiki/Exponential_growth
Wikipedia. (n.d.). *Graphics processing unit*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Graphics_processing_unit
Wikipedia. (n.d.). *Internet Control Message Protocol*. Retrieved 08 09, 2021, from https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol
Wikipedia. (n.d.). *IP Flow Information Export.* Retrieved 07 13, 2021, from https://en.wikipedia.org/wiki/IP_Flow_Information_Export
Wikipedia. (n.d.). *Iperf*. Retrieved 07 09, 2021, from https://en.wikipedia.org/wiki/Iperf
Wikipedia. (n.d.). *Isaac Newton*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Isaac_Newton
Wikipedia. (n.d.). *IS-IS*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/IS-IS
Wikipedia. (n.d.). *Matrix Multiplication*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Matrix_multiplication
Wikipedia. (n.d.). *Microservices*. Retrieved from https://en.wikipedia.org/wiki/Microservices
Wikipedia. (n.d.). *MTR (software)*. Retrieved 07 09, 2021 , from https://en.wikipedia.org/wiki/MTR_(software)
Wikipedia. (n.d.). *NETCONF*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/NETCONF
Wikipedia. (n.d.). *Network Architecture Search*. Retrieved from https://en.wikipedia.org/wiki/Neural_architecture_search
Wikipedia. (n.d.). *Nyquist frequency.* Retrieved 07 16, 2021, from https://en.wikipedia.org/wiki/Nyquist_frequency
Wikipedia. (n.d.). *Open Shortest Path First*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Open_Shortest_Path_First
Wikipedia. (n.d.). *Recurrent Neural Network*. Retrieved June 11, 2021, from https://en.wikipedia.org/wiki/Recurrent_neural_network

Wikipedia. (n.d.). *Reinforcement Learning*. Retrieved from https://en.wikipedia.org/wiki/Transfer_learning
Wikipedia. (n.d.). *Representational state transfer*. Retrieved from https://en.wikipedia.org/wiki/Representational_state_transfer
Wikipedia. (n.d.). *Residual Neural Network*. Retrieved June 11, 2021, from https://en.wikipedia.org/wiki/Residual_neural_network
Wikipedia. (n.d.). *Service Mesh*. Retrieved from https://en.wikipedia.org/wiki/Service_mesh
Wikipedia. (n.d.). *Simple Network Management Protocol*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/Simple_Network_Management_Protocol
Wikipedia. (n.d.). *Software as a service*. Retrieved from https://en.wikipedia.org/wiki/Software_as_a_service
Wikipedia. (n.d.). *Structured programming*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Structured_programming
Wikipedia. (n.d.). *Tensor Processing Unit*. Retrieved June 2, 2021, from https://en.wikipedia.org/wiki/Tensor_Processing_Unit
Wikipedia. (n.d.). *Transfer Learning*. Retrieved from https://en.wikipedia.org/wiki/Transfer_learning
Wikipedia. (n.d.). *YANG*. Retrieved 07 06, 2021, from https://en.wikipedia.org/wiki/YANG
Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., . . . Perrault, R. (n.d.). *The AI Index 2021 Annual Report.* Retrieved from arXiv: https://arxiv.org/abs/2103.06312

Using SCTE 224 To Increase Advertising Revenue

A Technical Paper prepared for SCTE by

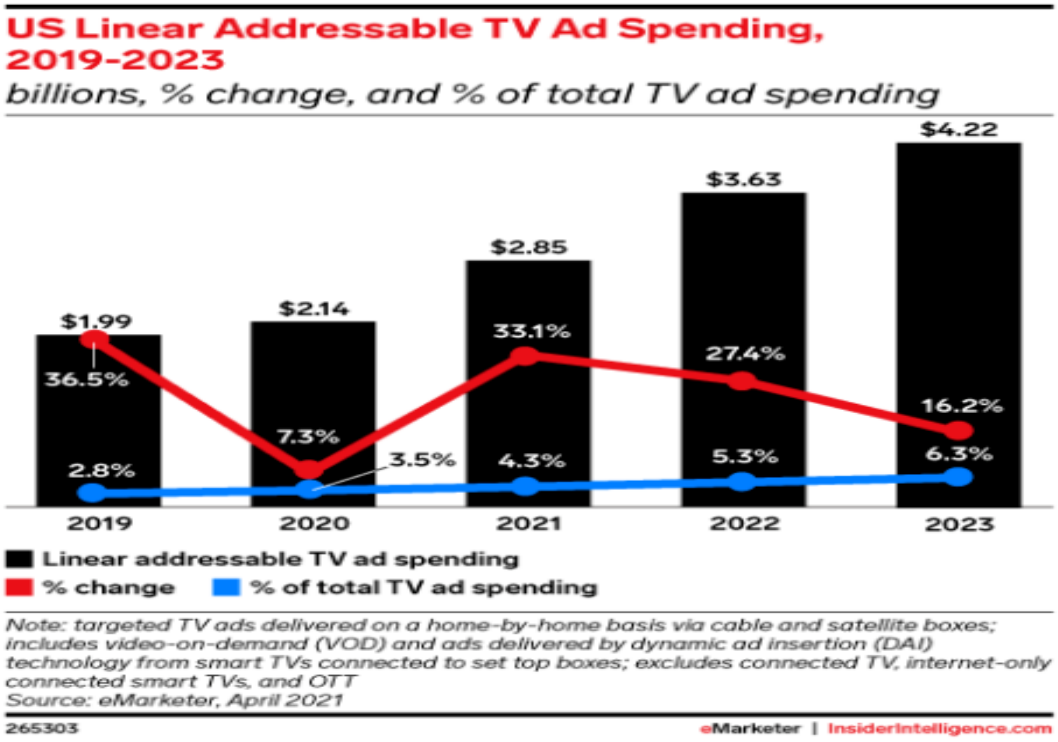
Gregg Brown
Senior Product Manager
Comcast Technology Solutions
Atlanta, GA
678-736-0480
Gregg_Brown@comcast.com

Stuart Kurkowski, PhD
Distinguished Engineer
Comcast Technology Solutions
1899 Wynkoop St, Denver CO.
303-503-2680
Stuart_Kurkowski@comcast.com

Neill Kipp, Comcast Technology Solutions

1. Introduction

The popularity and adoption of the SCTE 224, Event Scheduling and Notification Interface (ESNI) is opening new use cases where the protocol is a great fit. One new and exciting use cases is for addressable advertising where content providers and operators can use static, national ad inventory which are more targeted to the viewers and therefore potentially generate higher CPMs. Addressable advertisements, in this context, means replacing advertisements sold on broad age/gender demographics with advertising sold on more specific audience definitions. This evolution in ad avails has both content providers and operators excited for revenue growth potential in this relatively untapped market. According to eMarketer, U.S. addressable TV advertisement spending is expected to grow 33.1%, 27.4%, and 16.2% respectively in 2021, 2022, and 2023 and will eventually represent 6.3% of the total TV ad spending.



SCTE 224 has proven itself as an efficient and effective means for machine-to-machine communication of out-of-band (OOB) linear rights management. Additionally, combining of SCTE 224 with SCTE 35 to trigger the in-band signaling allows precision execution of linear rights for content substitution and addressable advertising management. ESNI is well known as the solution of record for content embargos and alternate content, however, content providers are now using ESNI for addressable advertising,

The ESNI protocol is perfectly well suited to communicate rules and policies at an audience-based level, thereby providing a substrate to implement addressability. Two critical requirements within the addressable advertising workflow are 1) identifying which slots within the content provider's 14-15 minutes per hour are addressable and 2) conveying ad information specific to those slots. ESNI is the key mechanism to communicate slot schedules for addressable slots versus national advertisements. ESNI also communicates specific rules for the addressable inventory on behalf of the content provider to the

distributors' Advertising Decisioning Service (ADS). ESNI therefore facilitates appropriate ad avail decisioning, providing information for ad spot inclusion and exclusion.

ESNI objects are extensible markup language (XML) messages with relevant fields for advertising such as ViewingPolicy actions for ad inclusion, exclusion, and ADS directives. These ESNI messages are managed with a representational state transfer (REST) interface for exchange between the operator and content provider. SCTE 35 markers in the stream trigger these breaks as they have done in the past. The out-of-band ESNI execution components then link these in-band SCTE 35 markers with the ESNI instructions.

2. Addressability Background

Let's start with a high-level description of addressable advertising: Linear-based television advertising (cable or satellite TV ads) has been essentially sold the same way for years. The operator/distributor typically has the ability to sell 2 to 3 minutes of commercials per hour per program – historically called the local avails – and the content provider sells the rest of the ad inventory per program per hour – roughly 14 to 15 minutes of commercials. In this workflow, the local avails inventory has been classified as addressable ads, targeted and sold at a local level vs. ads sold by the content providers which are national ad campaigns that are "locked and loaded" to be seen by everyone watching that particular program.

The buzz around the industry is the transition of taking those 14 to 15 minutes of content provider-owned national inventory and converting it into "addressable" ad inventory. Addressability has both content providers and operators interested because of the potential revenue and growth of this relatively untapped market.

There are also a few industry consortiums focused in this area, such as Go Addressable and Project OAR that are working to tackle both business and technical challenges to grow the space. Additionally, the SCTE Digital Video Subcommittee working group five recently released a best practices document around using SCTE 224 for advertising (SCTE 268, Operational Practice on the Usage of SCTE 224 for Advertising Information).

3. SCTE 224 Context

The SCTE 224 Event Scheduling and Notification Interface (ESNI) is an XML-based standard that provides a defined protocol for carrying machine-to-machine metadata for video. There are five basic constructs within SCTE 224, as shown in Figure 1. Each of these constructs work together to provide a content provider with a means to convey video rights for content replacement as well as advertising instructions on the operator side of the workflow. These five constructs are Media, MediaPoints, Policy, ViewingPolicy, and Audiences. We describe each of these here, and then tie them all together with the underpinning for the advertising use case in the following section.

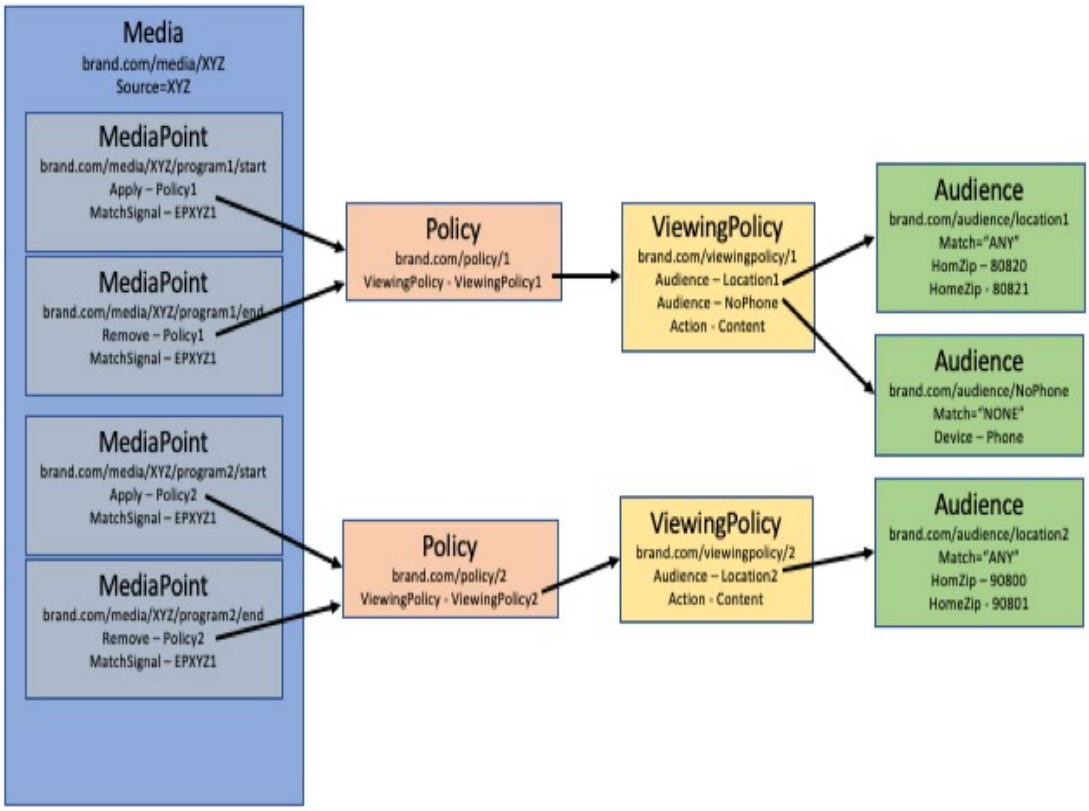


Figure 1 – SCTE 224 Constructs

3.1. SCTE 224 Media

The Media object is a top-level container representing a linear channel whose primary function is to carry all the MediaPoints, so it contains an ordered list of MediaPoint elements as shown in Figure 1. The Media object also contains a few key elements like a description and source for the linear channel it represents.

3.2. SCTE 224 MediaPoint

The MediaPoint object describes a point in the Media when a decision needs to be made or an action needs to be taken. These points in time can either be time-based (i.e., the presence of a @matchTime attribute in the MediaPoint) or SCTE 35 in-band signal based for frame accuracy. The signal-based MediaPoints contain a MatchSignal element with XPath matching logic to link the MediaPoint to the presence of the in-band signal. Signals can be reused, because MediaPoints also have an effective/expires window constraining when the MediaPoint can be evaluated.

When a MediaPoint is triggered, based on time or signal, it can either "Apply" or "Remove" one or more Policy objects which effect the state of the linear playout. See the Policy object description below for more details. MediaPoints that "Apply" a policy do so until another MediaPoint explicitly "Removes" that Policy or they time out based on the duration indicated in the "Apply" statement.

3.3. SCTE 224 Policy

A Policy object is nothing more than a container for defining a set of ViewingPolicy elements to be acted upon based on this Policy being "Apply" or "Removed" from the Policy stack. The "Apply" or application of a Policy means putting that Policy on that Media's stack via first-in-last-out queue. All policies currently on the stack that affect a particular Media are aggregated together, so multiple Policies can be affecting the state at one time. The removal of a Policy then takes it off that stack and out of the state of that Media. SCTE 224 has explicit rules about how to manage the Policy queue in a SCTE 224 execution engine.

3.4. SCTE 224 ViewingPolicy

The ViewingPolicy object is the key SCTE 224 object that associates one or more actions to an audience. These "Actions" can range from directing an audience to alternate content, restricting trick mode, or restricting resolution. For ads specifically, these actions can contain information about the Advertising Decisioning Service (ADS) to use for a particular audience, or various advertisement conflicting rules for a particular audience. The key to a ViewingPolicy is that if the "Audience" criteria is met, then the action must be taken. The SCTE 224 maintains a list of actions, many of which apply to the ad use case such as allocation and break owner, allowed within the ViewingPolicy object, many of which are specific to addressable advertising

3.5. SCTE 224 Audience

The Audience object is a set of characteristics that define a subset of viewers based on certain aspects of their device type (tablet, phone, etc.); device characteristics (local storage, mobile, etc.); or location information (such as zip codes, postal codes, latitude/longitude, market areas); or even receive categories such as Distributor or Virtual Integrated Receiver Decoder (vIRD). Audience objects can contain other Audience objects, making for compound Audiences. Additionally, logic to associate a client with an Audience is based on matches of ANY, ALL, or NONE of the characteristics outlined, for easily including or excluding specific characteristics. For example, you can say Match="ANY" for a list of zip codes to characterize the audience within that area, or Match="NONE" to characterize an audience outside that area.

3.6. SCTE 224 Example Logic

So now let's take those five objects from Figure 1 and run them through a scenario. A video signal acquisition system (SAS) sees an in-band SCTE 35 signal. It calls a signal decisioning system (SDS) to figure out what it should do. When the SAS calls the SDS it tells the SDS which source it was on, what time it saw the signal, the binary signal, and the client characteristics. If it could talk, it would say something like "I just saw the signal 'UhJeasfojoihe23edde' on source XYZ, at 1:00pm and I am encoding for zipcode 80820." The SDS then looks through all the out-of-band SCTE 224 and its Media to find the one for that source (i.e., XYZ). Once it finds the correct Media, it examines its MediaPoints to find those that fall within the designated time window. Once it has the list of MediaPoints, it evaluates each one to see if there is a match with the signal. For a MediaPoint that matches, it either performs an "Apply" or "Remove" of the associated Policy. It then goes from that Policy to the ViewingPolicy, where it determines whether a match exists between the audience and the designated zipcode, based on the Match criteria. If yes, the SDS would return the "Action" of the ViewingPolicy to the SAS for that audience. In the case of alternate content, for instance, it might tell the SAS that it needs to switch over to another source and start encoding the alternate source.

4. Addressable Advertising Specifics

How does SCTE 224 support this new dynamic with addressable ads? The exciting news is that the SCTE 224 standard is set up to communicate rules and policies at an audience-based level, which is exactly what addressable advertising needs to make it work.

Within the addressable advertising workflow, the first critical decisioning point is for the content provider and operator to know which advertisements are addressable, within the content providers' 14-15 minutes of ad avails per hour. To do that, SCTE 224 is used as the communication protocol between the three parties. The content provider's ad schedule is converted into SCTE 224 and then distributed throughout the value chain in SCTE 224 format. Figure 2 shows the video and SCTE 224 information being sent from the content provider to the operator and the Advertising Decisioning Service.

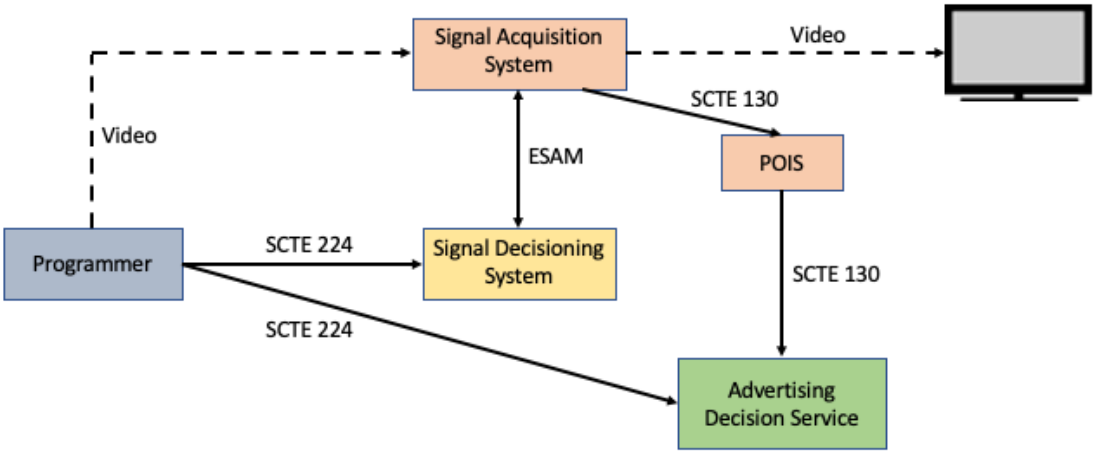


Figure 2 – SCTE 224 Advertising Use Case

4.1. Identifying Addressable Slots

This SCTE 224 message set is sent to the operator's signal decisioning system (SDS) to work with the signal acquisition system (SAS), which is watching for the SCTE 35 in-band signal in the video. When the SAS sees an SCTE 35 marker for the start of an avail break/pod, it calls the SDS to see if that is an addressable ad slot or not. The SDS logic then looks for the corresponding MediaPoint to "Apply" the policy and then validate the operator against the ViewingPolicy Audience and resulting action. For example, if the operator is ABC and the "ABC" is in the content provider's audience, and the action is "LinearDAI (Digital Ad Insertion)", then it is addressable. The SDS then will return the "LinearDAI" back to the SAS as the decision response, which means the SAS now knows that it is an addressable advertising slot, so it then calls the Placement Opportunity Information Service (POIS) to get an advertisement to play. If, on the other hand, the operator is "DEF" and is not in the Audience, the SDS would return a "noop," to mean "no opportunity," which tells the SAS to do nothing and/or keep doing what it was doing, which results in it playing out the baked-in advertisement. It should be noted that there could be other actions like "SignalDeletion".

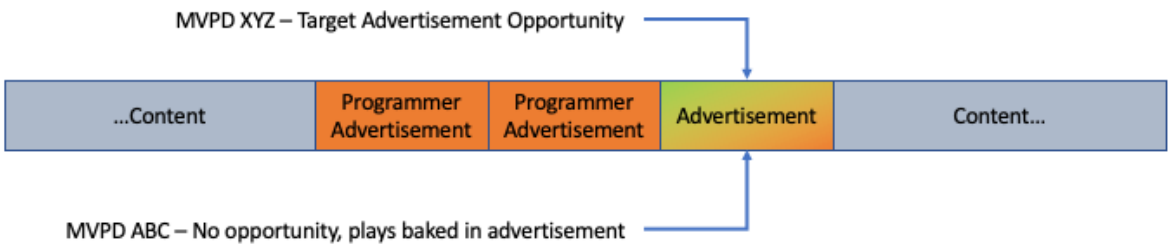


Figure 3 -- Identifying Slots

This might sound like a capability that SCTE 35 has today, but the benefit of out-of-band SCTE 224 is that it is unlimited in capacity and ability to match audience and action pairs. Because the signaling happens out-of-band, it does not impact the frames in the video by trying to cram everything into a SCTE 35 signal, so it greatly reduces the SCTE 35 work required with the video. In fact, it allows a single, simple SCTE 35 marker to be sent to multiple operators and have multiple meanings, which results in individual playout payloads to multiple operator recipients. Because content providers innately serve multiple operators and audiences, this is a highly desirable feature.

4.2. Addressable Advertising Rules

The second major addressability use case for SCTE 224 is providing specific ad rules for the addressable inventory on behalf of the content provider to the operator and its Advertising Decisioning Service (ADS). There are many advertisement rules, but two constitute the main rules conveyed from the content provider to the operator. One involves which ADS to use, whether there are campaign/order codes associated with that slot that the ADS needs to know. The second involves advertisement inclusion and exclusion, or whether there are conflicting rules, based on what advertisements the content provider already has slotted in its linear feed. In this case, a common language like SCTE 224 can facilitate the appropriate advertising placement with correct ad decisioning rules applied. Each of these types of advertisement rules can easily be carried by SCTE 224 and conveyed to an operator by or for the content provider, on a per-operator basis. Below are more details about these rules.

4.2.1. ADS Rules

Because SCTE 224 information is sent out-of-band and can be used differently for each operator, a content provider can set up different Advertisement Decisioning Systems for each operator. The content provider can use a SCTE 224 MediaPoint that has a Policy for audience “Comcast,” for instance, and action “use FreeWheel” and another operator with audience “Charter” and action “use Google DFP”. The SCTE 224 actions have been expanded in 2020/2021 to handle just this use case. This means that with a single, simple SCTE 35 in the video identifying the ad break, a content provider can tell different operators to use different ADSs.

Additionally, because the content provider now has individual direction capability for different operators, they can direct more than just *which* ADS to use. They can also include varying order numbers or varying campaign attributes that are different for each operator. This can result in a tighter coupling of addressable advertisements and the operator, leading to potentially higher revenue upsells.

4.2.2. Advertisement Conflicting Rules

Again, because the SCTE 224 signal path is out-of-band, it is not limited in size or what it can carry, compared to a SCTE 35 in-band signal. As a result, a content provider can include extensive metadata about an advertisement – or, more importantly, about what targeted advertisement(s) cannot be used, to avoid placing ad within the same ad break that conflict with one another. Because the content provider knows about the baked-in advertisements in the feed, it can provide rules to the operator or the operator's ADS about what advertisements would fit great in the pod -- it can say which advertisements to include. Or, a content provider can say which type or category of advertisements *not* to include (to exclude), because they would counter the previous or consequent advertisement that had been identified as addressable.

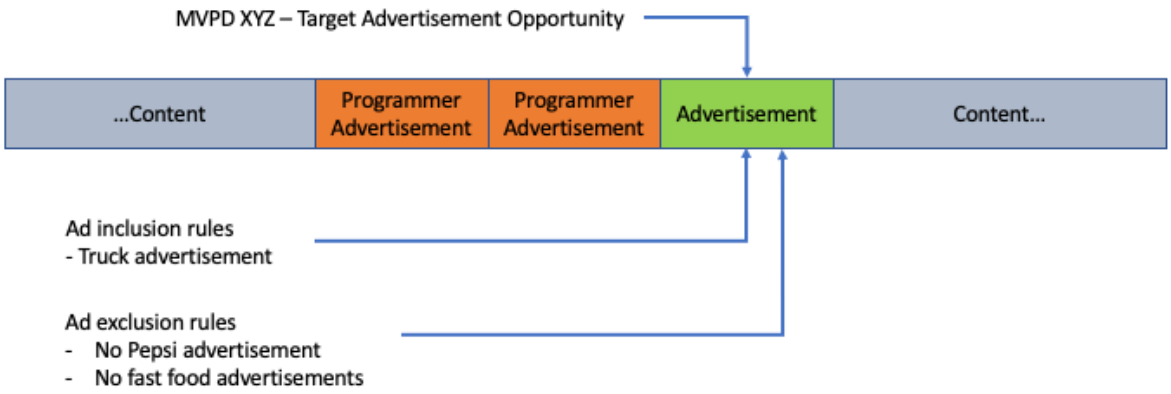


Figure 4 -- Conflicting Rules

Examples of these rules include, but are not limited to, the following:

1. Showing a specific number of advertisements or limiting the number of ad spots for a specific product group
2. Specifying a minimum spacing between advertisements belonging to a specific product group
3. Specifying that an ad spot may only be replaced by campaigns that's a member of a specific product category
4. Excluding particular products or categories that might not align with other advertisements in the advertisement pod.

Some real-world application of these types of rules includes not wanting to have a Brand A, a soda ad, and Brand B, another soda ad, appearing in the same advertisement pod. Or showing a national car advertisement followed by a local car dealer advertisement, or not showing an insurance advertisement in every pod.

4.3. Advertising Metadata Information (Intelligent Ad Insertion)

Because of its aforementioned large capacity and flexibility, SCTE 224 can also carry a great deal of information about the advertisements or the content around the addressable advertisement. Examples of these additional use cases include:

1. Scene information - Often ad campaigns are set up on overarching show genre, category, or age group, but if a content provider could convey more information at the scene level, it could be that

much more valuable for the slot. For example, what was last viewed by the audience in the last scene before the break: Maybe it was a car chase. Tying the next advertisement to that visual context, and scheduling into the next appropriate advertisement break a car dealer advertisement would more maximize effectiveness. This linkage could lead to greater CPM values.

2. Actors or other action-based metadata – Similar to the above example, product placement or metadata derived from action within a show (either live or pre-recorded) could be used to drive more intelligent ads and higher CPMs. Examples of this could include knowing the actor or additional scene information.
3. Ad Creative metadata – Using metadata from the previous ad or other ads from the advertisement pod could lead to a more intelligent ad break and higher CPM.
4. Technical ad metrics – Data such as ad duration could be used to ensure ad breaks are not missed, which could result in loss of revenue for the content providers and operators.

5. Conclusion

Ultimately, the benefits of using SCTE 224 to implement addressable advertising are plentiful and could enrich the ad environment for both content providers and operators/distributors. A short list includes:

- Addressable ad slots can be identified uniquely for different operators
- Different ADS can be supported by different operators
- Inclusion and exclusion rules can be conveyed machine-to-machine
- Different inclusion an exclusion rules can be executed for different operators
- Advertisements can be enriched with scene, actor, or creative metadata in ways that strengthen visual ties and could improve overall effectiveness/CPMs.

All of those benefits come with reduced complexity within SCTE 35, because they all can be done with a single simple SCTE 35 trigger; the rest is carried in the SCTE 224 Audience, ViewingPolicy, and Policy construct. This provides a content provider with control and execution within its linear feed. No more multiple versions of a video, just to carry different SCTE 35 markers to different operators, or multiple complex SCTE 35 markers into the video, leading to confusing and problematic interpretations by operators.

The addressable advertising eco-system is still highly complex. It involves and needs much more than SCTE-224 to solve its many nuances, but hopefully we have provided a glimpse into how much it plays a critical role in enabling addressable advertising — for both content providers and operators.

Abbreviations

ADS	Advertising Decisioning Service
CPM	cost per thousand
DFP	DoubleClick for Publishers
ESNI	Event Scheduling and Notification Interface
IRD	Integrated Receiver Decoder
OOB	out-of-band
POIS	Placement Opportunity Information Service
(Project) OAR	Open Addressable Ready
REST	representational state transfer

SAS	signal acquisition system
SCTE	Society of Cable Telecommunications Engineers
SDS	signal decisioning system
vIRD	virtual integrated receiver decoder
XML	Extensible Markup Language
Linear DAI	linear digital ad insertion

Bibliography & References

eMarketer, Linear addressable TV ad spending will grow 33.1% this year, https://www.emarketer.com/content/linear-addressable-tv-ad-spending-will-grow-33-percent-this-year

FreeWheel, https://www.freewheel.com/

Google DoubleClick for Publisher (DFP). https://www.google.com/ads/publisher/

On Addressability, http://www.onaddressability.com/

Project OAR, https://projectoar.org/

SCTE 35, Digital Program Insertion Curing Message for Cable, 2020. https://www.scte.org/standards-development/library/standards-catalog/scte-35-2019/

SCTE 224, ESNI, Event Scheduling and Notification Interface 2021. https://www.scte.org/standards-development/library/standards-catalog/ansiscte-224-2021/

SCTE 268, Operational Practice on the Usage of SCTE 224 for Advertising Information, https://www.scte.org/standards-development/library/standards-catalog/operational-practice-on-the-usage-of-scte-224-for-advertising-information/

W3C XML Base (Second Edition). W3C Recommendation 28 January 2009. http://www.w3.org/TR/xmlbase/

W3C XML Schema Part 2: Datatypes Second Edition. W3C Recommendation 28 October 2004. http://www.w3.org/TR/xmlschema-2/

W3C XML Path Language (XPath) 2.0 (Second Edition). W3C Recommendation 14 December 2010. http://www.w3.org/TR/xpath20/

Water Can Run, But It Can't Hide

PNM Finds Soaked Cables

A Technical Paper prepared for SCTE by

Kathy Fox
Vice President, XOC
Comcast
Kathy_Fox@comcast.com

Jason Rupe, Principal Architect, CableLabs

Tom Williams, Distinguished Technologist, CableLabs

Jay Zhu, Senior Engineer, CableLabs

Ron Hranac, Chair, SCTE Network Operations Subcommittee

Nathan Zedan, RF Lab Manager, Comcast

James Kolcun, Director of Field Operations Standards, Comcast

Larry Wolcott, Fellow, Comcast

1. Introduction

The cable industry started taking advantage of proactive network maintenance (PNM) nearly a dozen years ago, and has shared results at many previous SCTE Cable-Tec Expos – so what's new in PNM for 2021? Water. Not the liquid itself, but what it can do to our subscriber drop plant and the services we provide over that plant. Operators have long been haunted by coaxial cable water ingress – since the very first days of the industry. Water ingress is nothing new, but it now has its own special tool in the PNM toolbox. In the grand scheme of things, that tool is a way to sharpen the focus on enhancing the customer experience and network performance, which is where PNM tends to take center stage. (Even, and perhaps especially, when it's raining.)

This paper reviews the progress of water detection, location, and severity assessment in the cable plant. The authors explore the background, motivation, theory and provide an outline for operators to evaluate their networks. In addition, there are field and lab examples which clearly illustrate the importance to customer experience, using customer testimony and speed test results as points of validation. Lastly, the authors provide information to help operators determine the future impacts when considering DOCSIS 4.0's features, such as extended spectrum (ES) DOCSIS and full duplex (FDX) DOCSIS.

2. Background

It has been said that one of the most versatile pieces of test equipment available to the cable industry is the spectrum analyzer. These are instruments that display signals in the frequency domain, and have been used for decades to install, validate, and troubleshoot service on cable networks.

Starting November of 2012, the DOCSIS 3.0 specification was expanded to include spectrum analyzer-like functionality in cable modems (CMs). This feature is known as full band capture (FBC) and is supported by most DOCSIS 3.0 and all 3.1 CMs. Since this time, most cable operator-deployed modems now have the FBC spectrum analysis capability.

This was an important moment for cable operators, creating the opportunity to automate a long-time manual process, known as sweeping. Prior to this time, technicians were required to manually connect broadband test equipment to take measurements of the RF plant. In addition to significantly improving operational efficiency and reducing costs, FBC allows cable networks to be monitored constantly, without a technician being present. This is important for diagnosing intermittent issues which can occur at odd times, often as the result of temperature or weather changes.

Among the earliest recognized impairments (Figure 1) was the standing wave, or more accurately, amplitude ripple which is caused by standing waves. These are characterized by a periodic, or predictably repeating waveform in the frequency response, which may be sinusoidal or scalloped. Standing waves can be classified by a number of parameters including their periodicity and magnitude.

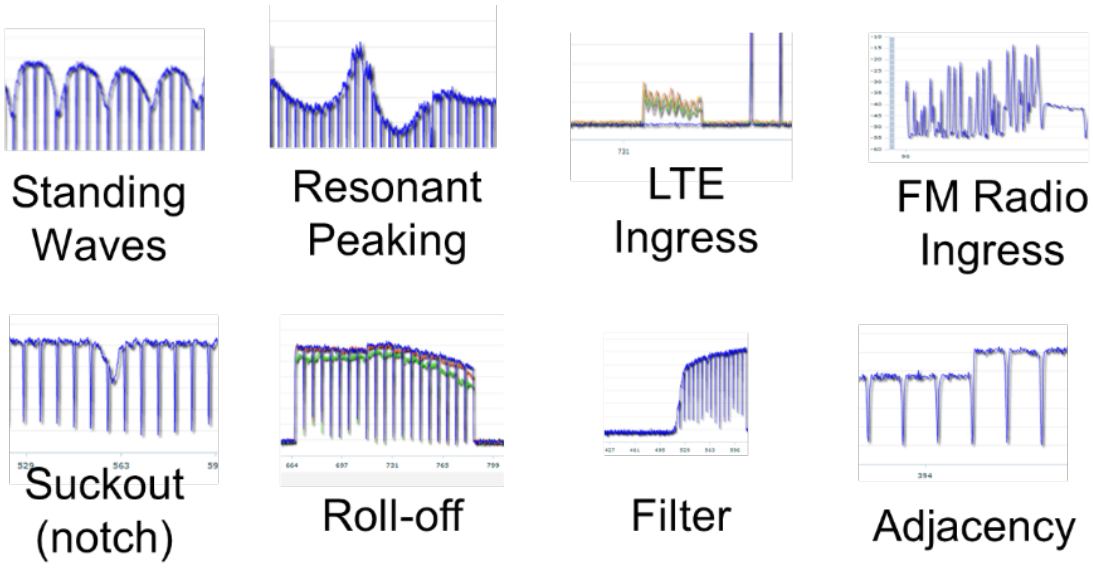


Figure 1 - Examples of impairments found using FBC (Source: Comcast)

As illustrated in Figure 1, there are several common RF impairments which can be detected and classified using FBC. However, there are a few impairment types which do not clearly fall into these general impairment categories. This typically occurs in the case of multiple problems resulting in a compound impairment. The detection software algorithms can become ineffective, sometimes detecting one or the other.

3. Cable RF Spectrum Fundamentals and Water

Some of the basic concepts of cable RF spectrum are illustrated in Figure 2. The horizontal axis represents the frequency spectrum, starting from 5 MHz and ending above 1000 MHz [1000 MHz is the same as 1 gigahertz (GHz.)] RF signals are precisely modulated and transmitted at specific frequencies, and their quality is measured using a variety of metrics: RF power, carrier- or signal-to-noise ratio, frequency response, and more.

For instance, we measure RF power (signal strength) using the decibel millivolt (dBmV) rather than watts. This is because the range of RF signal power in cable networks is very large, so, expressing those numbers in units of watts gets unwieldy. When comparing two values, we use units of the decibel (dB) because it represents a ratio, although it is logarithmic. One simple rule-of-thumb to remember when using decibels is that a 3 dB change in signal level represents a doubling or halving of RF power. For example, 50 dBmV is twice the power of 47 dBmV.

Notice in Figure 2 that all the RF signal levels across the spectrum are similar, making a nice flat line across the peaks of the signals. However, in Figure 3 there are significant power variations at different frequencies within the spectrum. When things are working properly, the levels should be relatively flat and sometimes may have a tilt in one direction or the other. The overall amplitude-versus-frequency performance of the spectrum is known as the “frequency response.”

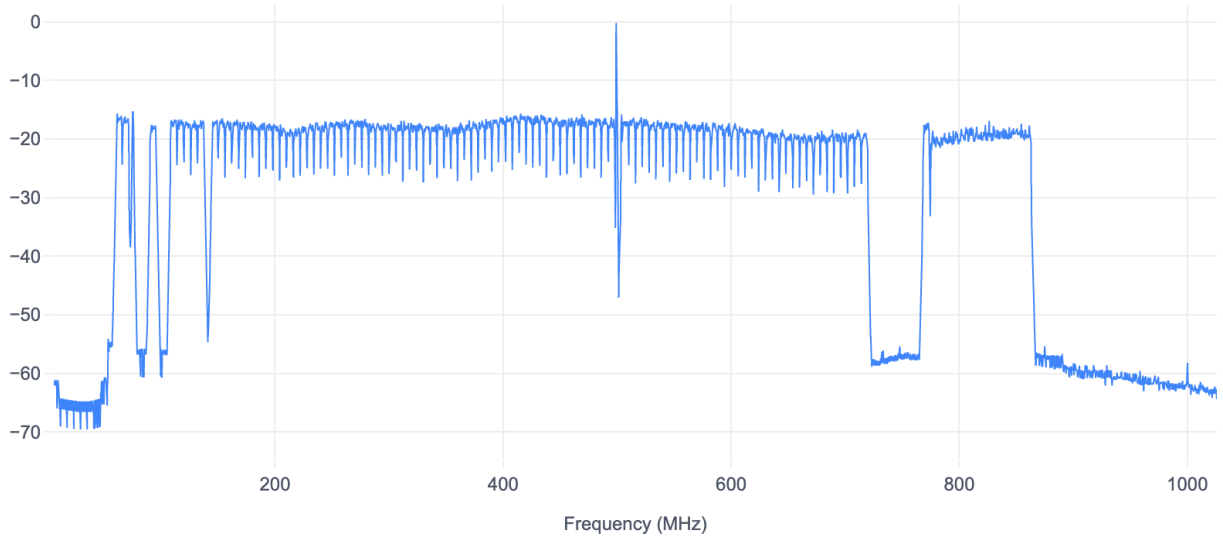


Figure 2 - Unimpaired frequency response

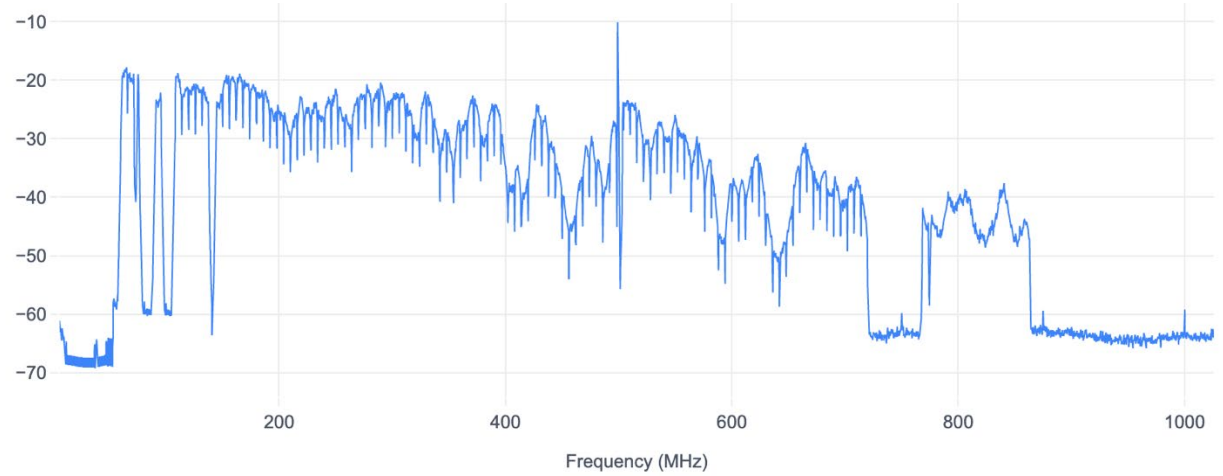


Figure 3 - Water impaired frequency response

When water enters coaxial cable, several things happen to create the distinctive frequency response shown in Figure 3. Why does water in coaxial cable have that effect on RF? The presence of water in the cable's dielectric changes the dielectric constant, which changes the velocity factor, characteristic impedance, and attenuation (see the Appendix for more information on the characteristics of coaxial cable). Further complicating the water-related degradation is the fact that the water is not uniformly distributed throughout the length of the cable. That, in turn, results in randomly distributed, localized variations in the cable's velocity factor, impedance (think micro-reflections) and attenuation, causing a non-periodic shape in the frequency response.

The severity of this problem will depend on the amount of water present in the cable and other factors such as temperature and system RF levels. These problems have been observed to coincide with rainy weather and tend to be variable, sometimes completely clearing when the water drains or evaporates. The

amount of customer impact can be measured with downstream receiver power levels, tilt, per-channel RxMER, and codeword errors or packet loss, which may inform the repair prioritization.

To avoid confusion, another related but different impairment you should be familiar with is known as an amplitude ripple, commonly referred to as a "standing wave." This is especially important because the standing wave is somewhat like our water signature, but there are subtle and not-so-subtle differences. Standing waves are also caused by impedance mismatches and the resulting micro-reflections, but water is not present. Because there's no water to add *random* attenuation, the signal bounces and attenuates in a predictable manner. In the case of a standing wave, a repeating and *periodic pattern* can be seen in the frequency response. Standing waves tend to have a sinusoidal wave shape, but can sometimes have a sharp, scalloped appearance. A standing wave may affect all or part of the RF spectrum. These problems tend to be constant (non-variable) or change very little. The changes in standing waves are subject to environmental influence such as wind or temperature, which can influence the mechanical properties of the plant.

It is most common for our drop cables and taps to be impacted by the presence of water, but it could be feeder or distribution cable affecting multiple locations. Drop cables are easily damaged by squirrels chewing on the jacket and shielding and is very common in some areas. Hardline is also subject to animal chews, radial cracks and holes caused by all manner of hostile forces. It is important to distinguish between drop and hardline because these are sometimes two different repair categories, each requiring a different type of technicians to fix the problem. Generally, a drop cable signature will be common to all devices within a single location and would be repaired by an install/repair tech or business partner (contractor). Larger plant issues would be repaired by a network maintenance technician and can disrupt service for a larger segment of our network. The latter often requires additional attention to scheduling and notification to help limit the negative impacts to customers.

From the customer's perspective, excessive RF signal attenuation is typically experienced as diminished quality and reliability of their internet or video experiences – or, in some cases, it renders those services unusable. What gets affected depends on the specific frequency which is impacted by the impairment. Conventional DOCSIS, video or other system signals can be used to evaluate the severity of the problem. However, given the transient nature of water in our cable systems, these types of problems can be temporal and associated with weather. Therefore, time and environmental components can be used to help with predictability. For example, additional resources may be allocated to a service area in advance of a rainy season.

3.1. Water Migration - Peripheral Damage

In addition to the cable, which is often a primary victim of water ingress, it's also common for the water to migrate and damage peripheral components. When additional network elements are damaged, multiple problems can become compounded and worsened. Among the most common examples are taps, splitters, splices, block splices and all the different filters and pads installed in the drop network.

3.1.1. Water Damaged Tap

In the following example, a water-soaked drop was the primary source of water ingress. However, the tap was physically located at a lower elevation on the pole than where the water entered the cable. With water accumulating over time and the influence of gravity, the water eventually migrated into the tap. A closer look at Figure 4 clearly shows water droplets in the upper left and lower right corners of the tap faceplate. The circuit also shows rust and other signs of corrosion. The subsequent frequency responses were captured before and after the faceplate was replaced. The FBC spectrum in Figure 5 represents a typical high-frequency roll-off starting around 500 MHz, becoming dramatically worse at 750 MHz (nearly 30

dB). In this case, the OFDM channel was significantly impacted, causing severely degraded service performance. Figure 6 shows the frequency response improvement after replacing the water-damaged faceplate.

Figure 4 - Water-soaked tap faceplate, water droplets visible

(Courtesy of James Medlock, Akleza)

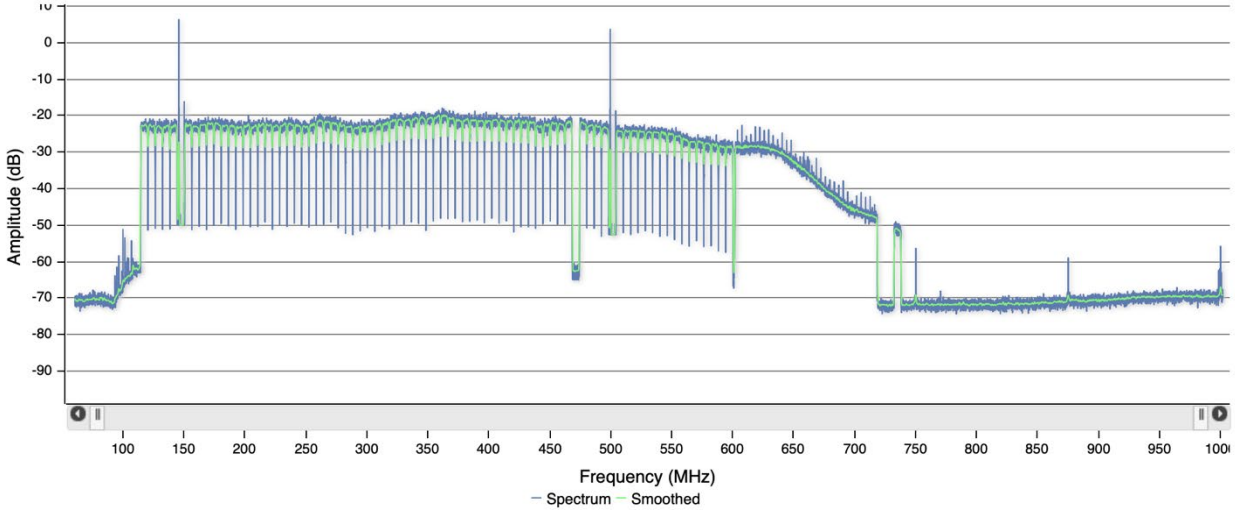


Figure 5 - Roll-off frequency response of water damaged tap port
(Courtesy of James Medlock, Akleza)

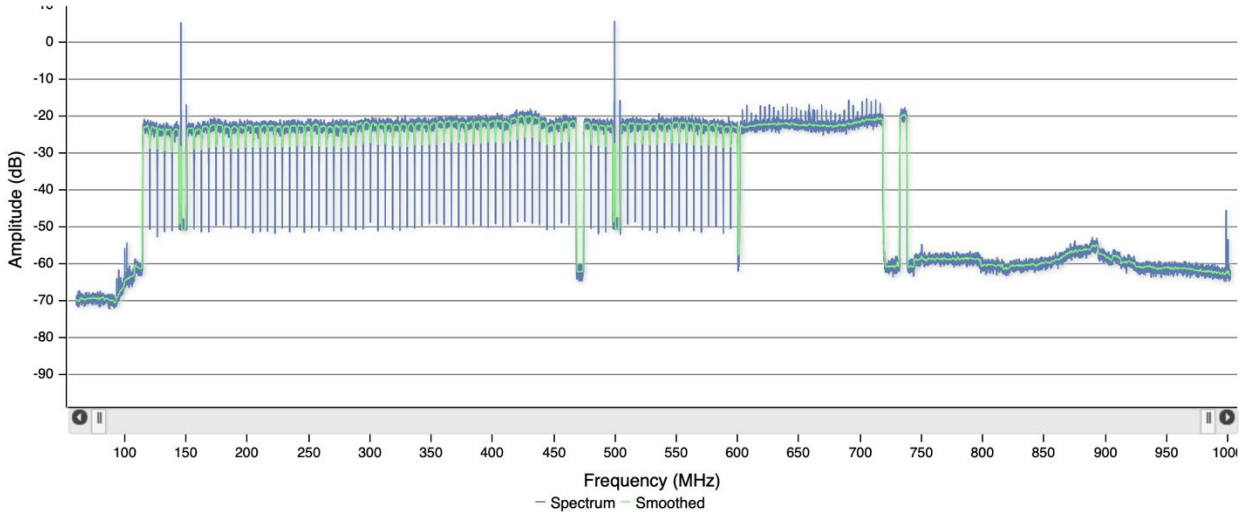


Figure 6 - Flat frequency response after faceplate replacement

(Courtesy of James Medlock, Akleza)

3.1.2. Water Damaged MoCA Filter

Virtually any passive or active network element can become subject to water migration. In some cases, the water can indirectly affect the network elements, resulting in unexpected impedance mismatches. The following example shows an in-line MoCA point-of-entry filter located at the ground block. Like many other types of filters, these passive devices are sealed against water ingress. However, when the drop cable jacket is compromised, the water can easily migrate along the center conductor or dielectric, where there is no water barrier. In the example shown in Figure 7, the filter became filled with water, froze, and expanded, causing the press-fit housing to become separated. The entire assembly was recovered from the field including the drop cable, filter, and ground block. Upon inspection, the Series 6 drop cable was damaged near the tap-side fitting (Figure 8) which is highly consistent with rodent chew marks, commonly seen on drop cables (Figure 9). In this case, there was no water immediately present at either connector interface. However, when a vacuum was applied to one end, the water quickly migrated and became evident at the connector (Figure 10).

The condition and integrity of the outer jacket is critically important to protect the cable plant from water ingress. While evaluating several damaged filters, water ingress and freezing could readily be attributed as the cause. In each example, the water ingress point could be located. Figure 11 shows more examples of typical jacket damage.

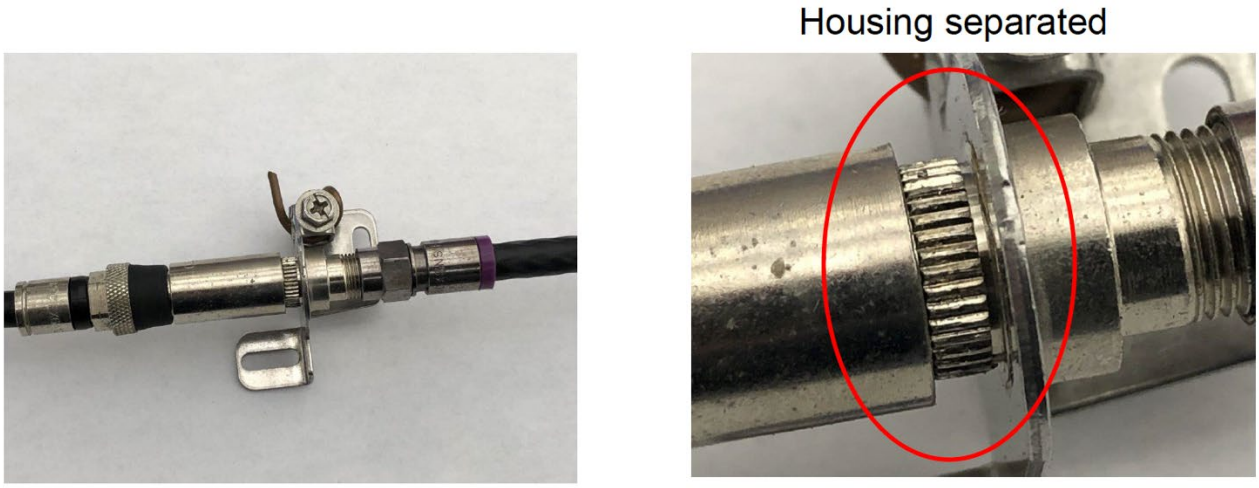


Figure 7 - MoCA filter housing separation

(Courtesy of Skip Palinkas, PPC)

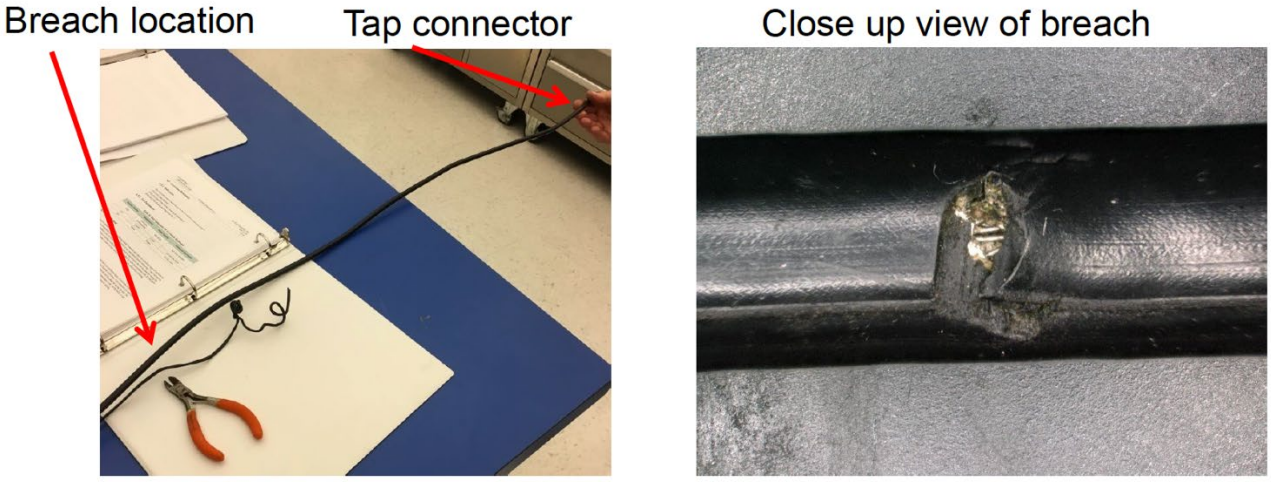


Figure 8 - Coaxial jacket breach near tap-side connector

(Courtesy of Skip Palinkas, PPC)

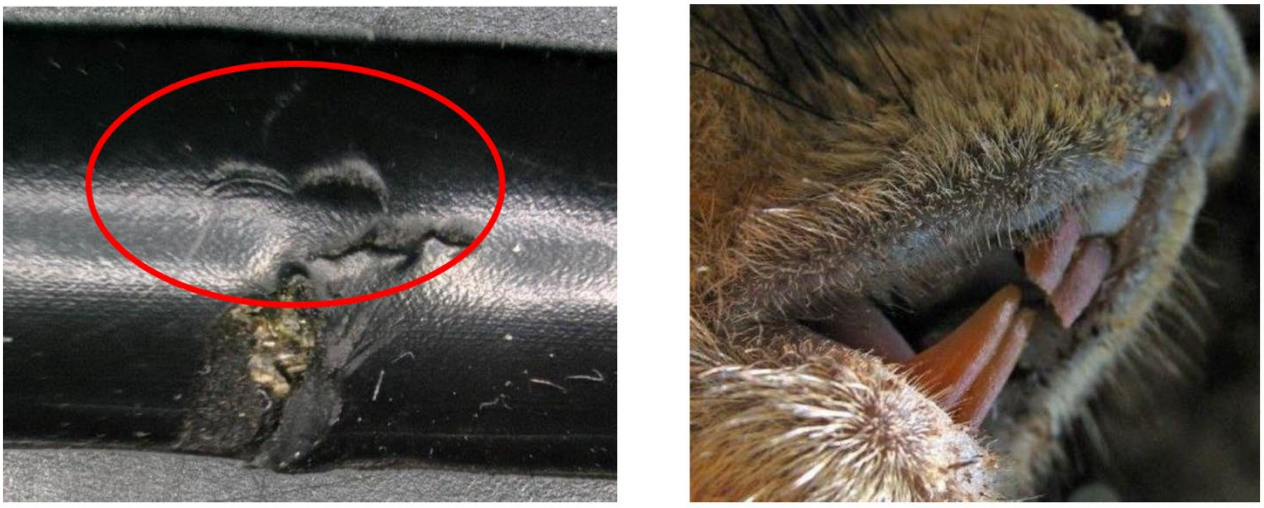


Figure 9 - Jacket breach compared with rodent teeth

(Courtesy of Skip Palinkas, PPC)

Connector removed from MoCA GB

Same connector after a vacuum pulled.
Demonstrates drop is saturated

Figure 10 - Water visible after vacuum is applied

(Courtesy of Skip Palinkas, PPC)

Figure 11 - Rodent chew compared to elongated cut, common coaxial damages

(Courtesy of Skip Palinkas, PPC)

3.2. Test Results

A number of field trials were conducted in 2020 resulting in a large number of cable samples recovered from the field. In one trial, over 100 drop cables were located and replaced, providing a substantial sample group. Other control groups of bad drops were also brought in from the field. In the latter group, the damaged drops were not necessarily associated with water.

In the example, Figure 12 shows the frequency response of a water-soaked drop cable (bottom) compared with the same type and length of new, unimpaired cable (top). Pockets of severe attenuation can be observed.

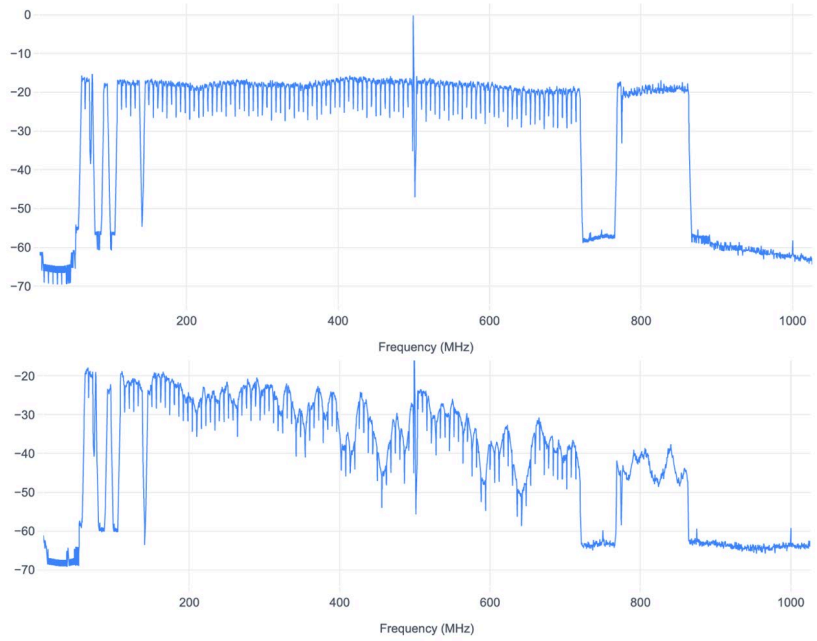


Figure 12 - New drop cable (top) compared to water-soaked cable (bottom)

These two cables were analyzed with test equipment including a speed test which closely approximates the experience a customer would have. A number of these tests were run and Figure 13 shows a typical result. The top value of 1262.7 Mbps download speed is consistently achieved using a new 95-foot RG6 drop cable. Then, when using the same type and length of cable with water damage, a speed of 179.3 Mbps is achieved. This is significantly below the provisioned speed of 1200 Mbps, delivering only 14% of the provisioned performance.

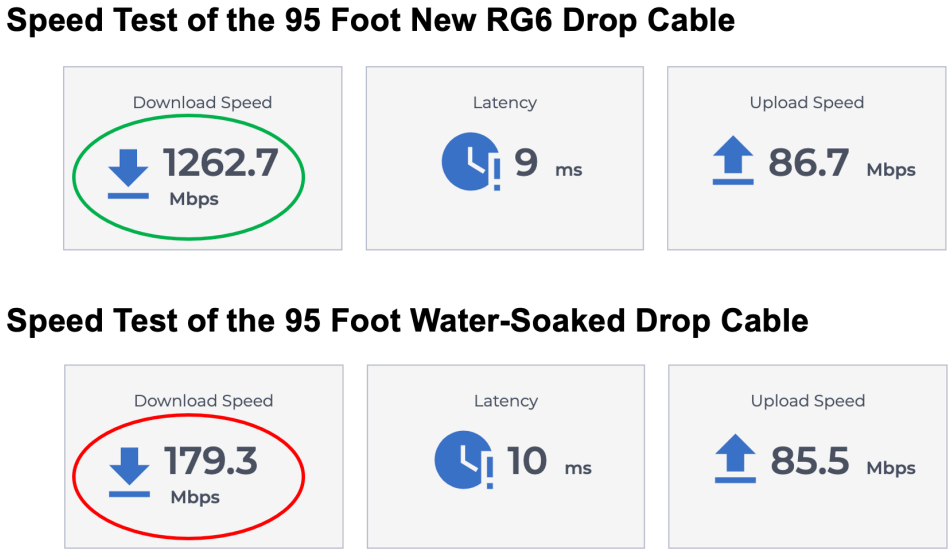


Figure 13 - Speed test comparison of new vs. damaged drop cable

To examine the influence of temperature, the same damaged cable was frozen (Figure 14) at -10 degrees Fahrenheit. After freezing, the frequency response was measured (Figure 15) and the attenuation greatly improved, having well over 25 dB improvement at certain frequencies.

At the same time, speed tests were run, and the results are show in Figure 16. When the damaged cable is frozen, the speed test results improved dramatically. A speed of 1078.4 Mbps was achieved, reaching nearly the same speed of a brand-new cable. Then, within minutes, the cable thawed and was retested at 68 degrees Fahrenheit. When the frozen water returned to a liquid state, diminished speeds returned. In this example, a paltry 70.9 Mbps was the peak download speed.

Figure 14 - Freezing the water-soaked drop cable

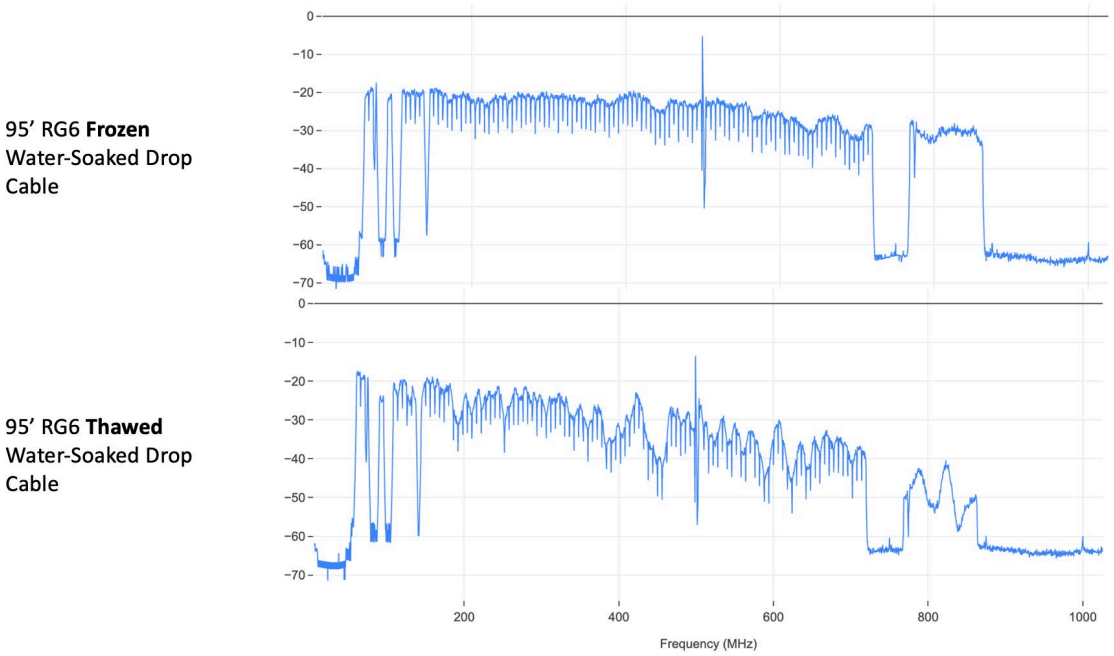


Figure 15 - Frozen (top) compared to thawed (bottom) frequency response

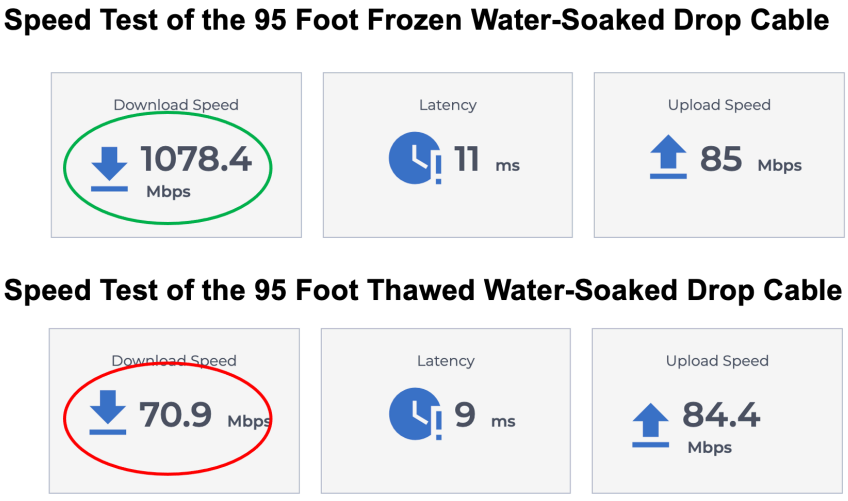


Figure 16 - Frozen (top) compared to thawed (bottom) speed test

4. Customer Impact

4.1. About Customer Experience

As previously discussed, water-soaked cables are a regular occurrence in most cable systems. Regardless of underground or overhead construction, cables can become damaged or otherwise deteriorate, allowing

water to ingress and eventually migrate through the length of cable. Depending on the cable's exposure to the elements, the impact on the customer experience can be highly variable.

4.2. Customer Experience – A Typical Example

In the presence of water damage, it's common to hear from our customers that their service was poor and unreliable. In the following example, FBC was used to identify an individual subscriber drop that had water damage, specifically by the unique signature in the displayed frequency response. As Figure 17 shows, the response has a non-periodic wave shape, and attenuation increases dramatically at higher frequencies. Remote polling of the modem showed that most of the downstream SC-QAM signals had poor performance, as shown in Figure 18. The upstream was relatively unaffected.

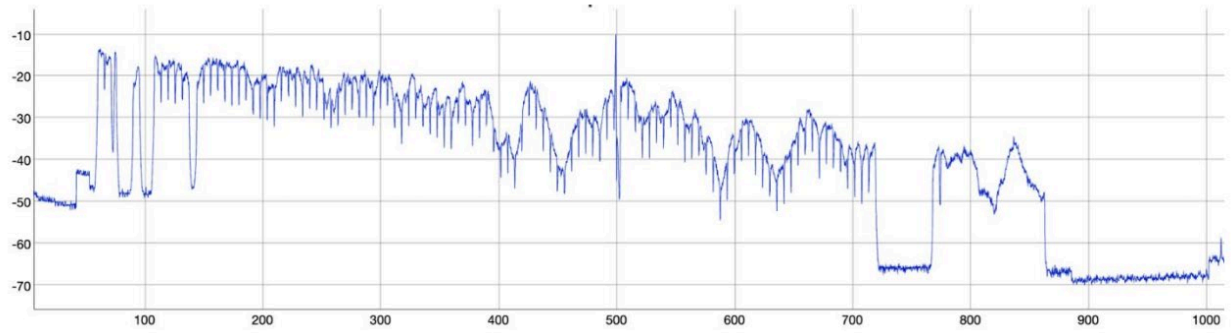


Figure 17 - Water damaged cable frequency response, prior to repair

Note the non-periodic wave shape in the FBC response in Figure 17 and the higher attenuation at higher frequencies. This example occurred when abrasion damaged the cable's jacket, allowing water to enter the cable.

Figure 18 - Impaired downstream SC-QAM power and RxMER levels, per channel

As seen in Figure 18, most of the downstream SC-QAM signals have low signal level and degraded RxMER, indicated in red shaded boxes. The upstream was relatively unaffected.

Technicians went to the subscriber location and were able to find and fix the problem without having to enter the premises. Figure 19 shows the coax jacket, which had been damaged by abrasion from the electrical service drop. The damaged coax jacket allowed water to enter the cable and travel inside of the cable all the way to the ground block. Figure 20 shows water coming out of the connector at the ground block end of the drop.

Figure 19 - Damaged coax jacket where water was able to enter the cable

Figure 20 - Water coming out of the end of the connector at the ground block

The fix was to replace the subscriber drop from the tap to the ground block. Figure 21 shows the FBC screen shot after the new drop was installed, and Figure 22 the post-repair SC-QAM performance.

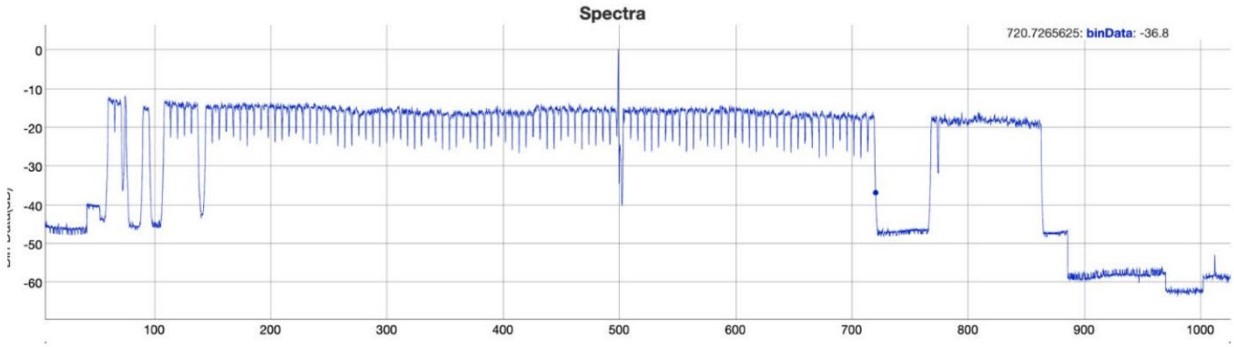


Figure 21 - FBC response after drop cable was replaced from the tap to the ground block

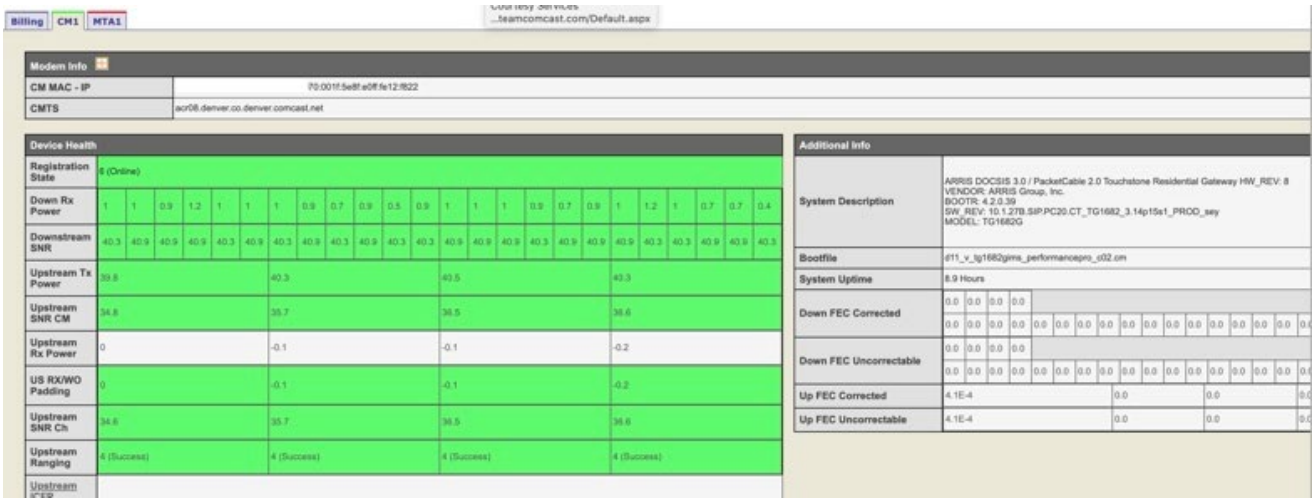


Figure 22 - 40+ dB RxMER, signal performance after drop cable replacement

4.1. Environmental Influence

The impact of temperature and thermal influence on cable networks is well-known. When the cable plant is structurally intact, the system is designed to cope with hot and cold temperatures. However, when the characteristic impedance is compromised, things can become predictably unpredictable.

4.2. Severity Assesment

Fortunately, as illustrated in Figure 23, our fundamental DOCSIS signal quality measurements are excellent for determining the customer impact. The water causes a directly observable degradation in downstream signal power, which can result in degraded receive modulation error ratio (RxMER) and ultimately poor, unreliable performance. Figure 23 enumerates the basic order for conducting a customer impact and severity assessment.

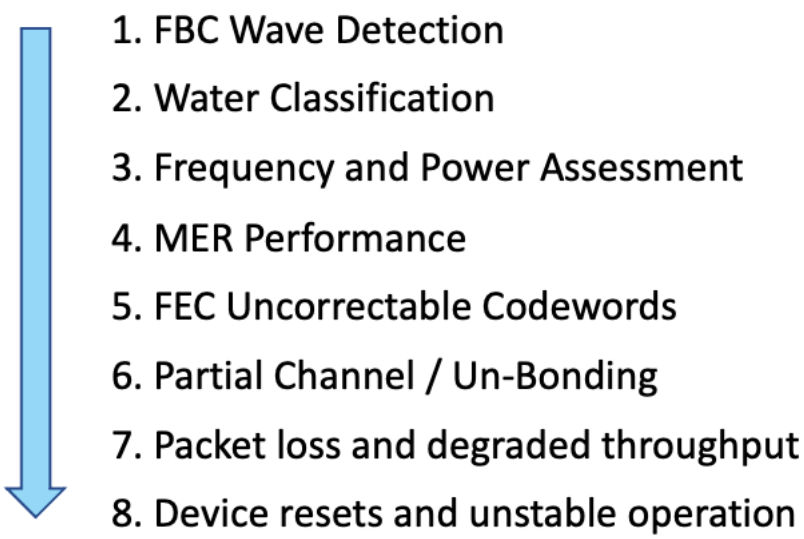


Figure 23 - Severity assessment

Beginning from step 3 in Figure 23, it's common for operator tools to provide some if not all the information. It's fair to say that operators have already been dealing with the outcome of water-soaked cables. Unfortunately, lacking the detection and classification of water damage, the repairs can be inconsistent and often-times, unpredictable. And when the technician has some information about the likely cause of the problem, knowing what to look for is easier so troubleshooting can be faster, and repairs are more likely effective.

4.3. Proactive Repairs

By adding steps 1 and 2 in Figure 23, operators can proactively identify and attribute water damage as the cause. One of the common effects of water damaged cables is a progressive decay of service quality. This is influenced by a number of factors including the amount of moisture, freezing, thawing, heating and cooling. These environmental influences contribute to water migration and accumulation. In some circumstances, these damaged cables can be detected, located and repaired prior to affecting the customer's service.

5. Water Wave Detection Methods

This section covers the methods defined for automatically differentiating between impedance mismatch-related standing waves and water in a coax cable. When it was discovered that there were visually discernable differences between the two impairment types, work began to algorithmically differentiate between the two impairment types. With a few known results to start with, we have tested these methods enough to develop them. After applying the methods to known test results, one known wet drop was added to the PNM test rack at CableLabs and further confirmed after time, after some drying of the cable, too.

5.1. Generalities

Figure 24 shows a typical standing wave in downstream spectrum capture data on the top, while Figure 25 shows the same for a water-soaked cable. The second plot in each figure shows the respective spectrum data after inverse Fourier transform (IFFT) filtered autocorrelation time domain values, and the last graphs in each are cumulative distribution functions (CDFs) of the time domain values obtained by ordering the values after the IFFT from largest to smallest and taking the cumulative values for each observation. Note the differences between the standing wave and water wave plots. First, the waves in the spectrum data are scalloped and repeating in the standing wave case (standing), but do not follow closely any repeating pattern in the water-soaked cable case (wet). Looking at the transformed data in the middle plots of each figure, the standing case appears to have a strong peak, whereas the wet case is more spread out. Translating these data into CDF plots at the bottom of each figure, we see that the standing case has an initial spike and then a more gradual curve up, whereas the wet case starts lower and has a steeper initial climb.

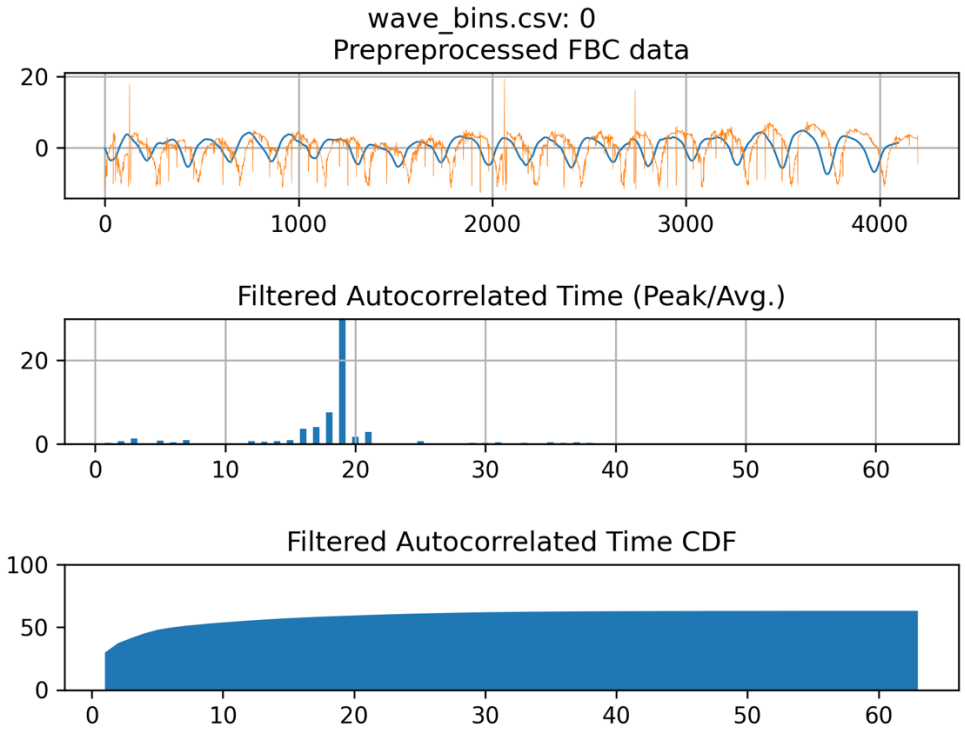


Figure 24 - Standing wave spectrum capture plot, IFFT transformed time domain plot of that same spectrum data, and a CDF of the time domain data

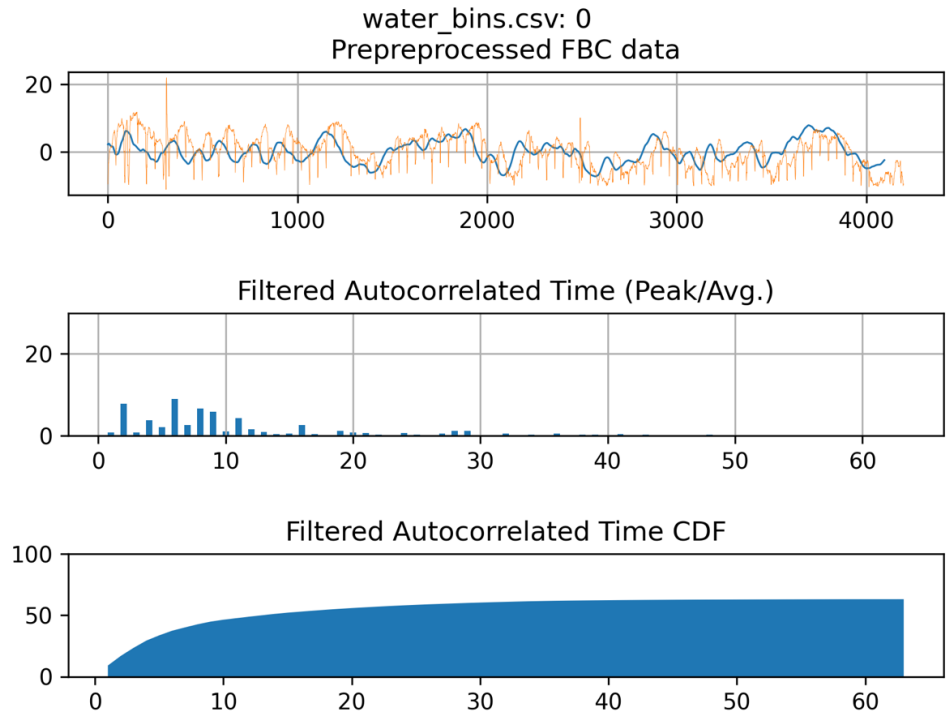


Figure 25 - Water wave spectrum capture plot, IFFT transformed time domain plot of that same spectrum data, and a CDF of the time domain data

5.2. Test 1: Comparisons of a known wet drop and a like section of drop cable.

We obtained a drop from the field which we could visually confirm the existence of water in the drop. We then created a new drop using the same cable length and type for comparison in the tests that follow.

For another indication of the existence of water in a coax cable, see the S_{12} measurements for a normal unimpaired section of coaxial cable (Figure 26) versus the same type of cable that has been affected by water (Figure 27). Note the loss as a function of frequency is rather smooth in an unimpaired cable, but it is not smooth at all in the case of a wet cable. For a treatment of S-parameters including S_{12}, see the PNM point of view document on full duplex DOCSIS® or Ron Hranac's *Broadband Library* article on the subject (https://broadbandlibrary.com/a-quick-look-at-s-parameters/).

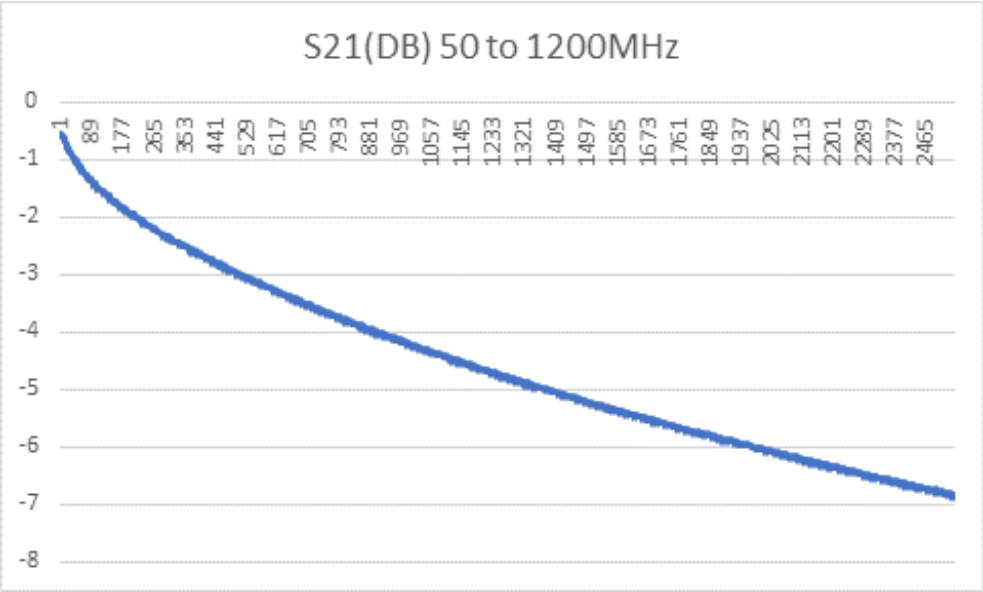


Figure 26 - S_{21} for a normal, unimpaired drop

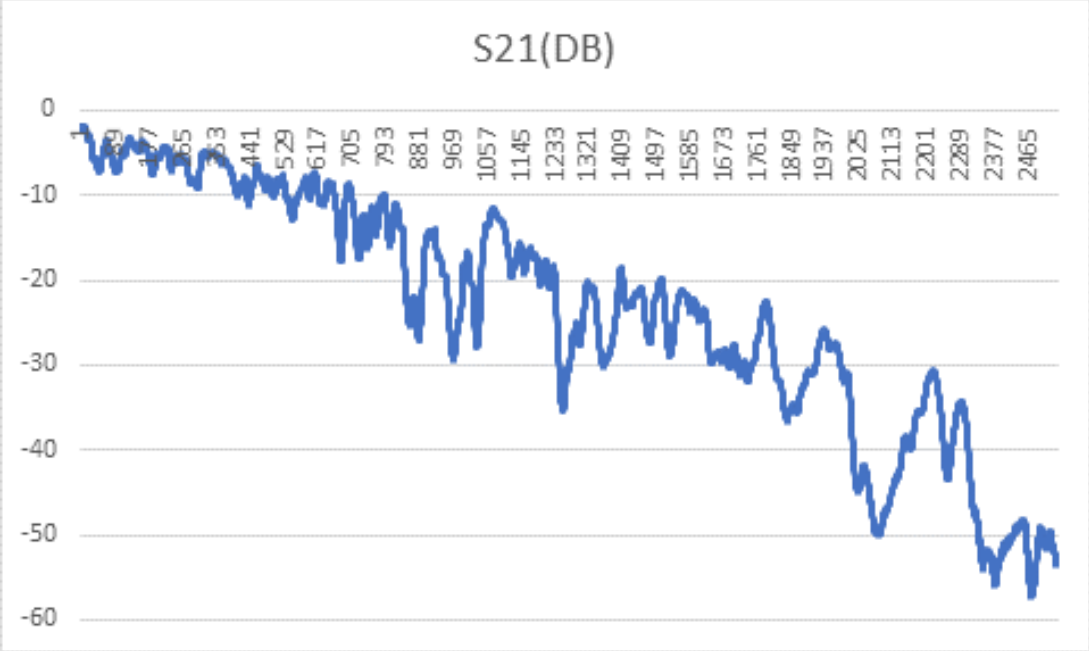


Figure 27 - S_{21} for a water-soaked drop

Next, we examine the group delay of the wet cable, as shown in Figure 28. Note that the group delay is uneven to a small degree over the measurable frequencies, and then at about 1500 MHz measurement isn't possible so the plot shows much higher variability. Ignoring the portion of the plot that is noise and not measurable above 1500 MHz, we see that group delay variation appears in wet cables.

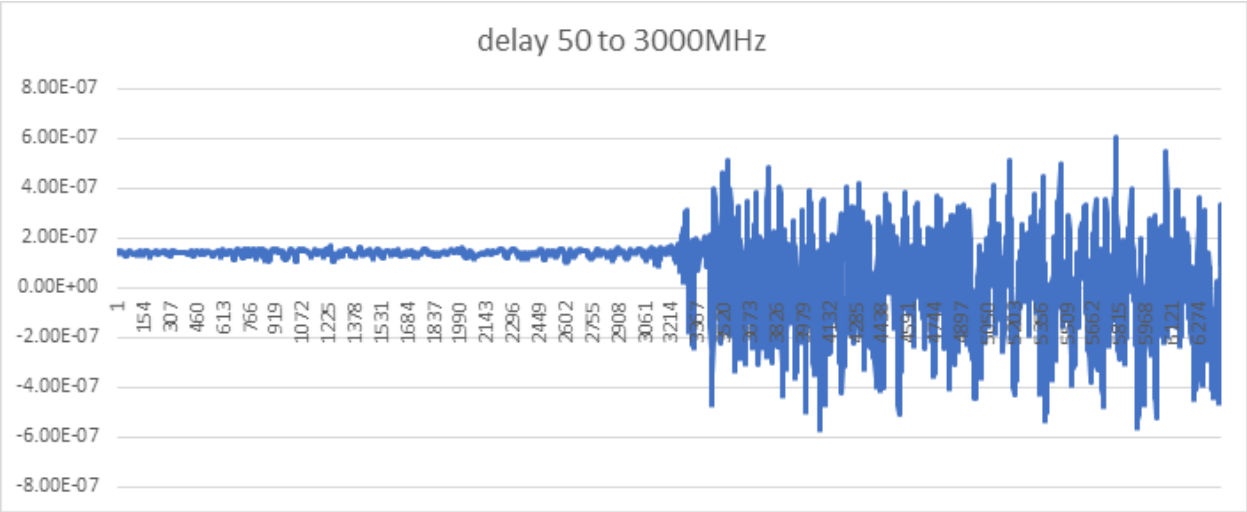


Figure 28 - Group delay for a water-soaked cable

Next, we apply a time domain reflectometer (TDR) to a pair of cables, one wet (WWI cable) and the other in good health (100 ft Series 6), both the same type and length of cable. Here we display the S_{11} and S_{22} values which are reflection coefficients. We see that more energy is reflected in the wet cable compared to the dry cable. Note that we aren't claiming that the entire length of the wet cable is filled with water; we have no way to measure how much of the cable is water soaked and to what degree. We are only showing that some amount of water in the cable will appear differently in reflection measurements (S_{11} and S_{22}).

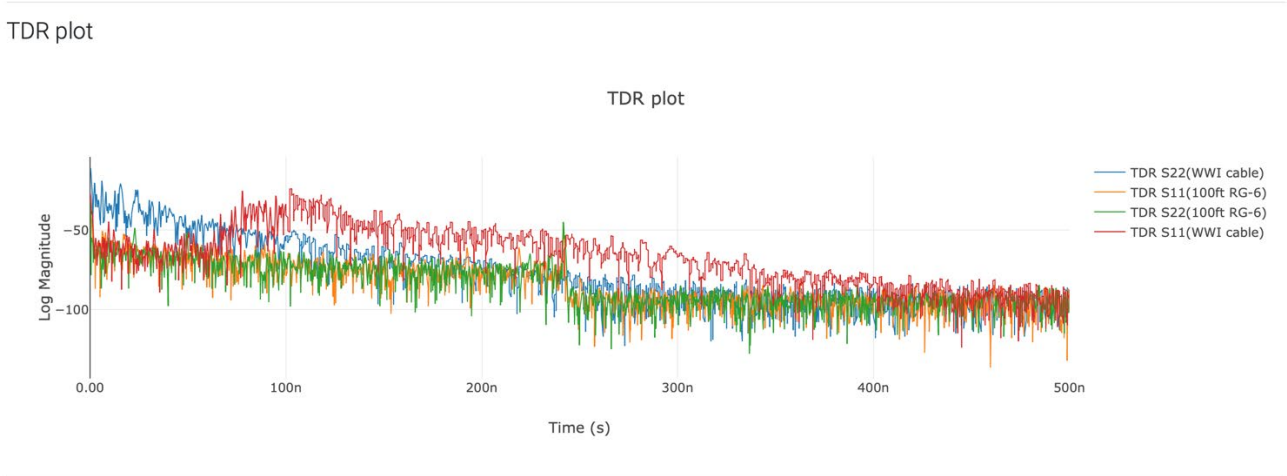


Figure 29 - S_{11} and S_{22} measurements of wet (WWI cable) and dry cable, both 100 feet of Series 6 cable.

Next, we compare the same two cables' S_{11} group delay values. Note again that the wet cable shows variability over the measurable frequencies before about 1.5 GHz, and lots of noise over the higher frequencies that are not reliably measurable. But the unimpaired cable shows nearly flat across the entire

plot, at both lower and higher frequencies. Clearly, wet cable has reflective properties, and impacts group delay as well.

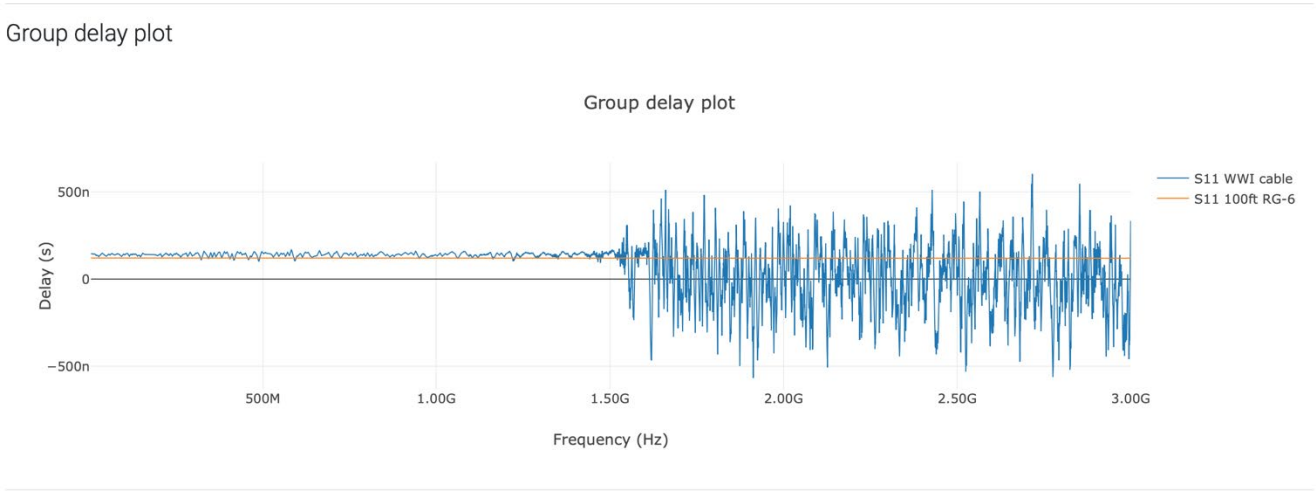


Figure 30 - Group delay plot of S_{11} values for a water-soaked cable (WWI cable) and a clean, unimpaired, dry cable, both 100 feet of Series 6 cable

Looking at the magnitude S-parameter values for these two cables, shown in Figure 29, we see several differences.

- Comparing S_{11} values, the wet cable's values (blue) are higher and slightly more variable than the dry cable's (orange).
- Comparing S_{12} values, the wet cable's values (green) are highly variable and are lower values than the very stable and higher values for the dry cable (straight red line).
- Comparing S_{21} values, the wet cable's values (purple) are again highly variable and are lower values than the very stable and higher values for the dry cable (straight brown line).
- Comparing the S_{22} values, the wet cable's values (pink) are higher and a bit less variable than the values for the dry cable (grey).

While any of the S-parameters can be used to differentiate a wet cable from a normal dry cable, S_{12} and S_{21} appear to show the difference most clearly.

Note: We have not tested and shown a cable with a standing wave for comparison. A follow up step would be to include other types of cable impairments to see if S-parameters can be used to differentiate between different types of impairments. But because the S-parameters can technically define whether a cable is impaired or not, the exercise would be only to determine if S-parameters can be used to differentiate between different types of impairments.

Our primary concern with this test is to find data we can utilize to differentiate between an impedance mismatch-related standing wave and a wet cable. Both are impaired, but a wet cable may be more difficult to spot visually in the field, yet likely isolated to a small span of hard line or a single drop and may be easy to isolate using FBC in CMs.

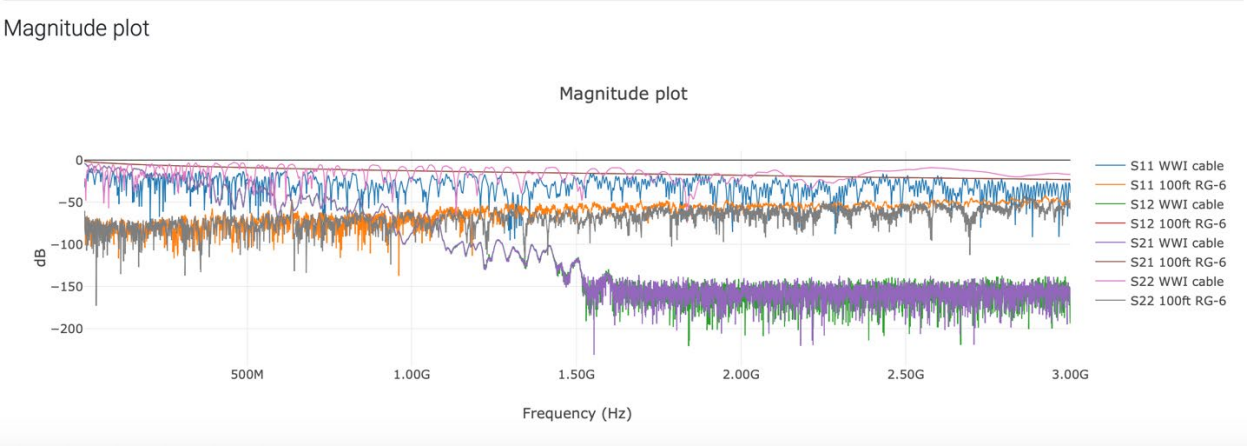


Figure 31 - S-parameters for a wet cable (WWI cable) and a dry, unimpaired cable of the same length and type: 100 feet of Series 6 cable

5.3. Methods for differentiating water from standing waves in coax

The previous test results suggest a few competing methods for determining whether a coaxial cable has been impaired by water intrusion versus is damaged in a way to cause a standing wave. In this paper, we show three promising methods.

We outline two early methods here, and show evidence of their utility, with the expectation that future tests will allow us to compare the effectiveness of these methods and perhaps develop improved methods. Both methods rely on spectrum data obtainable from the CM but can be applied to spectrum data from other sources. An advantage of these simple methods is that they rely on common modem spectrum capture data, and complex data. But we intend to research other methods and data sources for comparison and to improve the reliability of our methods. The intent of these two methods is to identify impairments that can be quickly found and repaired with a clear net positive impact on service and plant health.

Both methods explained here rely on removal of spikes and other noise in the data through smoothing methods, and then the samples are normalized for processing.

A third method is explained here as well, which takes a TDR-like approach and applies a simple threshold, which is consistent with other methods in use, and will be easy for technicians to follow.

5.4. Tom's IFFT method

CableLabs' Tom Williams suggested a simple IFFT of the spectrum data, then to manually look for the clear difference in the time domain data. Group delay in the wet cable should reveal more energy later. Tom's method is essentially as follows.

1. Take a spectrum response, then filter, flatten, and interpolate to remove high level responses such as pilot signals, and low-level responses such as unused spectrum.
2. Apply an IFFT to produce an impulse response to look for dispersion.
    a. If one or two narrow lines are found, this indicates a standing wave.
    b. If the time response is distributed beyond one or two lines, water in the cable is indicated.

3. To improve detection, apply autocorrelation to the time response data. Measure the impulse response coefficients relative to the DC term, remove the two largest coefficients, and remeasure the coefficients. Search for a small drop from these two coefficients to the third as an indicator of water.
4. Search for a downward tilt in the spectrum plot, from low frequencies to higher frequencies. If this tilt is not intended in the plant design (which should be very rare), this is another indication of water in the line.

Tom further tested his method on several suspected CMs whose data were obtained from the field but not all confirmed as standing wave versus wet. Using the suspected impairment categories, he plotted the tilt and an error ratio improvement relative to DC, then drew a line between the categories to form a function that can serve as a threshold for differentiating the two impairment types. The method is updated then as follows.

1. Remove tilt from spectral response. Record the tilt.
2. Perform an IFFT to get time domain data.
3. Select time samples <65 to eliminate the tallest wave response.
4. Record error ratio improvement relative to DC term.

Figure 32 shows a plot of the resulting tilt and ratio statistics from 10 standing waves (as seen in the spectrum data manually) and 10 wet waves (also as seen in the spectrum data).

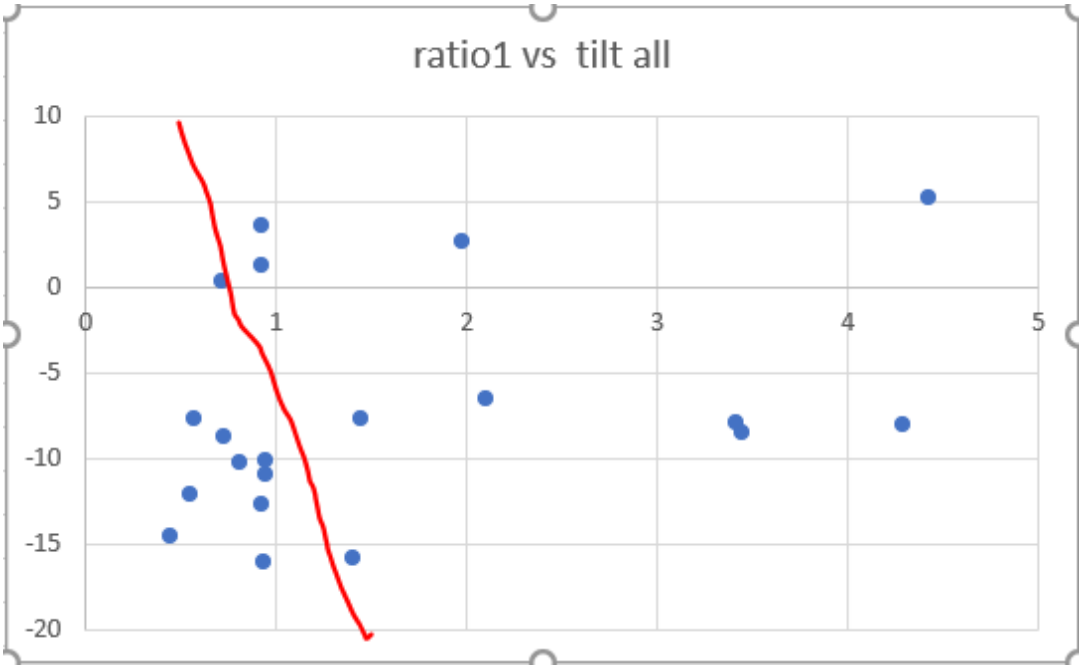


Figure 32 - Error ratio versus tilt, showing a red line that differentiates between a wet cable (left and below) versus a cable with a standing wave (above and right)

Because these 20 spectrum captures are from CMs that have not been confirmed to have standing waves versus water waves, we can only say that they were impaired, and that the two impairment types were

visually clustered. Further, these 20 results were picked from a much larger pool of CM responses for their obvious impairments and strong visual differences in the two types of impairment patterns (standing versus wet).

To be certain of the effectiveness of this method, more field data must be collected, and confirmation of the cause should be conducted if the difference is important.

5.5. Jason & Jay's method

CableLabs' Jason Rupe and Jay Zhu, purposely working independently from Tom Williams to come up with a different method, developed a method very similar to Tom's method and extended it to better enable computer programs to differentiate between the two impairment types, relying on proven machine learning and statistical methods to identify that the spectrum data shows an impairment. Once an impairment is discovered to exist, we want to determine whether it is a wet cable or a standing wave primarily.

This method extends off the first by calculating a CDF, then fitting a curve to the resulting data, and using the resulting parameters to differentiate between a standing wave and a water wave. Recall Figure 24 and Figure 25 which showed the CDFs and the initial jump in a standing wave versus the slower climb of the water wave data. By fitting a function to the CDF, we obtain parameters that describe the desired CDF pattern which can then be mathematically compared to rules that will be statistically determined later from field data and can be initially determined using the same data used in Tom's method. The method steps are as follows. This procedure picks up after the spectrum data are cleaned, and an IFFT is performed to obtain the time domain data, which are used in the procedure.

1. In this step, we preprocess the data for additional clean up, and apply Tom's method to get the autocorrelated magnitude values.
2. Calculate the CDF from the time domain data. This requires taking the IFFT results from the spectrum data, in the time domain, and sorting from largest to smallest, then calculating the cumulative of the current and all previous observations for each observation.
3. After calculating the CDF, we fit an exponential curve of the form $a*e^{(-b*x)}+c$ and keep the three parameters.

The parameters of the curve fit are then used to characterize the captured spectrum data from each CM. In the figures that follow, we show a few plots of data from CMs, some with standing waves and others with water waves. Figure 33 is a set of plots from the water impaired drop obtained from the field and then placed in the PNM lab at CableLabs. The data were captured months after placing the drop cable in the lab, and after some drying has taken place. Note that the spectrum plot at the top left almost shows a repeating scalloped pattern, and the filtered autocorrelated time domain values show much energy in a narrow area. This cable may appear almost like a standing wave. This result suggests that a cable that cycles between being heavily water intruded and drying out some may actually look like a standing wave at times. Collecting data over time might confirm the issue, but a CM with this severe of an issue, with neighbors who indicate no issue, should be quick to repair with a drop replacement, regardless of the cause.

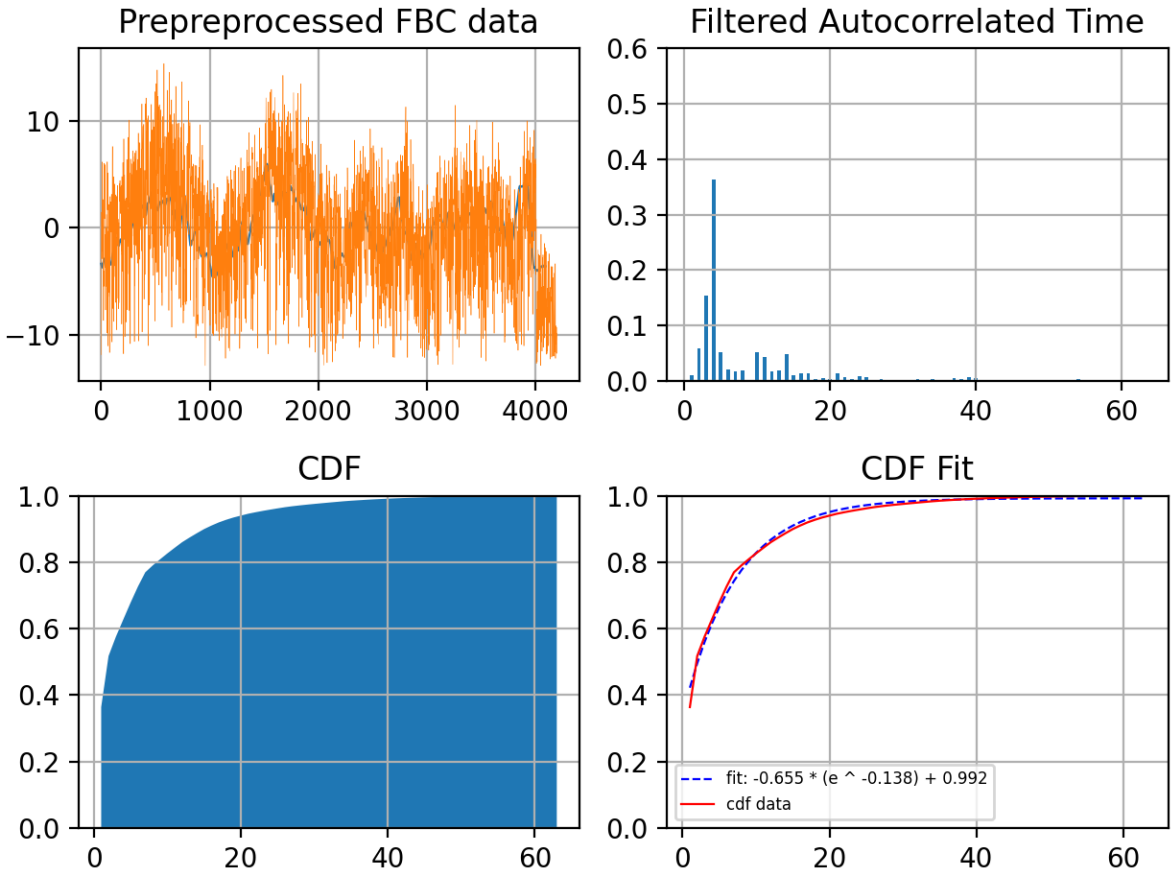
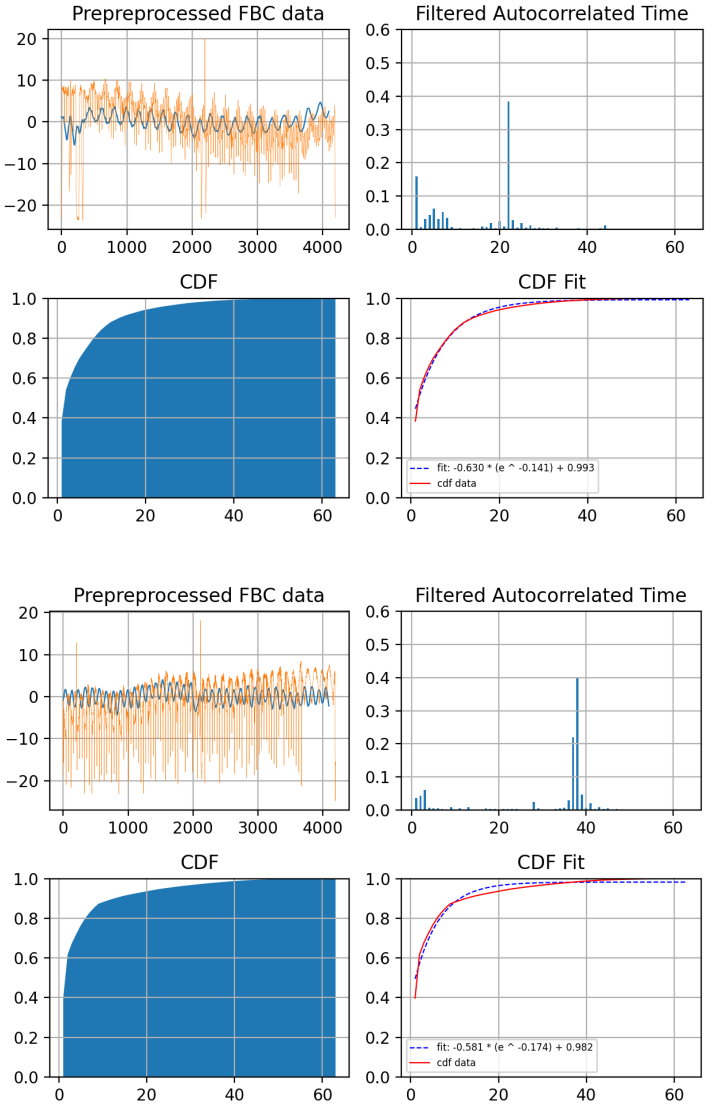


Figure 33 - A confirmed wet drop from the field which has been installed in the PNM lab at CableLabs, and the CM on the end of the drop used to gather spectrum data then processed with the CDF curve fitting method

Figure 34, and Figure 35 show a few CMs with standing waves, and water waves respectively, all sampled from the same 10 used in Figure 33 of Tom's method. Note that, as expected, the standing waves show a sharper CDF curve than the water wave CMs do. While subtle, we expect that the CDF parameters will show this difference and allow us to create, through a large sample of confirmed cable sections, a statistical model that can, based on the parameters of the fit, assign a likelihood of water intrusion versus a standing wave in the cable.



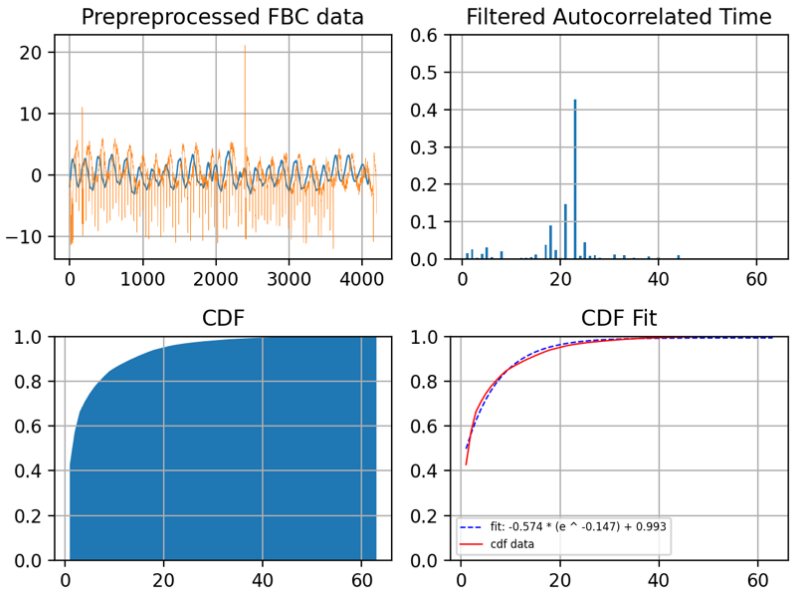


Figure 34 - Three selected suspected CMs with apparent standing waves

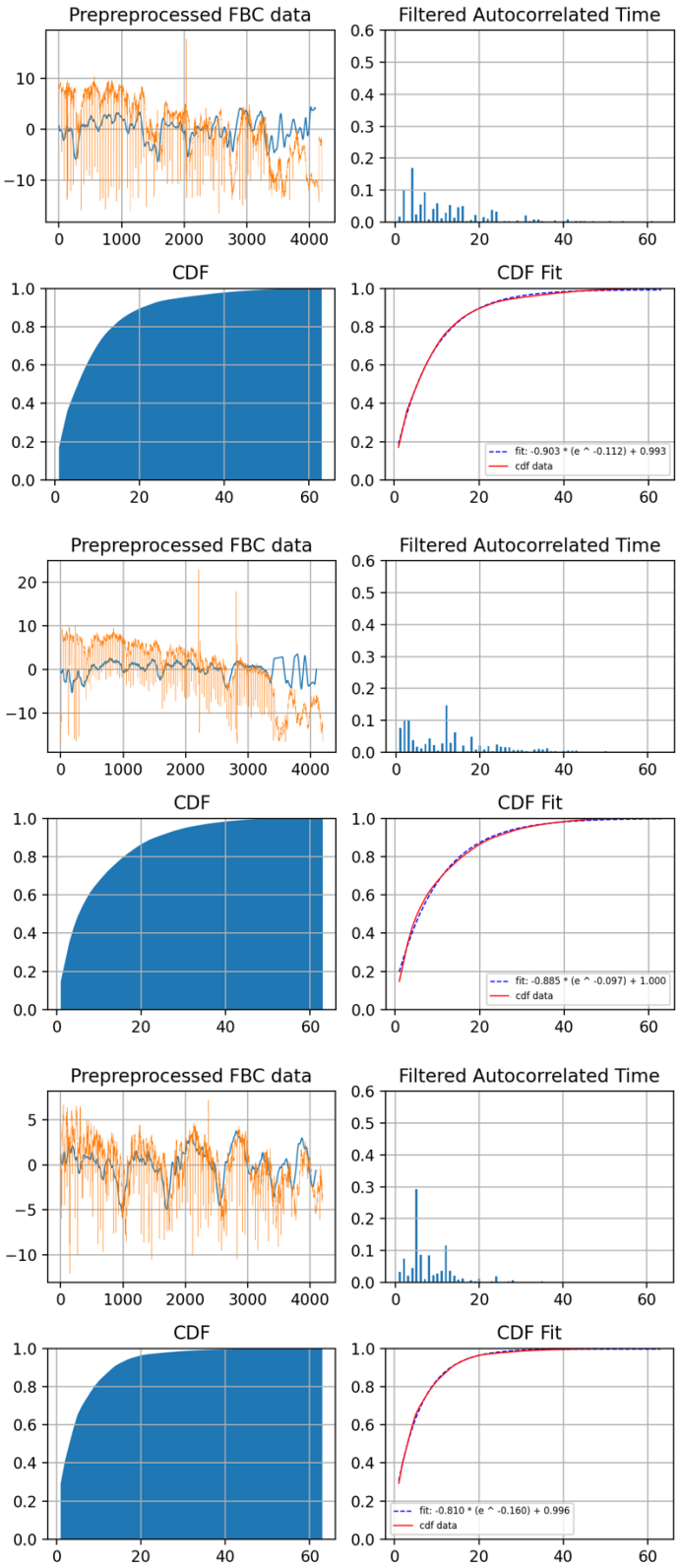


Figure 35 - Three selected suspected CMs with apparent water waves

Extracting the three fit parameters from each of the 20 CMs (10 each believed to have standing waves or water waves), plotting these in 3D in Figure 36, and coloring the point for each CM by the believed class of impairment each has, we see a clear pattern again. Note we show the 3D plot twice in the figure, from different angles, to better show the delineation.

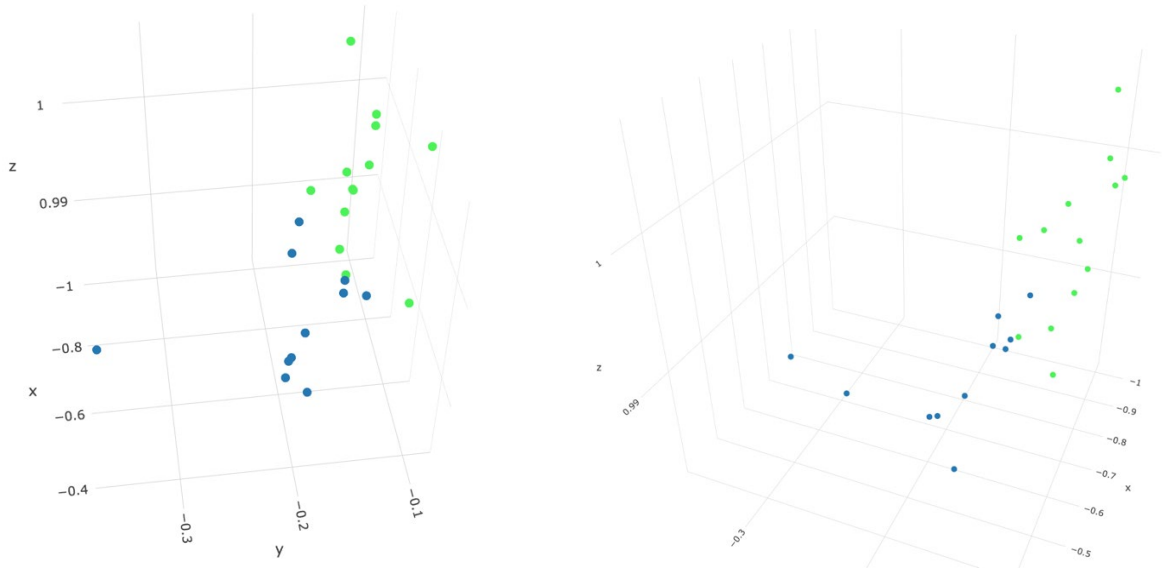


Figure 36 - Curve Fit Plot of Wave Parameters

Figure 36 is a plot, from two perspectives of the three parameters from curve fitting of the 10 CMs with standing waves (blue) and 10 CMs with apparent water waves (green); X=a, Y=b, and Z=c, with X, Y, Z being the plotted values for the parameters a, b, c in the curve fit.
Note that the CMs with suspected water waves tend to have higher (less negative) Y=b values, higher z=c values, and often smaller (more negative) x=a values than those labeled to have standing waves. This pattern can also be seen in the data shown in Table 1, from these same 20 CM spectrum captures (one from each of 20 CMs identified as having either a standing wave or a water wave in the spectrum data) plotted in Figure 33.

Table 1, exponential function fit parameters from the CDFs of time domain data obtained though IFFT method applied to the spectrum capture data from 20 separate CMs, selected from a large group of field data as indicating an impairment, 10 with suspected water waves, and 10 with standing waves. Fitting function: a * e ^ (-b * x) + c.

Table 1 – Exponential function fit of 20 CMs

a	*b*	*c*	*Type*
-0.33276052	0.12582213	0.99626715	Wave
-0.62065159	0.15781587	0.99200869	Wave
-0.60244084	0.17260292	0.9840218	Wave
-0.69048994	0.12123575	0.9928308	Wave
-0.33739765	0.11461487	0.99332933	Wave
-0.73274074	0.3368255	0.98163125	Wave
-0.41528807	0.1360504	0.99606652	Wave
-0.66520054	0.1371799	0.99343827	Wave
-0.60399443	0.16658992	0.98946338	Wave
-0.66897829	0.15190012	0.99240217	Wave
-0.89719314	0.09721336	1.00034371	Water
-0.96202471	0.1512644	0.98994818	Water
-0.86697703	0.12554029	0.99425639	Water
-0.74403271	0.14132146	0.99422719	Water
-0.91005527	0.09920093	1.00122611	Water
-0.78745142	0.09696355	1.00375195	Water
-0.82397893	0.10764474	0.99936019	Water
-0.87474122	0.1210256	0.9909223	Water
-0.9056323	0.09884677	1.00366479	Water
-0.61743961	0.09148825	0.99392027	Water

5.6. Larry's Method

One of most common methods for field technicians to validate the presence of water is by use of a time domain reflectometer (TDR). It is common for a TDR to transmit a fixed-width impulse and capture the reflected response, or echo. The response can be analyzed, in conjunction with the known parameters of the cable (Appendix A) and typically display the fault distance(s). In the case of water reflections, the time domain impulse response shows a distinctive signature compared to a singular point of damage.

By using the full spectrum amplitude bins with some additional processing, the functionality of the TDR can be approximated. Figure 37 shows 3 FBC traces, the top is unimpaired, middle is water damage and bottom are a typical amplitude ripple caused by a standing wave.

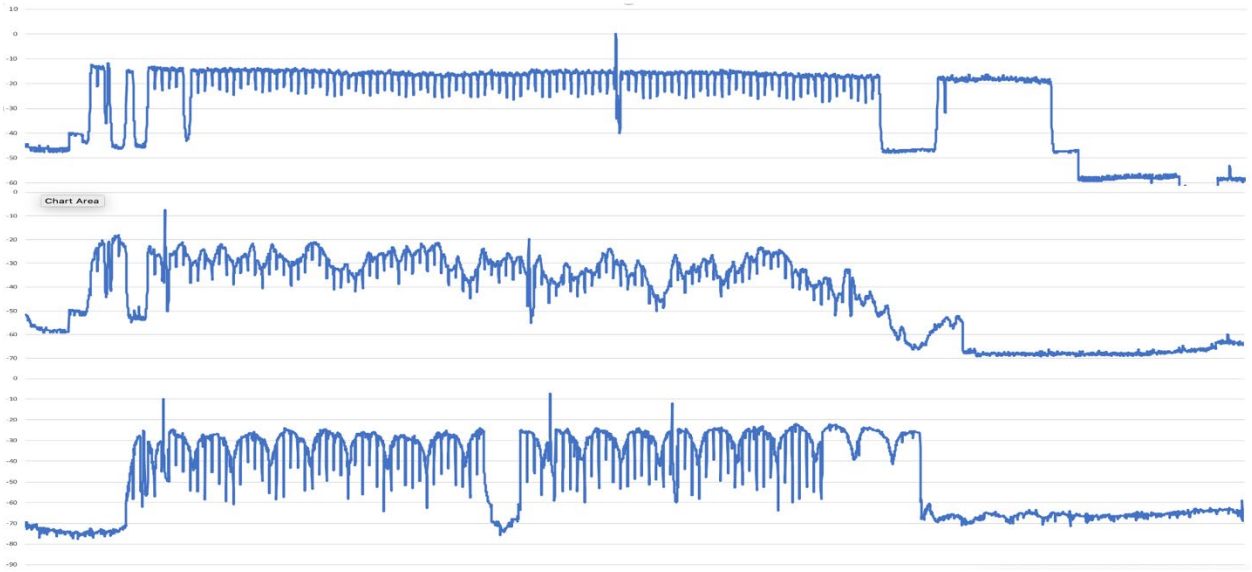


Figure 37 - FBC samples compared, unimpaired (top), water damage (middle), standing wave with amplitude ripple (bottom)

A bit of pre-processing the amplitude bins produces an improved result. The next step illustrated in Figure 38 demonstrates the frequency spectrum with guard band and vacant spectrum being interpolated, or "filled in." In these examples, a simple linear interpolation is done between the SC-QAM and OFDM channel alpha region. This is done to minimize the effect of unoccupied spectrum that would otherwise result as noise in the result. Figure 38 shows unimpaired, water damage and amplitude ripple (top to bottom), after interpolation between the known signal energy.

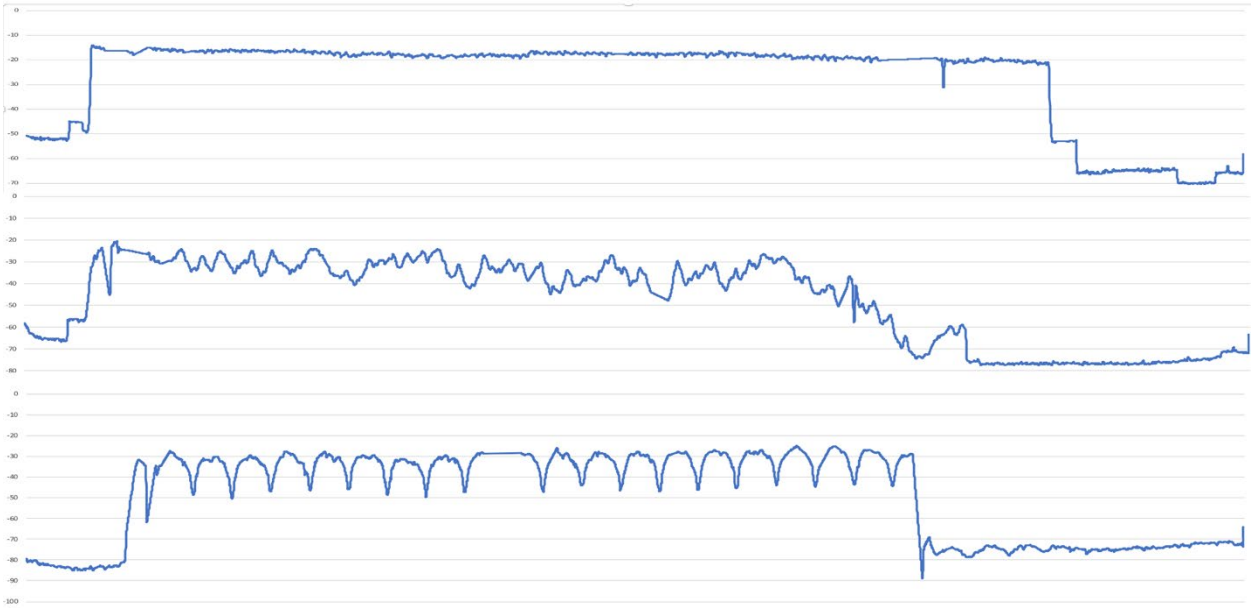


Figure 38 - Interpolated FBC samples with guard bands and vacant spectrum "filled"

After interpolation, the unoccupied spectrum can be cropped, leaving a contiguous block of signal energy, removing any remaining noise floor spectrum. Figure 39 shows the center 4096 bins after removing the other bins. Again, similar to Figure 37 and Figure 38, the top trace is unimpaired, center contains water and bottom is an amplitude ripple caused by a standing wave.

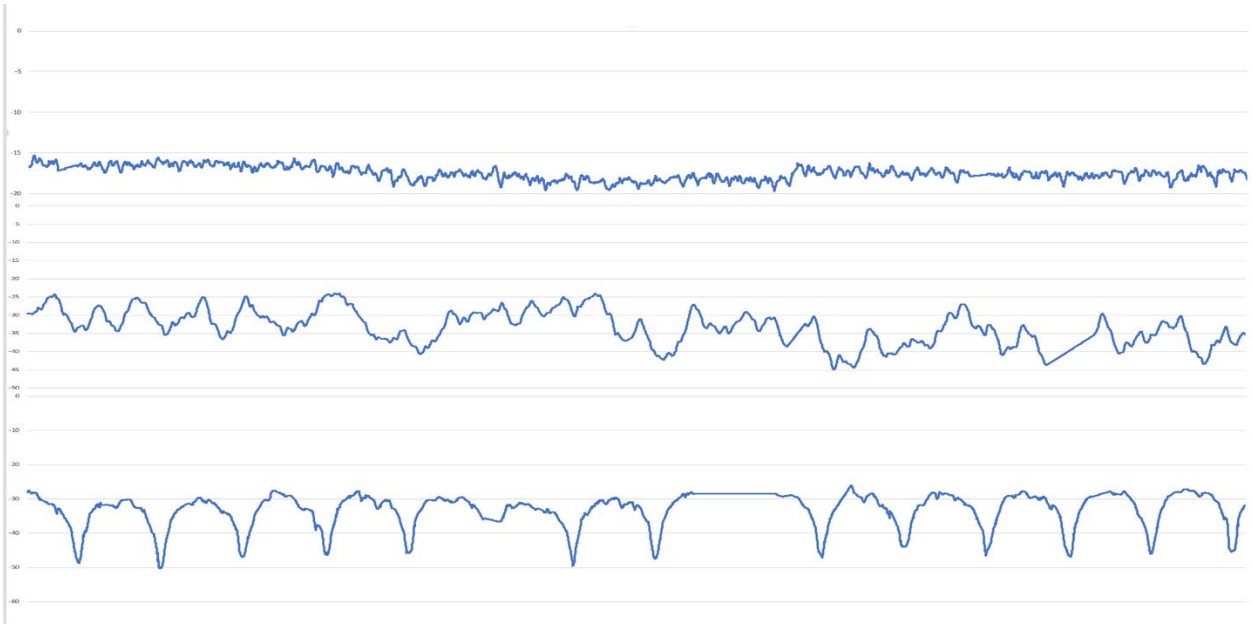


Figure 39 - Samples are cropped to 4096 bins of occupied spectrum

After some pre-processing, the amplitude bins can now be projected to the complex plane, using digital signal processing (DSP). The intent is to perform a Fourier transform to render a time domain impulse response, like the previously discussed TDR. However, the full band capture bins are represented as logarithmic magnitude, meaning the in-phase and quadrature (I&Q) have already been summed and squared. While it is impossible to recover the actual I&Q values, minimum phase assumptions are sufficient derive a reasonable impulse response. This is achieved by instantiating a complex number with zero phase component, uniformly for each bin. Then using an IFFT, the impulse response is obtained (Figure 40). Notice that the top unimpaired trace shows no echo, center trace shows energy spreading and bottom has a singular echo response.

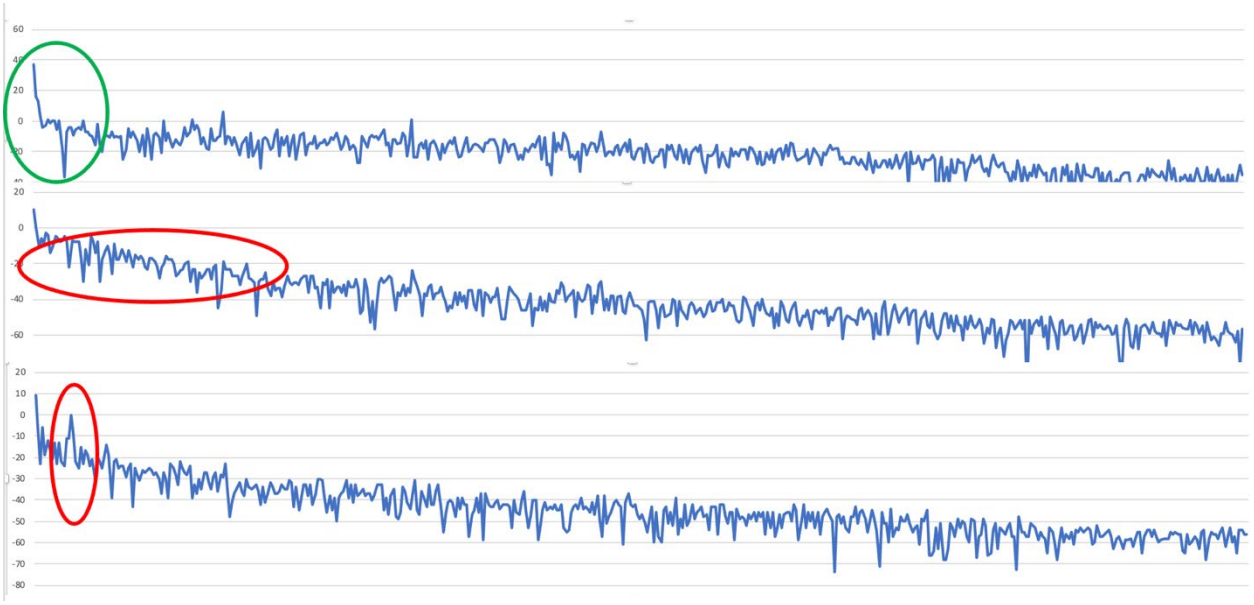


Figure 40 - Log magnitude bins are converted to time domain using zero-stuffed IFFT technique, using minimum phase assumptions

Finally, when multiple time domain impulse responses are overlayed in Figure 41, the spreading becomes more obvious. The red line indicates a threshold that correctly distinguishes between water and amplitude ripple on 100% of the examples provided.

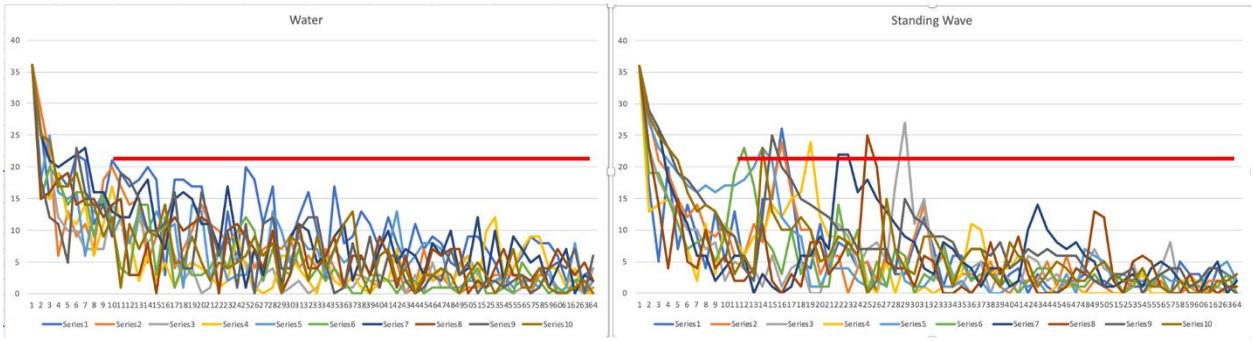


Figure 41 - Time domain spreading in water (left), and peak threshold detection (right)

6. Operational Practice Consideratons

Of the samples recovered in the field trials, virtually all the water damage could have been avoided. There are a number of recommendations discussed below to help reduce the occurrence of water damage.

6.1. Materials Selection

When installing the drop system, consideration of the cable and components are key in the performance and life expectancy of the system along with the methods and practices used for installing it. Our focus for this section will be on how to maintain a weather tight drop cable network.

Let's begin with material selection. When selecting RF cable ANSI/ SCTE 74 2011 Specification for Braided 75 Ohm Flexible RF Coaxial Drop Cable can provide the information needed to ensure that you select the correct cable for the proper use and purpose. Two of the key components of the standard will be jacket construction, designed for either aerial or underground and flooding compound for both aerial and underground cables. Before we get too far, we should discuss flooding compound. Flooding compound in both aerial and underground cable is meant to help preserve against rapid degradation due to corrosion only. It is not intended as a self-repair component and therefore any damage to the jacket of aerial or underground cable should be replaced.

Connectors used should be of the 360-degree compression style with integrated weather seals to limit water migration. The other component that we need to maintain a weather tight, moisture-proof drop cable network would be weather seals for RF port connections. These port seals may be integrated as part of the connector or a separate piece. In either case it is important that we are using the proper seals whenever the connection may be exposed to outside elements or fluctuations in temperature or areas of high moisture such as basements, garages, crawl spaces, pedestals, lock boxes, house boxes etc.

6.2. Installation Practices

Now that we have the materials let us look at how to install them to ensure the integrity of the drop system for years to come. The points that we will cover are those specific to the weatherproofing of the drop system and is not meant to reflect all considerations when installing the drop system.

When preparing the cable for installation we need to be careful with the tools we use and how we use them as damage to the jacket can easily happen. Let's start with proper fitting, preparation, and installation. Using the correct prep tool for the cable size and fitting style as well as ensuring that the tool is sharp will ensure a good fit between the connector and the cable being used. Next is the compression of the connector. For this to be successful we need to be sure we are utilizing the correct compression tool for the fitting being used and that it is in good working condition. A visual inspection of the compression shall be made to ensure that there is a complete and even compression of the fitting to the cable.
RF Port seals, when installing the seal, we need to ensure that the seal extends past the threads of the RF port and contacts the smooth portion of the barrel connection. If the seal is not integrated with the connector the seal should be installed so that the leading edge of the connector is in direct contact with the seal. The seal does not need to be compressed between the connector and the body of the component of the barrel connector to be properly installed.

Regarding the cable itself there are several practices we need to consider as well. Water will follow the route of the cable, riding the exterior jacket looking for a point of entry and/or flowing on the interior of the cable. Therefore, we utilize the practice of installing drip loops along the pathway of the cable to provide a means to displace the exterior water. Drip loops should also be utilized to ensure that any connections or drop components are always higher than the lowest point of the cable. Utilizing gravity to keep all moisture away from connections and components.

Care must be taken on how we attach the cable and methods used for attachment so that the jacket is not damaged. Attachments such as clips and hangers should not have hard or sharp edges that could damage

the jacket during installation. In addition, care must be taken during installation as a slip of a tool or improper use can also result in damage to the jacket of the cable.

Something as basic as removing the messenger from aerial cable, if done incorrectly, can also lead to damage of the jacket. By utilizing the vertical pull method, not the horizontal “wishbone” method, the jacket will remain intact during the removal process. Extra care must be taken when preparing a mid-span drop so as not to damage the jacket during messenger wire separation.

When installing aerial cable, the pathway of the cable should be such that the cable does not come in contact with objects that could wear against the jacket and compromise the integrity over time. If such contact cannot be avoided, the use of a protective barrier such as tree guard should be used. With underground cable the best way to protect the jacket is the use of conduit. This would protect the jacket from hard sharp objects below the surface that may damage the jacket, provide a means of protection if there is any digging in the area and finally protect the cable at the critical points where it enters or leaves the ground. If conduit is not used to protect the complete path of the underground cable, then it is a must to protect the cable at the points where it enters and exits the ground. This can be done by using smaller sections of conduit or U Guard. This protection should extend 8 inches below the ground and should extend to at least 3-4 feet above ground.

6.3. Inspection for Damage

We should always be performing a visual and tactile inspection of the cable. We need to be on the lookout for cable that does not look or feel right and completing a further inspection based on these observations. All knicks or cuts in the cable jacket can and in time will permit a pathway for water intrusion and therefore must be dealt with. Remember flooding compound is not meant to be a self-repair method. Whenever you see physical signs that water has entered the cable, this may be corrosion, discolored center conductor or moisture, the cable and / or components must be considered compromised and correctly remedied.

6.4. Repair vs Replace

Replacing the damaged section of the drop system should be the preferred method. By repairing or splicing of the damaged section we cannot be sure that there is not still the presence of moisture that will continue to degrade the cable.

6.5. Cable Handling and Storage

How we store the cable prior to use can also have a significant impact on its performance as well. If possible, all cable should be stored inside of the vehicle prior to use. If this is not possible care must be taken to properly protect the cut end of the cable from taking on moisture as it is exposed to the environment. The cable should also be protected from unintended damage from other items that may be stored around it.

7. Future

Water entering our cables is a common problem with unpredictable impacts to our customers. However, rain being the most common cause is somewhat predictable using national or local weather data. We can imagine a future opportunity to get ahead of water related problems by proactively repairing damaged cables in advance of rainy seasons with prolonged precipitation. In addition to weather, other causes such as sprinkler systems can have predictable periodicity and should also be considered.

Another important observation about this type of impairment is that it affects high-frequency spectrum, worse than low-frequency spectrum. This could influence an operators DOCSIS evolution strategy and how to prioritize cable replacements. For example, if considering using RF spectrum above 1.2 GHz, and operator might decide to increase the priority of replacing these damaged cables.

8. Conclusions

Proactive network maintenance has come a long way in the past 11 plus years but has more valuable but hard work ahead.

As CableLabs focused on achieving accurate impairment detection using PNM, SCTE has been developing operational practices that help operators make PNM actionable and affordable. Partnering with Comcast, over 100 water-damaged subscriber drop cables were recovered from the field. The cable samples were then sent to Comcast's Physical and Environmental lab for testing and characterization of the cables' physical condition and RF parameters. The result of this effort is a new, comprehensive understanding of how water, rain, and freezing impact our networks' coaxial cables.

Using spectrum captures from CMs in the field, we created approaches for differentiating water waves from standing waves (amplitude ripple). The unique frequency response signature created by water in a subscriber drop cable provides an easy way to quickly identify affected drops remotely, without a truck roll. Once a water-damaged drop has been identified, a technician can be dispatched to replace it. The result is a new PNM tool, and more efficient field practices to come.

A strong standing wave as well as a strong water wave both can have significant impact on a customer's service, so both must be addressed. But knowing the difference is important. While on the surface it may seem that knowing the difference between a standing wave versus a water wave in the coax plant is a secondary concern, the difference has at least two key important
values.

1. Knowing that the impairment is a water wave versus a standing wave tells the technician what to look for, and where to look for it. A drop affected by water can quickly be replaced once identified. Water-damaged passives can as well, and technicians can look for corrosion, water, nicks in cable, and other plant failure modes that are the causes for the water wave indicated. In contrast, an echo cavity that creates a standing wave may be harder to find and will be indicated by different failure modes in the cable plant.

2. Water in the cable plant can get worse with time. Therefore, early detection is an opportunity to be truly proactive. Removing the problem before it impacts service is best. And if you can identify and remove the problem before it worsens, even better. Early detection affords the opportunity to remove a small problem from the network before severe damage in amplifiers, taps, and other components happens. A quick fix early avoids a lengthy, more costly fix later.

We presented three methods in this paper for finding water in drops using RF spectrum data. We intend to collect validated field data and perform a comparison of the methods in the future, and report on the benchmarking results in an update of the Primer for PNM Best Practices document published by the PNM working group at CableLabs, and expect to reflect the best methods in a future SCTE NOS Working Group 7 field practice. Work is underway now to develop a PNM benchmarking tool for general comparison of PNM methods, allowing us to provide a benchmarking data set to certify and test any packaged PNM algorithm, starting with these water wave methods. A test report will be the output, based on the number of false positives, false negatives, and potential severity weighting of the results. Once

demonstrating the value in this use case, we expect to offer the same benchmarking approach for other PNM methods.

This work was the result of continuous improvement in how the industry develops and uses PNM. Working closely among operators, vendors, CableLabs, and SCTE, we were able to shift to a DevOps approach to PNM, with continuous cycling between the operations in the field with research and development. While simultaneously collecting experience and validating ideas in the field, the tools were developed in software, using the PNM test methods available, to create this new tool for improving network quality and customer services.

We intend take this work further to align the causes with various cable types, ages, and other factors to determine whether certain cable types in certain environments or use cases are better than others, and what this means for future cable plant deployments, maintenance, and the useful lifetime of coax cable plant.

9. Appendix

9.1. Characteristics of coaxial cable

Coaxial cable is a two-conductor transmission line. One of the conductors is called an inner or center conductor and is "…surrounded by a concentric conducting shield, with the two separated by a dielectric (insulating material); many coaxial cables also have a protective outer sheath or jacket. The term 'coaxial' refers to the inner conductor and the outer shield sharing a geometric axis."[1] See Figure 42.

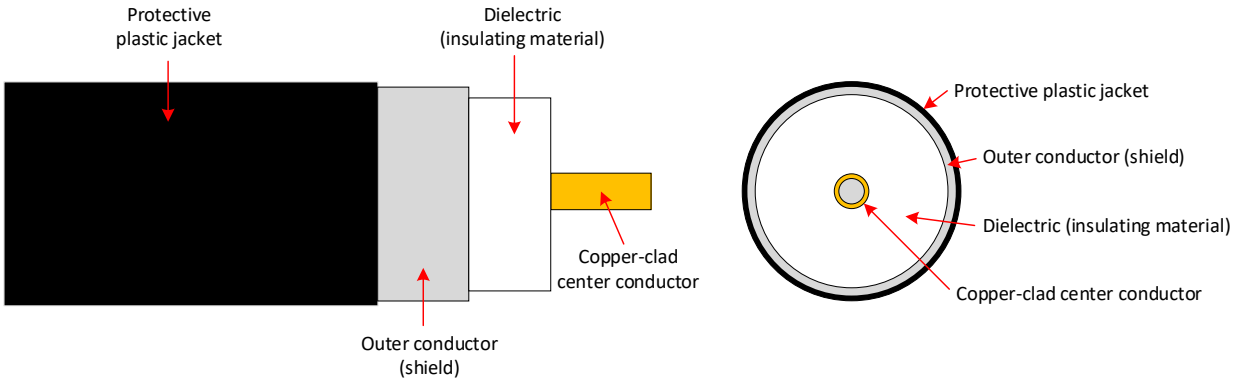


Figure 42 - Coaxial cable side view (left) and end view cross-section (right)

There are two major types of coaxial cable found in cable networks. One type is known as hardline cable and is used in the distribution plant that is attached to utility poles or buried underground. The center conductor is typically copper-clad aluminum (but can be solid copper in some applications), and the shield an aluminum alloy. The name comes from the semi-flexible solid tube-like outer conductor (shield). Hardline cables distribute RF signals throughout the community being served by the cable operator, and in many cases also carry 60 volts to 90 volts AC to power nodes and amplifiers. Figure 43 shows two examples of hardline coaxial cable.

[1] From Wikipedia: https://en.wikipedia.org/wiki/Coaxial_cable

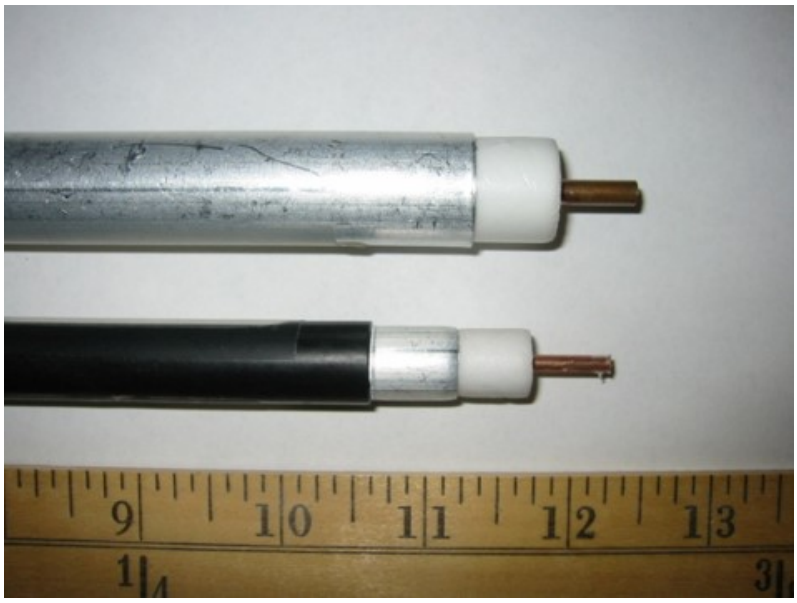


Figure 43 - Examples of hardline coaxial cable: unjacketed 0.750 inch diameter (top) and jacketed 0.500 inch diameter (bottom)

The second type is a smaller diameter, flexible coaxial cable used for the subscriber drop, which is that part of a cable network between the hardline distribution plant and the customer premises equipment inside the home. The center conductor is typically copper-clad steel, and the shield a combination of Mylar-backed aluminum tape and braid. See Figure 44.

Figure 44 - Series 6 coaxial cable with the end prepped for installation of a connector

In both hardline and subscriber drop cables used in cable networks, the dielectric is a closed-cell gas-injected foam. The protective jacket can be polyvinyl chloride (PVC) or polyethylene (PE) plastic, depending on application.

The following summarizes some of the electrical characteristics of coaxial cable. When any of these parameters deviates from desired nominal values, the performance of the coaxial cable and the cable network can degrade.

9.1.1. Impedance

Generally speaking, *impedance* is the combined opposition to current in a component, circuit, device, or transmission line that contains both resistance and reactance. Impedance is represented by the symbol Z and is expressed in ohms. Impedance is further defined as the frequency domain ratio of voltage to current, $Z = E/I$. Impedance in an alternating current circuit, including RF, is a complex value and includes both resistance (the real part of complex impedance) and reactance (the imaginary part of complex impedance) – that is, both magnitude and phase. Impedance can be thought of as a way to describe the concept of AC resistance.

The *characteristic impedance*, Z_0, of coaxial cable is expressed in ohms, and is related to the outside diameter D of the inner or center conductor, the inside diameter d of the outer conductor or shield, and the dielectric constant ϵ (relative permittivity) of the insulating material (dielectric) separating the two conductors. Cable networks use coaxial cables with a nominal characteristic impedance of 75 ohms. As long as the characteristic impedance of the signal source, coaxial cable transmission line, and load or termination to which the cable is connected is the same (that is, 75 ohms), essentially all RF power from the source is delivered to the termination or load, except that which is lost to attenuation.

The following formula can be used to calculate the characteristic impedance of coaxial cable.

$$Z_0 = \frac{138}{\sqrt{\epsilon}} log_{10} \frac{D}{d}$$

where
Z_0 is the cable's characteristic impedance in ohms
ϵ is the dielectric constant of the insulating dielectric material
log is base 10 logarithm
D is the inner diameter of the shield
d is the outer diameter of the center conductor

The calculated characteristic impedance of 0.500 diameter hardline coaxial cable, assuming a dielectric constant of 1.32, a shield inner diameter of 0.452 inch, and a center conductor outer diameter of 0.109 inch, is 74.2 ohms. Anything that affects the cable's dielectric constant and/or the ratio D/d will change the impedance. Figure 45 illustrates the D/d relationship in coaxial cable. Figure 46 shows coaxial cable that has a kinked shield, resulting in a different D/d ratio, which in this case would cause the impedance at the point of damage to be reduced from what it is in the rest of the cable.

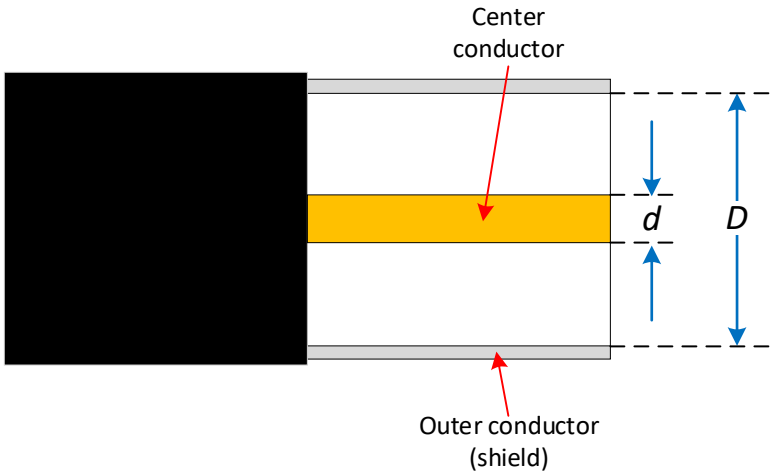


Figure 45 - Illustration of the D/d relationship in coaxial cable

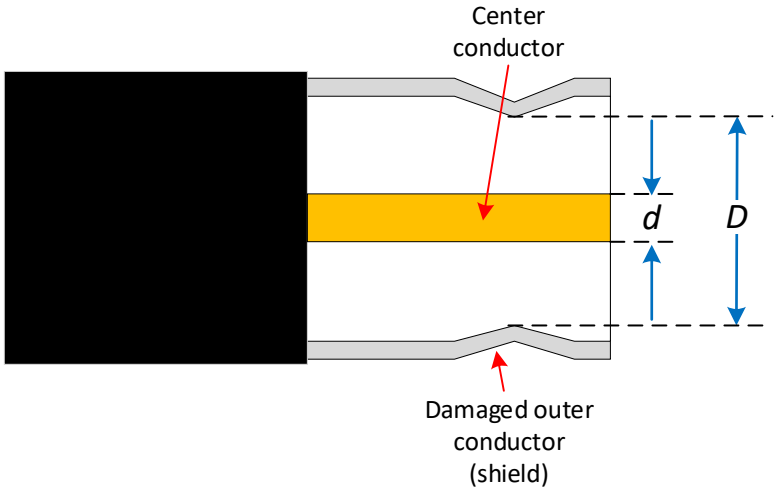


Figure 46 - Damage to the shield in this example changes the D/d relationship, resulting in a change of impedance at the point of damage relative to the rest of the cable

9.1.2. Attenuation

Attenuation (also called loss) is a decrease in the power of a signal or signals, usually measured in decibels. Expressed mathematically, $\alpha_{dB} = 10log_{10}(P_{in}/P_{out})$, where α_{dB} is attenuation in decibels, P_{in} is input power in watts, P_{out} is output power in watts, and $P_{out} < P_{in}$. When signal power is stated in dBmV, $\alpha_{dB} = P_{in}(dBmV) - P_{out}(dBmV)$.

According to *Modern Cable Television Technology, 2nd Ed.*,

> Signal loss (attenuation) through coaxial cable can occur through any of four principal means:
>
> - Radiation out of the cable due to imperfect shielding

- Resistive losses in the cable conductors
- Signal absorption in the dielectric of the cable
- Signal reflection due to mismatches between the cable and terminations or along the cable due to nonuniform impedance

Assuming signal leakage (radiation) from the cable is negligible and there are no significant impedance mismatches, resistive losses in the metallic conductors are the dominant contributor to attenuation, followed by signal absorption in the dielectric.[2] Coaxial cable attenuation is greater at higher frequencies than at lower frequencies, as shown in Figure 47.

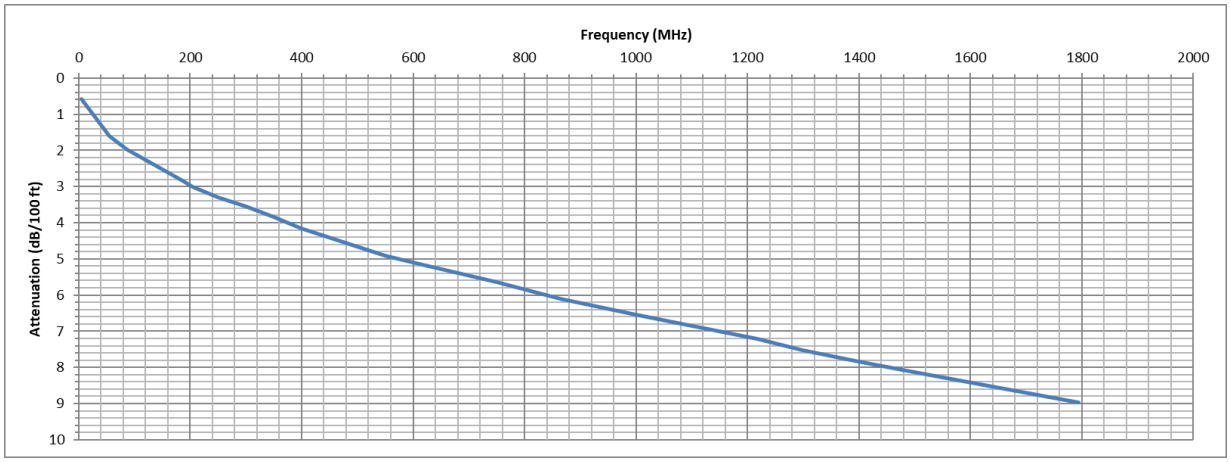


Figure 47 - Plot of typical attenuation in dB/100 feet for Series 6 subscriber drop cable, from 5 MHz to 1794 MHz

Coaxial cable attenuation in decibels changes about 1% per 10 °F temperature change (as the temperature increases, attenuation increases; as the temperature decreases, attenuation decreases).

For more information, see "Coaxial Cable Attenuation" in the Summer 2021 issue of *Broadband Library*.[3]

9.1.3. Dielectric constant

Dielectric constant is a parameter that applies to the dielectric in coaxial cable, and typically refers to relative permittivity.[4] Dielectric constant is related to coaxial cable's velocity factor using the following formula:

$$\epsilon = \frac{1}{VF^2}$$

[2] Note: An increase in the dielectric constant means a lower velocity factor, which increases signal absorption in the dielectric, and increases attenuation for a given size and impedance coaxial cable.
[3] https://broadbandlibrary.com/coaxial-cable-attenuation/
[4] This usage is considered obsolete by some standards bodies in favor of *relative static permittivity*.

where
ϵ is dielectric constant (relative permittivity)
VF is velocity factor

For example, the dielectric constant of hardline coaxial cable with VF = 0.87 is about 1.32. As mentioned previously, anything that changes the dielectric constant will change coaxial cable's impedance (and its attenuation).

9.1.4. Velocity factor

Velocity factor is the ratio – in decimal form – of the velocity of an electromagnetic signal propagating through coaxial cable to the speed of light in a vacuum. A common VF for hardline coaxial cable is 0.87, and for drop cable is 0.85. Mathematically,

$$VF = \frac{c}{c_0}$$

where
VF is velocity factor
c is the velocity of the electromagnetic signal traveling through coaxial cable
c_0 is the speed of light in a vacuum,[5] in the same units as c

Another formula for velocity factor is

$$VF = \frac{1}{\sqrt{\epsilon}}$$

where
VF is velocity factor
ϵ is the dielectric constant.

9.1.5. Velocity of propagation

Velocity of propagation (VoP) is velocity factor expressed as a percentage. For example, a VF of 0.87 equals a VoP of 87%. The latter means that RF signals propagating through the coaxial cable have a velocity that is 87% of the speed of light in a vacuum.

9.1.6. Return loss

When the impedance of a load or termination equals the characteristic impedance of the transmission line connected to that load, an incident wave is completely absorbed by the load. In the real world, there are no perfectly reflectionless loads, which means impedance mismatches exist. Impedance mismatches cause reflections. Reflected waves interact with incident waves to produce a distribution of fields in the transmission line known as standing waves. The presence of standing waves in coaxial cable can cause

[5] According to the National Institute of Standards and Technology, c_0 is 299,792,458 meters per second.

amplitude ripple in the frequency domain.[6] There are several ways to characterize the severity of impedance mismatches, among them is *return loss* (R).

Return loss (which is not the same thing as attenuation in the return or upstream spectrum of a cable network) is the ratio, in decibels, of the power incident ($P_{incident}$) upon an impedance discontinuity to the power reflected ($P_{reflected}$) from the impedance discontinuity. Note: When $P_{reflected} < P_{incident}$, return loss is a positive number.

Return loss is used to characterize network components such as active and passive devices, connectors, customer premises equipment, etc. Return loss has sometimes been used to characterize coaxial cable, although structural return loss is far more commonly used. The following is one formula that can be used to calculate return loss:

$$R = -20log_{10}\left(\left|\frac{Z_{device} - Z_0}{Z_{device} + Z_0}\right|\right)$$

where
R is return loss in decibels
log is base 10 logarithm
Z_{device} is the complex characteristic impedance of a device, in ohms
Z_0 is 75 ohms for cable networks

Hardline coaxial cable used in cable networks is typically specified to have a characteristic impedance of 75 ohms ±2 ohms, so the calculated return loss would be at least as good as about 37.4 dB.

9.1.7. Structural return loss

As mentioned in the previous section, return loss is one way to characterize the severity of impedance mismatches, especially in active and passive devices, connectors, and other components used in cable networks. *Structural return loss* (SRL) has been used for decades for coaxial cable, in large part because SRL deals with return loss at specific frequencies caused by evenly-spaced repetitive impedance discontinuities arising during the manufacturing process.[7] The following is from Technical Note 1069, "Testing CATV Cable to 1 GHz," published by Times Fiber Communications, Inc., in April 1999:

> As coaxial cable is manufactured, a number of variables can cause the impedance to change. Recall, the cable's impedance is a function of the cable's physical properties (conductor diameters, insulation's dielectric constant), and if any of these properties change, the impedance will change. For example, the dielectric material is extruded over the center conductor during the manufacturing process. As the dielectric is extruded, its diameter or dielectric constant can change and cause the impedance to change. This impedance change is extremely small and difficult to measure. If only one of these impedance changes occurs in the cable or if they occur at random intervals, the return loss will be good; but due to manufacturing processes, there may be many evenly spaced

[6] The term *standing wave* is often used to describe *amplitude ripple*, although technically speaking amplitude ripple is not the same thing as a standing wave.
[7] Pulley diameter and spacing, non-uniformity of line speed through extruders, vibrations, and other factors contribute to the creation of periodically-spaced, almost microscopic physical dimension variations in the center conductor, dielectric, and shield during the manufacture of coaxial cable.

> impedance changes and return loss problems will arise. Reflections from these evenly spaced impedance changes add together at a frequency corresponding to a half wavelength spacing. Although each impedance change may be very small, when they all add together, they cause a return loss "spike." These spikes can be narrower than 200 kHz. The return loss from these impedance changes is called the structural return loss because the impedance variations are due to structural nonuniformities in the cable.

In the past, a broadband sweep generator in conjunction with a variable bridge and variable termination were used for coaxial cable SRL measurements. As the operating bandwidth of cable networks increased beyond about 600 MHz, the aforementioned method could no longer provide accurate results. New measurement techniques were developed, based on the use of a fixed bridge, network analyzer, a set of calibration standards, and calculations to determine the SRL. A test procedure is described in the standard ANSI/SCTE 03 2016 Test Method for Coaxial Cable Structural Return Loss. The following is an excerpt from ANSI/SCTE 03.

> The purpose of this procedure is to provide instructions to measure cable structural return loss (SRL). The cable impedance as a function of frequency is calculated from a vector (magnitude and phase) return loss. The average of this impedance across the desired frequency range is the "cable reference impedance." The structural return loss is calculated from the cable impedance as a function of frequency and the cable reference impedance. This may be automated, but requires a vector network analyzer, and may be subject to errors due to the cable connection.

Figure *48* shows an example of a measurement of average impedance for a reel of coaxial cable. The top trace is the measurement from one end of the reel of cable, and the bottom trace is from the other end. Figure *49* shows a measurement of the same reel of cable's vector return loss from both ends of the reel.

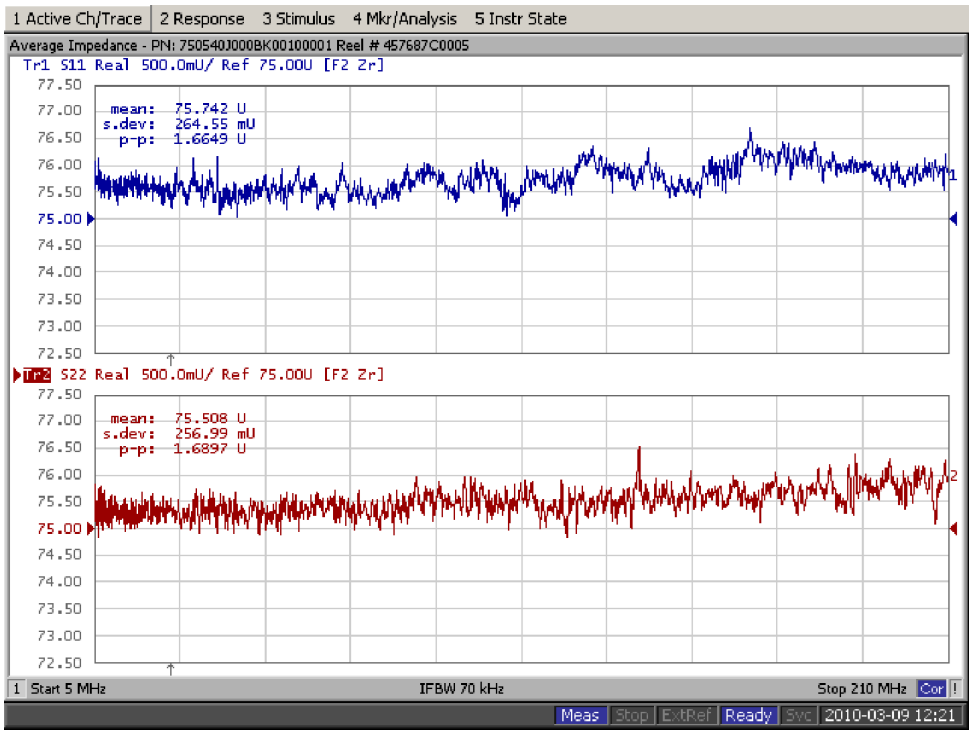


Figure 48 - Example measurement of the average impedance of a reel of coaxial cable, using the method described in ANSI/SCTE 03 2016. (Screen shot courtesy of Amphenol Broadband Solutions.)

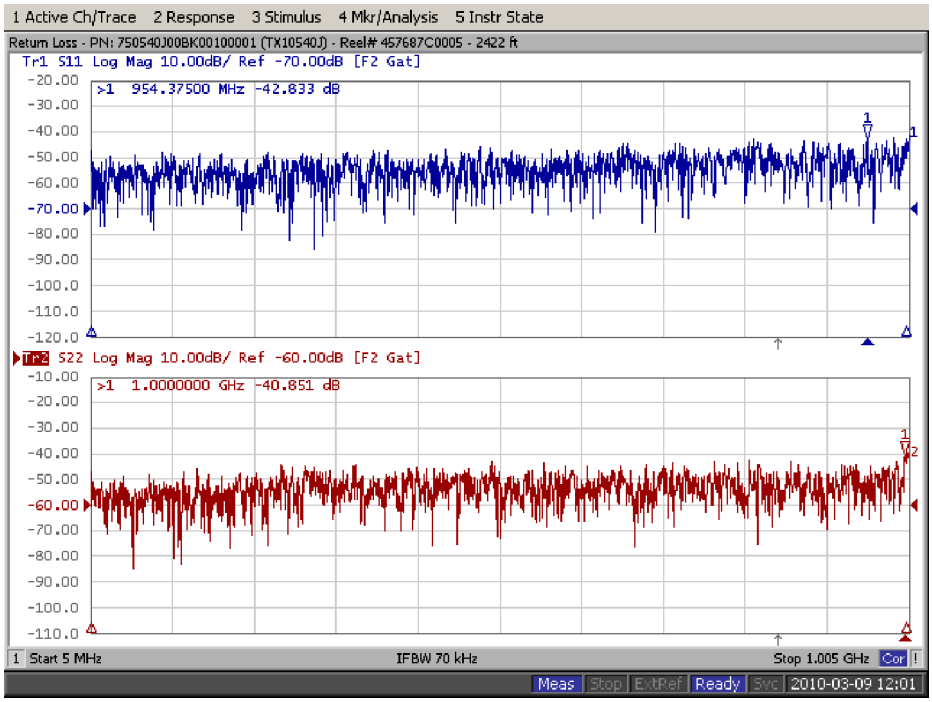


Figure 49 - Measurement of return loss of the same reel of cable in the previous figure. (Screen shot courtesy of Amphenol Broadband Solutions.)

Figure 50 shows the calculation of SRL from the parameters in Figure 48 and Figure 49.

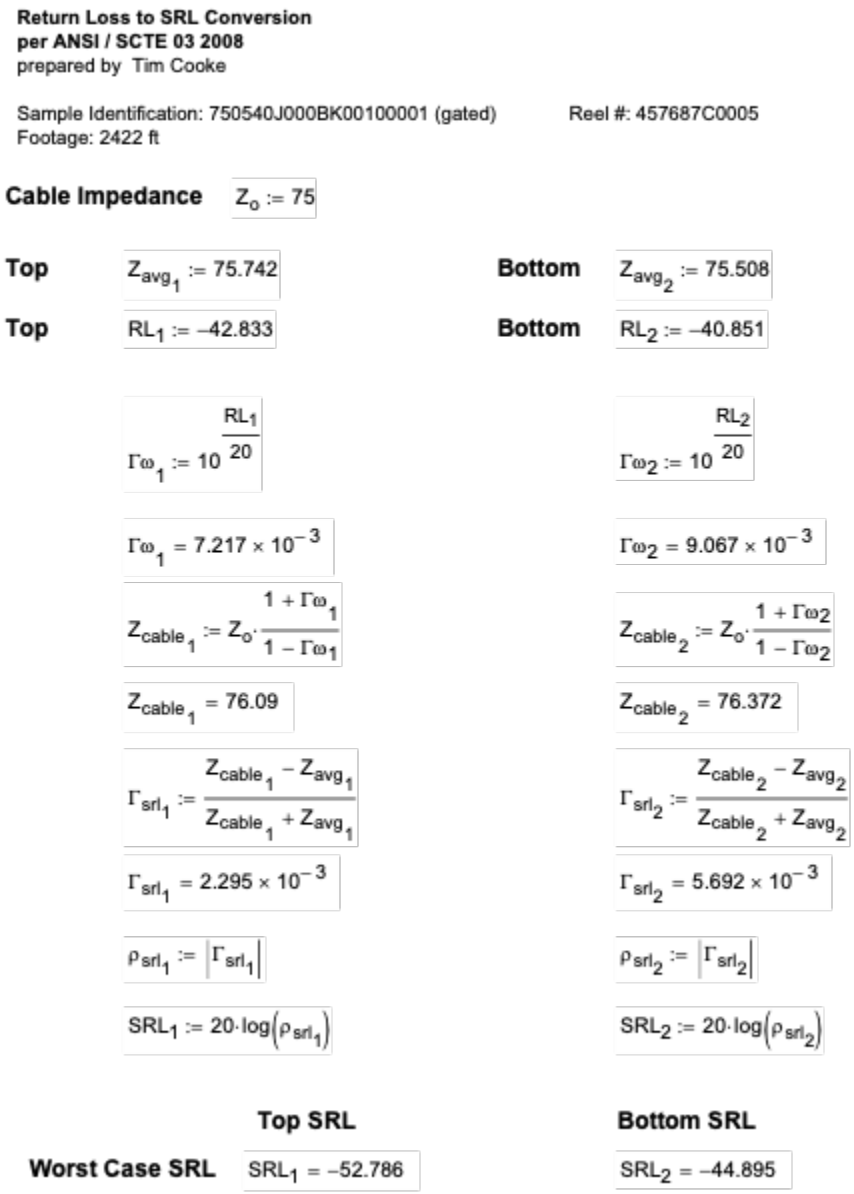


Figure 50 - Calculated worst case SRL for the reel of cable discussed in this section. (Courtesy of Amphenol Broadband Solutions.)

9.1.8. DC loop resistance

Loop resistance – more accurately, DC loop resistance – is a parameter usually specified in ohms per 1,000 feet, and is important for cable network powering purposes. Typical published DC resistance values for 1,000 ft. of 0.500 hardline cable are 1.35 Ω for the center conductor (measured end-to-end), 0.37 Ω for the shield (also measured end-to-end), and 1.72 Ω for the loop resistance. For loop resistance, imagine shorting one end of a 1,000 ft. length of cable, and measuring the DC resistance between the center conductor and shield from the other end.

What's important here is that the resistance values are at DC – the resistance one would measure with a conventional ohmmeter – and not at the frequencies of the RF traveling through the coaxial cable. Direct current travels through the entire cross section of a conductor. Alternating current, which includes RF, travels on and near the surface of a conductor, a phenomenon known as skin effect.

Abbreviations

AC	alternating current
ANSI	American National Standards Institute
CDF	cumulative distribution function
CM	cable modem
dB	decibel
dBmV	decibel millivolt
DC	direct current
DOCSIS	Data-Over-Cable Service Interface Specifications
DSP	digital signal processing
FBC	full band capture
FDX	full duplex [DOCSIS]
ft.	foot or feet
GHz	gigahertz
I	in-phase
IFFT	inverse fast Fourier transform
kHz	kilohertz
log	logarithm
LTE	long term evolution
MHz	megahertz
MoCA	Multimedia over Coax Alliance
OFDM	orthogonal frequency division multiplexing
PE	polyethylene
PNM	proactive network maintenance
PVC	polyvinyl chloride
Q	quadrature
R	return loss
RF	radio frequency
RxMER	receive modulation error ratio
SC-QAM	single carrier quadrature amplitude modulation
SCTE	Society of Cable Telecommunications Engineers
SID	spectral impairment detection
SRL	structural return loss

TDR	time domain reflectometer
VF	velocity factor
VoP	velocity of propagation

Bibliography & References

- Data-Over-Cable Service Interface Specifications DOCSIS 3.0 Operations Support System Interface Specification CM-SP-OSSIv3.0-I20-121113 (Cable Television Laboratories) [Note: Full band capture was first introduced in this version of the OSSI specification]
- Primer for PNM Best Practices in HFC Networks (DOCSIS® 3.1), CM-GL-PNM-3.1-V02-210114
- Hranac, R., et al, "Full Band Capture Revisited" SCTE Cable-Tec Expo 2020
- Hranac, R., "A Quick Look at S-Parameters" *Broadband Library*, Winter 2019 (https://broadbandlibrary.com/a-quick-look-at-s-parameters/)
- Hranac, R., "Coaxial Cable Attenuation," *Broadband Library*, Summer 2021 (https://broadbandlibrary.com/coaxial-cable-attenuation/)
- National Institute of Standards and Technology, speed of light (https://www.nist.gov/si-redefinition/meet-constants)
- Technical Note / 1069 "Testing CATV Cable To 1 GHz," Times Fiber Communications, Inc., April 1999

What It Takes to Automate Operations at Scale

Coupling Strategic Growth Analytics with Automated Methods for Real-Time Scalable Network Planning

Dr. Sung-eun Kim
Sr. Network Planning Engineer
Cox Communications, Inc.
Atlanta GA, 30328
sung-eun.kim@cox.com

Richard Brown
Sr. Manager Network Engineering
Cox Communications, Inc.
Atlanta GA, 30328
richard.brown@cox.com

1. Introduction

Understanding the traffic consumption trend is one of the essential elements in network capacity planning. We have seen an exponential traffic growth for long periods and a huge step function increase during the COVID-19 pandemic period. In the first part of the paper, we present the historical traffic growth over the past years, especially emphasizing the impact of COVID shelter in place in various ways.

The next part of the paper, we will examine the ongoing analysis, innovation, strategic direction, and how all those factors and components materialize as actionable plans using our patent-pending planning tools. We will also examine how our scalable suite of tools was a key driver in staying in front of COVID, as we were able to make fast decisions with near-real-time data in the ever-evolving early days of COVID-19 lockdowns in the United States

2. Deep Analysis

Internet peak traffic has increased exponentially over the years. Figure 1 shows the internet peak traffic increase over the past 10 years. Peak traffic is calculated based on the aggregated 95th percentile peak traffic volume measured at the northbound Gigabit ethernet ports in each CMTS device. It includes residential and business customer traffic and a new customer growth on top of the organic traffic growth.

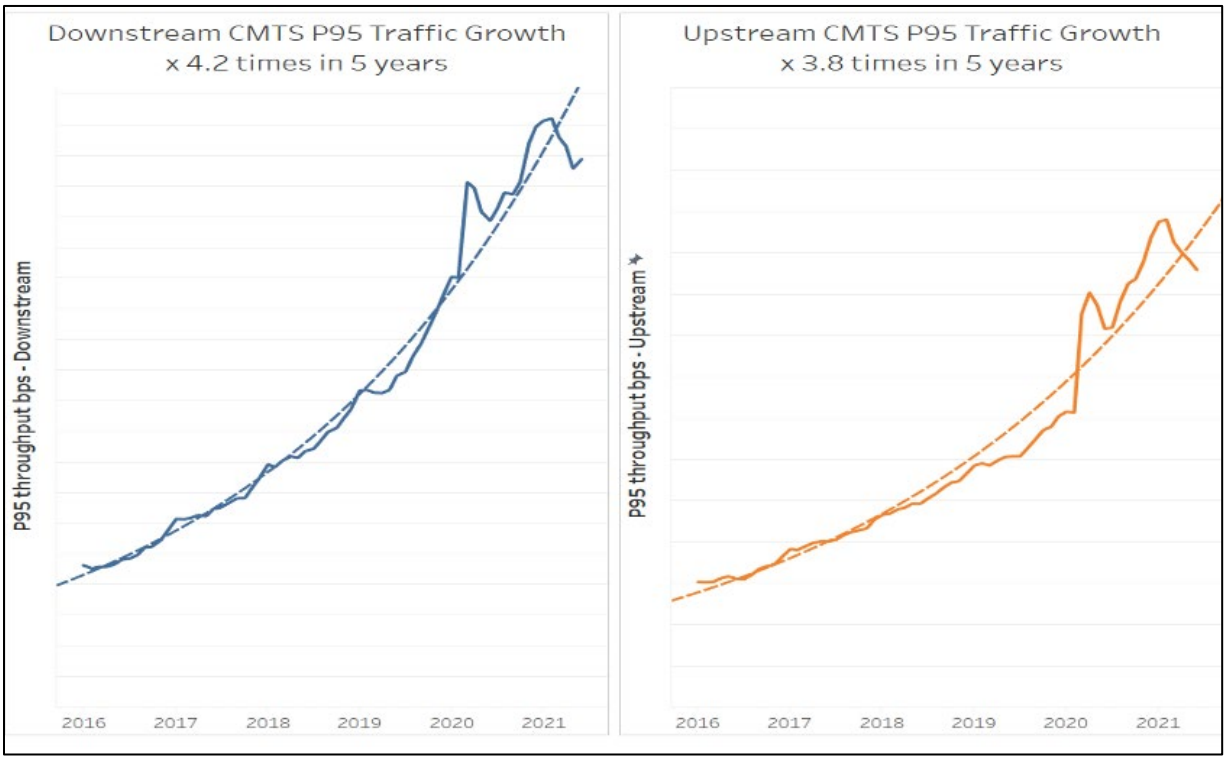


Figure 1 - Aggregated Peak Traffic Growth

Figure 2 shows the year over year (YoY) peak traffic growth rate. It had increased by 40-50% YoY growth until 2016 when major streaming service providers re-encoded the contents with optimized encoding algorithms and reduced the bandwidth usage. Since then, we have seen the traffic YoY growth rate to be about 30% until COVID-19 shelter in place started in March 2020.Traffic increased dramatically in a short period of time when shelter in place began. Upstream traffic growth is higher than downstream growth. Cox networks experienced 22% growth downstream and 34% growth upstream in 2 weeks after COVID-19 stay at home started.

As shown in Figure 2, upstream traffic grew about 70% and downstream traffic grew about 67% in April 2020 from April 2019. Downstream traffic growth continued to decline after it hit the highest point in April 2020. YoY growth rate since March 2021 significantly reduced because the latest value in the YoY calculation is compared to the surgical value after COVID-19 in 2020.

Figure 2 - YoY Peak Traffic Growth Rate

Figure 3 presents peak hour usage increase relative to pre-COVID-19 in February 2020. It compares the historical YoY growth trend and the actual growth rate. The recent downstream growth rate is smaller than the historical growth pattern. Upstream traffic continues to be growing and the recent trend shows 15% points above the historical YoY growth rate.

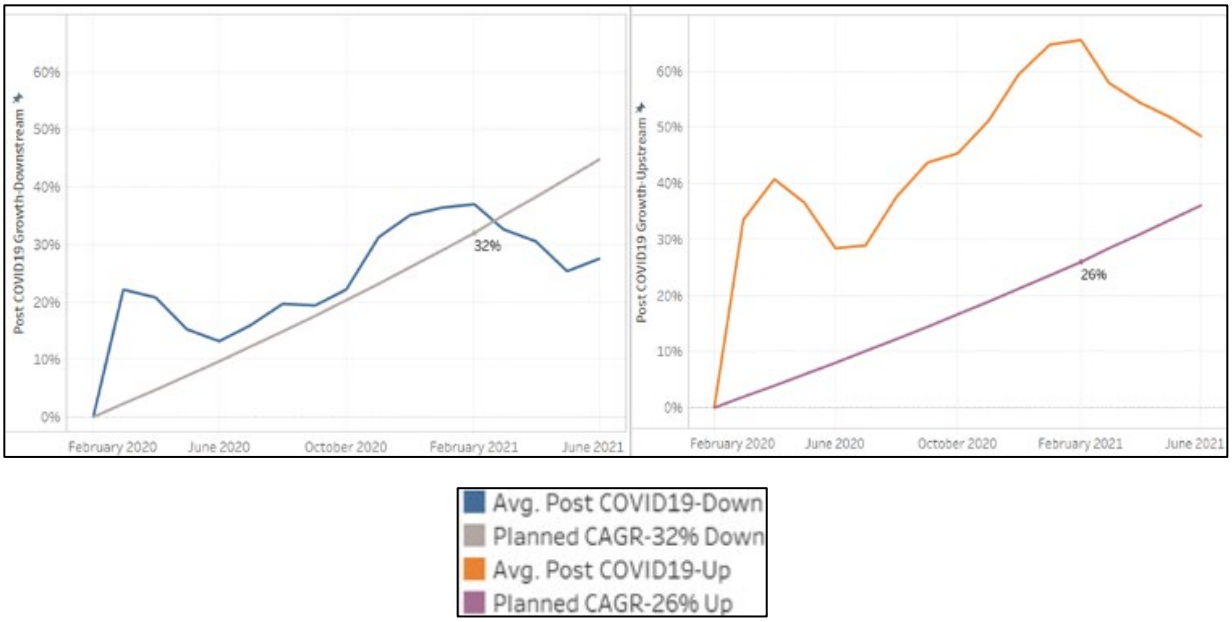


Figure 3 - Post COVID-19 Traffic Growth: Actual vs Expected

During COVID-19, the residential upstream traffic usage pattern has changed due to WFH (work from home) and SFH (study from home). In addition, upstream peak hours of residential customers have shifted to the daytime as shown in Figure 4. Recently it has returned to late evening, especially when the summer started, however, the daytime upstream usage pattern is still higher than pre-COVID-19.

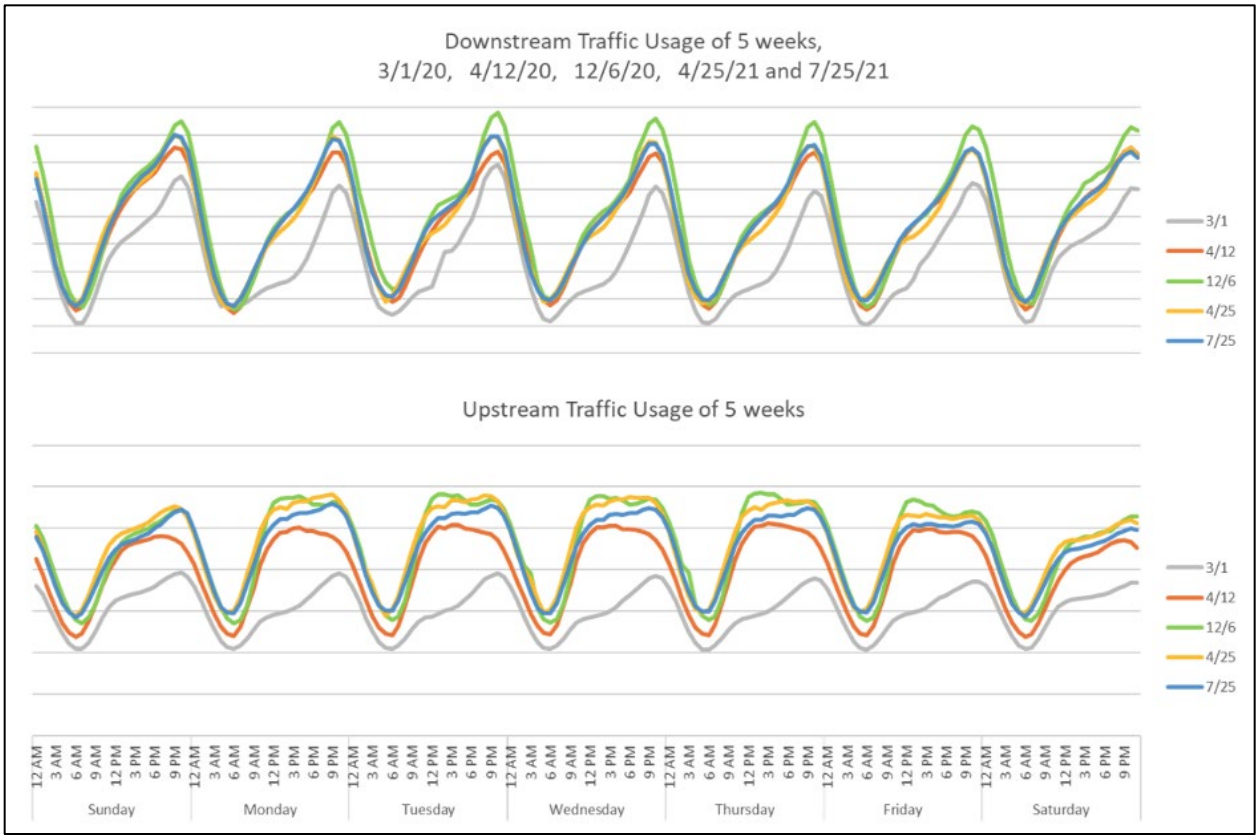


Figure 4 - Traffic Usage Pattern by Time of Day

The residential customers' peak traffic ratio of downstream to upstream is presented in Figure 5. It has grown to almost 19:1 since downstream traffic growth rate is higher; however, it reduced to 16:1 because upstream growth rate is higher after COVID-19 started. Currently the network has plenty of downstream capacity and relatively lower upstream capacity. Since Cox is building a mid-split plant, the capacity ratio of downstream to upstream will support the demand substantially.

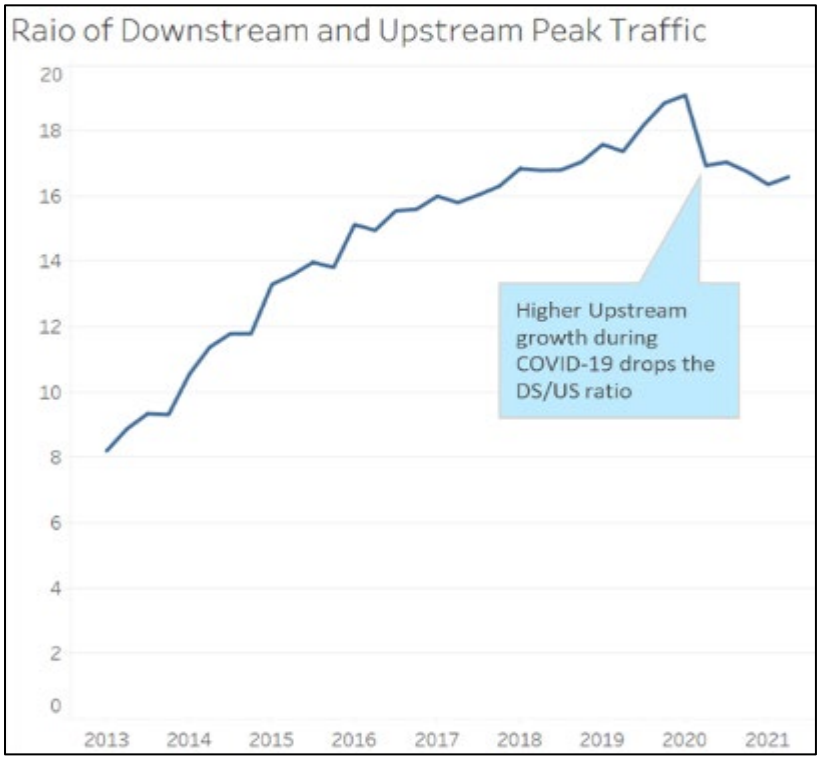


Figure 5 - Ratio of Downstream and Upstream Peak Traffic

Figure 6 displays the traffic category mix since March 2020. There has been a significant increase in WFH, and moderate increases in video, web, gaming and social since COVID started. The overall traffic remains dominated by video and web. Peak usage growth was still driven by traditional drivers such as OTT (over the top) video. In early 2021, WFH traffic increased up to 400% from the baseline of March 2020. Recently it went down to a 150% increase from the baseline.

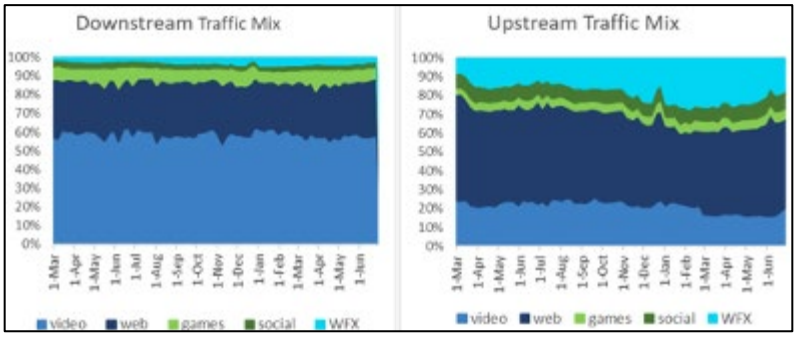


Figure 6 - Traffic Mix by Category

It would be interesting to see how business customers and residential customers consume network bandwidth differently. Figure 7 presents the average usage pattern of business and residential customers by the time of day. Business customers' usage peak hours are from 10AM to 4PM, while residential customers' peak hours are from 7PM to 10PM. Business customers consume less downstream traffic and more upstream than residential customers. Business customers' usage ratio of downstream to upstream is 5:1, while residential customers' usage ratio is about 16:1 as of time of authorship (August 2021).

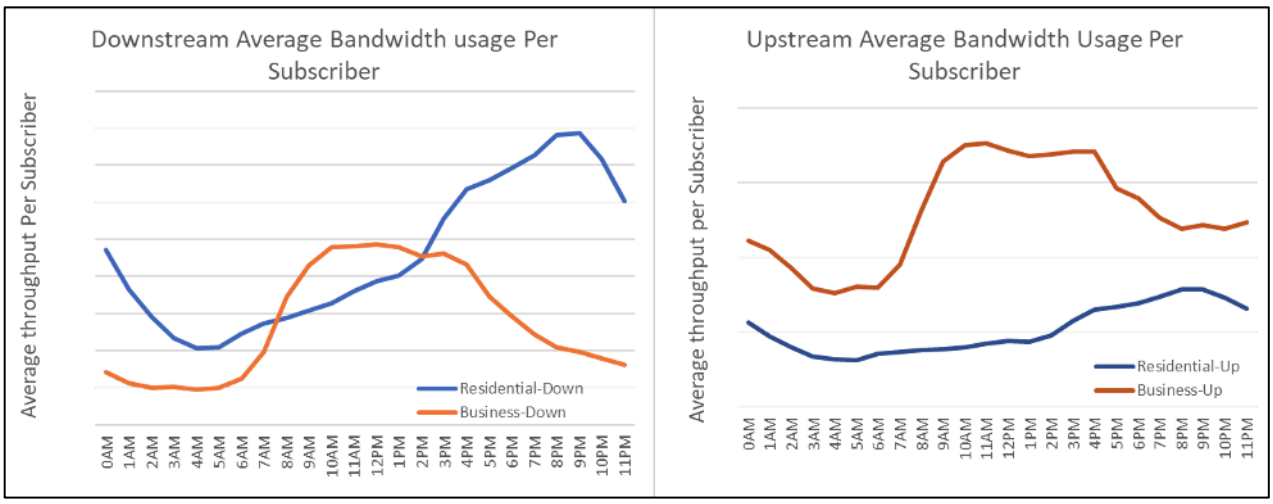


Figure 7 - Usage Pattern of Residential and Business Customers

Figure 8 shows the average usage throughput per customer by different tier speed, with the most popular tier, the first and second highest tier and the lowest speed tier. Basically, the average time of day usage patterns are almost the same across all of the different tiers. Figure 8 shows an example of the average throughput by speed Tier. The higher speed tier consumed more bandwidth, but not proportional to the maximum speed it can use.

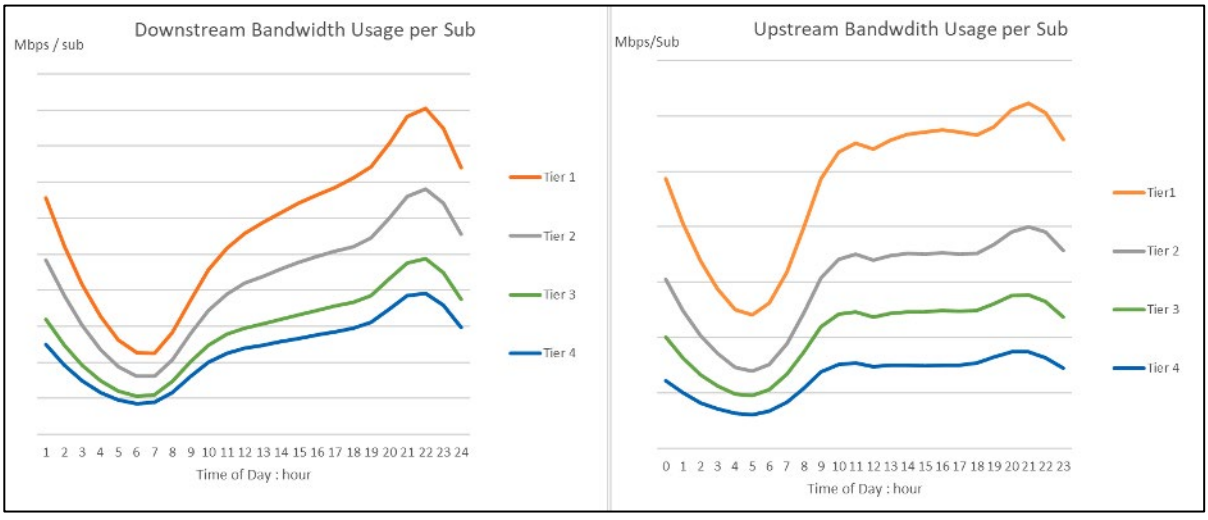


Figure 8 - Average Throughput of each Customer

3. Modeling and Planning

Translating network analysis into actionable information and data is the next facet of our teams. Implementing insights from our network analysis, in partnership with other industry experts such as CableLabs, has cultivated into a robust automated pipeline for the access network, and automated connectivity with other areas of network planning, such as backbone planning tools. These tools are developed on Cloudera, but built modularly for simpler migration as Cox's data strategy shifts

These modeling efforts were designed leveraging big data concepts and presented at SCTE in their infancy a few years back. As a recap, prior to the current methods, a series of manually generated Excel workbooks were used to infer a static CAGR-based growth rate for planning. These workbooks were first automated into workflows, connecting source data into a single schema, then logically joining and cleaning data. Within the workflow are tools designed to statistically identify anomalies, using kernel density and analysis of variance (ANOVA) to detect and alarm issues in the data. In addition to triggering external teams to research root cause and correct data issues, if possible, there are alternate "corrected" data tables generated. The corrected data uses interpolation methods to correct the gaps and errors until any corrected data loaded into the base (untouched) tables.

The final dataset, consisting of telemetry, topology, geospatial, upstream and downstream load, capacity, product and revenue data all on a time series basis, are loaded with node and market-level keying. That weekly utilization table is then used to infer node-level forecasts using an array of statistical and machine learning methods, specific to each node's specific growth pattern(s). From there, each node, with all its detailing attributes and forecasted load, are passed through a robust rules-based optimizer to determine the growth path, again, on a by-node basis.

This rules-based optimizer, dubbed the capacity response engine (or CRE) is one of the most important apprentices of the planning tools. The CRE can apply "if then" based business logic, simply applying business rules to grow node capacity, or it can use constraints such as budget, costs, manpower, ROI, customer counts, and other customizable variables to return strategic-based network growth views; and it can deploy combinations of business logic and constraint-based functions. With the data and information about the node, the CRE uses the current node topology and the various attributes of the node to determine the best series of node actions to both manage congestion, as well as selecting the shortest and/or most economic path to fulfil Cox's strategic network initiatives.

With the CRE completed, applying node actions to the 10 plus years of forecast, the results are used to create reports, as well as used to create the Volume Model (VM). The VM is our resource management model that feeds directly into our cost models. Here, we determine the license and physical equipment needed to grow the network. These outputs are also used to determine field headcount resources and critical facilities needed, with automations to link that data to their respective models, as well.

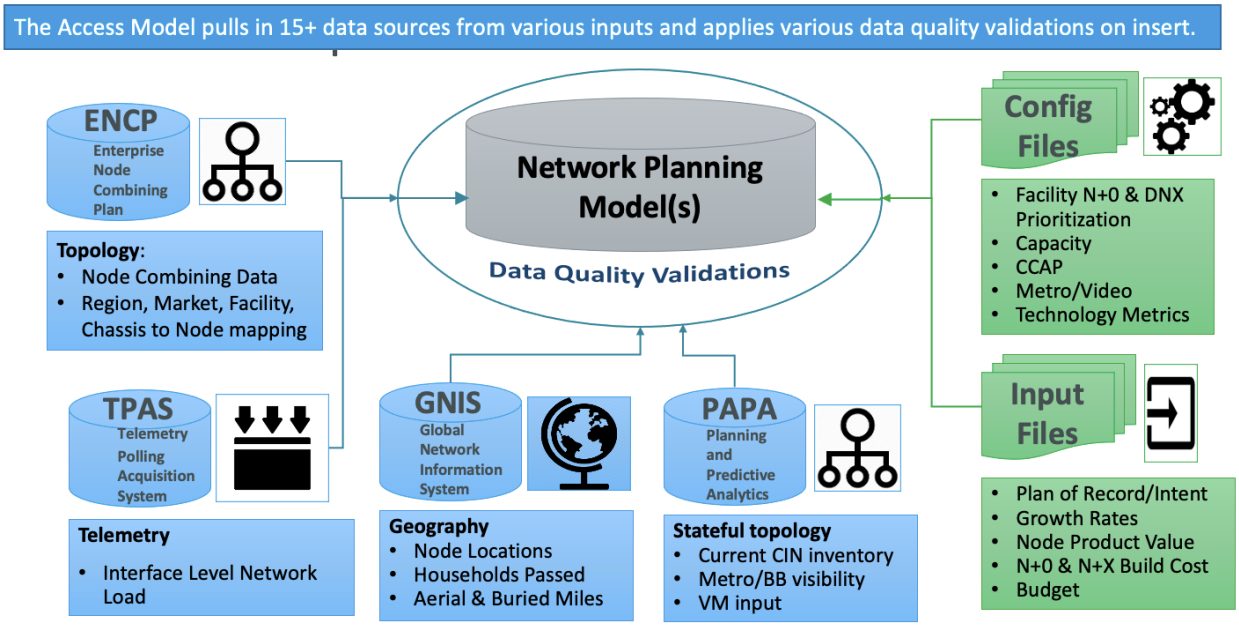


Figure 9 - General Overview of tooling

4. Business Continuity and COVID

Coupling our analysis and understanding of network and customer behavior has been vital for planning the network and keeping in front of the growth triggered by COVID. In the early days of COVID, in the beginning of lockdowns, we experienced a shift from the norm where customer-heavy nodes saw drastic decreases in usage while residential nodes saw a full year of growth, >30% year over year, in a few weeks' time frame (reference Figure 2).

With the robust pipeline of automated processes, we were able to marry the rapidly evolving data and analyses related to COVID and utilize insights from those analyses to quickly spin up forecasts to help us understand what "could" happen, as well as how we would need to grow the network to respond, thanks to the CRE. Data was used in creating hundreds of scenarios, and vital in determining courses of actions needed to address the elevated number of highly congested nodes* (see Figure 3). As well, leaders were able to assess costs near real time and make decisive decisions to keep up with this new network demand.

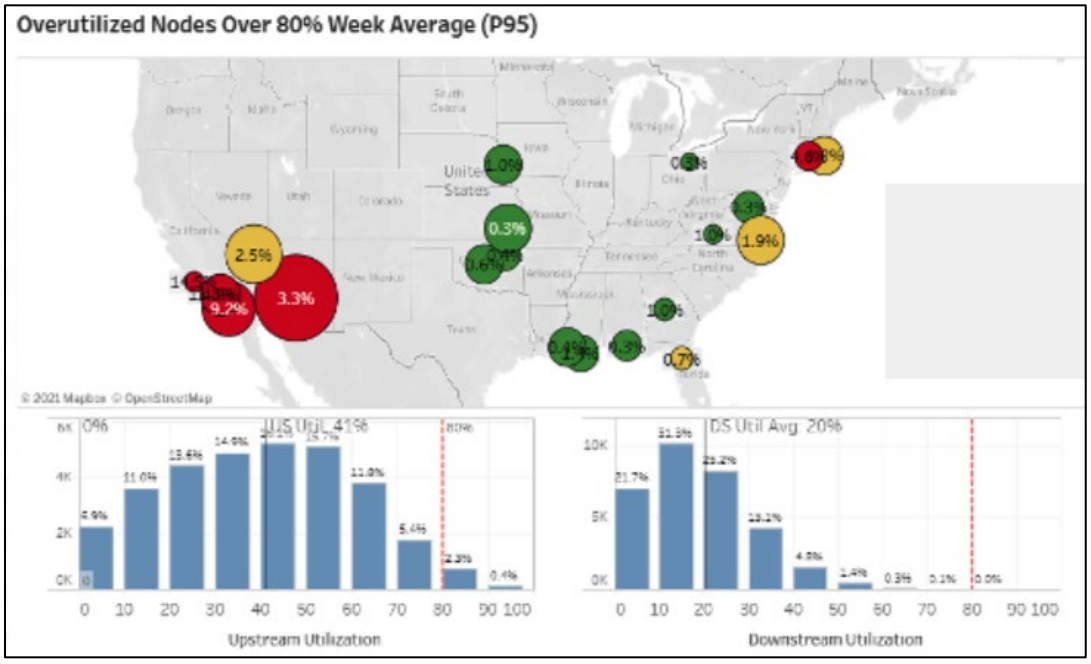


Figure 10 - COVID Related Network with High Congestion Nodes

Understanding the increased congestion was just a small portion of what was required. Next, we needed to be able to forecast and predict where we would see sustained growth, where we could expect less growth, and when we should see the effects of COVID decay from our growth curves. With our pipeline in place, we were able to increase the accuracy of our forecasts by implementing what we had learned so far, as well as reviewing and implementing assumptions regarding what we 'think' may happen. Of course, forecasting models weren't always 100% correct. Since we didn't have prior pandemic data to inform our models on the impact to network usage and node growth, we had to cycle input data in as fast as we could obtain it. From there, the models were updated on an ongoing basis. We were then able to add and control the COVID related variation by calculating a coefficient for the increased growth over what was expected (i.e. COVID growth over baseline expectation) then adding them into our Auto Regressive Integrated Moving Average models

(ARIMA, expressed: $Y_t - Y_{t-1} - \mu = \phi_1(Y_{t-1} - Y_{t-2} - \mu) + a_t - \theta_1 a_{t-1}$) and controlling with future assumptions and COVID decay with indicator variables (sometimes referred to as dummy variables or one-hot encoding in computer science). Figures 11 – 14 show some of the evolution of our assumptions, as well as the increased accuracy as we integrated more and more data after the mass lockdowns started.

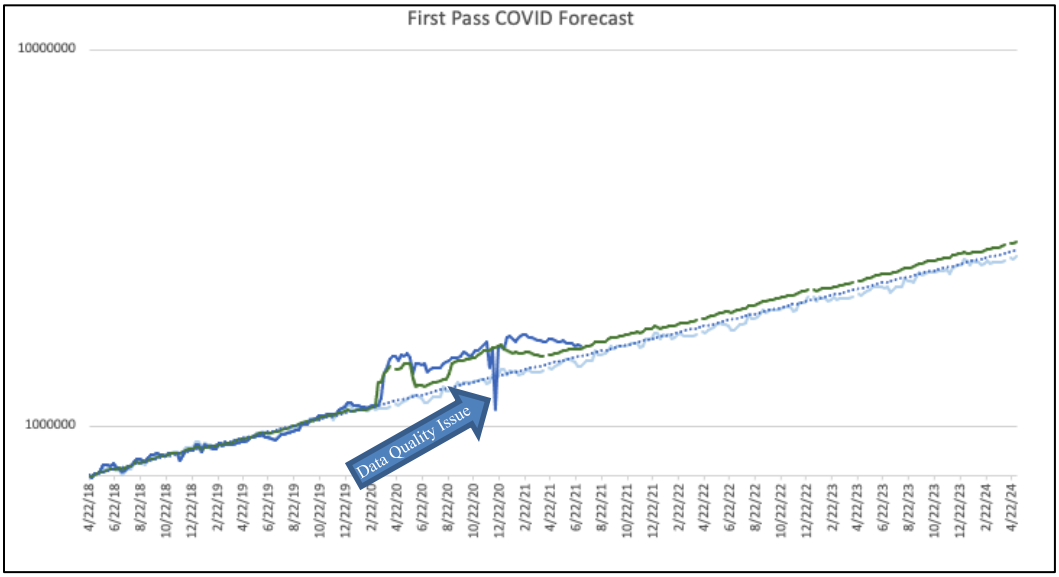


Figure 11 - Forecast post COVID W/O Forecasting COVID Effect

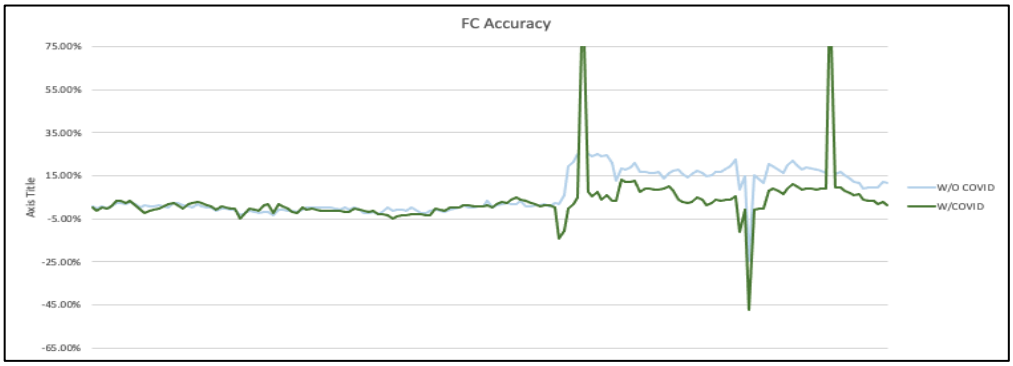


Figure 12 - First Pass Forecast post COVID

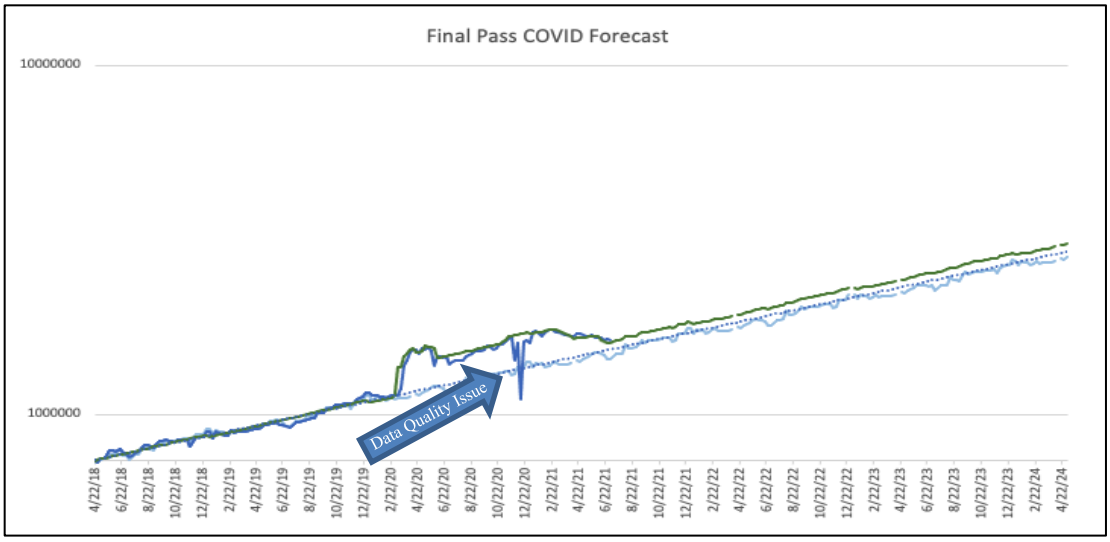


Figure 13 - Forecast post COVID with COVID Effect Forecasting

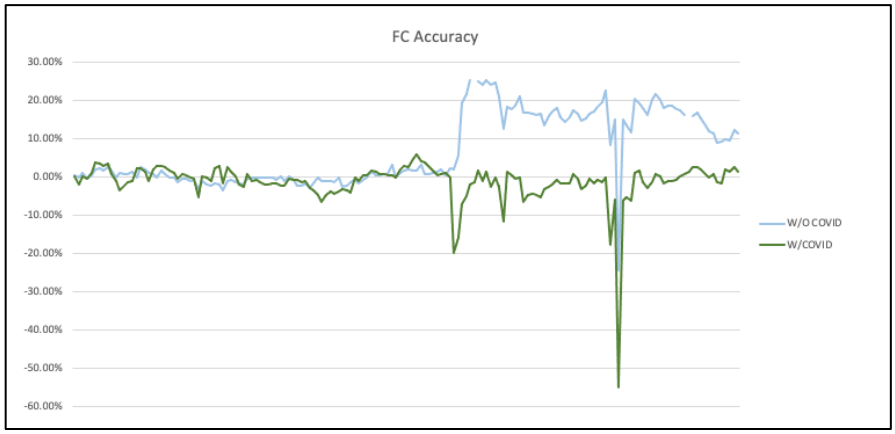


Figure 14 - Final Pass forecast with improved accuracy

With that, we were able to update and publish the models regularly. These published models are output into reports or used to generate costs and resource models, which are used in decision making about the direction of growing and planning the network. At first, we were updating the forecasts models weekly and communicating the findings to the broader teams. Despite the robust tasks of computing all the predictive models on a by-node basis, our scalable pipeline was always up to the task, and our planning teams were able to respond quickly. When we would see large shifts in the forecasts due to variations in the COVID related traffic, we would push the forecasts through the CRE to the outcome for our near-to-short-term plans.

This led to a need for complete network/model reworks of the full 2021 calendar year node growth plans, directly effecting outside plant and critical facilities executional plan of record. Utilizing our tools, we were able to re-work about year in just a few months, leaving more time for our field partners to adjust and react to the updates. To put this in perspective, our BAU (business as usual) policies require 18 months advanced planning “locks”, worked one quarter at a time over a period of a quarter. Achieving a

re-work of majority of the plan in the midst of the plan's deployment required solutions that not only work at scale, but can also be quickly adapted to unforeseen events, such as the COVID pandemic.

But just adjusting the plans also wasn't enough to address this congestion. Our engineering partners engineered solutions to address congestion; network solutions we had never modeled or deployed before. As these solutions were defined, we were able to model them. These solutions included spectrum actions, such as the implementation of 5th carrier and SOFA (sub split orthogonal frequency-division multiple access), to add fast and lower cost relief to highly utilized nodes. Being able to model these solutions as they were developed meant our leaders could measure the cost and benefits within days of the details being finalized so they could be integrated into planning for testing and deployment. Figure 15 below is an example of the output after a scenario run, implementing 5th carrier and SOFA to shoe the decreased need in other more invasive and expensive actions.

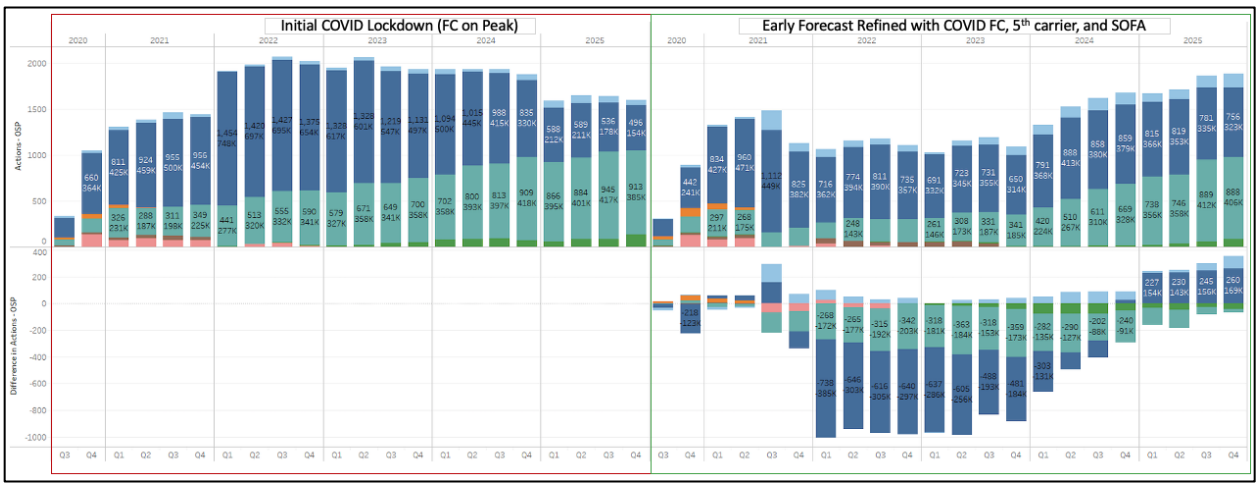


Figure 15 - Planning Scenario Implementing 5th carrier compared to predecessor

Conclusion

Cox's advanced network planning tools are where the rubber meets the road. A few years ago, the idea of these tools in their early developmental stages were presented at SCTE. Since then, we have been maturing the tools and increasing the capabilities and functionalities. Over the past few years, our planning teams have continuously developed and refined our tools to meet the needs of the business. In an extremely robust application, much of our analysis and planning components are automated, and automated with the flexibility to adapt quickly to changing business needs. The 2020 pandemic lookdowns tested our tools and infrastructure, but thanks to the scalable architecture, Cox was able to deliver refined plans to keep our customers connected.

Abbreviations and Definitions

ANOVA	analysis of variance

ARIMA	auto-regressive, integrated, moving-average
BAU	business as usual
CRE	capacity response engine
HDP	Hortonworks Data Platform
Highly Congest Node	Nodes with sustained congested p95 utilization >80%
OOT	over the top
SFH	school from home
SOFA	sub split orthogonal frequency-division multiple access
WFH	work from home
YOY	year over year

Bibliography & References

(1) *Texas A&M Department of Statistics, Dr. Simon Sheather, 2018*

What's Smart About Smart Power?

Modernizing the Power Grid and HFC Networks: Power Outage Notifications and Advanced Sensing

A Technical Paper prepared for SCTE by

Dr. Robert F. Cruickshank, III
Gridmetrics™ R&D Liaison
Cable Television Laboratories, Inc.
Louisville CO, 80027
+1-703-568-8370
r.cruickshank-c@gridmetrics.io

1. Introduction

This paper calls the cable broadband industry to action to investigate and capitalize on the increasingly important intersections between grid/utility power and cable network power. Managing the intersections of power and communications across the landscape and throughout our respective infrastructures—from communities to forests—is fundamental to improving network reliability and cost-effective operations. Additionally, working with organizations responsible for emergency response, corporate security, situational awareness, public safety, and business resilience is fundamental to preserving our way of life and preventing catastrophic loss of life and property.

The paper begins with motivational examples related to grid outages caused by severe weather and cyber-attacks, their exponential rise, and troubling forecasts for increasing numbers of issues. A deep dive into the failures of supply and demand during the February 2021 Texas Power Crisis includes a spatiotemporal summary of issues in Houston and sheds light on the global nature of the many grid operational challenges ahead.

The background on powering the grid and how it connects to the Hybrid Fiber-Coaxial (HFC) network includes a discussion of the connection points of HFC power supplies and the possibility for fatal back-feeding from misbehaving power electronics in the growing base of solar photovoltaic inverters, battery walls, and electric vehicles. The sea change of implications of the grid incorporating renewable energy sources and transitioning from 1-way central-station delivery of power to two-way flows among distributed energy resources are discussed.

The Gridmetrics™ Power Event Notification System (PENS™) is introduced as the most capable, fastest, and lowest latency system for monitoring the massive sensor-starved grid edge. Also discussed is Gridmetrics' role in U.S. Department of Energy (DOE), Office of Cybersecurity, Energy Security, and Emergency Response (CESER), Cybersecurity for Energy Delivery Systems (CEDS) R&D program. The new American National Standards Institute (ANSI), Society of Cable Telecommunications Engineers (SCTE) Standard 271 2021 is discussed along with a few of the plethora of anticipated operational applications in the field.

Readers focused on how to better access utility power data, improve relationships with utility ecosystems, and create new business opportunities will benefit greatly from this paper and the accompanying presentation and panel session.

2. Motivation: Operational Costs and Network Reliability

Issues resulting from outages and poor power quality affect everyone and directly impact cost and service reliability. Outages and poor power quality issues are on the rise and are expected to increase.

2.1. Outages

At best, outages are expensive and disrupt cable broadband operations and the customer experience. At worst, outages lead to property damage, human suffering, and lives lost. Outages from 2002 to 2020 are shown in Figure 1 [1].

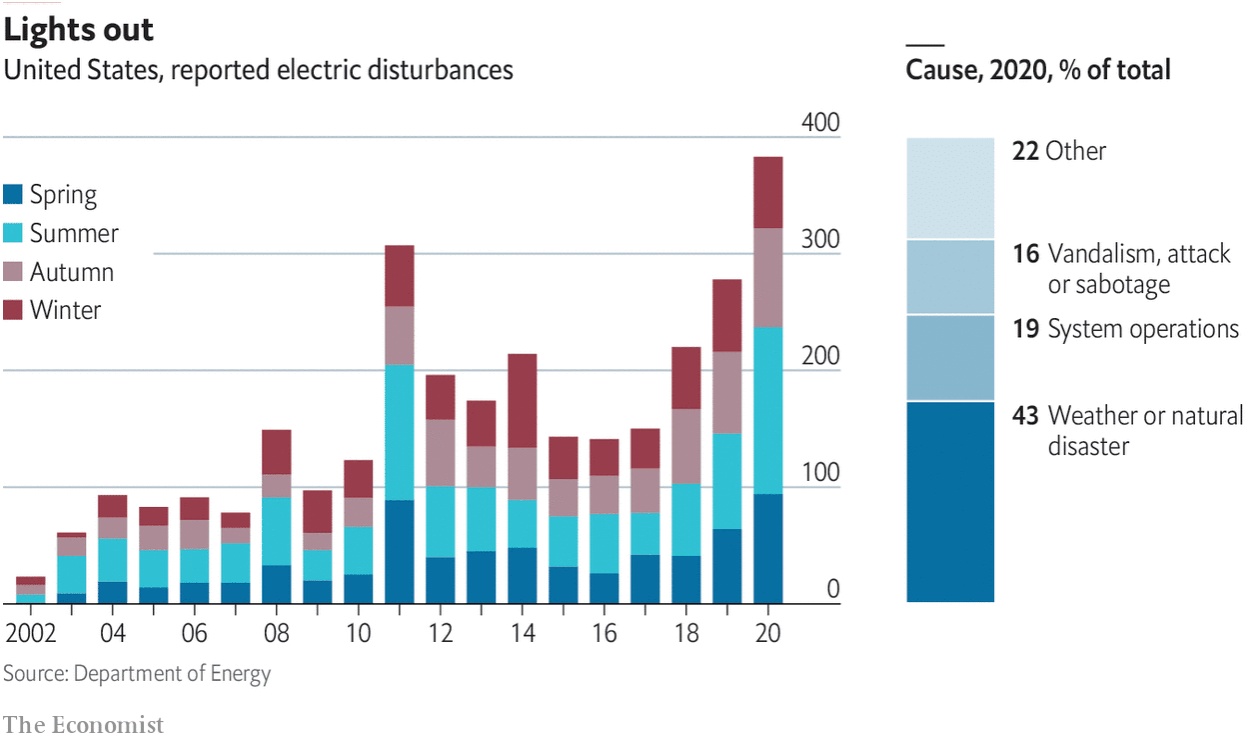


Figure 1 – U.S. Power Outages 2002 – 2020

Outages most often occur at the grid edge [2], and utilities are way behind the cable broadband industry in terms of being mostly unaware of an outage until customers call to report a loss of service. While both industries have made advancements in automatic outage detection and declaration, cable's battery-backed broadband networks are far superior to utility wireless mesh networks in terms of latency, loss, throughput, and jitter—and hence are the best alternative for monitoring grid voltage, phase, current, and on/off status.

2.2. Power Quality

Poor power quality is a silent and stealthy foe that is largely unmonitored. The American National Standards Institute establishes nominal voltage ratings and operating tolerances for electric power systems in ANSI C84.1-2016 provided by the National Electrical Manufacturers

Association [3]. The upper and lower acceptable voltage limits are 105% and 95% respectively, and there are additional considerations for the frequency, intensity, and duration of voltage excursions [4].

Within normal operating voltages, the customer experience is acceptable. Above or below voltage limits, issues arise with billing, resilience, safety, and equipment longevity. For example, high voltages lead to higher energy usage and higher electric bills and may damage capacitors on electric motors widely used in refrigeration, heating, air conditioning, and pumping water--and low voltages lead to overheating and a shortened lifespan of motors.
Unfortunately, utility "smart meters" are not up to the challenge of reporting costly power quality issues such as voltage spikes and sags. By design, all smart meters in the U.S. are limited to 240 VAC split-phase leg-to-leg measurements, providing no measurement of the performance of ground or neutral circuits. In every cable shop throughout the world, there is likely a prominently displayed section of melted in-home coaxial cable that serves as a warning to technicians of the fatal perils of unbonded neutrals and intermittent grounds.
While more than half of U.S. households have smart meters deployed, the capability is often unused in daily operations due to bandwidth limitations in the backhaul mesh communications infrastructure that results in bottlenecks in utilities receiving and processing smart meter data.

3. February 2021 Texas Power Crisis

The February 13–17, 2021 North American Winter Storm Uri, was a major winter and ice storm that had widespread impacts across the United States, Northern Mexico, and parts of Canada. The storm resulted in over 170 million Americans under various winter weather alerts and caused blackouts for over 9.9 million people in the U.S. and Mexico, most notably in the 2021 Texas power crisis [5]. The blackouts were the largest in the U.S. since the Northeast blackout of 2003 and resulted in economic costs, human suffering, lives lost. Losses were greater than Hurricane's Harvey and Ike.

3.1. Severe Storms

The Texas Power Crisis came about as a result of three severe winter storms sweeping across the United States from February 10–20, resulting in massive electricity generation failures, and resultant shortages of water, food, and heat [6]. Nearly than 4.5 million homes and businesses were left without power for four days of freezing darkness [7]. At least 210 people were killed directly or indirectly, with some estimates as high as 702 killed as a result of the crisis [8].

3.2. Failure of Generation and Fuel Supply

As shown in Figure 2, the arctic temperatures resulted in the failure of 48% of electricity supply as generators and fuel supplies for generators froze and were unable to operate to meet the unprecedented rising demand of heating loads. The February 15–18 gap between (colored) supply and demand depicts massive declines in hundreds of natural gas and coal-powered generators along with lesser declines in wind and increases in solar generation.

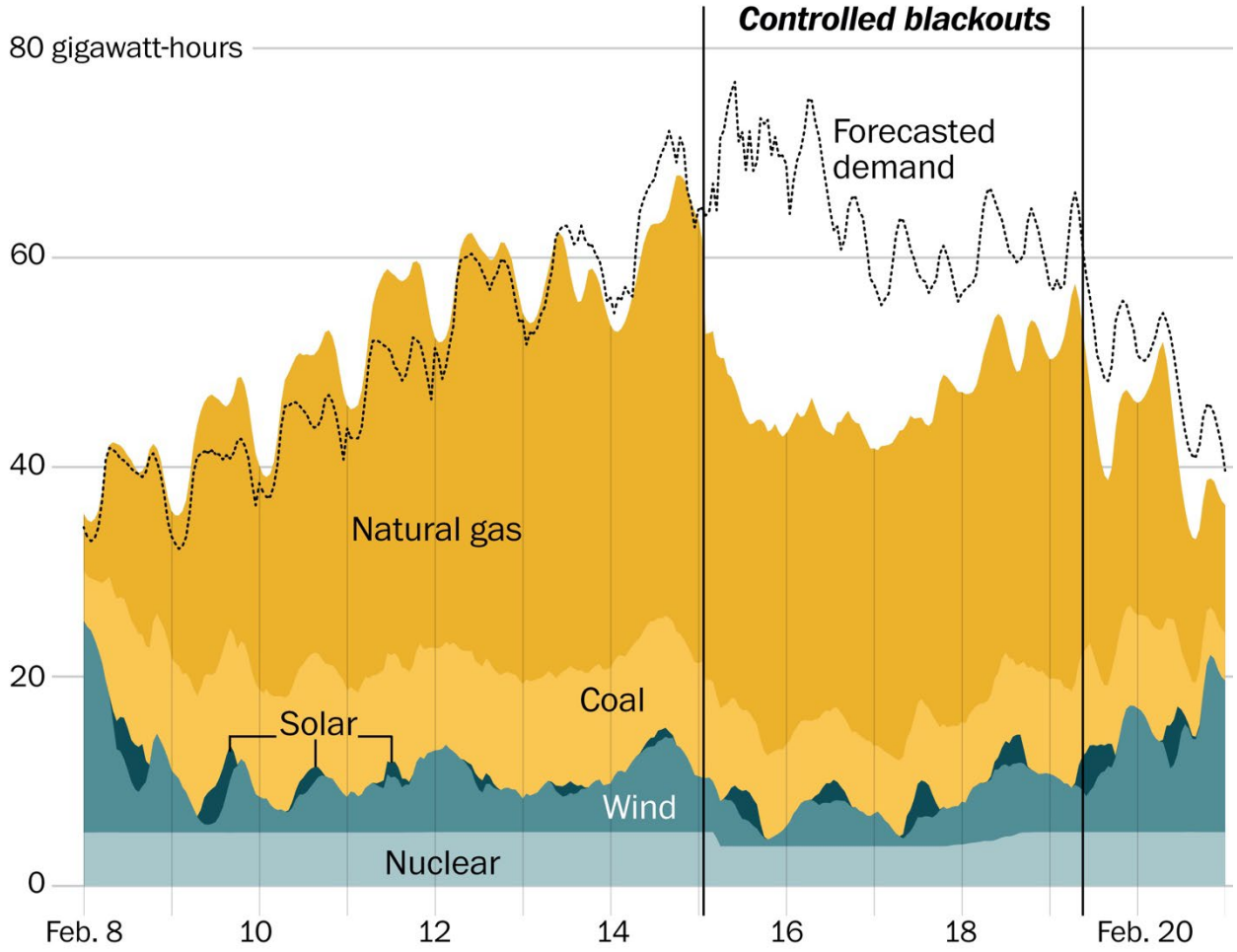


Figure 2 – Texas Electricity Supply and Demand February 8-20, 2021

The observations presented in Figure 3 are based on measurements from over 10,000 HFC power supply sensors distributed across the low-voltage distribution grid in the Houston, TX area. The sensors provide a voltage reading (± ~1.2 volts) at a particular latitude and longitude at 5-minute intervals. For the purposes of this analysis, power sensors have been aggregated and mapped to the U.S. National Grid, (USNG) a standard 1 km x 1 km square. The color scale at right is an index that denotes outage duration, where red denotes sensors out all three days.

3.3. Spatial and Temporal Analysis

In the data, both macro and subtle patterns can be seen as the widespread power outage and subsequent recovery unfold. The exact cause of an outage at any particular sensor cannot be inferred directly from the data—i.e., whether the outage was due to downed lines or due to operator-controlled power shutoffs (aka load shed)—but analyzing temporal and spatial behavior of voltages and power conditions reveals a new, independent lens to view and understand the severity and duration of outage events across the city and within particular neighborhoods. In addition, analyzing voltage trends prior to outage events, as well as residual voltage readings on de-energized lines during outage events, highlights an urgent need for better real-time situational awareness of the distribution grid to anticipate localized grid stresses and to manage utility

worker and public safety. A special thanks goes to Dr. Scott Clearwater and the team at CableLabs for the analysis.

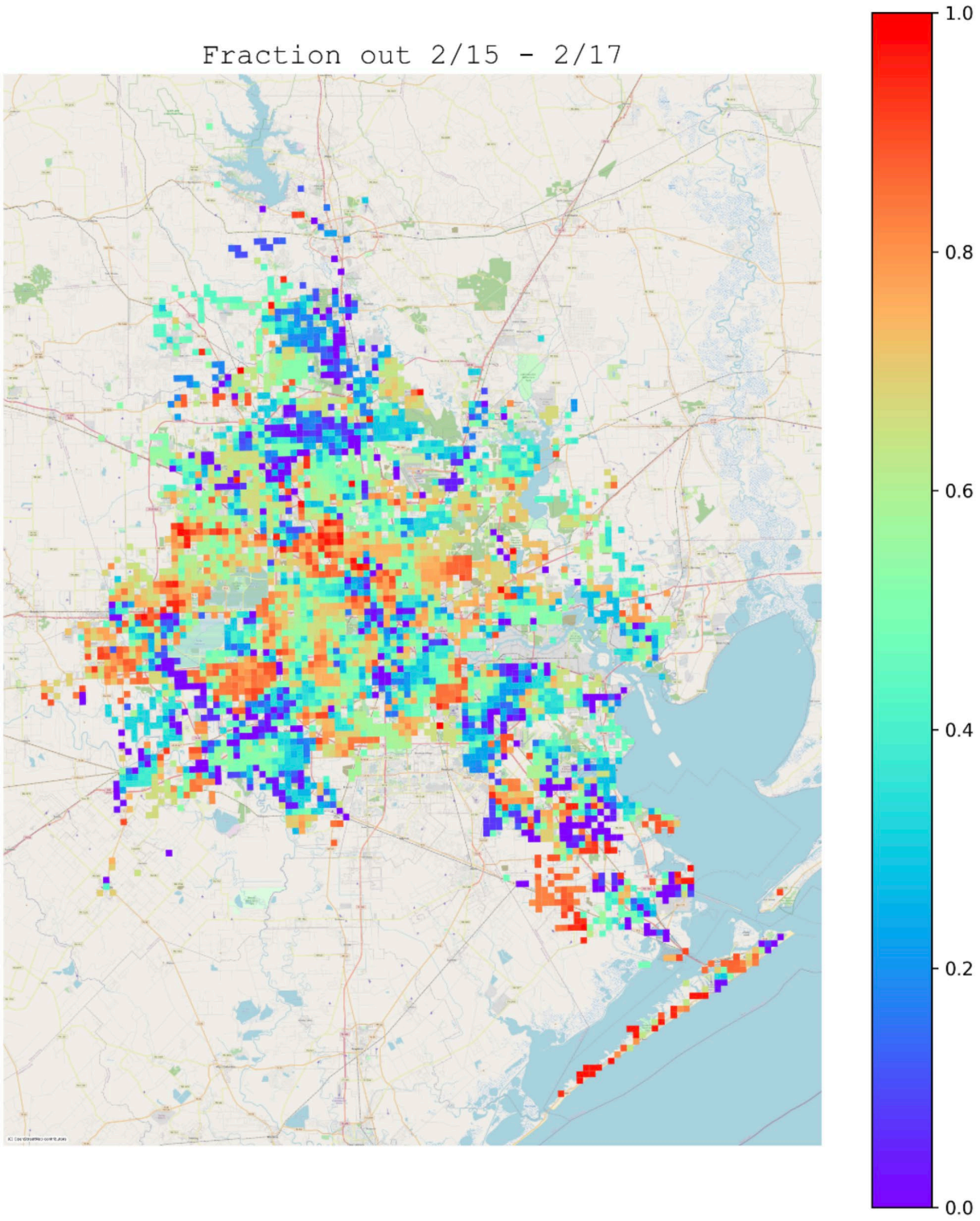


Figure 3 – Spatiotempotral view of Texas Power Crisis, February 15-20, 2021

3.4. Missing Focus on the Demand Side

The February Texas Power Crisis along with two more Texas blackout close calls in April and June resulted in a focus on improving the nexus of natural gas and power generation, better public communication, more power plant weatherization/protection, and new power transmission projects. Demand-side work has been ignored and missing to date is a focus on shoring up the grid by aggressively managing load [9].

4. Background on the Grid and HFC Powering

Situational awareness of the electric power grid is gaining in importance with the increasing number of power generators, power consuming devices, and power infrastructure failures. In the United States, 200,000 miles of well-instrumented high-voltage transmission lines make up the grid core, or backbone. However, lower voltage, local distribution lines connect the remaining 96.5% of the grid, accounting for 5.5 million miles of poorly instrumented grid infrastructure as shown in Figure 4. [2].

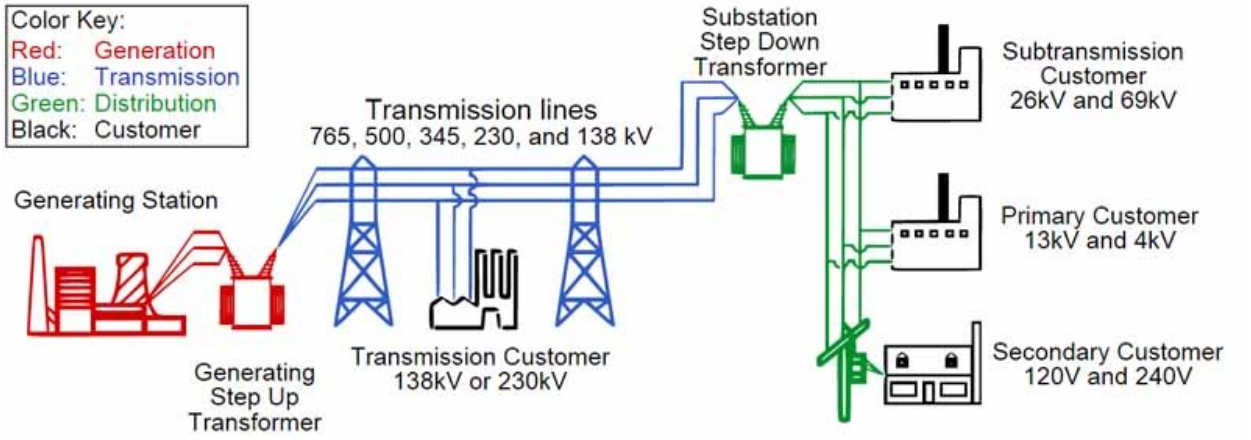


Figure 4 – Power Grid Schematic

The connection points for HFC power supplies, electric vehicle (EV) chargers, and photovoltaic systems are located in the distribution grid shown in green in Figure 4.

5. Solutions for Monitoring the Grid and HFC

There are inevitable requirements to get new, real-time instrumentation deployed into the last miles of the distribution grid, and the broadband industry already has a great head start.

5.1. Gridmetrics™ and PENS™

Gridmetrics is the premium supplier of power event notifications. Gridmetrics was born at CableLabs and inspired by 2017-2021 conversations with the National Renewable Energy Laboratory. Gridmetrics makes grid insights available via the Power Event Notification System (PENS™). PENS Aggregates unique data from ~300,000 sensors in HFC power supplies and provides an unmatched observational view of the state of power in the last mile of the distribution grid. PENS alerts are available via email, Esri, and an API for use by emergency

response, public safety, FEMA, DHS, business resilience, and other users. The nationwide footprint of Gridmetrics sensors is shown in Figure 5, where red the notes outages.

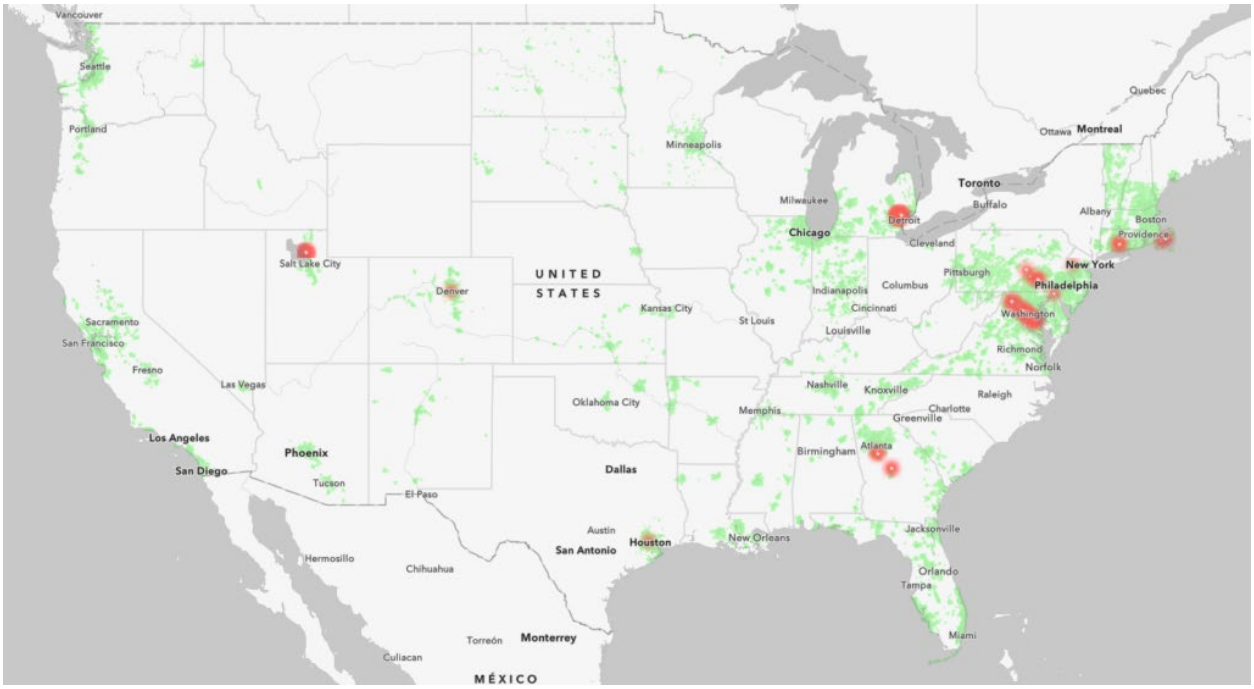


Figure 5 – Gridmetrics™ U.S. Map

Figure 6 shows an example of a PENS Alert for an area to the West of Tallahassee, Florida. Note how the event, the population, and Gridmetrics sensors are alongside the edge of a wildlife area.

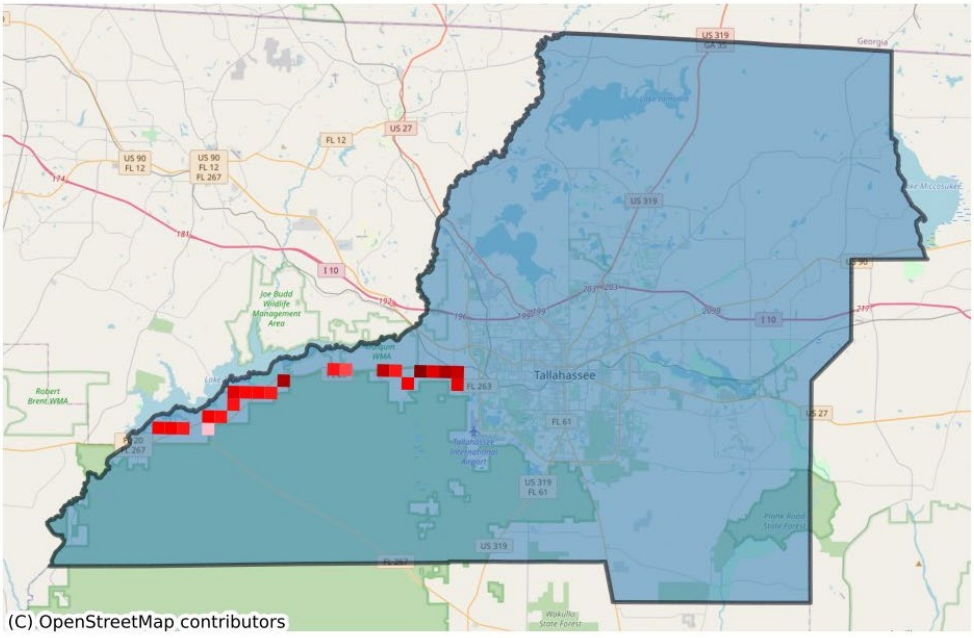


Figure 6 – Example PENS™ Alert

Figure 7 provides an additional example of Gridmetrics sensors tightly aligned with population. Darker green denotes denser populations and darker red denotes outages affecting more people.

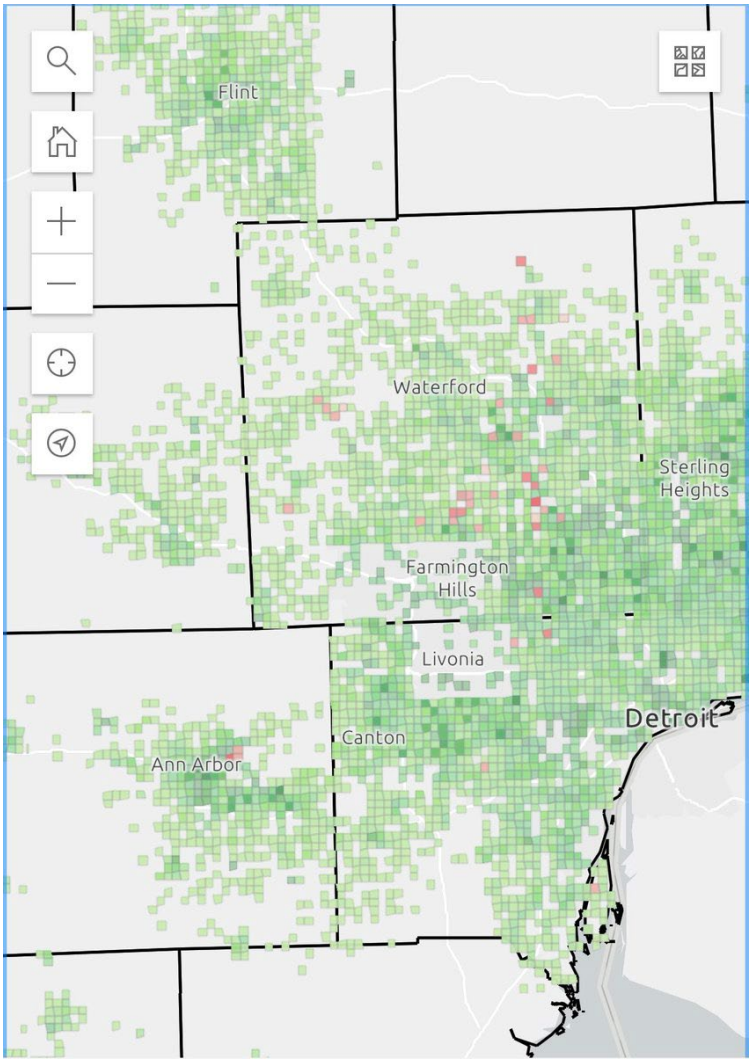


Figure 7 – July 21, 20201 PENS™ Detroit Outage

Today, there is no comprehensive, independent source for power event insights. Most solutions offer insights only at the county level. As mentioned previously, PENS provides insights using the USNG 1km x 1km grid overlay projection. In addition, most solutions offered by utilities and other entities provide updates only every 15 minutes. PENS scans the broadband sensor network every 5 minutes looking for events.

PENS Email Alerts are unique in that they provide: 1) Initial Alert, 2) Update Alerts over time, and 3) Closing Alerts to indicate service is restored. The combination of the three types of Alerts allows for automatic opening and closing of network incident tickets in different operations centers within a utility, broadband provider, emergency responders, etc. Interested users may sign up for alerts and more information at www.gridmetrics.io.

Gridmetrics is part of the Esri partner network and provides a Feature Service Layer view in Esri Marketplace [10]. For existing Esri users, the set-up is simple, and a no-fee trial version is available.

5.2. SAGA - Situation Awareness of Grid Anomalies

Another powerful solution for monitoring the power grid and HFC networks as SAGA, Situational Awareness of Grid Anomalies. SAGA is a $3M, 3-year project to develop near real-time cyber-physical resiliency through machine learning—using existing infrastructure. The U.S. Government funding for SAGA came in 2019, after several joint proposals were developed by the National Renewable Energy Laboratory (NREL) and CableLabs, submitted, and evaluated by the U.S. Department of Energy and the Advanced Research Project Agency.

SAGA identifies anomalous behavior using cable broadband's secure out-of-band in-service network and rapidly detects cyberattacks dynamically and in near real-time through machine-learning. The SAGA learning-aided low-voltage estimation framework is shown in Figure 8.

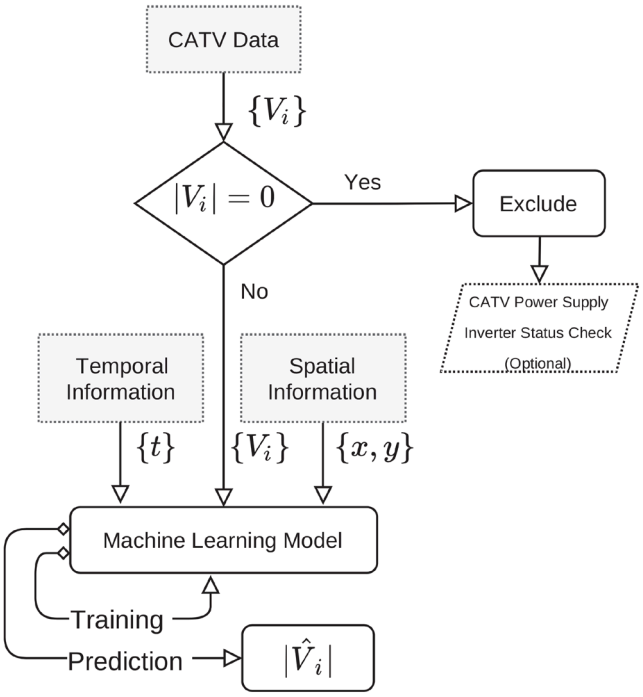


Figure 8 – SAGA learning-aided low-voltage estimation framework

The goals of SAGA include demonstrating a disruptive technology for power system data analytics using existing infrastructure while providing commercialization and product feature roadmaps. SAGA builds upon NREL's extensive collection of power system state estimation and mapping tools and integrates the growing set of Gridmetrics 'situational' data. SAGA assimilates other time-series geospatial data such as weather and cyber-physical phenomena, distribution infrastructure maps, and tax lots. To ensure the ongoing evolution of SAGA and global impact

on operational efficiencies and network reliability, lessons learned inspired the development of a new next-generation ANSI/SCTE grid and HFC sensing standard.

6. New U.S. National Standard: ANSI SCTE 271 2021

Motivated by the successes of PENS and SAGA and recognizing that existing sensing capabilities are out of date, the new ANSI SCTE 271 2021 standard specifies requirements for Power Sensing in Cable and Utility Networks. SCTE 271 specifies additions, without replacing prior sensor specifications developed in the late 1990s during the development of DOCSIS 1.x.

SCTE 271 specifies how to monitor HFC and Grid for voltage and current anomalies and send *raw* waveforms to the cloud for further processing. The importance of the ability to observe, communicate, and then cloud-compare multiple high-resolution traces of voltage and current in real-time cannot be overstated. Secure backhaul of streaming continuous point-on-wave (CPOW) power observations is a quantum leap beyond phasor measurement units that assume a sine wave and then compress and distort all data before backhauling, as shown in Figure 9 (note the phaser approximation completely misses and misrepresents the voltage spike anomaly).

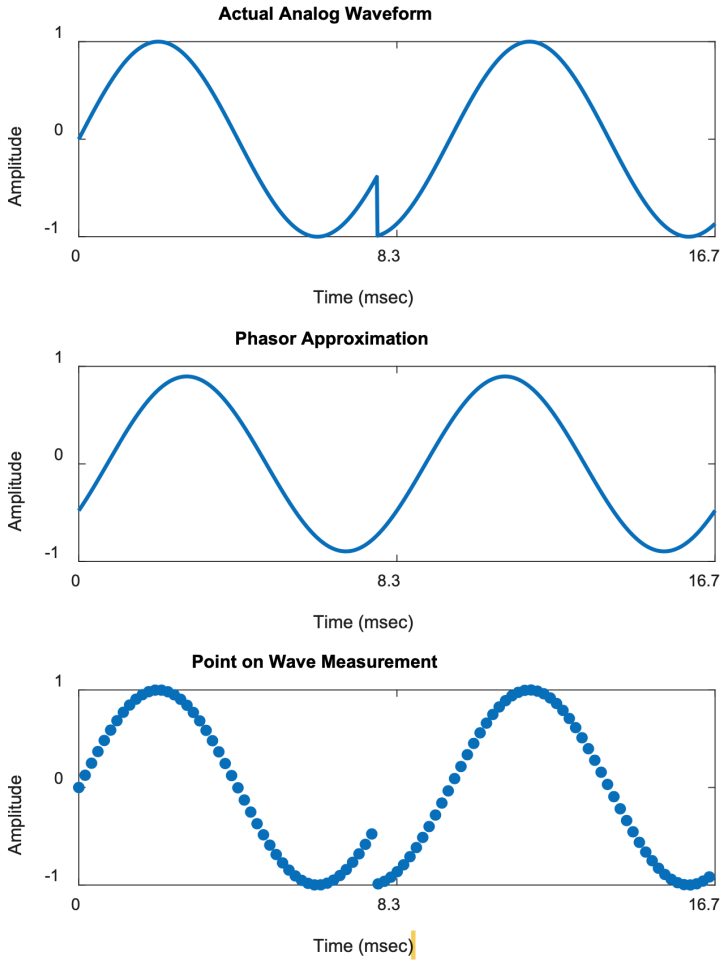


Figure 9 – SAGA learning-aided low-voltage estimation framework

SCTE 271 measures the 60/75/90 VAC quasi-square wave HFC network and the 120/240 VAC power grid. Goals include: 1) reducing network element reboots, outages, and issues that cause wildfires, and 2) improving the customer experience and the lifespan of the HFC, Customer Premise Equipment, and the grid. SCTE 271 can be used to identify voltage and current highs, lows, fluctuations, as well as outage and voltage sages that indicate grid congestion—that can be used to actively manage load and increase use of renewables.

The following requirements are included in SCTE 271:

1. If voltage or current is sensed, it ***shall*** be measured with a precision of 0.002 per-unit (0.2% of nominal value), e.g., +-0.24 volts at 120 VAC
2. If CPOW capture is provided, the sampling rate ***shall*** be a minimum of 10k samples/second = 166 samples/period at 60 Hz (more better)
3. If observation timestamp provided, the resolution ***shall*** be <= 1 microsecond. Clock accuracy ***shall*** be <= ½ microsecond, which is ~ 1/100th of a degree at 60 Hz
   (1 sec / 60 cycles) * (1 cycle / 360 degrees) * (1 degree / 100) = .46 microsecs
4. If configurable remote reporting is provided, control plane ***shall*** enable a) a 1-time poll reply, b) continuous replies and/or c) fixed interval replies
5. If a communication plane is provided, it ***shall*** use IETF/APSIS YANG model and SSL or TLS for authentication & encryption. No SNMP is required.

The new capabilities specified SCTE 271 are expected to unleash a plethora of opportunities for proving value in advanced grid sensing that helps modernize and manage the grid by predicting catastrophes and advancing grid state estimation with visibility to two-way electricity flows.

6.1. The Aging and Failing Grid: Predicting the Next Catastrophe

Figure 10 shows additional examples that reinforce the need for and benefits of advanced grid sensing. Both images were taken from *in-service* conductors and were made possible by developing and applying grid sensing capabilities in Australia in the wake of catastrophic wildfires. The image at left is a wire-rope conductor that is unravelling mid-span. Imagine looking up from a chairlift or gondola, that you are riding on, and seeing the wire-rope unravel! The image at right is a flat “Licorice” drop cable that is often used in direct burial applications.

Figure 10 – Examples of failing in-service conductors

6.2. The Changing Grid: Advancing State Estimation by Monitoring Two-Way flows at the Grid Edge

The grid edge is constantly changing. From with a historical perspective, in 1882, Thomas Edison installed generation in New York City and London using coal fired power plants, which use the thermoelectric Rankine cycle—and haven't changed all that much. All electricity delivered from power plants travelled outward through the grid to customers; this is referred to as a one-way delivery of central station power. And that's how the grid operated for the next hundred years; there was organic load growth in the sense that, new appliances, housing subdivisions, power substations were developed, but behavior was mostly predictable, and the grid had one normal state. It was either on or it was off.

Then came the Public Utility Regulatory Policies Act in 1978, which for the first time allowed non-utility generators to market their power to utilities. Suddenly, anybody could build and operate a generator, not just utilities. Fast forward to today where every state and nation has renewable portfolio standards, trying to achieve, say, 30% renewables within five years, 50% renewables within 10 years, etc. But that means generators are everywhere. And now we have not just one-way flow, we've got two-way flows at the grid edge! If the solar panels on your house make more energy than you're using, then grid electricity is not flowing into your house—it's flowing out of your house. And that's great, but the problem is, the sun and the wind are intermittent and variable. They're uncertain and forecast error is on the rise in the face of climate change and severe weather.

So, we're not certain what the production and demand at the edge is going to be; we do our best to forecast it, but we don't know for sure and that makes these two-way flows, even more unpredictable. The net effect is, it's much harder to "see" and estimate what's going on in the grid today—than in the old grid with one normal state—where you could easily tell if it was on or off based on customer call volume. Today, there's an unlimited number of dynamic normal states and that thwarts detection of non-normal conditions—caused by something really being wrong. For example, is a voltage sag or spike in a neighborhood just because a cloud went over the sun or because the sun came out again—or is it the result of a coordinated outside-in cyber-attack against thermostats, HVAC controllers, "smart" inverters at the edge with the intent of taking down the core?

Today, we struggle to identify failures and cyberattacks. It's difficult to detect cyberattacks and here's why: If you found the keys to a car and you were mischievous, you might wait until midnight, and then, lurk around and figure out which car those keys fit—very quietly and unobtrusively. Once you had access and you were in, then you'd wait until the owners left and then you'd steal the car. And that's the fear with cyberattacks; attackers constantly try to get in and might make only small disturbances—until they know that they're in. But the disturbance they create can be so small that we won't be able to detect their presence until they come back and do something terrible and bring down the grid. And we all know in the pandemic and severe weather, that the reliability of the network is extremely important, and we simply cannot have energy and communications networks go down—ever!

7. Conclusion

As the number of distributed energy resources and two-way electricity flows rapidly increase, the aging grid infrastructure elements that are supposed to keep the grid safe are failing and causing unprecedented loss of life and property. While the enormity of the electric power grid is such that in the U.S. alone, the 5.5-million-mile distribution network is long enough to reach the moon nearly 21 times—the performance of the last mile of the grid is sparsely monitored and hence unable to be optimally managed. Gridmetrics sensor readings fill the immediate need to augment utility supervisory control and data acquisition systems by rapidly improving the monitoring of the secondary distribution portion of the grid.

The growing and evolving Gridmetrics data set is available to aid in monitoring and managing the secondary distribution networks that make up the last mile of the grid. The locations of specific anomalies worthy of investigation are available for use in the utility ecosystem including emergency response, public safety, the Federal Emergency Management Agency, the Department of Homeland Security, business resilience, and other entities.

Through maintenance and repair efforts, infrastructure aging, wear, and tear—and local weather—the location and severity of anomalies will change over time, supporting the case for the real-time Gridmetrics API and real-time data feeds. Through collaboration with the utility ecosystem and sharing best practices for anomaly detection and classification, it is expected that anomalies that foretell of impending infrastructure failures and safety issues, and high-risk for loss-of-life can be identified. In addition, the criteria used to identify anomalies can be expanded, refined, and validated to achieve maximum benefit from Gridmetrics data.

Data from next generation broadband power quality sensors can help pinpoint existing portions of the grid that can be inspected for high, low, and fluctuating voltages—and high-impedance faults—which can cause unsafe conditions, poor customer experiences, and premature failures of customer equipment. In addition, outage data from broadband sensors can be correlated with existing data sets to create a more comprehensive understanding of distribution network performance and frailties. Combining insights from utility supervisory control and data acquisition systems and Gridmetrics data will help improve network reliability, resilience, and safety.

Abbreviations

ANSI	American National Standards Institute
API	application programming interface
CEDS	Cybersecurity for Energy Delivery Systems
CESER	Office of Cybersecurity, Energy Security, and Emergency Response
CPOW	continuous point-on-wave
DOE	U.S. Department of Energy
EV	Electric vehicle
HFC	Hybrid fiber-coaxial
NREL	National Renewable Energy Laboratory
PENS	Power Event Notification System
USNG	United States National Grid

Bibliography & References

1. https://www.economist.com/graphic-detail/2021/03/01/power-outages-like-the-one-in-texas-are-becoming-more-common-in-america
2. https://www.scientificamerican.com/article/what-is-the-smart-grid/
3. https://webstore.ansi.org/standards/nema/ansic842016
4. https://www.spgsamerica.com/information/acceptable-voltage-ranges
5. https://time.com/5939633/texas-power-outage-blackouts/#:~:text=5%20Million%20Americans%20Have%20Lost,on%20Feb.%2015%2C%202021.
6. https://www.cnn.com/2021/02/18/weather/texas-winter-storm-thursday/index.html
7. https://www.nytimes.com/live/2021/02/17/us/winter-storm-weather-live
8. https://web.archive.org/web/20210718121413/https://www.buzzfeednews.com/article/peteraldhous/texas-winter-storm-power-outage-death-toll
9. https://www.utilitydive.com/news/fix-texas-electricity-and-hurry/603159/
10. https://www.esri.com/en-us/arcgis-marketplace/overview
11. https://www.wsj.com/articles/pg-e-knew-for-years-its-lines-could-spark-wildfires-and-didnt-fix-them-11562768885
12. https://www.wsj.com/articles/this-old-metal-hook-could-determine-whether-pg-e-committed-a-crime-11583623059

When Physical Layer Simulation Gets Real

Next-Gen Network Modeling

A Technical Paper prepared for SCTE by

Ramya Narayanaswamy
Sr. Manager
Comcast, NGAN Enterprise Data and Analysis
Virtual Location, PA, 19148
(215) 286-2634
Ramya_Narayanaswamy@cable.comcast.com

Karthik Subramanya
Engineer 4
Comcast, NGAN Technical Research
Virtual Location, PA, 19148
(267) 260-2289
Karthik_Subramanya@comcast.com

Dr. Richard Prodan
Comcast Fellow
Comcast, Next Gen Access Networks
1401 Wynkoop, Suite 300, Denver CO 80202
(720) 512-3742
Rich_Prodan@comcast.com

Larry Wolcott
Comcast Fellow
Comcast, Next Gen Access Networks
1401 Wynkoop, Suite 300, Denver CO 80202
(720) 512-3643
Larry_wolcott@cable.comcast.com

1. Introduction

There are many ways to model advanced broadband networks, and a growing number of ways to simulate their behavior, based on available information about performance characteristics. This paper, co-authored by Comcast's Ramya Narayanaswamy, Karthik Subramanya, Richard Prodan and Larry Wolcott, will explore the intersection of theoretical modeling, practical proactive network maintenance (PNM) and modern data science – a potent combination that is well suited for the sophisticated modeling and simulation needs of cable 10G networks.

Using traditional RF reflection measurements like scattering parameters / S-parameter matrices, fed by real-world field PNM (Proactive Network Management) data into a graph topology, the authors will show how advanced 10G networks, and particular bidirectional signal flows, can be simulated for both existing and proposed networks.

Having a graph topology of the network, along with cable and component specifications, provides a means to apply transmission line theory as never before. Most cable operators have this RF and reflection data readily available for their systems. This simulation of cable systems, as-built, can predict end-to-end performance which can be compared against actual measured performance using PNM tools. This technique can provide a method for the evaluation of full-duplex (FDX), extended spectrum (ES) and traditional RF transmission performance of existing plant, as well as experimental designs. Beyond that, it can solve for fault detection, and many other previously unsolvable RF mysteries within our cable universe.

2. Background of Network Modeling

Scattering parameters or S-parameters are commonly used to model the electrical behavior of radio frequency (RF) communications networks when stimulated by electrical signals. The S-parameters are used to form a scattering matrix or S-matrix, that collectively describe the behavior of a network or circuit at different frequencies. These parameters are convenient for high-frequency RF design because they are easily measured by standard lab equipment such as a vector network analyzer (VNA). Coaxial S-parameters are further described in section 7.2.

There are alternatives to S-parameter modeling, which can be used to understand the transfer function of communications networks. Analysis of the frequency response including signal levels transmitted from the node and received by the cable modem or vice versa is described. This approach allows the use of log magnitude values versus frequency from component spec sheets to model the transfer function of each section of transmission line (i.e., cable) terminated at either end with an impedance (i.e., a tap, node, or cable modem) characterized by its magnitude return loss versus frequency. Note that this formulation provides the complex frequency response with only the scalar amplitude versus frequency of the cable transmission line and the magnitude return loss and insertion loss versus frequency of the tap terminating impedances. This provides a convenient alternative to S-parameter and T-parameter measurements, which are difficult to measure on existing networks.

3. Network Modeling and Simulation

3.1. Physical Layer (PHY) Model

Traditional hybrid fiber/coax (HFC) cable distribution networks have been built as tree and branch networks consisting of a fiber node connecting multiple cascaded amplifier coax cable sections. Each section connects to a series of multiport taps transmitting signal to and receiving signals from drop cables

to customer premise equipment. An example of one coax branch of a conventional Node + N HFC architecture shown in Figure 1. The node span contains multiple amplifier spans, each with multiple taps between amplifiers.

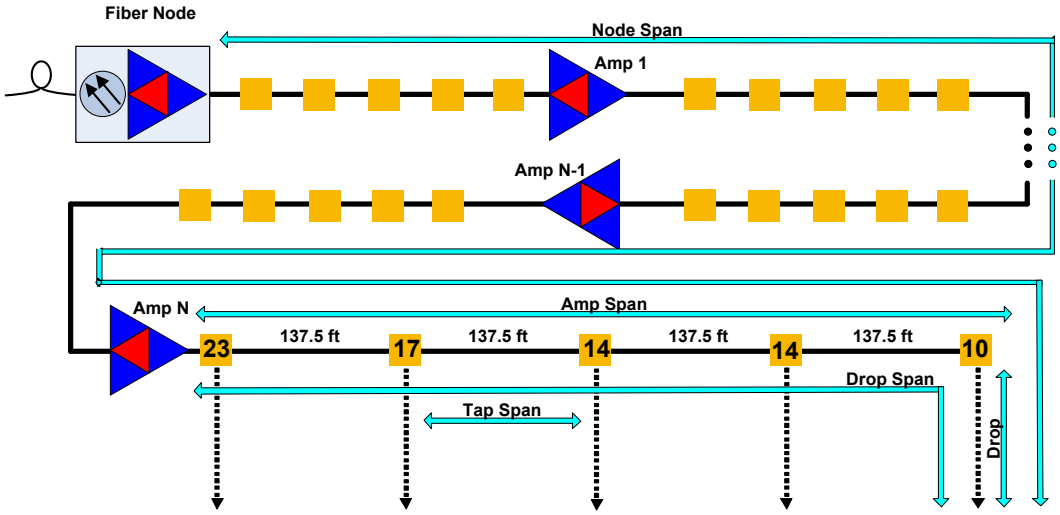


Figure 1 - Conventional node plus N network architecture

This conventional architecture provides two-way signal transmission on separate spectral bands using frequency division duplex operation. Each amplifier section uses diplex filtering to separate upstream transmissions toward the node in the narrower lower frequency band (typically 85 MHz or less) from downstream transmissions from the node in the much wider upper frequency band (up to 1.2 GHz). Such diplex filtering prevents two-way transmission within the same bandwidth. Each multiport tap contains a directional coupler that diverts a portion of the downstream signal to the drops connected to the tap ports and injects the upstream signals present on the tap ports toward the node. The directivity of the directional coupler prevents upstream signals from propagating in the downstream direction or from diverting to other drops upstream from that tap port.

The design of a hybrid fiber coax (HFC) network requires specifying the individual components and their values to obtain consistent signal levels across all the subscriber drops. These include:

- hardline cable (trunk and express feeder) and drop cable type and parameters
- tap parameters (insertion loss, tap loss, return loss, isolation)
- downstream and upstream node and tap equalization type, frequency range, and tilt value
- tap spacing
- drop cable length
- amplifier output (level and tilt)

Analysis of the frequency response including signal levels transmitted from the node and received by the cable modem or vice versa is described. This approach allows the use of log magnitude values versus frequency from component spec sheets to model the transfer function of each section of transmission line (i.e., cable) terminated at either end with an impedance (i.e., a tap, node, or cable modem) characterized by its magnitude return loss and insertion loss versus frequency.

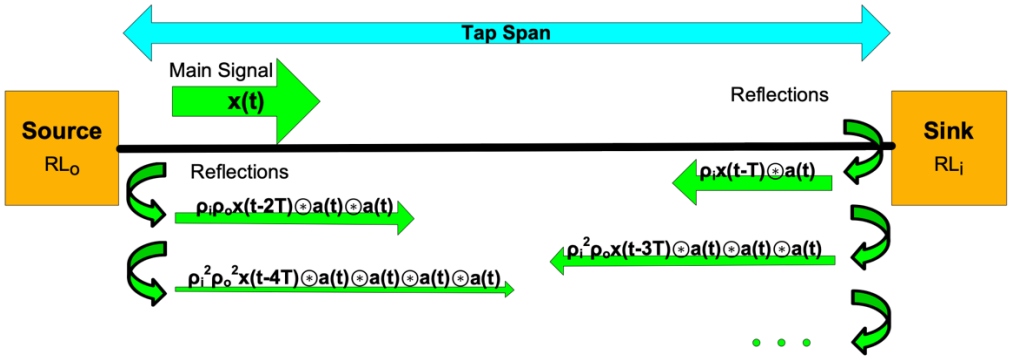


Figure 2 – Signal reflections in a cable between adjacent taps

Consider a signal transmitted downstream from a tap output to the adjacent tap input as shown in Figure 2. with amplitude response $A(f)$ and linear phase response which has the corresponding impulse response of the cable denoted by $a(t)$. The transmitter at the signal source has (nearly) matched impedance to the drop cable but with a return loss RL_o (dB). In general, the return loss is also a function of frequency $RL(f)$. The signal traverses the cable to the tap with propagation delay T which has a (nearly) matched impedance to the cable with return loss RL_i (dB). A portion of the signal equal to the reflection coefficient $\rho_i = 10^{-RLi/20}$ is reflected back to the source, which in turn a portion of the reflected signal equal to the reflection coefficient $\rho_o = 10^{-RLo/20}$ is re-reflected back toward the tap, and so on ad infinitum. This can be represented as a sum of the incident signal x(t) and the infinite series of reflections each delayed by the round trip (i.e., twice) the propagation delay T of the cable.

The same analysis applies to a drop cable section between a tap port and the cable modem F-connector port where the attenuation model and propagation delay are specified for the drop cable instead of the hard-line cable and the input and output return losses are specified for the tap port and the cable modem F-connector port respectively. The network simulation will require each of these transmission lines to be cascaded from the node to each cable modem attached to the network.

3.2. Network Topology as a Graph

Cable operators typically have some electronic form of system design, usually rendered by computer aided design (CAD) systems. These electronic representations of the network are used throughout the lifecycle of the network, from design to construction, upgrades, maintenance and so forth.

One of the major goals of the graph system is to map the physical and logical topology of the access network into a graph database that will allow querying paths between the physical and logical objects of the access network. There is no single dataset or database that currently has the required data compiled with common keys, instead the data is distributed across different units of the org among databases currently in use for services with differing goals. The lack of common keys and differing identifiers and definitions for the same objects in different databases present a challenge. The initial goal is to achieve mapping 95% of an access network node in the graph. To overcome this challenge, we develop and implement algorithms that check for availability of an access network object in multiple databases and when found confirm matches to the same vertex, when not-found provides an edge to the nearest vertex with a score that provides the strength of the attribution.

The goal is to map the following path: Device → Household (HH) → Tap → Amp → Buss Leg → Node using the following 3 data sources:

1. *CAD* provides detailed data on the physical design topology of serviceable addresses. Attributes reported include model, type, and location coordinates for CMTSs, Nodes, Taps, Amps and Serviceable Addresses: This is used for HH → Tap → Amp → Buss Leg → connections
2. The node combining plan *(NCP)* provides billing information presenting currently active services serving subscribers with activated devices for each location. Providing for Device → HH connections
3. *CM Topology* provides CMTS registry, and the devices connected to the CMTS at the given moment. Used to validate online inventory from Device → CMTS

While CMTS naming is conventional and CM Topology can be matched to NCP rather easily, the main difficulty is identifying the activated addresses in physical design topology in CAD and matching them with the addresses from the billing data. The matching algorithm consists of the following steps:

Step 1: Filter the CAD dataset for HH's using a node ID and account numbers with corresponding Addresses and Devices in the network data aggregator

Step 2: Use an enhanced fuzzy matching algorithm on the address fields of CAD data and network data aggregator to get high confidence matches for same addresses.

Step 3: Use lon/lat coordinates as reported in CAD data and network data aggregator to match records based on geographical proximity. We match coordinates that are closest to each other using a threshold for maximum distance.

Step 4: When no match can be made between CAD and network data aggregator records, get the records for all the taps with their coordinates attached to the node from CAD. Match all the network data aggregator records to the nearest tap using lat/lon coordinates. All the addresses from network data aggregator gets a match at the end of this step, all the unmatched addresses from CAD are tagged as inactive/passed addresses.

Step 4: Merge all the matched addresses to CM Topology data using media access control (MAC) addresses. Anything unmatched on left and right side are tagged accordingly and will be passed off to related teams for further review.

3.2.1. Graph Topology Data Structure

The graph data used in the simulation consists of a set of vertices/nodes that represent the hierarchical elements of the network, and a set of edges that represents the relationships in this hierarchy. A high-level example is illustrated in Figure 3. Figure 5 shows the details of the graph structure which contains a cable modem (Device), equipment location (HH), drop cable (Drop) and tap (Tap) which are connected. Figure 4 shows a more complex example of the CMTS topology mapping to the RF active ports.

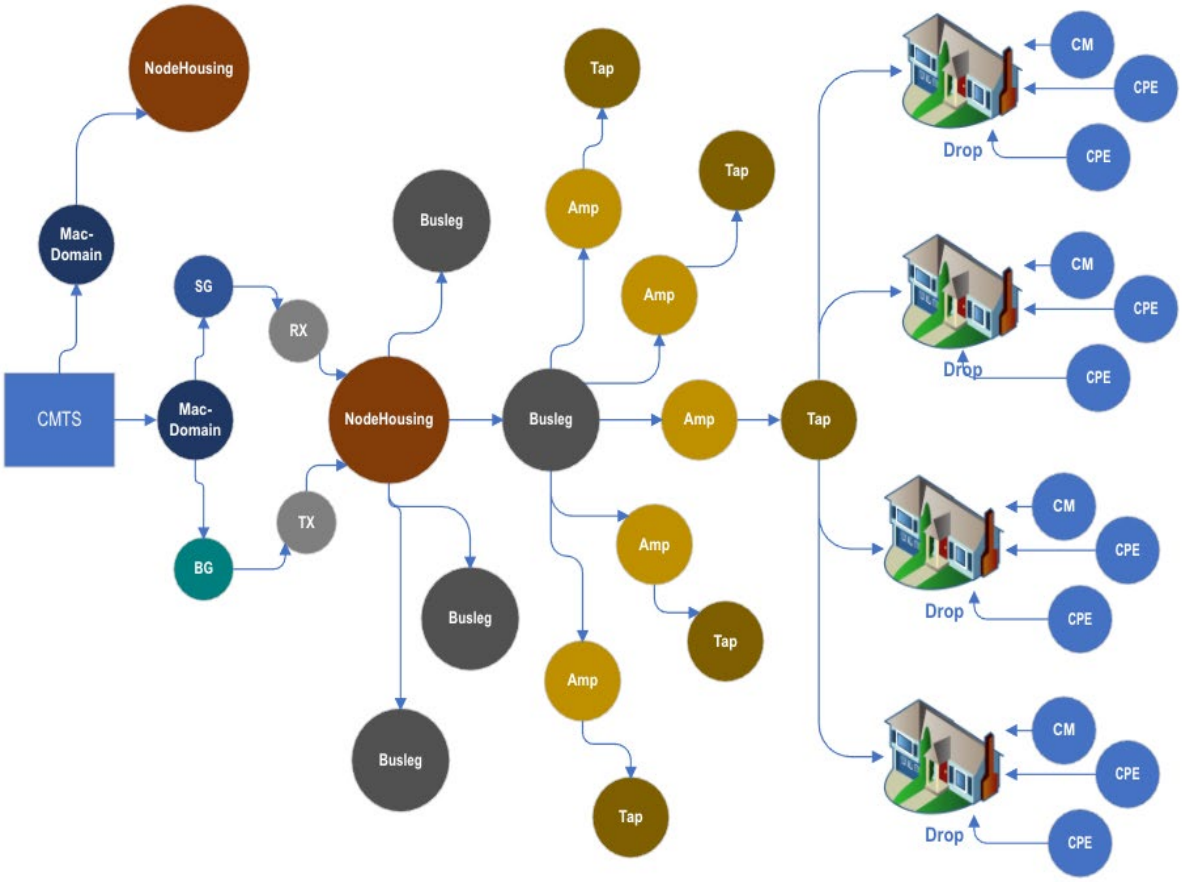


Figure 3 – High-level graph model example

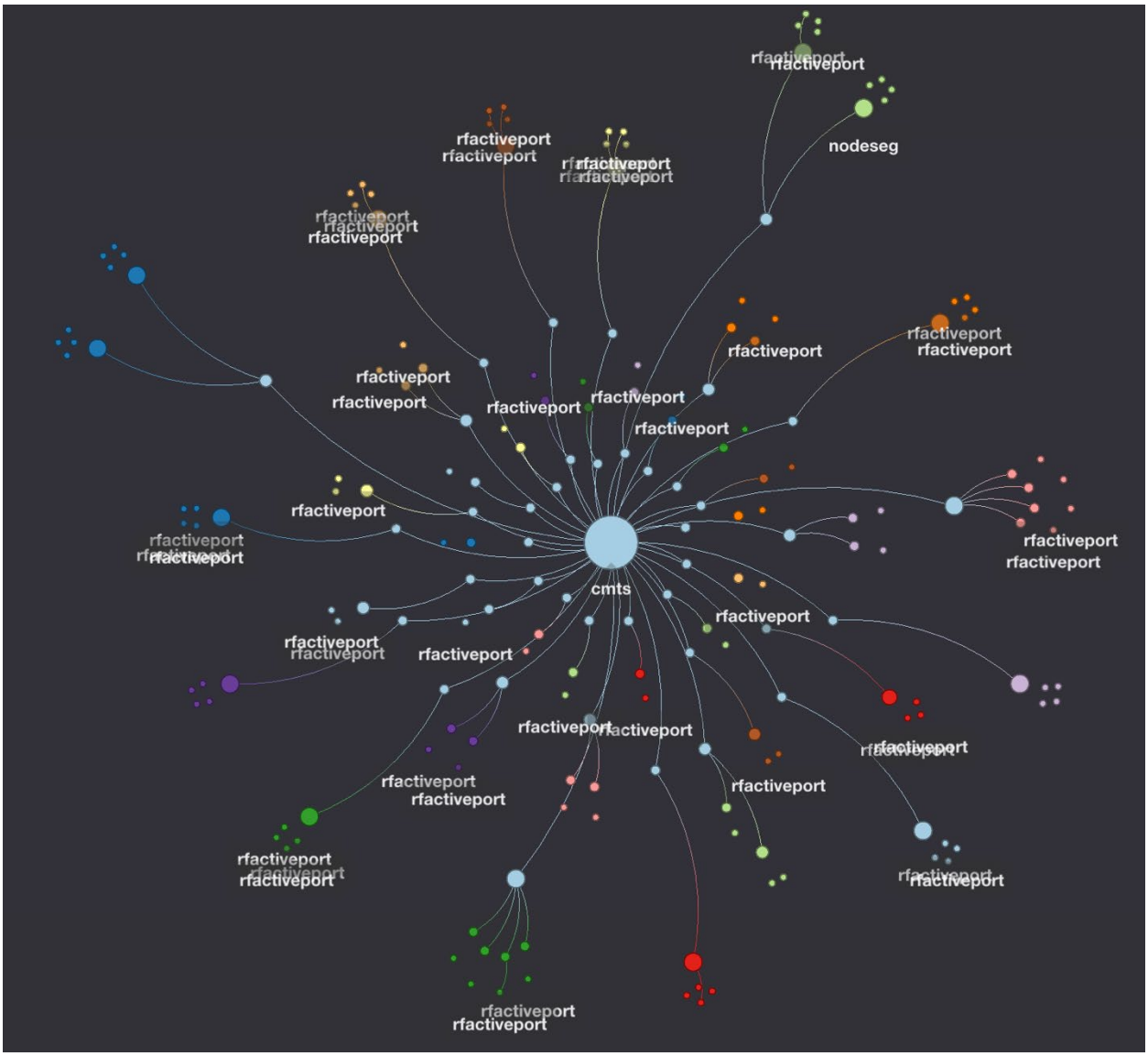


Figure 4 – Example graph terminating at Rfactiveports/busleg

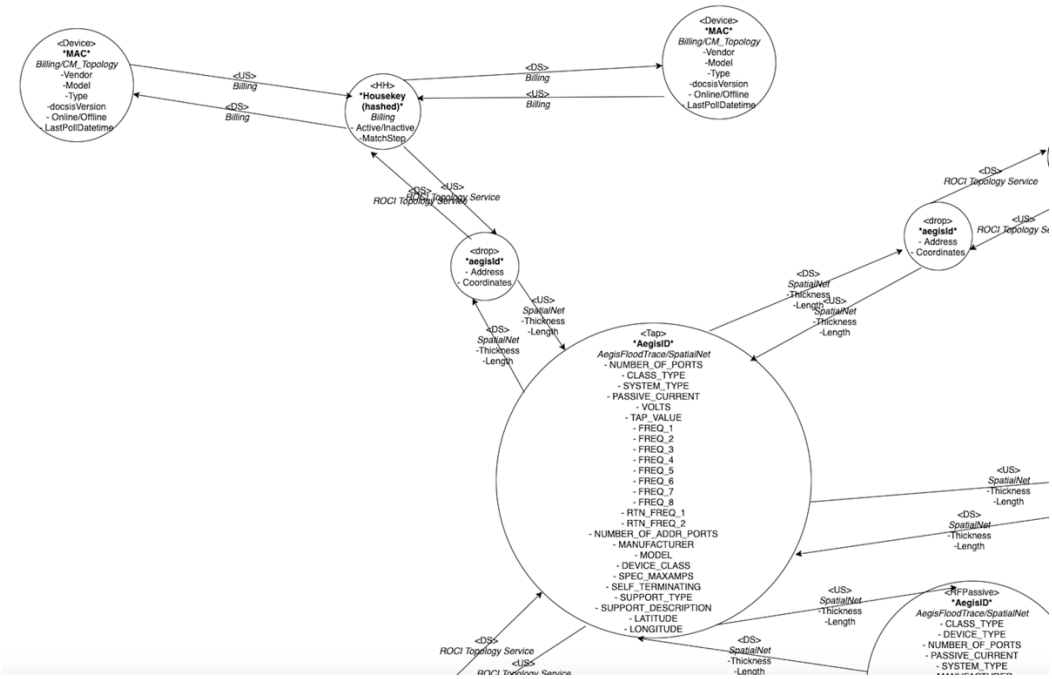


Figure 5 – Detailed graph model (tap, drop, modem)

1. Vertices
    a. The graph vertices are elements from the Division to a site, a CMTS in the site down to the individual devices in a subscriber's home.
    b. Division, Region, Site, CMTS, Headend Tx/Rx. US/DS Port, MAC Domain, Service/Bonding Group, Clamshell (node housing), Bussleg, Active (Amp), Passive, Tap, Design Address/Drop, Billing Address/HH, and Device.
    c. Labeled as:
    'division', 'region', 'cmts', 'headendTx', 'headendRx', 'usport', 'dsport', 'macDomain', 'servic eGroup', 'bondingGroup', 'clamshell' is the node housing, 'bussleg' is the fiber node in the node housing, 'amp', 'passive', 'tap', 'drop', 'billingaddress', 'device'
2. Edges
    a. Edges exist between vertices in the graph that denote the relationship between the vertices.
    b. The edge type depends on the data source that creates the edge, the context of the edge in terms of direction, or physical nature of the relationship between vertices.
        i. US - an upstream edge. This exists from device upstream through the flood trace all the way to MAC Domain and directly from MAC Domain to the CMTS.
        ii. DS - a downstream edge. This exists from the CMTS directly to Mac Domain down through the flood trace to device.
        iii. NCP_US - this is an edge created based on a confidence metric that we can map NCP reports of CMTS US ports to RX to the correct MAC Domain and bonding groups.
        iv. NCP_DS - this is an edge created based on a confidence metric that we can map NCP reports of CMTS DS ports to TX to the correct MAC Domain and Service Groups.

v. resides_in - Used for busslegs and optical rx/tx that physically exist inside a node housing.

3.3. Software Simulation

After reviewing the physical layer model and graph topology, we now have sufficient context to begin the simulation. Each of the previously discussed models and network elements are implemented as software components. The software objects are used to instantiate the network topology and gather the required parameters to build transmission line segments. These transmission line segments can then be cascaded to simulate a complete network that is redrived from the network design and specifications.

3.3.1. Database of Component Specifications

The graph network topology provides information that identifies the equipment types, which are used to assemble the required specifications. The product specifications are usually provided in the form of human-readable documentation such as portable document format (PDF). These specifications need to be converted to a database format that can be referenced by the software. To support this simulation exercise, the scope of equipment was limited and generalized to reduce the amount of document translation required. The list network elements and their required specifications are as follows:

- **Cables** - Diameter, length, structural return loss, nominal velocity of propagation (NVP) and attenuation vs frequency
- **Taps** – Return loss vs. frequency, attenuation vs. frequency, tap loss, port-to-port isolation, port-to-tap isolation, tap-to-tap isolation
- **Conditioners and pads** – Attenuation vs. frequency
- **Amplifiers** – Return loss vs frequency, gain vs. frequency
- **Splitters (and directional couplers)** – Through loss vs. frequency, tap loss vs. frequency, return loss vs. frequency

3.3.2. Network Object Model

The software object model looks a lot like the network physical model in terms of equipment and their connections. All the typical network elements have corresponding classes which are extended from a parent class of `NetworkElement`. The abstract class of `NetworkElement` contains all the functionality that is common to all network elements, such as their graph relationships (parent-child or node-edge), attenuation vs. frequency and return loss vs. frequency.

The abstract class of `NetworkElement` is extended by all the subclasses that provide their extended implementations. Our model contains `Cable`, `Node`, `Splitter` (inclusive of Directional Coupler), `Tap`, `Amplifier` and `CableModem`, which each have unique characteristics beyond the abstract `NetworkElement`. The `NetworkElement` is an abstract class which is extended by several inheriting subclasses including cables, taps and so forth. An example of this class diagram is shown in Figure 8.

3.3.3. Transmission Line Object Model

Figure 6 shows a common schematic symbol for a coaxial transmission line. The source impedance is ZS, characteristic impedance of the transmission line is Z0 and sink impedance is ZL.

Having the instantiated network elements, equipment specifications and connection topology (Figure 7) allows for transmission lines to be modeled and simulated. These are the minimum parameters required to perform a simulation using the proposed model: cable length, cable velocity of propagation, cable attenuation vs. frequency, source output return loss vs. frequency and sink input return loss vs. frequency.

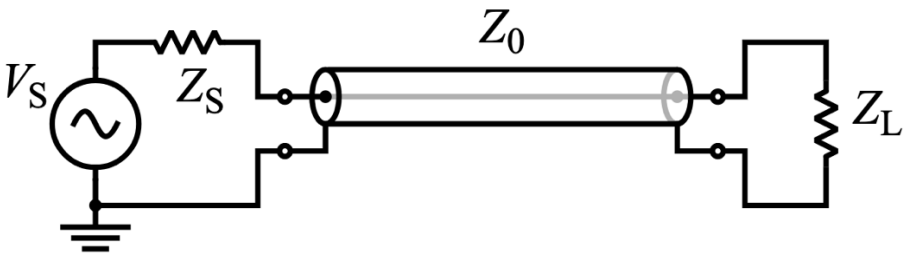


Figure 6 – Schematic symbol for a coaxial transmission line

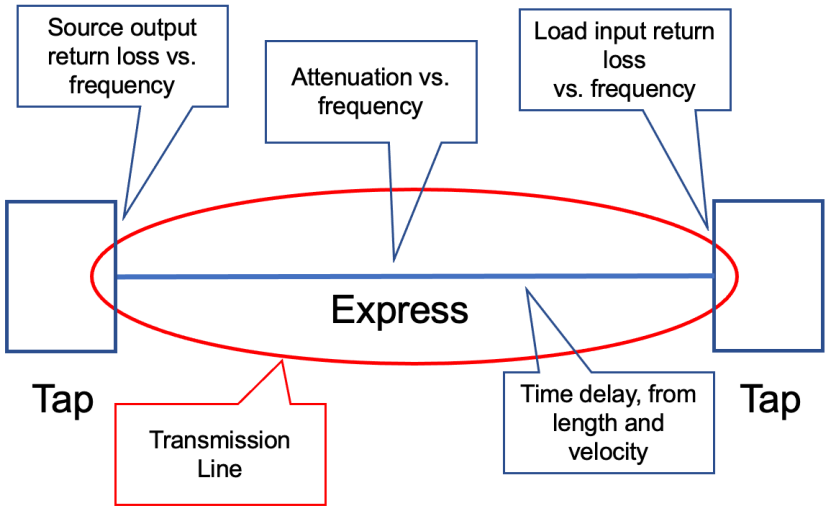


Figure 7 – Parameters required for S21 equivalent of a transmission line

To support the assembly and cascading of the transmission line parameters, a transmission software package is created to provide the implementation. Our TransmissionLine objects are constructed by traversing the fully assembled network graph, extracting the connected network elements (example, Figure 7) and copying the required parameters (Figure 7) to the `TransmissionLine` object.

3.3.4. Simulation of a Transmission Line

Transmission line transfer function is described earlier by Dr. Prodan in Figure 2. Recall that there is an infinite series of reflections interacting with incident signals. The series of reflections might imply that a recursive software routine could be useful. Fortunately, because of attenuation, the loss associated to the infinite recursions does eventually become insignificant (or having a practical limit). Because of this, there is a math shorthand which can be used, known as the Euler summation. This is a handy way to accelerate the summation of an infinite series.

for a length of trunk cable with cable propagation delay T, cable amplitude response A(f), and tap input/output port return loss RL = -10 Log(ρ) where ρ = reflection coefficient, that the transfer function H(f) for the tap span cable transmission line is given by:

$$H(f) = \frac{A(f)\, e^{-j2\pi fT}}{1 - A^2(f)\, 10^{-\frac{(\mathrm{RLi+RLo})}{20}} e^{-j4\pi fT}}$$

This formulation provides the complex frequency response with only the scalar amplitude versus frequency of the transmission line and the magnitude return loss versus frequency of the terminating impedances. This avoids the need to measure complex valued S-parameters versus frequency for each component. It also avoids the conversion to t-parameters for transmission frequency responses that would be cascaded to compute the end-to-end transfer function between any two points in the network.

3.3.5. Running the Simulation

3.3.5.1. Instantiating the Network Graph

To be useful in the network simulation, the network objects need to be instantiated and assembled dynamically, driven from the graph model discussed in section 3.2 and section 3.3.

Fortunately, coaxial cable networks adhere to a tree-and-branch structure, so graph models are easily traversed without concern of recursion or circular referencing (circular dependencies). This cardinality is strictly enforced by the `NetworkElement` object having only a singular, mandatory parent with zero-to-many child relationships. Each graph vertex can be used to instantiate the object from its respective class using a factory software pattern. The unified modeling language (UML) diagram in Figure 8 helps illustrate the factory pattern.

Converting to-and-from graph and tree-branch networks may seem like an anti-pattern. However, a good way to think about the network topology is having a physical vs. logical duality. The graph model better serves the network's logical model with properties such as multiple combining of ports, service groups and MAC domains. At the same time, HFC's physical network elements are designed and built in a tree-branch structure with elements such as nodes, cables, splitters, and taps.

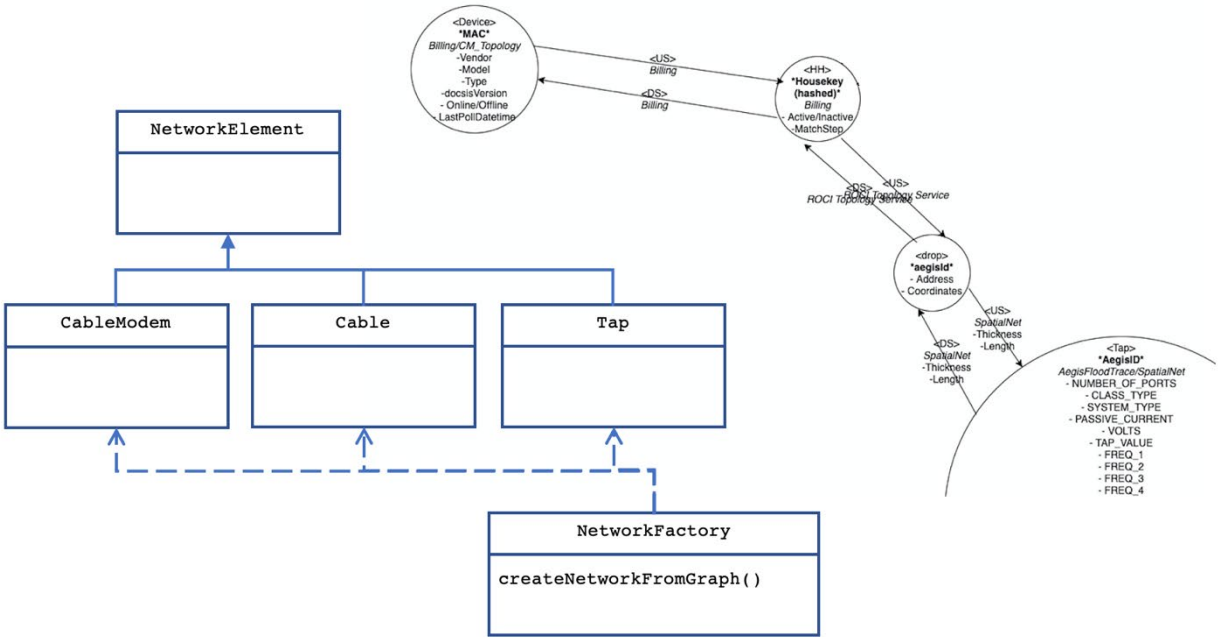


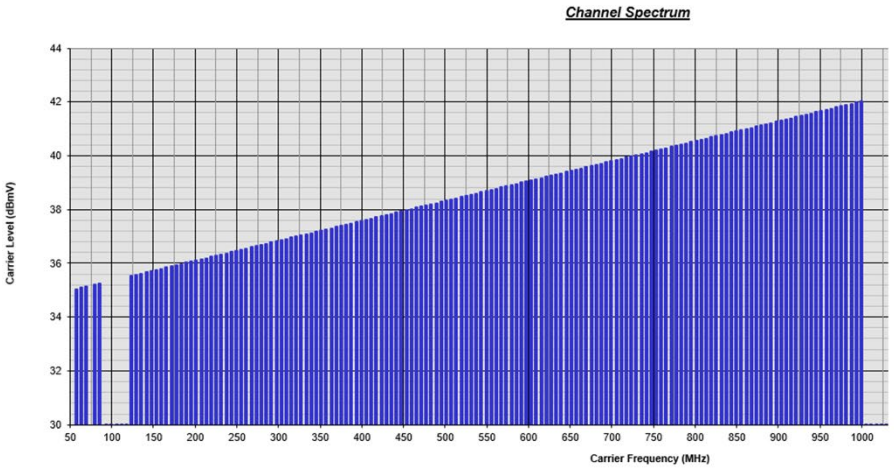
Figure 8 – UML diagram of network factory from graph

3.3.5.2. Operational Power Profiles and Launching

When simulating the transmission of RF signals on a cable network, the power levels and tilt are required. There are several different variables that go into operational launch power configurations. Especially when considering existing cable networks, different types of equipment, lengths, and types of cables, taps and architecture all influence the launch power selection. At Comcast, nearly 50 different variations of operational power launch profiles exist. As a practical matter in the simulator, the top 6 profiles will be generalized, allowing for additional power profiles to be added as needed.

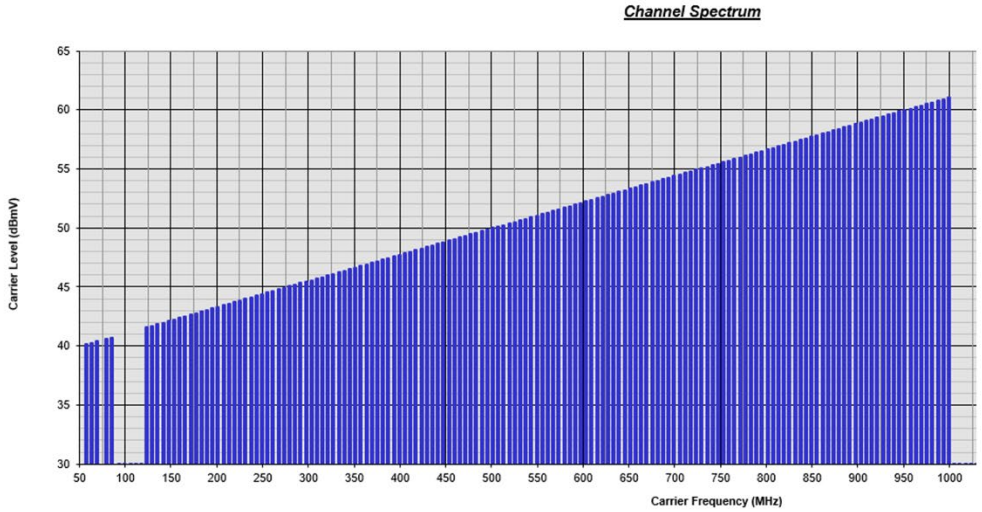
The operational launch power is expressed as power levels at given frequencies, resulting in a positive tilt. This tilt is designed to accommodate the anticipated attenuation of the passive components. When designing the systems, the operational tilt is optimized to deliver a flat RF spectrum at the end of the line(s). Additional tilt can be removed by conditioning and equalizing the signal, which are also accommodated in the software model, described later.

Figure 9 and Figure 10 show two examples of operational power profiles. The tilt is measured as the difference between the 6 MHz channel power measured at 57 MHz and 873 MHz respectively. Notice that the frequency spectrum is fully occupied up to 1 GHz. In both examples, the full-spectrum tilt is higher than the specified profile tilt value. When comparing the two figures, note the power per-division on the y-axis is 2 and 5, Figure 9 and Figure 10 respectively.



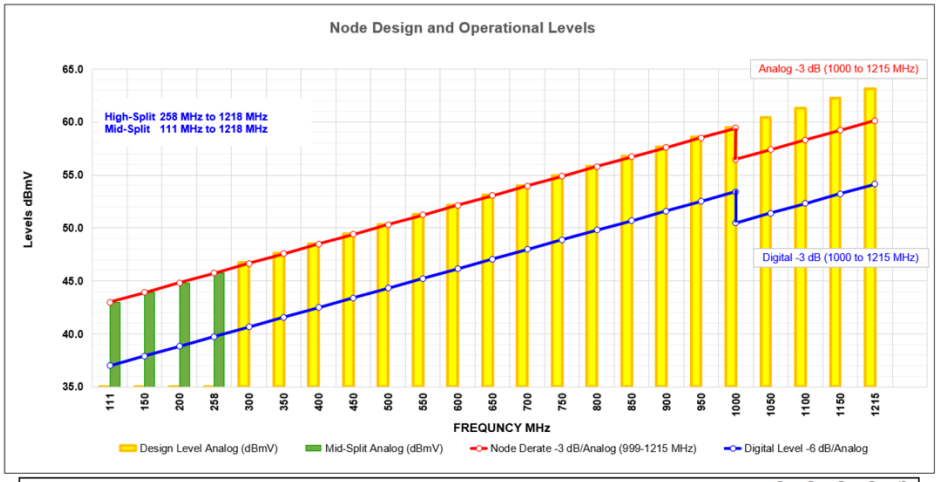
Frequency	RF Level
57 MHz	35 dBmV
549 MHz	39 dBmV
753 MHz	40 dBmV
873 MHz	41 dBmV
999 MHz	42 dBmV
Tilt @ 750 MHz	5 dB
Tilt @ 870 MHz	6 dB
Tilt @ 999 MHz	7 dB

Figure 9 – Operational power profile with 6 dB of tilt



Frequency	RF Level
57 MHz	40 dBmV
549 MHz	51 dBmV
753 MHz	56 dBmV
873 MHz	58 dBmV
999 MHz	61 dBmV
Tilt @ 750 MHz	15 dB
Tilt @ 870 MHz	18 dB
Tilt @ 999 MHz	21 dB

Figure 10 – Operational power profile with 18 dB of tilt



MHz	111	150	200	258	300	350	400	450	500	550	600	650	700	750	800	850	900	950	1000	1050	1100	1150	1215
Design Level Analog (dBmV)	43	44	45	46	47	48	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
Operational Node Derate -3 dB/Analog (999-1215 MHz)	43	44	45	46	47	48	48	49	50	51	52	53	54	55	56	57	58	59	57	58	59	60	61
Operational Digital Level -6 dB/Analog	37	38	39	40	41	42	42	43	44	45	46	47	48	49	50	51	52	53	54	51	52	53	54

Figure 11 – Operational power profile with 1 GHz derating of -3 dB

Operational power profiles become further complicated when considering designs using frequency spectrum above 1 GHz. In contemporary system designs, RF output ports can be set with a 19 dB linear tilt from 258 MHz to 1,215 MHz (measured from CTA carrier center frequency to CTA carrier center frequency).

In this example (Figure 11), the node operating levels include a 3dB derate for all signals above 1000 MHz and for each RF output port set to:

- 43 dBmV analog-equivalent NTSC level at 111 MHz. Single-channel, 6-MHz 256-QAM level shall be 6 dB lower, or 37 dBmV at 111 MHz for initial mid-split deployments.
- 46 dBmV analog-equivalent NTSC level at 258 MHz. Single-channel, 6-MHz 256-QAM level shall be 6 dB lower, or 40 dBmV at 258 MHz.
- 60 dBmV analog-equivalent NTSC level at 1000 MHz. Single-channel, 6-MHz 256-QAM level shall be 6 dB lower, or 54 dBmV at 1000 MHz.
- 61 dBmV analog-equivalent NTSC level at 1,215 MHz. Single channel, 6-MHz 256-QAM level shall be 6 dB lower, or 55 dBmV at 1,215 MHz.

3.3.5.3. Channel Plans and Lineups

For modeling purposes, it's typical to use a standard channel plan and fully populate the RF spectrum model with 6 MHz power equivalent SC-QAM channels. In Figure 9, Figure 10, and Figure 11, it can be seen that the operational power profiles are often expressed as such. However, the simulation will be required to support all common forms of cable signal transmissions such as SC-QAM, OFDMA, SCTE-51 out-of-band (OOB), analog channels and continuous wave (CW) pilots. This makes it possible to simulate and compare channel lineups and power loaded scenarios that are found in deployed networks. Figure 12 shows 2 examples of different channel lineups. The top graph shows a simulated RF spectrum, fully loaded with SQ-QAM from 57 MHz to 1 GHz. The bottom graph shows the same frequency spectrum, mixed with SC-QAM and OFDM which are sparsely loaded. In this example, there are 5 channel blocks of SC-QAM and 1 block of OFDM.

Figure 12 –Examples of channel plans and lineups

3.3.5.4. Echo Transfer Function

Consider a signal transmitted downstream from a tap output to the adjacent tap input as shown in Figure 2 with amplitude response $A(f)$ and linear phase response which has the corresponding impulse response of the cable denoted by $a(t)$. The transmitter at the signal source has (nearly) matched impedance to the drop cable but with a return loss RL_o (dB). In general, the return loss is also a function of frequency RL(f). The signal traverses the cable to the tap with propagation delay T which has a (nearly) matched impedance to the cable with return loss RL_i (dB). A portion of the signal equal to the reflection coefficient $10^{-RLi/20}$ is reflected back to the source, which in turn a portion of the reflected signal equal to the reflection coefficient $10^{-RLo/20}$ is re-reflected back toward the tap, and so on ad infinitum. This reflected signal can be represented as a sum of the infinite series of reflections each delayed by the propagation delay T plus multiples of the round-trip time (i.e. twice the propagation delay or 2T) of the cable.

3.4. Simulation vs. Measured Performance

Having the measured values that can be compared with simulated values enables several use-cases. Implementors can evaluate designs vs. specifications, locate significant impedance mismatches and identify unanticipated path loss or gain. In the simplest form, this can be used to improve the classification if impedance mismatches such as faulty or damaged cable and equipment. In more complex scenarios, new designs can be simulated and validated, comparing with pre-existing network designs.

3.4.1. PNM Spectrum measurements

As of November of 2012, the DOCSIS 3.0 specification was expanded to include spectrum analyzer-like functionality in cable modems. This feature is known as full band capture (FBC) and is supported by most DOCSIS 3.0 and all 3.1 cable modems. Since this time, most cable operator-deployed modems now have the FBC spectrum analysis capability. This offers a convenient tool for measuring RF spectrum performance at the location of the cable modems and CMTS.

3.4.2. Configuring, enabling, and measuring the RF spectrum

Information about controlling PNM spectrum measurements is provided in [CM-OSSIv3.1], Section D.2.4, "CM Spectrum Analysis Objects." The spectrum analyzer control parameters, defined by Table 1, are used to configure and enable the FBC. Most of the implementations have default values, so simply enabling the spectrum analyzer should result in usable information for testing the basic functions of TFTP and SNMP. The measurements are configured and enabled by using SNMP with the following MIB objects.

Table 1 - Spectrum analyzer control parameters

Attribute Name	Type	Access	Type Constraints	Units	Default
Enable	Boolean	R/W			False
InactivityTimeout	UnsignedInt	R/W	0..86400	seconds	300
FirstSegmentCenterFrequency	UnsignedInt	R/W		Hz	93000000
LastSegmentCenterFrequency	UnsignedInt	R/W		Hz	993000000
SegmentFrequencySpan	UnsignedInt	R/W	1000000..900000000	Hz	7500000
NumBinsPerSegment	UnsignedShort	R/W	2..2048	bins-per-segment	256
EquivalentNoiseBandwidth	UnsignedShort	R/W	50..500	hundredths of bin spacing	150
WindowFunction	Enum	R/W	other(0), hann(1), blackmanHarris(2), rectangular(3), hamming(4), flatTop(5), gaussian(6), chebyshev(7)		
NumberOfAverages	UnsignedShort	R/W	1..1000		1
FileEnable	Boolean	R/W			False
MeasStatus	MeasStatusType	R/O			
FileName	AdminString	R/W	SIZE(1..255)		

The spectrum measurement output is obtained either through SNMP or TFTP file transfer. Table 2 describes the structure of this file, including byte offsets for each of the discrete values reported by the test. See [CM-OSSIv3.1], Section D.2.4, "CM Spectrum Analysis Objects," for the MIB details for implementation.

Table 2 - Spectrum analysis response format

Element	Size
Size File type (value = 504E4D09)	4 bytes
Major Version (value = 1)	1 byte
Minor Version	1 byte
Capture Time	4 bytes
Channel ID	1 byte
CM MAC Address	6 bytes
FirstSegmentCenterFrequency	4 bytes
LastSegmentCenterFrequency	4 bytes
SegmentFrequencySpan	4 bytes
NumBinsPerSegment	2 bytes
EquivalentNoiseBandWidth	2 bytes
WindowFunction	2 bytes
Length (in bytes) of SpectrumAnalysis Data	4 bytes
SpectrumAnalysisData	BinAmplitudeFileData

Described in Table 2, the PNM spectrum measurements are expressed as log magnitude power, per bin. Bins represent a range of frequency spectrum, given a center frequency and frequency span. The frequency span is sometimes referred to as the resolution bandwidth (RBW) and is used to specify the width of the measured spectrum.

For further details about the PNM spectrum capture capabilities, refer to [PNM Best Practices Primer: HFC Networks (DOCSIS® 3.1)], Section 5.2.8, “How to Implement”.

3.4.3. Simulated launch of the RF spectrum

The input of the simulator begins with a channel lineup, which has an operational power profile. Then, as the mathematical models are applied and cascaded, a resulting amplitude vs. spectrum is calculated. The amplitude vs. spectrum can be calculated anywhere within the network graph, at the input or output of the transmission line segments. Figure 13 shows the output of the simulated launch of a channel lineup, using the 18 dB tilt operational power profile shown in Figure 10. The simulated channel lineup includes both SC-QAM and OFDAM channels, fully occupying a 57 MHz to 873 MHz channel plan.

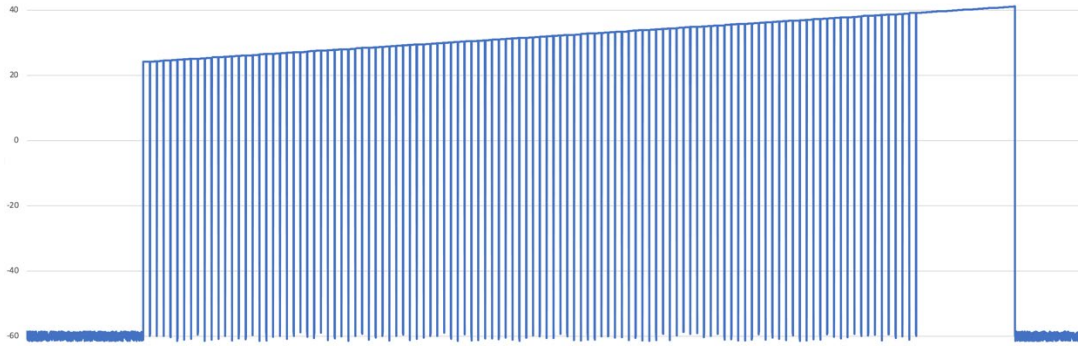


Figure 13 – Shows the simulated launch of an RF channel lineup with 18 dB of tilt

To simplify the comparison of the measured vs. simulated RF spectrum, the log magnitude power values should be scaled to the same RBW. This is a straightforward exercise using the following conversion equation. To change from a power spectral density (PSD) measured in RBW1 (i.e., dBmV/RBW1 MHz) to RBW2: PSD2 (dBmV/RBW2 MHz) = PSD1 + 10*LOG(RBW2/RBW1)

The process for generating simulated bins is as follows (including code examples). Note that this example does not contain any accommodating for spectrum shaping, such as a square-root raised cosine (RRC) filter. This may be a feature available in a future version.

1. Construct an array of bins using the frequency span divided by the target RBW. Our examples will use 6 MHz to 1026 MHz (1020 MHz total width) at 117187.5 kHz which results in 8704 total bins.

```
Bin[] bins = new Bin[(int)((endFreq - startFreq) / resolutionBandwidth)];
```

2. Initialize the bins with a simulated random noise floor.

```
for (int i = 0; i < bins.length; i++){

       bins[i].setLogPower(randomNoise);

       …
```

3. Enumerate the channels from the simulated launch plan and render the bins from each, interpolated at the center frequency from the operational power profile. Note that different channel types may have different rendering rules, such as SC-QAM and OFDM. For example, SC-QAM will have a simple roll-off factor and OFDM will have raised pilots.

```
bins[i].setLogPower((double) interpolatedBinPower);
```

3.4.3.1. Attenuation vs. Frequency

The linear attenuation versus square root frequency characteristic of coaxial cable is predictable. This useful property simplifies modeling the behavior of signal attenuation of the interconnecting hardline cable between taps or between a tap port and the terminating device on the drop cable.

Tap parameters including insertion loss, tap loss, return loss, and tap-to-output isolation are also specified as attenuation versus frequency. Figure 14 shows the proposed attenuation model (left) compared to the simulated output of the same model (right). Both charts show the attenuation vs. frequency on 100 meters

of series 11 coaxial cable, and they agree. Further, the chart in Figure 15 shows a channel lineup using the 18 dB launch profile (green trace), compared with simulated attenuation from 100 meters of P3 .875 (amber trace), P3 .500 (blue trace) and RG6 coaxial cables (grey trace). The attenuation parameters were extracted from the component parameter database discussed in section 3.3.1.

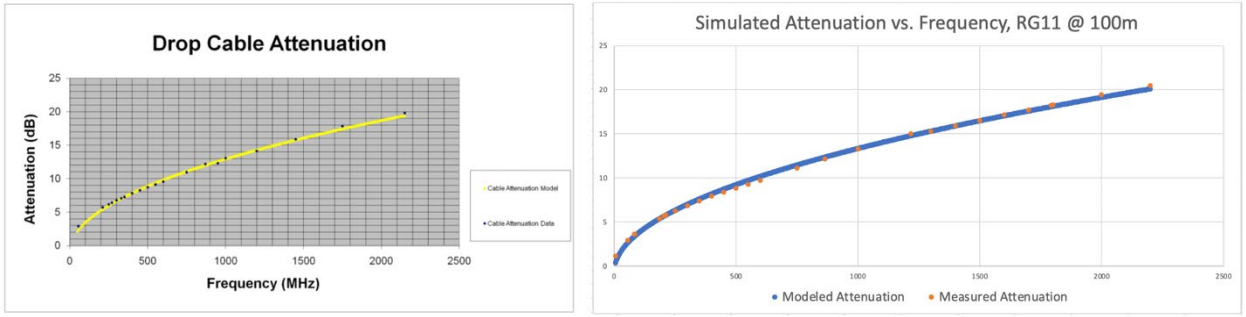


Figure 14 – RG11 attenuation model compared to RG11 simulated attenuation

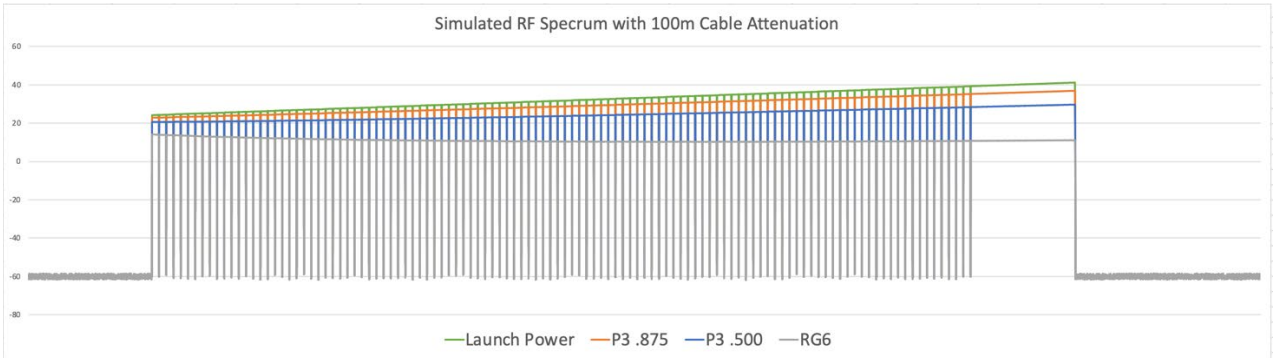


Figure 15 – Simulated launch RF compared to attenuation on P3 .875, P3 .500 and RG6 cables

3.4.3.2. Tap Equalization and Signal Conditioning

An important consideration for supporting the wider bandwidth on the passive fiber deep plant is the use of tap plug-in equalizers. Such equalizers provide either upward or downward attenuation (tilt) with increasing frequency. The amount for each tap is determined to provide approximately the same output levels versus frequency at all tap ports. An example of tap equalizer magnitude frequency (amplitude) responses is shown in Figure 16 and the simulator response in Figure 17.

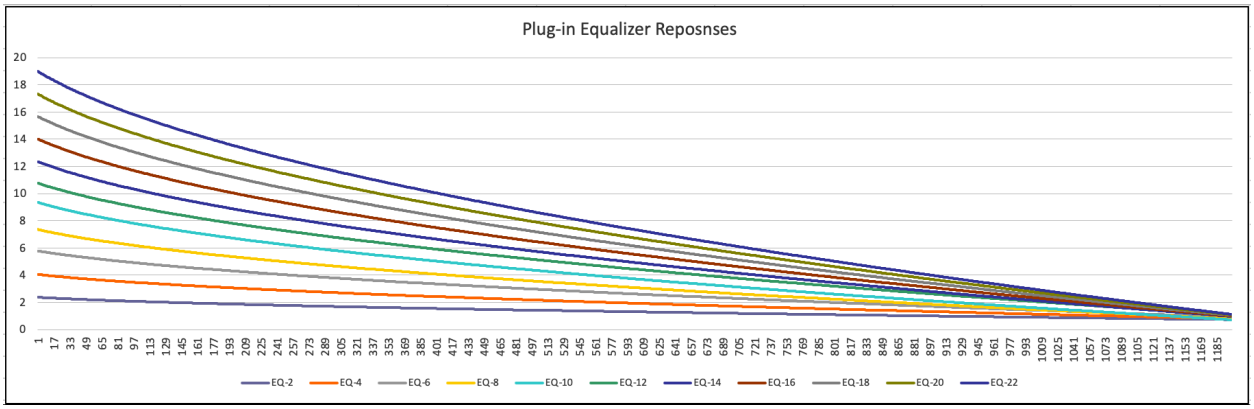


Figure 16 – Tap equalizer response vs. frequency

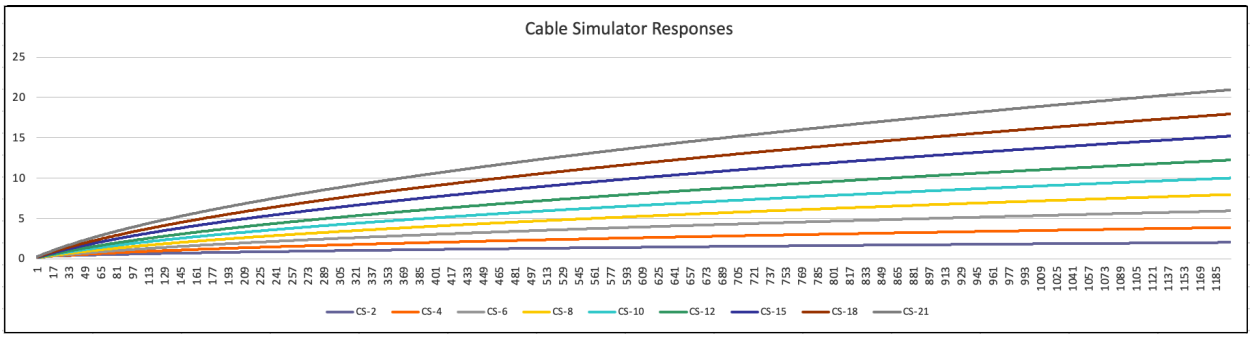


Figure 17 – Cable simulator response vs. frequency

3.4.3.3. Return Loss vs. Frequency

As seen in Dr. Prodan's model, tap parameters including insertion loss, tap loss, return loss, and tap-to-output isolation are also specified as attenuation versus frequency. Figure 18 shows an example of such tap specifications.

2-Way Maximum Insertion Loss (dB)										
Number of Ports (#-Way),Tap Value (dB)										
Frequency ≤(MHz)	2,4	2,7	2,10	2,12	2,14	2,17	2,20	2,23	2,26	2,29
5		3.9	2	1.8	1.3	1.2	1	0.7	0.7	0.7
10		3.6	1.8	1.6	1.2	1.1	0.9	0.6	0.6	0.6
50		3.5	1.8	1.6	1.2	1.1	0.9	0.7	0.7	0.7
100		3.6	2	1.8	1.3	1.2	1	0.8	0.8	0.8
450		4.2	2.4	2.1	1.6	1.5	1.3	1.1	1.1	1.1
550		4.2	2.4	2.3	1.8	1.5	1.4	1.2	1.2	1.2
750		4.2	2.4	2.4	1.8	1.6	1.5	1.3	1.3	1.3
870		4.3	2.7	2.7	2	1.6	1.6	1.4	1.4	1.4
1000		4.3	3	3	2.3	1.6	1.6	1.6	1.6	1.6
1218		5.2	3.6	3.6	2.7	2.2	2	2	2	2

Minimum Return Loss (dB)		
For ALL Tap Values		
Frequency ≤(MHz)	Minimum Return Loss In-Out (dB)	Minimum Return Loss Tap (dB)
5	16	16
10	16	17
750	16	17
1000	16	17
1218	16	16

2-Way Nominal Tap Value (dB)											Tap Value Tolerance ±(dB)		
Number of Ports (#-Way),Tap Value (dB)											Number of Ports (#-Way)		
Frequency ≤(MHz)	2,4	2,7	2,10	2,12	2,14	2,17	2,20	2,23	2,26	2,29	2	4	8
5	4	7	10	12	14	17	20	23	26	29	2	2	3
1000	4	7	10	12	14	17	20	23	26	29	2	2	3
1218	4	7	10	12	14	17	20	23	26	29	2.5	2.5	3.5

2-Way Tap-to-Output Isolation (dB)										
Number of Ports (#-Way),Tap Value (dB)										
Frequency ≤(MHz)	2,4	2,7	2,10	2,12	2,14	2,17	2,20	2,23	2,26	2,29
5		20	20	22	22	26	29	32	35	38
10		20	20	22	22	26	29	32	35	38
85		25	25	25	26	30	33	36	38	40
300		21	22	23	26	30	33	36	38	40
750		22	22	23	26	30	31	34	36	39
900		20	20	22	23	28	30	33	35	37
1218		20	20	20	22	25	29	31	33	35

Tap-to-Tap Isolation (dB)	
For ALL Tap Values	
Frequency ≤(MHz)	Isolation (dB)
5	20
10	25
85	27
300	27
750	23
1218	20

Figure 18 – Tap specifications for insertion loss, tap value loss, return loss, and tap-to-output isolation versus frequency for each tap value

4. Implementation and Results

4.1. Examples

When the simulation is implemented as described in the previous sections, a number of use cases can be satisfied. At the time of writing, the simulation software and analysis were just being completed so we will review 3 preliminary results. Further development will continue, including material described in Section 5. Future

4.1.1. Plant fault isolation

PNM continues to be one of the most important tools for operators to proactively monitor and maintain their networks. By enhancing PNM with an RF transmission line simulation, its classification and accuracy can sometimes be improved.

Among the most commonly detected PNM signatures using FBC are amplitude ripples, often caused by standing waves. Standing waves are caused by impedance mismatches somewhere between the node and the modem. Having a minimum of 2 impedance mismatches, energy reflects off the second mismatched impedance, back to the first, and then some of that energy continues downstream, but delayed. This causes an impact to the signal that appears as a standing wave in the spectrum amplitude. When observed in a spectrum display, the standing wave will appear as a periodic amplitude change. This is precisely what the physical layer modeling and simulation are doing. However, in an ideally designed and deployed cable plant, there should be no amplitude ripple.

This example (Figure 19) shows log magnitude frequency response obtained using the FBC feature of PNM. In this example, the PNM measurement data has 8704 total bins representing 6 MHz to 1026 MHz frequency spectrum. Each bin spans 117.1875 kHz, or .1171875 MHz of frequency spectrum. In this case, an amplitude ripple is clearly present throughout the occupied frequency spectrum. Then in Figure 20, the

network response is simulated including transfer function of the distribution cables, tap and drop cables. The simulation is configured to match the channel plan, launch-power profile and similar output resolution bandwidth of 117.1875 kHz.

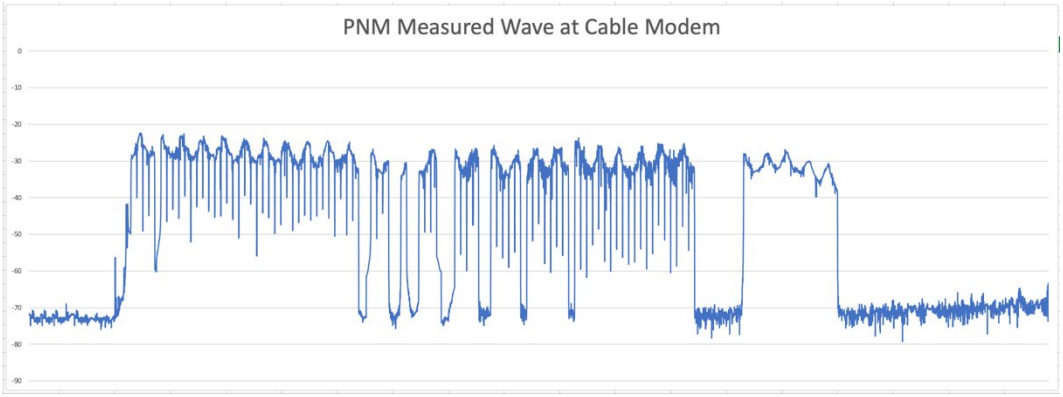


Figure 19 – PNM measurement of amplitude ripple from cable modem using PNM

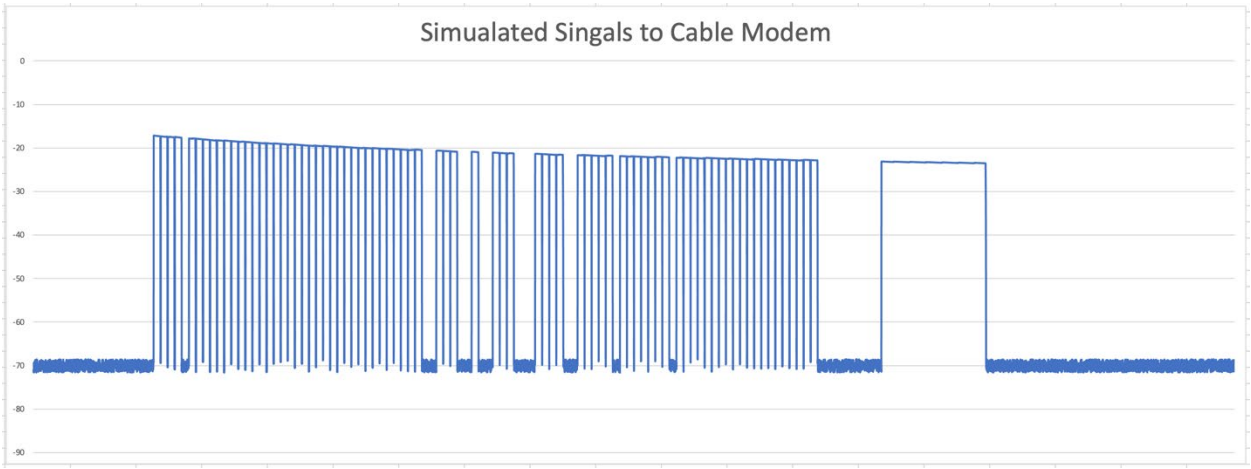


Figure 20 – Simulation of launched signal with cable, tap and drop attenuation

The resulting log magnitude bins from the PNM measurement and simulation cab be processed using common digital signal processing (DSP) techniques using fast Fourier transformations (FFT). The following results were achieved by projecting magnitude-only amplitude bins to the Complex number plane. This can be achieved in a number of ways, demonstrated here and discussed further in Appendix 7.1.

This method uses “zero stuffing” of the phase component which achieves a satisfactory result, assuming that the cable plant exhibits minimum-phase characteristics.

First, the bins need to be converted from logarithmic to linear form.

```
double power = Math.pow(10.0, (0.1 * decibels));
```

Next, an array of Complex numbers is instantiated by constructing each Complex value with the magnitude and a zero-value phase component.

```
Complex[] c = new Complex[linearBins.length];

for (int i = 0; i < linearBins.length; i++){
    c[i] = new Complex(linearBins[i], 0);
}
```

The resulting Complex arrays can then be used to perform Complex division, resulting in the frequency-domain Complex quotient seen in the top section of Figure 21.

```
Complex[] quotientArray = new Complex[complexArray.length];

for (int i = 0; i < complexArray.length; i++){
    quotientArray[i] = complexArray[i].divide(denominatorArray[i]);
}
```

Finally, using the inverse fast Fourier transformation (IFFT), a time-domain response of the Complex quotient is produced (bottom of Figure 21). In this case, the frequency and magnitude accuracy of the impairment is increased by referencing the impaired signal with the simulated signal, which are otherwise unknown to the PNM analysis.

```
Complex[] complexTimeDomain = fourier.transform(getPaddedArray(binsComplex),
TransformType.FORWARD);
```

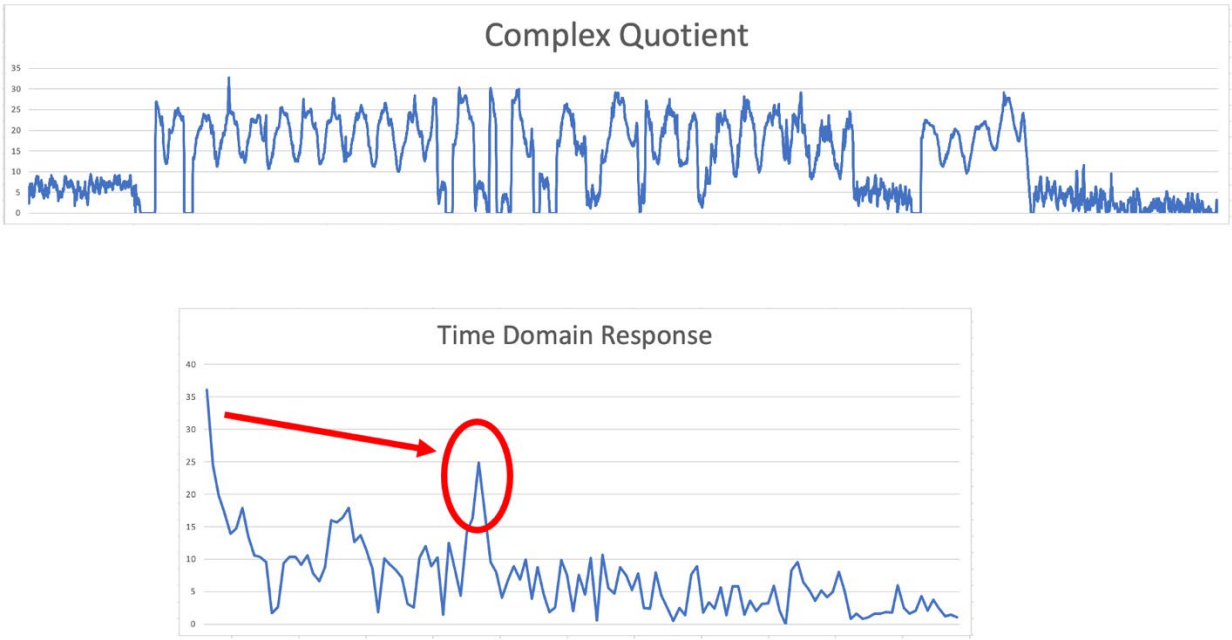


Figure 21 – Frequency(top) and time domain (bottom) analysis of Complex quotient

Having the improved frequency and time accuracy of the fault results in more precise fault lengths which can be calculated in either frequency or time.

This example was calculated, starting with the speed of light in feet per microseconds, rounded for simplicity to 983.571. The number is then divided by 2 for round-trip time and multiplied by the velocity factor of the drop cable, which is 0.85. Finally, the number is divided by the peak-to-peak ripple length of 19.28 MHz. The resulting calculation reveals that the faulty cable length is 21.68 feet, or perhaps the distance of the reflective fault from the cable modem F-connector.

4.1.2. System Design Verification

Another interesting and potentially valuable use case for simulation is to evaluate the quality and completeness of system designs. This simulation can predict end-to-end performance which can be compared against the design specifications. By traversing the network graph and comparing against PNM measurements, deviations and threshold violations can help identify inconsistencies in the designs.

4.1.3. In-home (drop) Amplifier Detection

The third example of the simulation is to help operators identify the presence of in-home (drop) amplifiers. These small amplifiers have been around for decades, used by operators and customers to add gain to the signal received by the modem. There are many different types of amplifiers including those with active and passive returns (upstream), passive, balanced and unbalanced active port configurations. In some cases, depending on equipment and configuration, it will be difficult or impossible to detect the presence of an amplifier. This method does not suggest an all-inclusive way of detecting every amplifier, rather it does propose a method that will work in some common configurations.

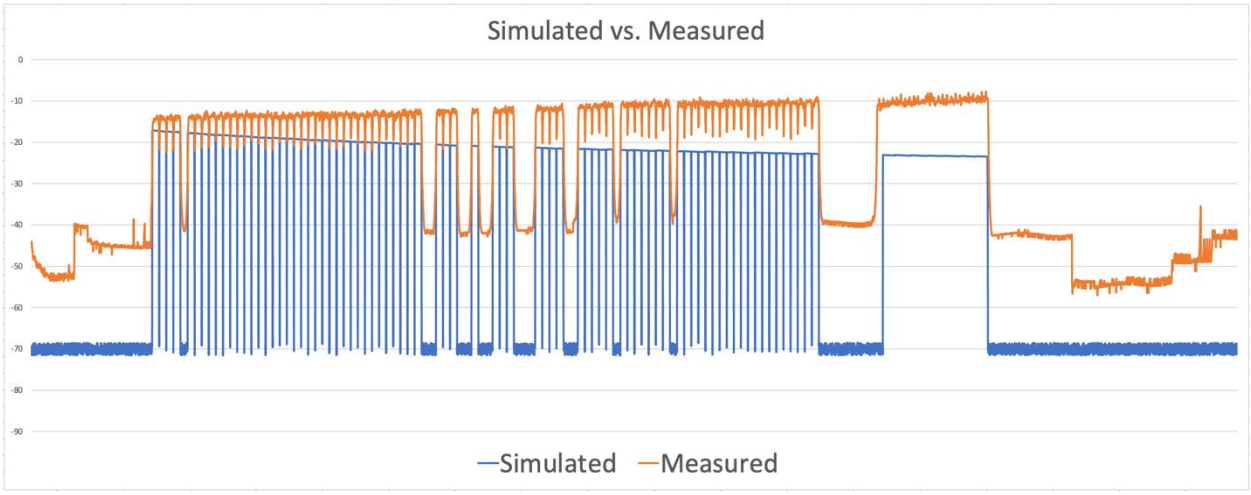


Figure 22 – Drop-amplifier detected: 15 dB gain measured vs. simulated

5. Future

5.1. FDX and ESD PNM extensions

Future versions of FDX DOCSIS technology will be supporting architectures beyond node+0. To support amplifier cascades in FDX, the amplifier technology will be equipped with echo cancellation. This complicates the physical model, which is well suited for this simulation exercise. There are several different amplifier design proposals, each making different claims about bit loading vs. distance, cascade depths and other significant performance implications.

ESD DOCSIS technology is far simpler to model, having very little change in the simulation other than the equipment specifications. Most of the cables already specify upper frequencies in the range of 2 GHz to 3 GHz. However, many of the active and passive components such as nodes, splitters and taps remain to be designed. When complete, adding their specifications to the existing simulator database is trivial.

In Figure 23, the simulated launch and attenuated signals can be compared when expanding from DOCSIS 3.1 to DOCSIS 4.0 frequency spectrum. The top chart is 5 MHz to 1 GHz of fully occupied spectrum using SC-QAM and one 192 MHz wide OFDM channel. The bottom chart shows the same lineup, adding 4 additional 192 MHz wide OFDM channels up to 1.8 GHz. This simulation was run using the 18 dB launch profile (green trace), compared with simulated attenuation from 100 meters of P3 .875 (amber trace), P3 .500 (blue trace) and RG6 coaxial cables (grey trace).

The TCP of the simulated node launch and cables are easily calculated by summing the log magnitude bins at the output of each.

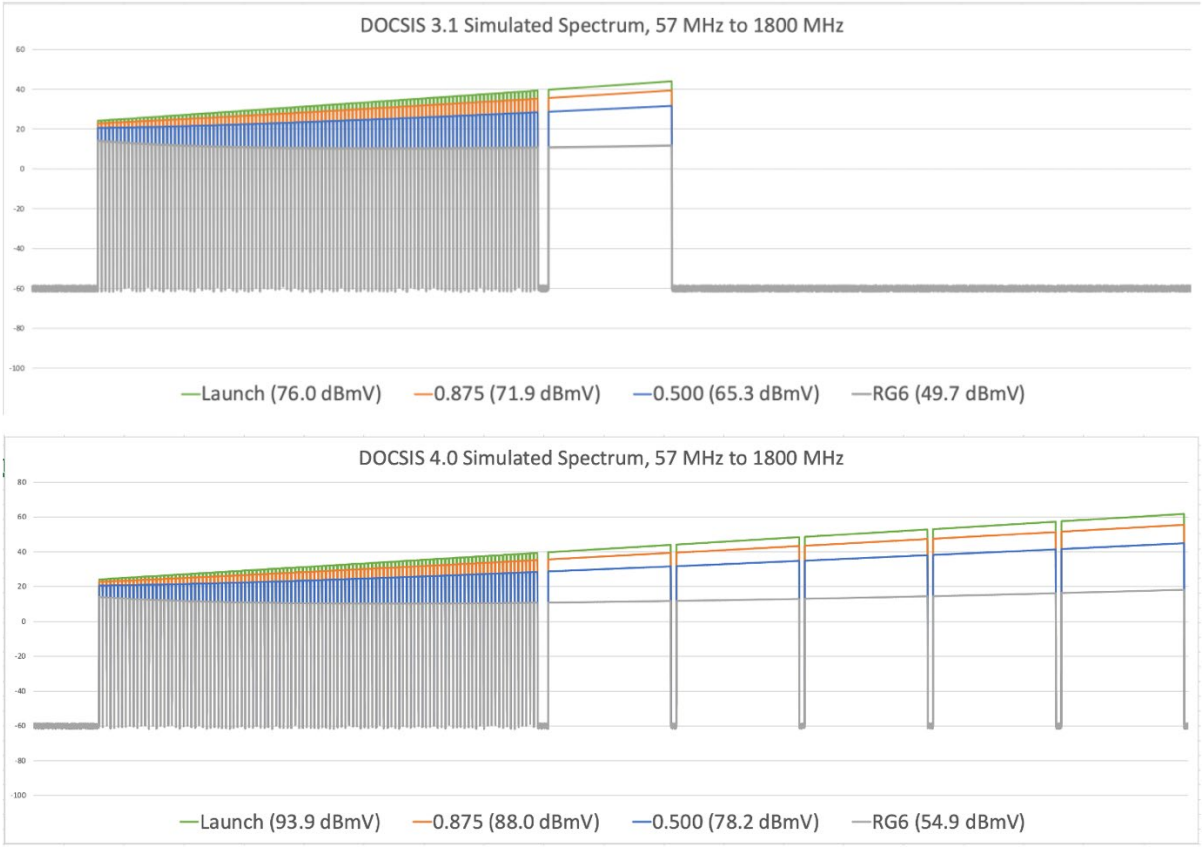


Figure 23 – Simulated spectrum of 1 GHz and 1.8 GHz compared, with TCP

Table 3 – DOCSIS 3.1 and DOCSIS 4.0 simulated output TCP compared

	18 dB Launch	**P3 0.875 (100m)**	**P3 0.500 (100m)**	**RG6 (100m)**
1 GHz	76.0 dBmV	71.9 dBmV	65.3 dBmV	49.7 dBmV
1.8 GHz	93.9 dBmV	88.0 dBmV	78.2 dBmV	54.9 dBmV

6. Conclusion

As our 10G networks continue to evolve and increase in complexity, so does our need to improve the sophistication of our modeling and simulations. This paper by Comcast's Ramya Narayanaswamy, Karthik Subramanya, Richard Prodan and Larry Wolcott, examine a practical approach to doing that. They prove that software, data science and graph theory can be used to advance the theoretical FDX and ESD physical-layer models previously proposed by Dr. Prodan.

By using a graph topology of the network, paired with PNM software, cable specifications, and data science, simulating cable networks just "got real". The simulation examples show that evaluating as-built

vs. simulated RF performance will be valuable for a number of use cases. Among them are the evaluation of full-duplex (FDX), extended spectrum (ES) RF performance. Also, traditional RF transmission performance of existing plant, as well as experimental designs. Lastly, it improves PNM fault detection, classification, and many other previously unsolved mysteries of RF.

7. Appendix

7.1. Phase recovery of magnitude-only measurements

Common digital signal processing (DSP) techniques exist to facilitate the analysis of digital signals. The PNM and simulation measurements discussed in this document are represented as magnitude-only power measurements. These measurements do not contain the constituent phase and amplitude information required to instantiate a Complex number, required for many DSP routines. There are several methods for projecting a magnitude value on to a complex plane, one of them is the Hilbert transformation.

Let G[k] be the minimum phase magnitude response. First convert magnitude to nepers using the natural (base-e) logarithm.

$$H[k] = \ln(G[k]) \qquad 0 \le k \le \frac{N}{2}$$

Next mirror the first half into the latter half (the latter half of the DFT corresponds to negative frequencies or negative times):

$$H[k] = H[N-k] \qquad \frac{N}{2} < k \le N-1$$

There are 3 steps to compute the Hilbert transform.

$$h[n] = \mathcal{DFT}\left\{ H[k] \right\}$$

$$= \sum_{k=0}^{N-1} H[k] e^{-j2\pi \frac{nk}{N}}$$

Then, multiply every positive time index ($n<N/2$) with $-j=e^{-j\pi/2}$ (or spin those complex values by -90°) and multiply every negative time index ($n>N/2$) with $+j=e^{+j\pi/2}$ (or spin those complex values by +90°). $h[0]$ and $h[N/2]$ (if N is even) should be set to 0.

$$h[n] \leftarrow \begin{cases} 0 & n = 0 \\ -j \cdot h[n] & 1 \le n < \frac{N}{2} \\ 0 & n = \frac{N}{2} \\ j \cdot h[n] & \frac{N}{2} < n \le N - 1 \end{cases}$$

Finally inverse transform that result and negate.

$$\phi[k] = -\mathcal{IDFT}\left\{ h[n] \right\}$$

$$= -\frac{1}{N} \sum_{n=0}^{N-1} h[n] e^{j2\pi \frac{nk}{N}}$$

$\phi[k]$ is the phase, in radians, of the minimum-phase system. The complex transfer function is:

$$G[k]\, e^{j\phi[k]}$$

7.2. About S parameters (Courtesy of Ron Hranac)

Consider a component or device being evaluated as a "network" with some number of ports N, and a characteristic impedance Z_0. For instance, a terminator can be considered a one-port network, as shown in Figure 1.

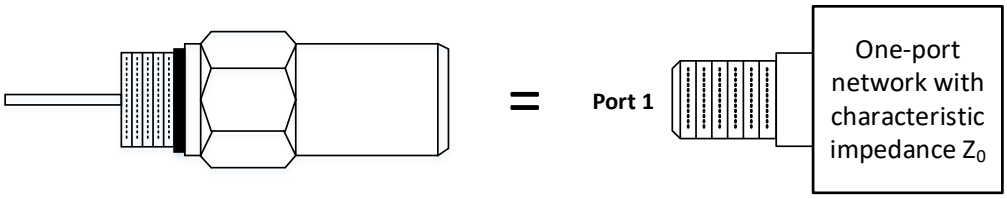


Figure 24 - Cable equipment chassis terminator represented as a one-port network.

Likewise, an in-line attenuator (or amplifier, tap, etc.) can be considered a two-port network, as shown in Figure 2.

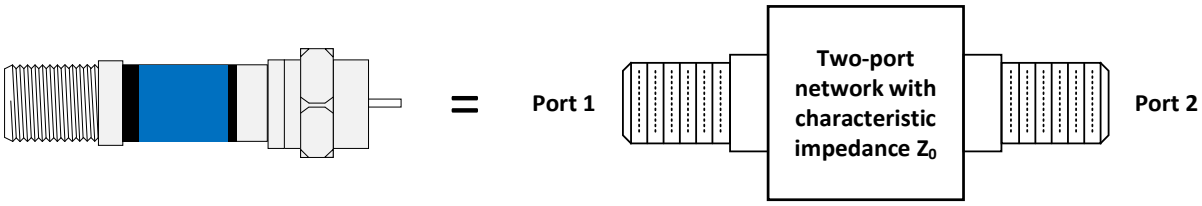


Figure 25 - In-line attenuator represented as a two-port network.

Among the metrics that can be used to characterize N-port networks are scattering parameters, or S-parameters. S-parameters are complex numbers usually expressed in the format S_{mn}, where m is Port 2 (the output port) and n is Port 1 (the input port). The remainder of this discussion focuses on two-port networks.

Assume that a two-port network such as an attenuator is being characterized. A test signal is applied to Port 1, and that signal measured at Port 2. From an S-parameter perspective, this gives us S_{21}, the forward voltage gain or transmission coefficient. See Figure 3.

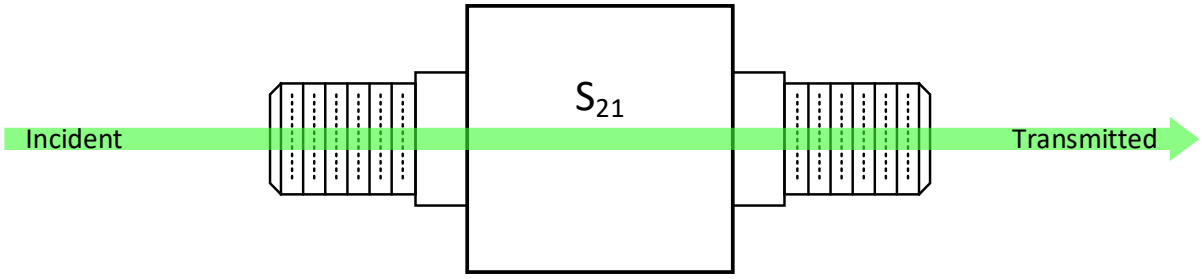


Figure 26 - Testing the two-port network from Port 1 to Port 2 gives the S-parameter S_{21}.

While the incident signal is being applied to Port 1, we can also measure any reflection from the two-port network. The reflection is a combination of reflections from Port 1's connector and all of the components inside of the two-port network. Here, the S-parameter is S_{11}, the voltage reflection coefficient for Port 1. See Figure 4.

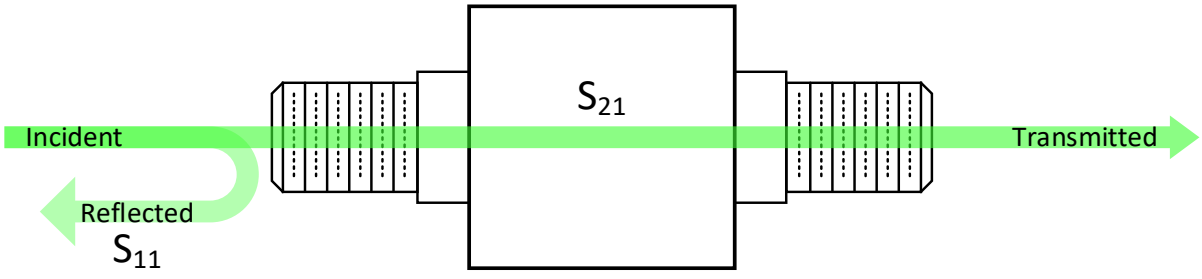


Figure 27 - Measuring the reflection from Port 1 gives the S-parameter S_{11}.

Note: When performing a measurement of S-parameters, the port(s) other than the one(s) being measured should be terminated in the network's characteristic impedance (typically 75 ohms for cable systems). An example is shown in Figure 5 for the case of a reflection measurement on Port 1, with Port 2 terminated.

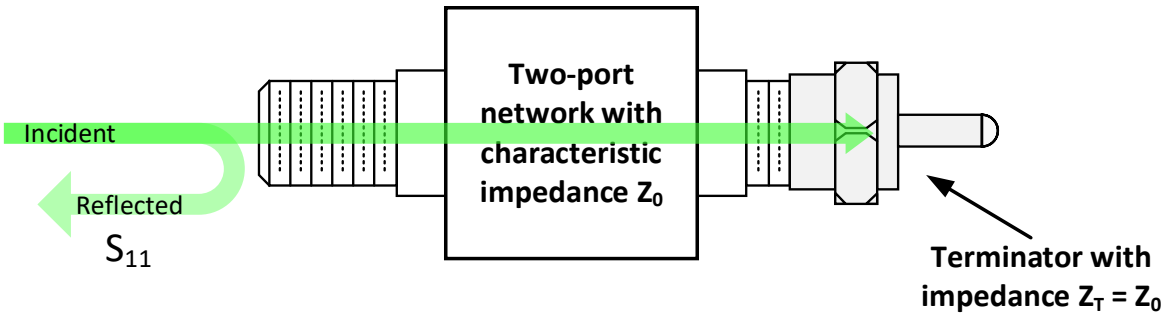


Figure 28 - It is common when measuring S_{11} to terminate Port 2 in an impedance equal to Z_0.

Next, apply a test signal to Port 2, and measure at Port 1. This gives us the S-parameter S_{12}, the reverse voltage gain or transmission coefficient. See Figure 6.

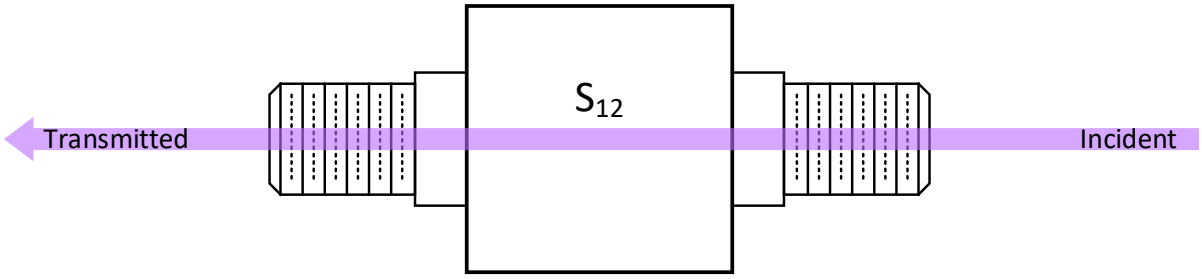


Figure 29 - Testing the network from Port 2 to Port 1 gives the S-parameter S_{12}.

Here, too, a reflection measurement can be made at Port 2, as shown in Figure 7. This gives us the S-parameter S_{22}, or Port 2's voltage reflection coefficient.

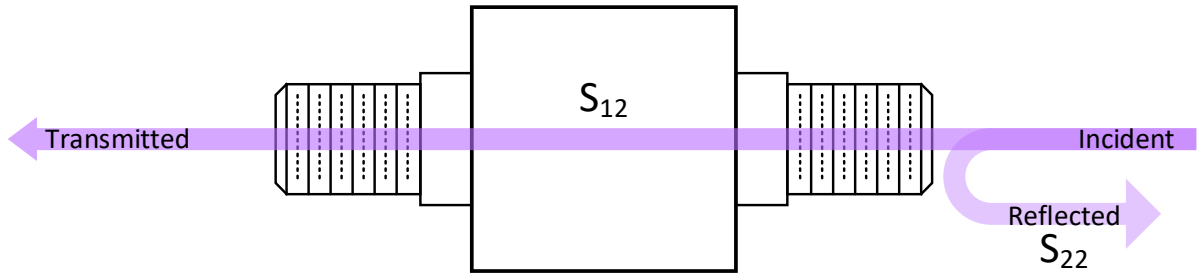


Figure 30 - Measuring the reflection from Port 2 gives the S-parameter S_{22}. In many cases Port 1 would be terminated in an impedance equal to Z_0 while performing this measurement (not shown).

In an operational FDX DOCSIS network, the FDX node's echo cancellation circuitry characterizes the network from the perspective of S_{11}, similar to the example shown in Figure 8.

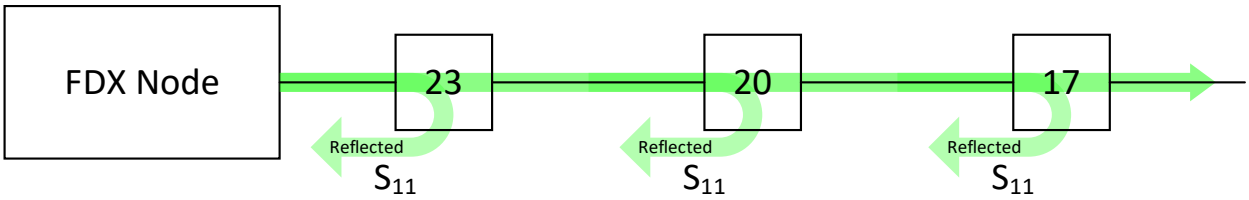


Figure 31 - FDX DOCSIS node S_{11} characterization of the cable network.

The following table summarizes the S-parameters just discussed.

Table 4 – Two-port S-paramters

Two-Port S-Parameters	
S_{11}	Port 1 (input port) voltage reflection coefficient
S_{12}	Reverse voltage gain or reverse transmission coefficient
S_{21}	Forward voltage gain or forward transmission coefficient
S_{22}	Port 2 (output port) voltage reflection coefficient

For S_{11} and S_{22}, the voltage reflection coefficient is the ratio of reflected voltage $E_{reflected}$ to $E_{incident}$. For S_{12} and S_{21}, the transmission coefficient is the ratio of transmitted voltage $E_{transmitted}$ to incident voltage $E_{incident}$.

Figure 9 shows the relationships of $E_{incident}$, $E_{reflected}$, and $E_{transmitted}$ when measuring from Port 1 to Port 2. Here, $E_{transmitted}/E_{incident} = S_{21}$ and $E_{reflected}/E_{incident} = S_{11}$.

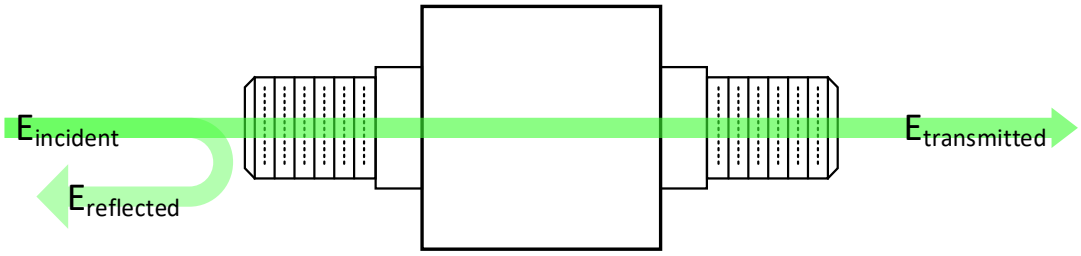


Figure 32 - Relationships of $E_{incident}$, $E_{transmitted}$, and $E_{reflected}$ when measuring from Port 1 (input port) to Port 2 (output port).

Going the other direction, from Port 2 to Port 1, $E_{transmitted}/E_{reflected} = S_{12}$ and $E_{reflected}/E_{incident} = S_{22}$. See Figure 10.

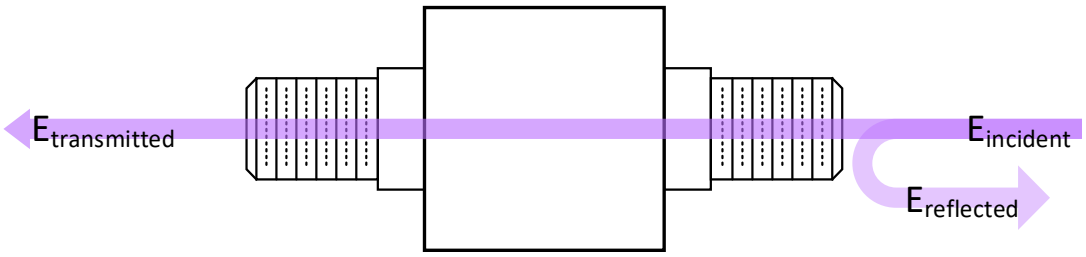


Figure 33 - Relationships of $E_{incident}$, $E_{transmitted}$, and $E_{reflected}$ when measuring from Port 2 (output port) to Port 1 (input port).

Cable operators are usually more familiar with characteristics such as gain, insertion loss, and return loss. Here are some common relationships derived from S-parameters (referenced to a two-port network, and assuming Port 1 is the input port and Port 2 is the output port).

Input return loss in decibels
$R_{in} = -20\log_{10}|S_{11}|$

Output return loss in decibels
$R_{out} = -20\log_{10}|S_{22}|$

Gain in decibels
$G_{dB} = 20\log_{10}|S_{21}|$

Insertion loss in decibels
$L_{dB} = -20\log_{10}|S_{21}|$

Abbreviations

CAD	computer aided design
CTA	Consumer Technology Association
CW	continuous wave
dB	decibel
dBmV	decibel millivolt
DOCSIS	Data-Over-Cable Service Interface Specifications
DSP	digital signal processing
ES	extended spectrum
FDD	frequency division duplex
FDX	full duplex
FFT	fast Fourier transformation
GHz	gigahertz
HFC	hybrid fiber-coax
HH	household
ID	identifier
IFFT	inverse fast Fourier transformation
IG	interference group
ISBE	International Society of Broadband Experts
MAC	media access control

MER	modulation error ratio
MHz	megahertz
NCP	node combining plan
NVP	nominal velocity of propagation
OFDM	orthogonal frequency division multiplex
OOB	out of band
PDF	portable document format
PHY	physical
PNM	proactive network maintenance
PSD	power spectral density
QAM	quadrature amplitude modulation
RBW	resolution bandwidth
RL	return loss
RRC	root raised cosine
SCTE	Society of Cable Telecommunications Engineers
SNR	signal-to-noise ratio
UML	unified markup language
VNA	vector network analyzer
VP	velocity of propagation

Bibliography & References

- A Closer Look at S-Parameters, Ron Hranac, 2019
- Full Duplex DOCSIS PHY Layer Design and Analysis for the Fiber Deep Architecture, Richard S. Prodan, Ph.D.
- PNM Best Practices Primer: HFC Networks (DOCSIS® 3.1) CM-GL-PNM-3.1-V01-200506
- Transmission Line, Wikipedia: https://en.wikipedia.org/wiki/Transmission_line
- Full Band Capture Revisited, Ron Hranac et al, SCTE Expo 2019

Why 6 GHz Standard Power Wi-Fi is the Game Changer for Residential Use in the US

A Technical Paper prepared for SCTE by

J.R. Flesch
Director, Advanced Technology, Home Networks
Commscope
3871 Lakefield Drive, Suwanee, GA 30024
jr.flesch@commscope.com

Charles Cheevers
CTO, Home Networks
Commscope
3871 Lakefield Drive, Suwanee, GA 30024
Charles.cheevers@commscope.com

Kurt Lumbatis, Commscope

Bryan Pavlich, Commscope

1. Introduction

We have an exemplary beachhead for better in-home Wi-Fi coverage with the instantiation of a low power indoor (LPI) effective isotropic radiating power (EIRP) spec (5 dBm/MHz power spectral density (PSD)) for access points (APs) by the FCC which promises a Gbps+ link budget for in-home data distribution (and scavenging) over 6E wireless networks. But the FCC also opened the door to standard power in that band (up to 36 dBm EIRP) for indoor use with the adoption of a spectrum coexistence scheme it refers to as automated frequency coordination (AFC). Using AFC and the higher power, we can now consider trunked indoor links which meet (or better) a 2.5 Gbps PHY for even large floorplan homes and perhaps enable mixed-power and mixed-band in-home mesh architectures. This enthusiasm is tempered somewhat by the observation that battery-powered clients are currently restricted by silicon to 20 dBm footprints at 6 GHz and as such, may determine extender density in the home for particular service mounts. However, alternating current (AC)-powered clients capable of power upticks (above LPI even if shy of standard power) – which include set top boxes (STBs), video streamers and gaming hardware – imply that we should be able to exploit link modulation and coding scheme (MCS) to fairly dense quadrature amplitude modulation (QAM) spectrum efficiencies for streaming and gaming services to these fixed endpoint clients – especially those of a heavy downstream data delivery bias.

This paper will examine the opportunities to be found in a standard power Wi-Fi 6E regime, discuss AFC implications (for cloud portal and AP endpoints) and suggest possible in-home architectural leverages of this substantial uptick in AP EIRP.

2. The LPI Reference Point

2.1 Availability of FCC-compliant Test Devices

6E consumer premise equipment (CPE) has begun to enter the market and it is obviously relevant to gauge the impact of LPI's allowable power in real scenarios which feature appropriate spatial stream (SS) scaling alongside the 6 dB client "power penalty" between client devices and APs -- and against path losses normally associated with various endpoint locations in a common floorplan. As in previous years, the 5300 square foot Commscope/Arris Wi-Fi house affords us the opportunity to test wireless links which emulate endpoint placements in homes up to that footprint in size and explore rate/reach and latency performance against these placements.

An available 6E-compliant smartphone was tested against a multi-band 4x4 AP to evaluate rate/reach throughout the Wi-Fi house. The results of these tests are shown below.

2.2 Early 6E Device Performance Data

Measurements were conducted early last year for nascent 6E devices (4x4 AP and 2x2 smartphone). These yielded the following measurements in the house at close operating range:

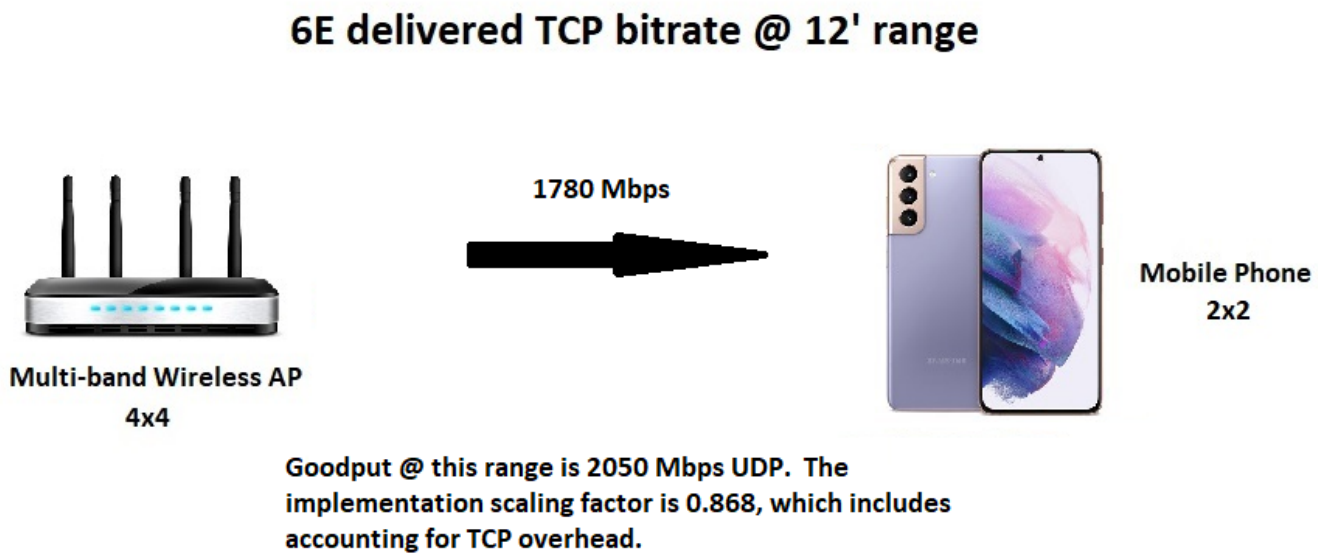


Figure 1 - Measured Close Range 6E LPI TCP Performance Between the AP and a Mobile Client

The testing was conducted across an open room at a radius of 12 feet or less (to establish a low path loss but far-field result which was expected to produce maximum MCS). The mobile was operating at backed off EIRP, relative to the AP, as mandated by the FCC. Free-space path loss (FSPL) at mid-band 6 GHz and that distance amounts to 60 dB (~49 of which gets you 3 feet off the antennae). Referring to the bitrate waterfall curve (goodput rate versus path loss) for 4x4 <-> 2x2 (albeit at equivalent powers, AP and client), we see that the link should support a goodput (bitrate accounting for Wi-Fi framing) of just over 2000 Mbps, with a TCP accounting moderating that to around 1800 Mbps or so (using 10% TCP overhead as a reference). Note that battery-powered mobile clients are likely limited even more -- to a maximum of 20 dBm EIRP -- which pushes the connection penalty to at least -7 dB relative to the AP. The operating points are captured on the waterfall curve for Wi-Fi 6 bitrates at LPI for 2x2 clients, leveraging the maximum 160 MHz channel bandwidth (note that the abscissa of the graph shows path loss off the antenna – which means you are actually down 49 dB at the leftmost point):

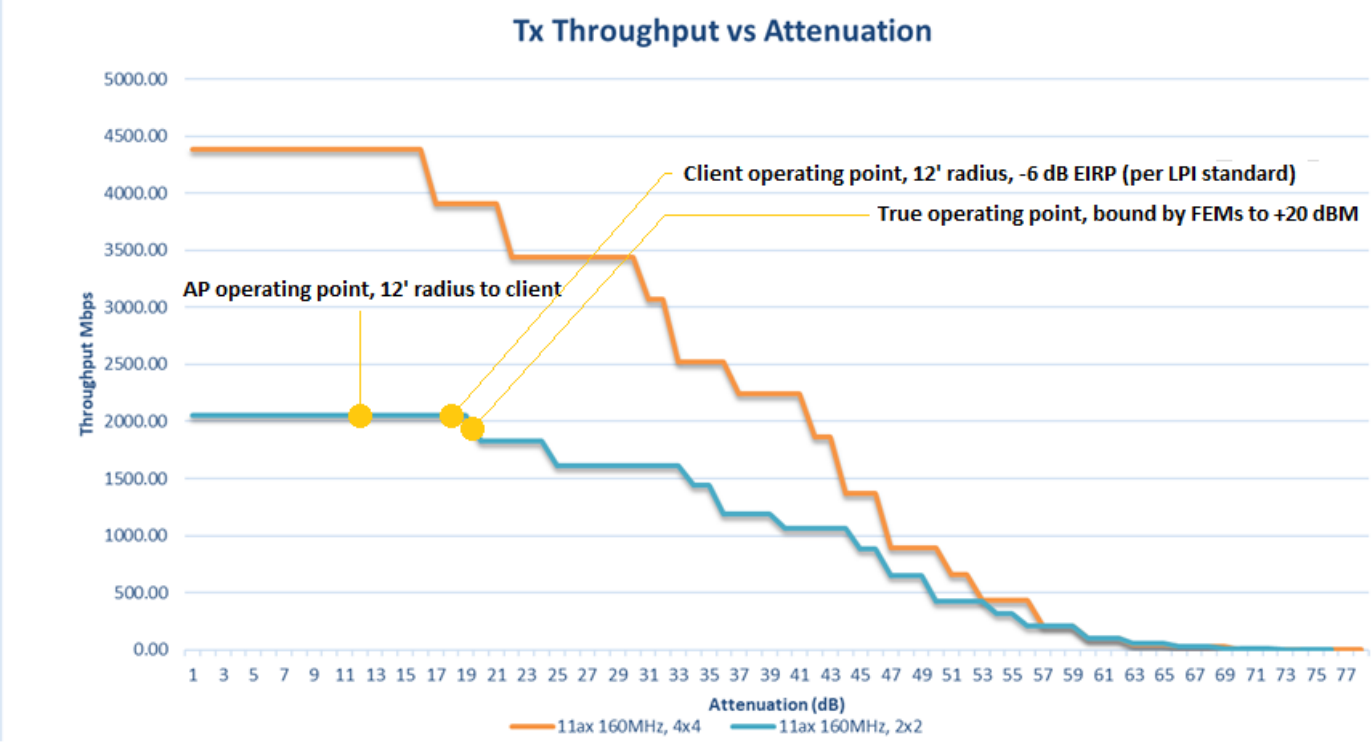


Figure 2 - Differential MCS Operating Points, Downlink and Uplink @ 12'

Note that the client – 6 or 7 dB down in power -- is operating at a mostly equivalent link budget to the AP at this close range, with a max compromise of one lower MCS step. This is not unusual behavior – and clients are almost universally the more restrictive on MCS setting than APs (certainly guaranteed when their transmit footprint is managed 6-7 dB below the AP).

2.3 Whole Home Performance

To get a feel for LPI coverage from an optimum AP location in the Wi-Fi house, the multi-band wireless router was placed in the central living room on the main floor. The mobile smartphone client was then moved room-to-room to measure data exchange rates and paint a coverage map for the house. Per housing floor, the TCP rates are shown below:

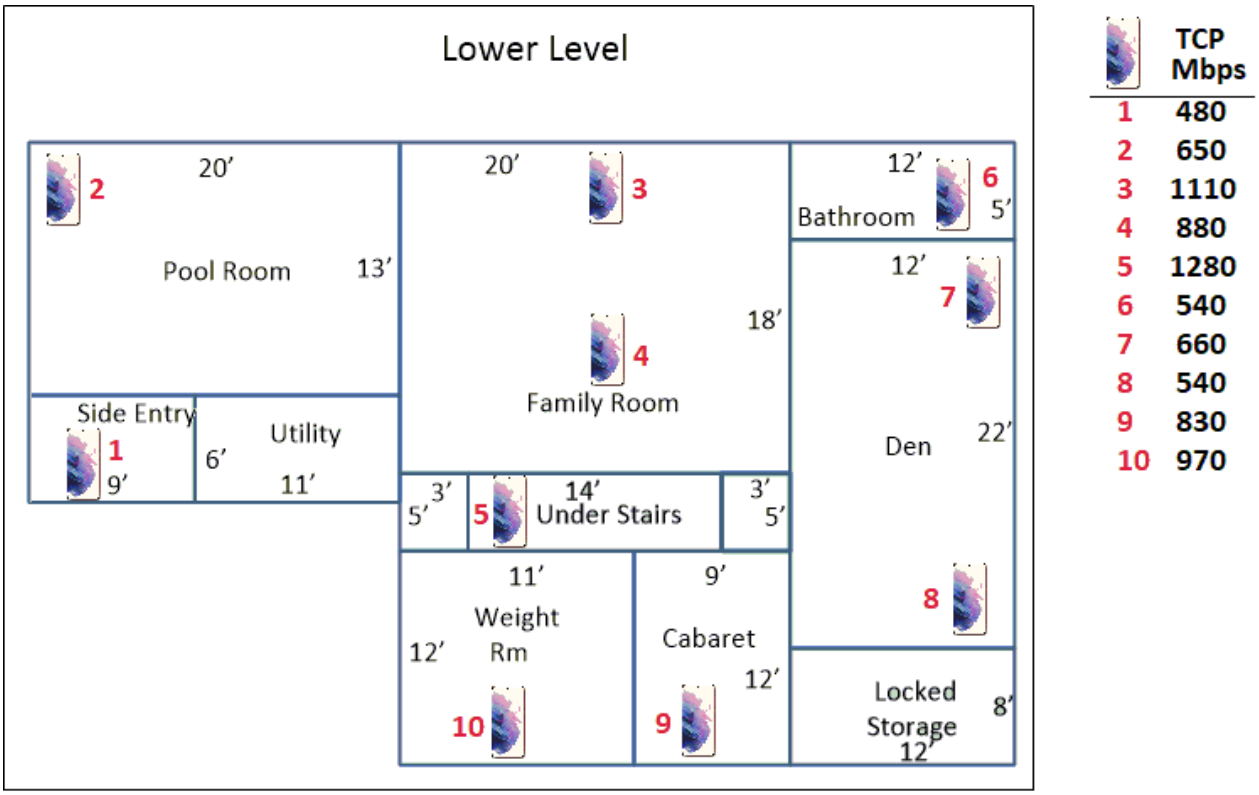
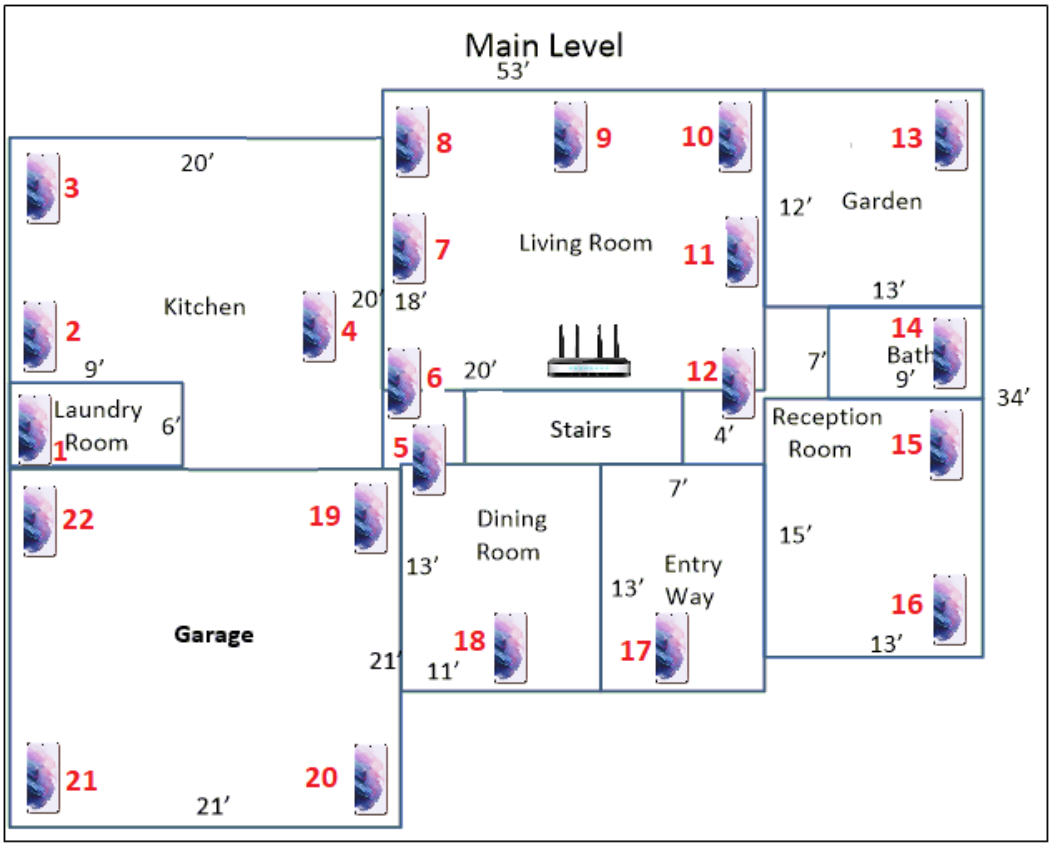


Figure 3 - Wi-Fi Hous Lower Level 2x2 Client Performance



	TCP Mbps
1	1310
2	1140
3	1330
4	1470
5	1450
6	1480
7	1780
8	1610
9	1750
10	1440
11	1550
12	1600
13	1360
14	1450
15	680
16	780
17	1140
18	1320
19	1300
20	680
21	850
22	870

Figure 4 - Wi-Fi House Main Level 2x2 Client Performance

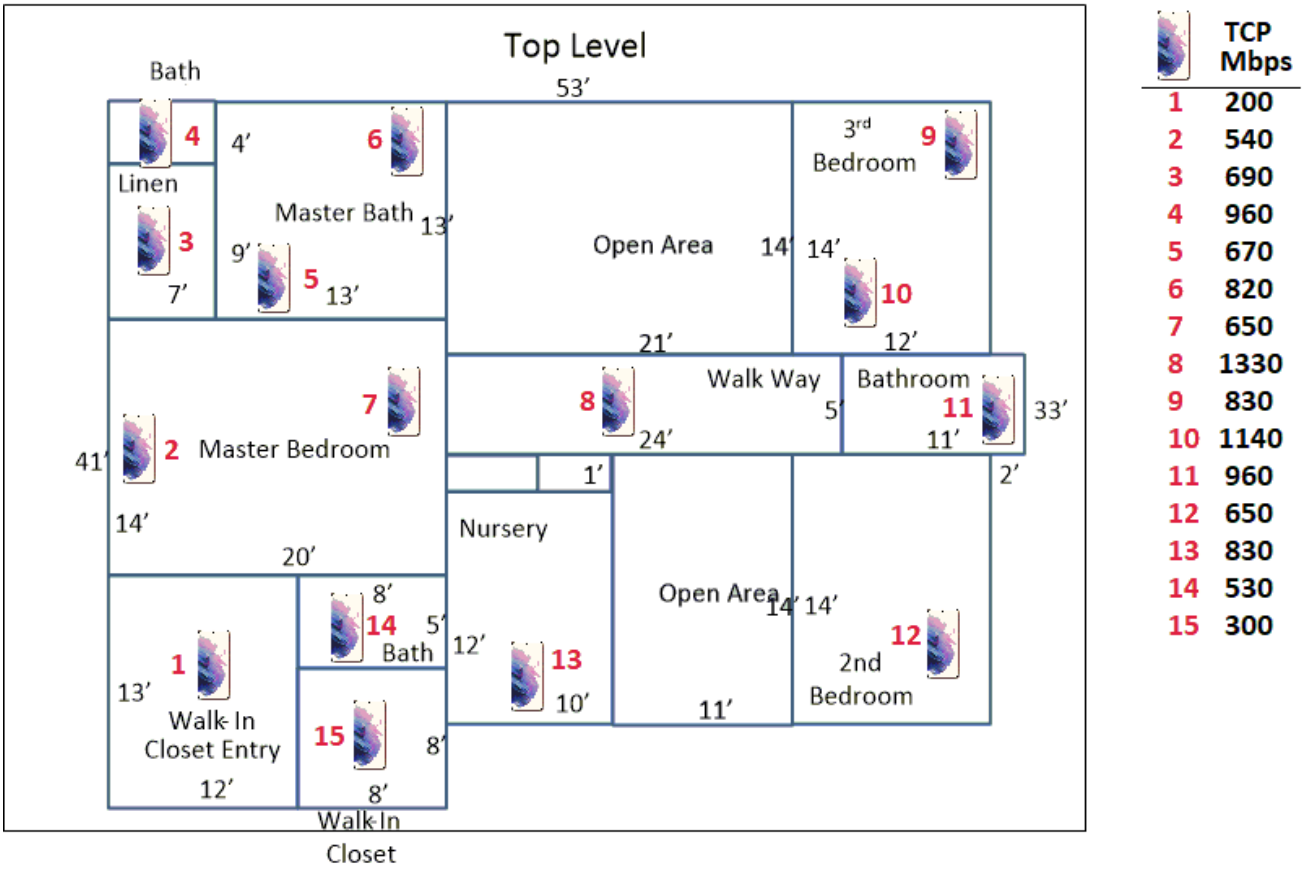


Figure 5 - Wi-Fi House Upper Level 2x2 Client Performance

These collections of data yield the following “heat maps” of the LPI-backed off mobile smartphone around the whole of the Wi-Fi house:

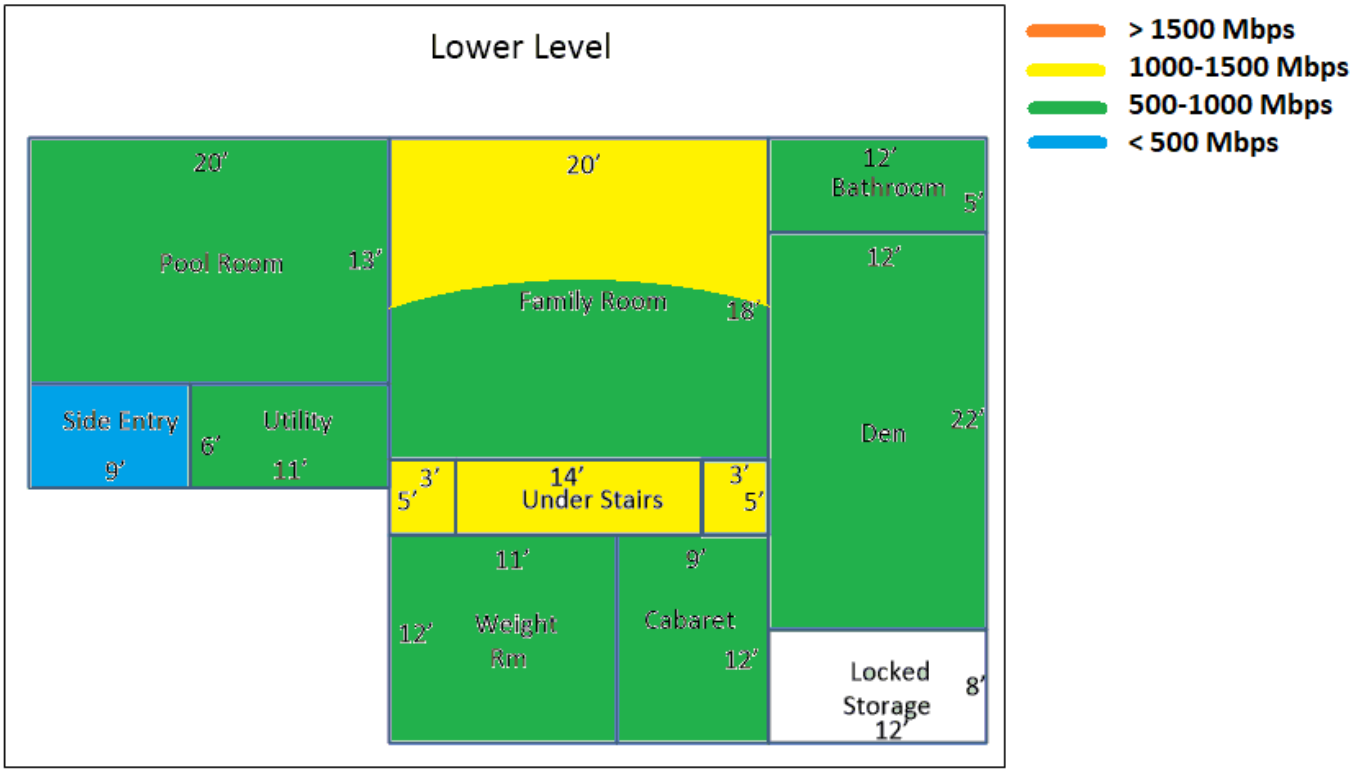


Figure 6 - Lower Level Heat Map

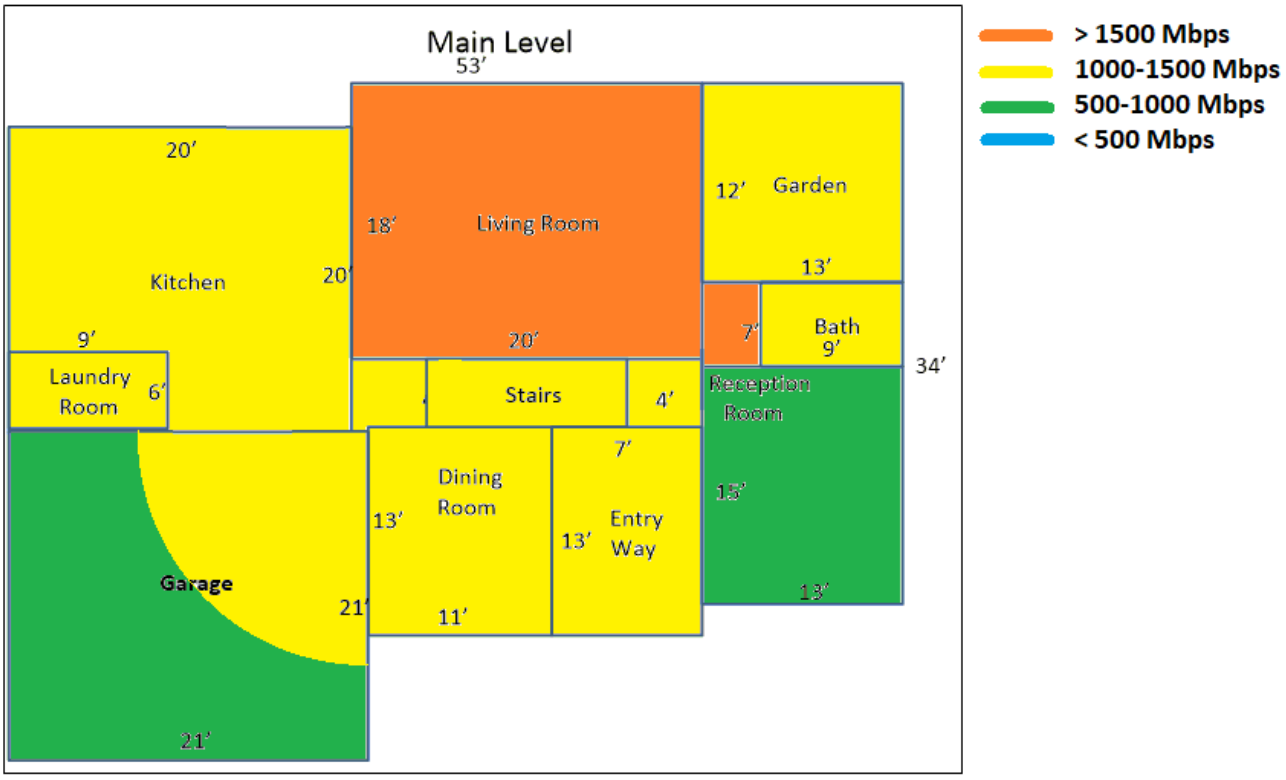


Figure 7 - Main Level Heat Map

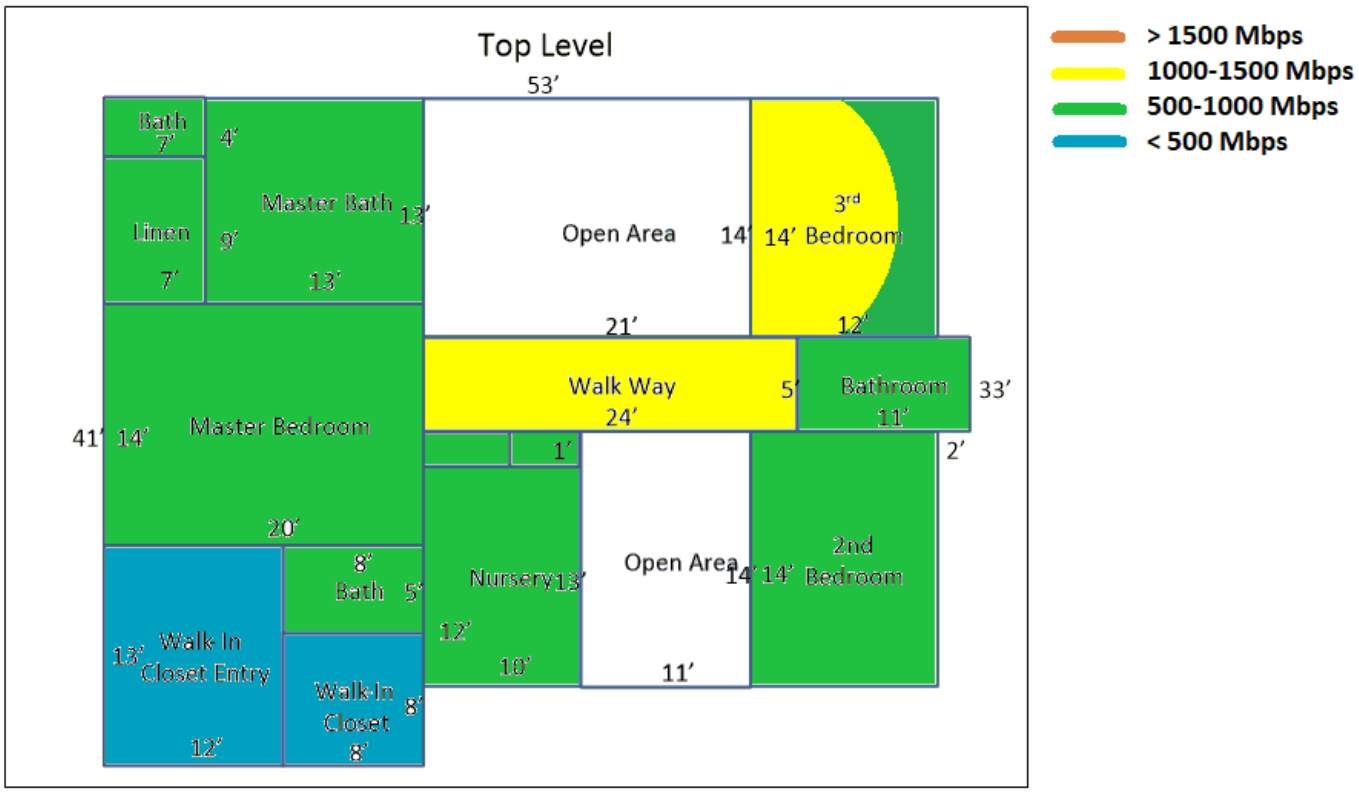


Figure 8 - Top Level Heat Map

Note that, on the lower level, there was a service null directly below the AP – likely a result of a compromised antenna gain pattern along the central vertical axis of that device. In any event, armed with these results, we can lay out expectations for coverage if we bump the AP power up to an AFC-approved 4W.

3. The Standard Power Prospectus

The fundamental tenet of higher EIRP is the notion that over identical path geometries, more transmit power means higher received power, better carrier-to-noise ratio (C/N) and better spectral efficiency (higher link MCS) which then sustains higher bitrate coverage over a greater service area.

3.1 Implications of the Power Gain

The bump from 27 dBm at 160 MHz BW in the 6 GHz band (LPI EIRP) up to 36 dBm is obviously significant. To put it in perspective, this implies a link linear throw increase of 2.8x to sustain at least the same MCS provided by LPI at a given service radius from client to selected AP or a service footprint up to almost 8x the coverage achieved at LPI for a selected reference bitrate – *provided the client can track the AP EIRP by no less than -6 dB*. (These are best-case numbers, essentially just FSPL without consideration for drywall and flooring – which, per transition, add lumped losses to the FSPL for a given radius). On the face of it, such an uptick suggests you can service a 10,000 square foot, multi-floor mansion from a single AP or gateway and achieve the same bitrate performance you would obtain throughout a 1500 square foot bungalow at LPI levels.

A couple of qualifying footnotes need to be appended at this juncture, however, to dampen and rationalize the more enthusiastic of these expectations.

3.2 Signal Propagation Considerations

The principal point of allowing standard power indoors (with an AFC system) is an acknowledgment that interior wall and flooring transitions will serve to attenuate and scatter 6 GHz Wi-Fi energy on its outbound journey. Further, depending upon exterior construction, building entry loss (BEL) anywhere from 6 to 30+ dB will constrain the radiation footprint on its inside-to-outside propagation. In prior testing at the Arris Wi-Fi house, multiple path tests indicated that, at mid-band 6 GHz, 4.5 dB drywall losses and 9 dB flooring losses could be inferred. (This will obviously vary somewhat with construction differences in other homes but seems a reasonable anchor in calculating available power at a given client location throughout the house, given the AP's own position and orientation.)

As a point of reference, it should be noted that these line-of-sight (LOS)-based calculations start to lose their predictive capability once you transition perhaps 3 walls/floors and get more than 40 feet from the AP – multipath effects not being accounted for (and one accumulates new signal energy vectors based on the dielectric thickness and size of every medium being traversed – as well as the sheer number of these transitions.) The best analysis involves a ray-tracing-based simulator (similar to the overlaid lighting effects summed in every computer-generated imagery (CGI) frame of a movie) but these can be expensive; lumped propagation channel models for in-home projection exist but require parameter tuning (related to the transition geometries associated with the main propagation path and the relative orientations of transmitter and receiver). All of the predictive work here is based on simple LOS analysis (to produce the most conservative link throw expectation); the bookmark is to just be aware of the sources of error in predicting more far-flung receive client behavior. Actual results at the range extrema would be better than our predictions.

Now, the setting of a goodput asymptote for a bitrate expectation versus path loss can overestimate the delivered signal footprint from AP to client if the spatial stream counts on both sides are equal. This is due to less-than-perfect antenna orthogonality and the effects of polarization mismatch (due to device orientation or path effects) between transmitter antenna farm and receiver antenna farm. In the cases where a 4x4 AP is lighting up various 2x2 clients, these diversity losses are washed out by what amounts to spatial oversampling available to the link by virtue of the spatial stream mismatch. However, this also means that the goodput expectations will scale to the least common denominator (2x2 in this case) – though the rate achieved will be closer to expectations in this scenario than for the case of a 4x4 device linking with another 4x4 device. (And all of this assumes fixed radiation patterns – i.e., no presumption of any type of antenna beam steering capability).

3.3 Qualifications for Standard Power Benefits

Using both the recorded LPI device performance at various in-home waypoints along with predictions for the same, and couple this with what amounts to a vernier or gain factor between expectation and measured performance at an open-look short link path, we can then begin to predict performance at standard power levels. Referring to our prior LPI measurements and projections at a small path loss in the Wi-Fi house yields the following observations: 1) Client spatial stream specification will define the rate multiplier (and hence, maximum asymptote) for delivered goodput bitrate performance; 2) Battery powered clients – with the constrained EIRP set by frontend module (FEM) thermal considerations (driving a more modest bias rail in pursuit of lower operating temperatures) – will not be able to track the increased output level of a standard power AP and hence these links will be compromised at large service radii as clients disconnect due to Wi-Fi framing loss; 3) AC-powered clients will benefit greatly from the

power uptick, if they can attain a 30 dBm EIRP (24 dBm conducted with a maximum antenna gain of 6 dBi – which amounts to a much less demanding conducted power budget than that provided by the typical -1 to 2 dBi gains of small form factor client devices). Put another way, if portable client electronics can achieve a 20 dBm footprint with a 0 dBi antenna gain factor, then fixed AC clients can bridge 6 of the 10 dB power gap from tracking standard power client to portable levels (30 dBm down to 20 dBm) and establish the necessary standard power with only antenna changes and 4 dB worth of additional conducted power.

It is worthwhile to also note that these range considerations anticipate full leverage of the 160 MHz BW available at 6 GHz; the client connect limitations become dramatically more concerning if narrower BWs are employed on the links (leading to reduced client EIRPs – potentially down to 12 dBm if the channel operates at 20 MHz BW).

These observations produce an interesting set of architectural permutations of in-home Wi-Fi networks, as we shall see. Standard power goodput bitrate achievable for both 4x4 and 2x2 links is captured immediately below, along with operating points for the various uplink and downlink power budgets:

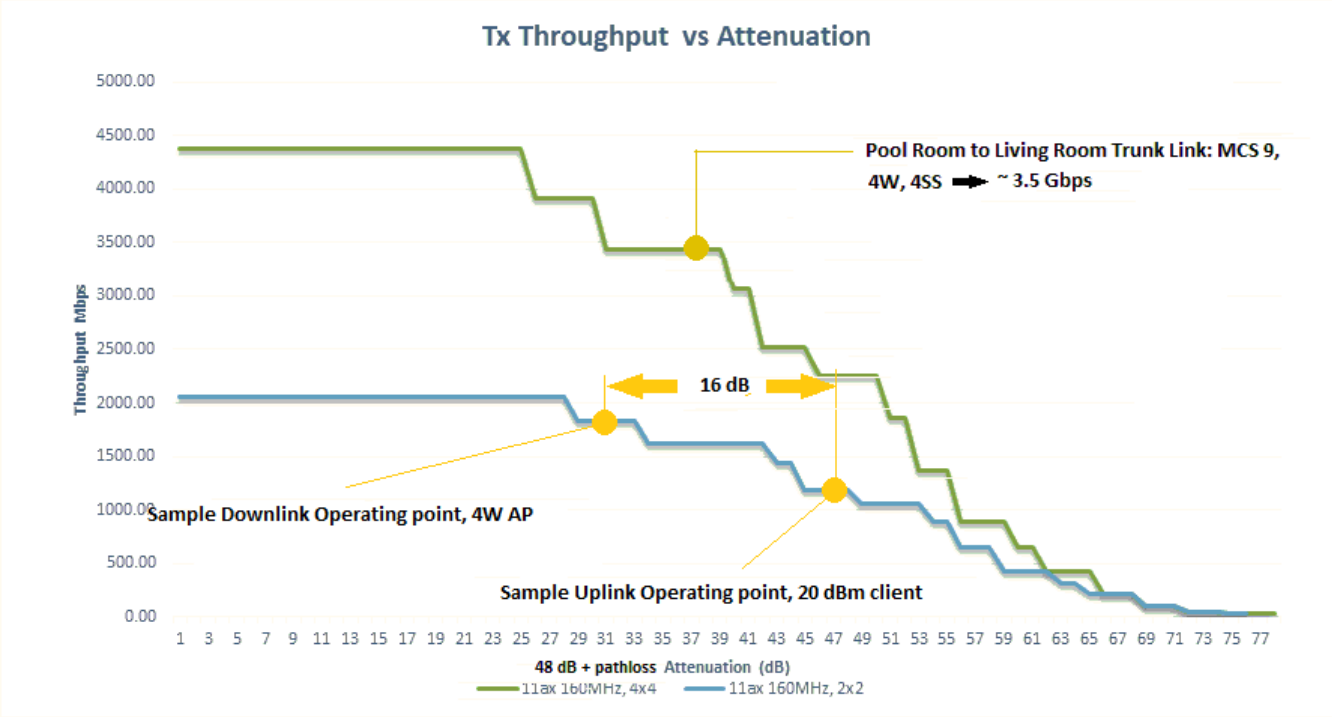


Figure 9 - Bitrate Waterfall vs Path Loss for 4W AP

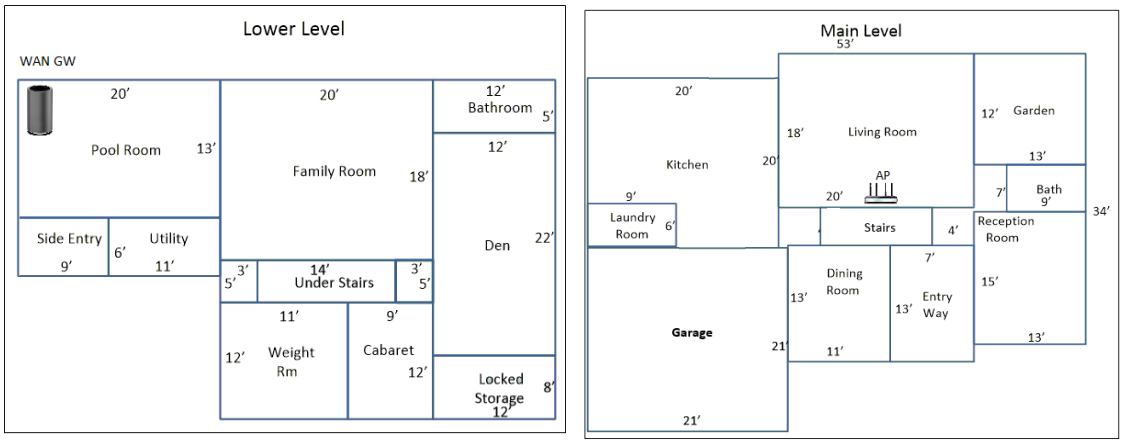


Figure 10 - Proposed Trunk Link Endpoints, Lower and Main Levels

As before, the ordinate axis assumes an antenna-launch loss of around 49 dB as a starting point and the values listed are essentially distance, in pathloss dB, from the antenna(e) to the client. Fundamentally, the move to higher power represents a push (reach expansion) on the pathloss axis by 9 dB:

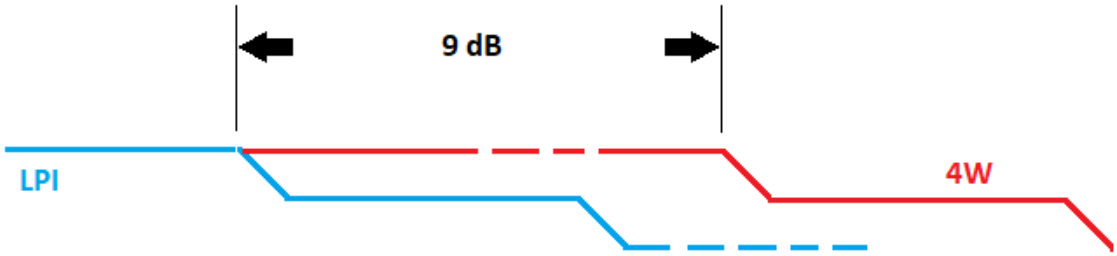


Figure 11 - Effect of Increasing EIRP on Delivered Bitrate

The constrained radiation footprint of portable, battery-powered client devices puts a damper on the more inflated of the coverage aspirations one would otherwise assign to a 4W EIRP, servicing AP, then. But the more robust (4x4) trunk performance and ability to reach more 1W-capable (largely fixed location) clients remains. The question arises: at what size in home floorplan would one want to invest in 4W APs, and what strategies on LPI and 4W mix might make sense? These considerations for in-home wireless service mounts are addressed in detail in section 7. But first, the machinery of AFC, its motivation and operational implications need to be understood (if only as background detail).

4. Potential for Interference

The operational "hall pass" granted Wi-Fi exploit of the 6 GHz spectrum at up to 4W EIRP levels comes (rightfully) with accommodation for exploits of that spectrum by existing communications infrastructure. It is in this accounting and policing of the asset by the FCC that we can arrive at fair use of 6 GHz, so it is worthwhile examining the known fixed wireless access (FWA) actors in the space.

4.1 The Contenders for Common 6 GHz Spectrum

While a reasonable directive for oversubscribed spectrum amounts to "listen before talk", the fact remains that simultaneous access to the same piece of the 6 GHz band for different unlicensed communication systems would be problematic absent some appreciation (and interdiction) of the opportunities for one system's transmissions to compromise the receptions required of an alternate system. The spectral region of concern here involves the 850 MHz worth of U-NII-5 and U-NII-7, where the FCC is now granting Wi-Fi the opportunity to extend its reach by leverage of AP power up to a 4W footprint. The contending system cases can be shown as follows:

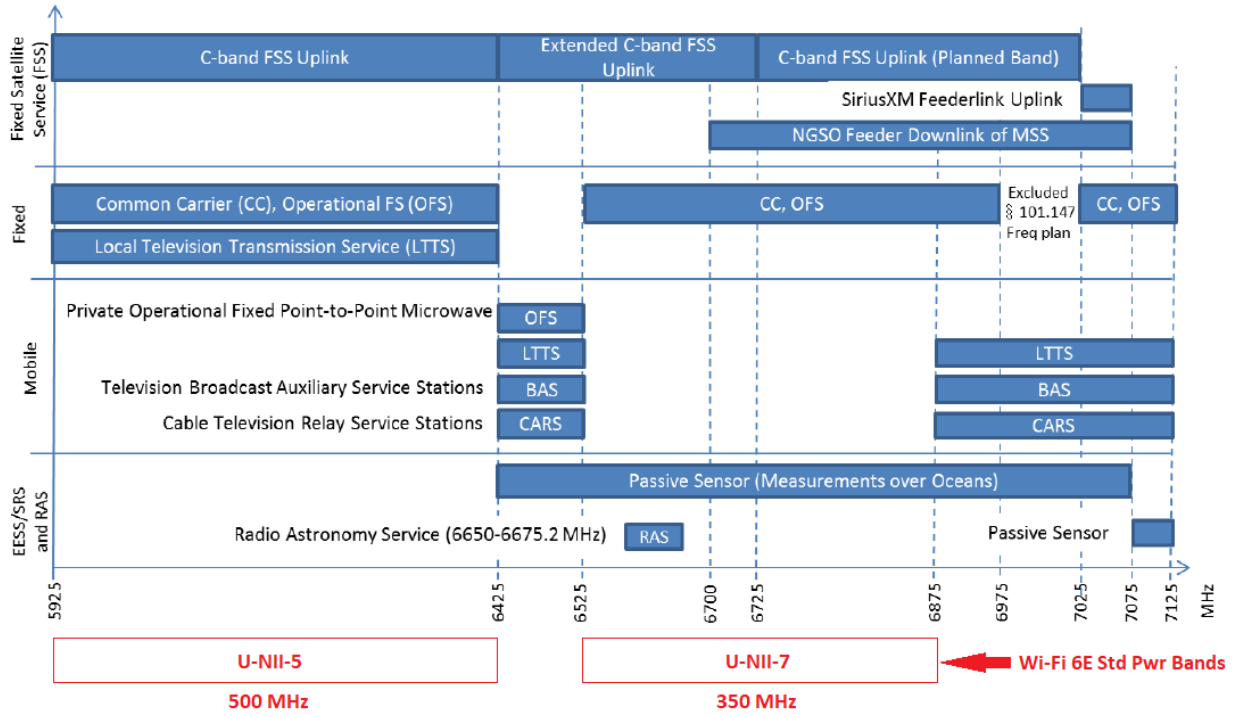


Figure 12 - Legacy FWA 6 GHz Spectrum Use

Note that the services described encompass both fixed service (FS) and C-band fixed satellite service (FSS) mounts, along with some mobile point-to-point (P2P) links as might be employed by news services with "on the scene" coverage. Given the predominant population of FS clients (which are also increasing in number, where C-band satellite growth shows a -2%/year reduction) and the fact that the much smaller number of mobile services (MS) have the ability to re-orient antennae to mitigate interference at the

deployment site, we will focus on the impact to FS infrastructure. However, low-horizon satellite positioning (in the norther latitudes) can potentially expose the satellite receivers to rogue, unlicensed ground link propagation paths at low elevation offsets (presenting much less than orthogonal look angles) and so an effort should be made in establishing the relative heft of 6 GHz indoor energy impinging on the geosynchronous equatorial orbits (GEOs) used by C-band satellites (whose uplink band overlays U-NII-5 and -7 as shown above).

4.2 Dismissing Satellite Interference Concerns

In the report *Frequency Sharing for Radio Local Area Networks in the 6 GHz band, revision 3,* prepared by RKF Engineering Solutions for the 6USC, there is an exhaustive simulation assigning 6 GHz devices (as either small cell or Wi-Fi devices) on a per-person basis, with historical representations on device use cases which serve to estimate the concurrent, accumulated radiation from all indoor standard power devices across the continental United States (CONUS) and estimate the interference-to-noise ratio (I/N) as presented to the GEO-parked C-band satellites in use. Fundamentally, duty cycles of use on high power devices, based upon inferred link capacities and data consumption per hour were assigned. The upshot of the protracted chain of assumptions was that ~ 1 billion devices could be involved across CONUS, with the instantaneous overlay of around 400,000 standard power ON cycles. (The report spells out all assumptions used in the simulation and the reader is invited to download the paper – which is available on the internet in the public domain.) For comparative purposes, the report also estimated the impact to satellite receivers of the legacy 6 GHz infrastructure (which obviously is already operating concurrently with C-band satellite without measurable negative findings). A key figure of the study shows the following impact to receivers on the GEO arc from both legacy systems and the expected high indoor use of 6 GHz standard power Wi-Fi:

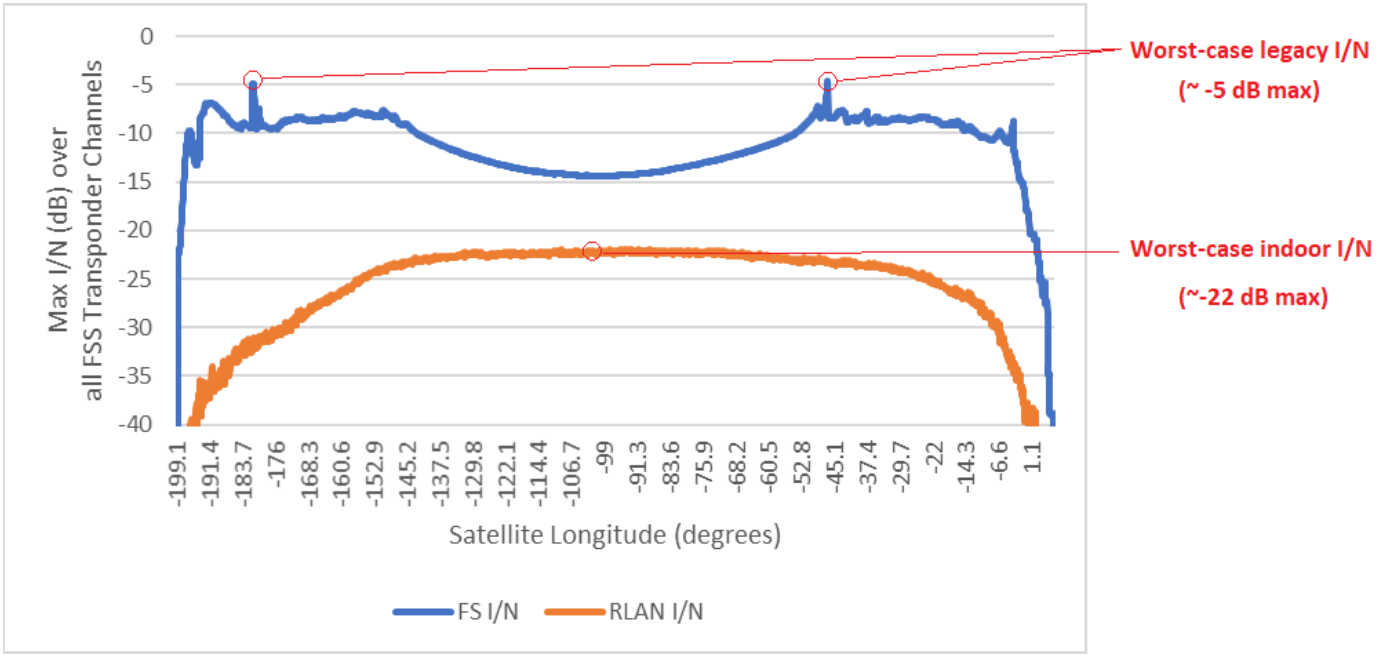


Figure 13 - Relative Obscurity of Potential 6 GHz Indoor Wi-Fi Interference to Satellites on the GEO Arc

The estimated 17 dB lower I/N due to indoor standard power Wi-Fi – versus what the satellites experience already from legacy 6 GHz outdoor links – is a telling graphic (and allows us to concentrate on interference implications to land-based FS links only). This consideration summons AFC exclusion zone calculation around these legacy endpoints and as such, will be addressed in the following section on AFC.

5. Band Management: Automated Frequency Coordination

5.1 The Premise

The intention of AFC to create an FCC-supervised gate into unlicensed indoor exploit of 6 GHz begs elaboration. In order to instantiate this higher performance gate, a specification has been developed which merges several databases, an exclusion zone calculator, the presumption of default LPI behavior by APs seeking higher EIRP, a restriction on indoor-only AP use and a messaging exchange function (between a cloud supervisory portal and those APs). The block diagram manifests as follows:

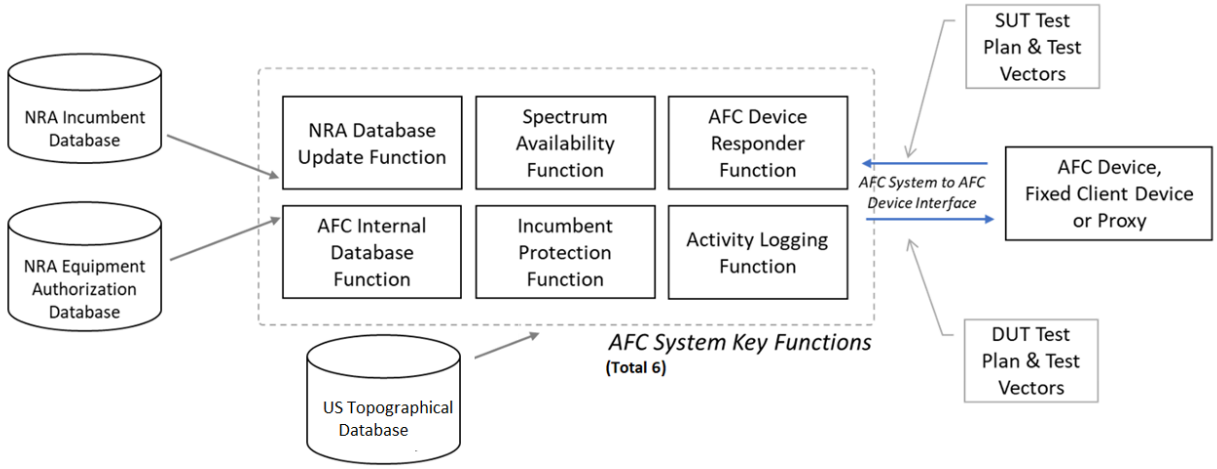


Figure 14 - AFC Block Diagram

As the diagram indicates, there have been six System Key Functions identified for the specification-in-progress, along with three reference databases. The intention of this organization is to support a bilateral data exchange with AP's soliciting the AFC for permission to operate at standard power; by and large, this exchange consists of a query from the AP which provides its latitude, longitude and operating height information to the AFC, which then references its databases for identification of the location of potential co-channel interference (CCI) targets and delivers a mask of available band and power back to the AP. The abbreviated schematic also points the way to certification – system under test (SUT) and device under test (DUT) test vectors at the point of data exchange (interface to the right in the above figure).

Key to all this is the presumption that, unless the appropriate provisioning message (parameterized permission) is received by the AP, its default operating EIRP envelope will be constrained to the LPI-specified 5 dBm/MHz PSD. Furthermore, the permission assigns a lease timeout period after which even approved standard power devices are required to revert to an LPI emissions envelope. (This can obviously be forestalled by periodic queries to the cloud AFC, which would then serve as renewals of the

standard power lease). The buried requirement here is that the AP must be able to access the AFC portal at least once per day (the currently mandated expiration timeout) – and must repeatedly guarantee its lease renewal in order to maintain permission to operate at a standard power envelope.

As of this writing, many implementation details for AFC remain unsettled (key among these being the hosting question for the application – which then begs additional work on whether the service is industry-sponsored or a private affair -- the latter, then, invoking appropriate subscription or licensing). As might be expected, multiple stakeholders in the venture are currently deeply involved in finalizing details, since the goal is to have an operational AFC some time in 2022.

5.2 Operational behavior

The following outlines the specific functional parsing associated with the cloud-based AFC (pertinent to block diagram above):

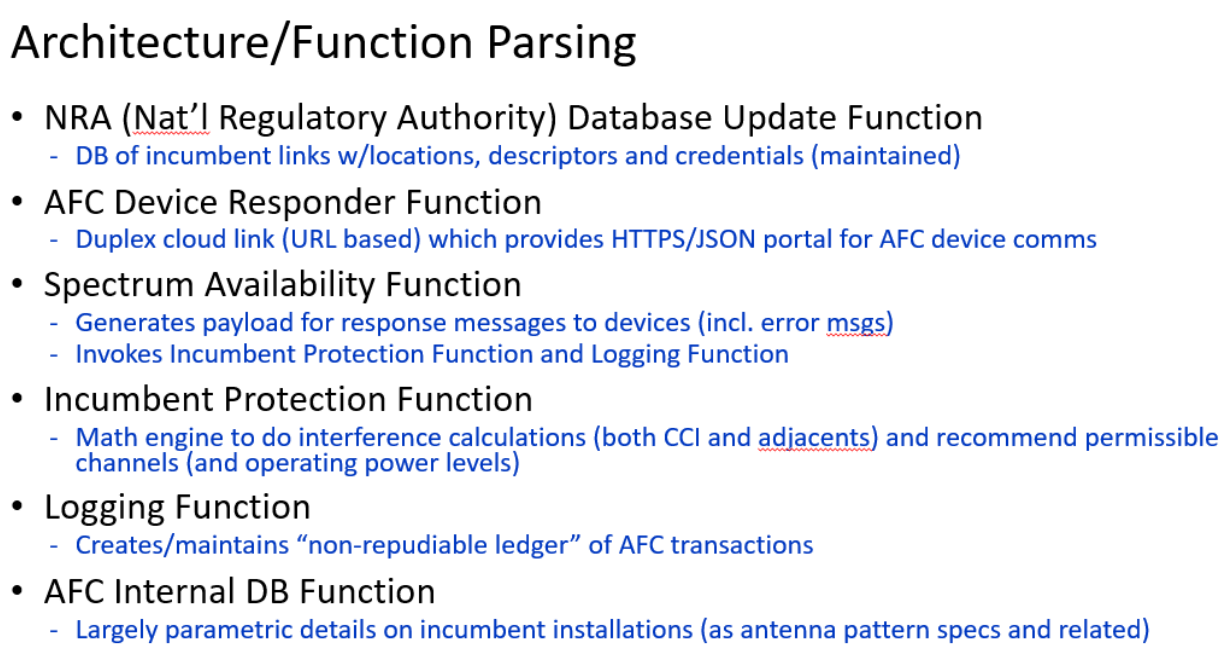


Figure 15 - System Key Function Responsibilities

Message exchange details provided below:

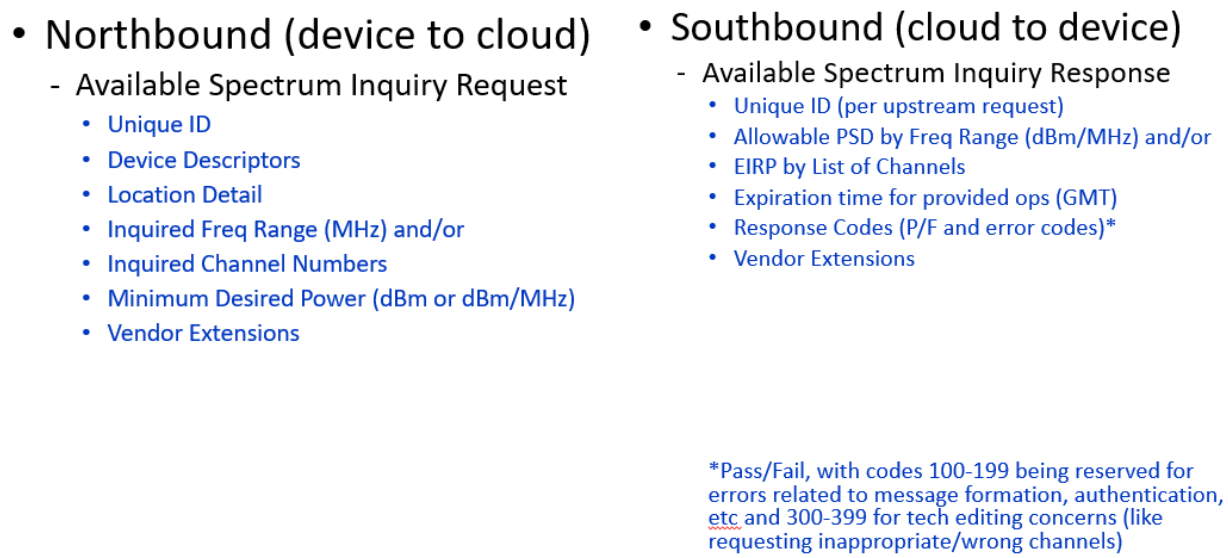


Figure 16 - Messaging Details Between AP and Cloud AFC

Additional messaging clarification is provided by the following:

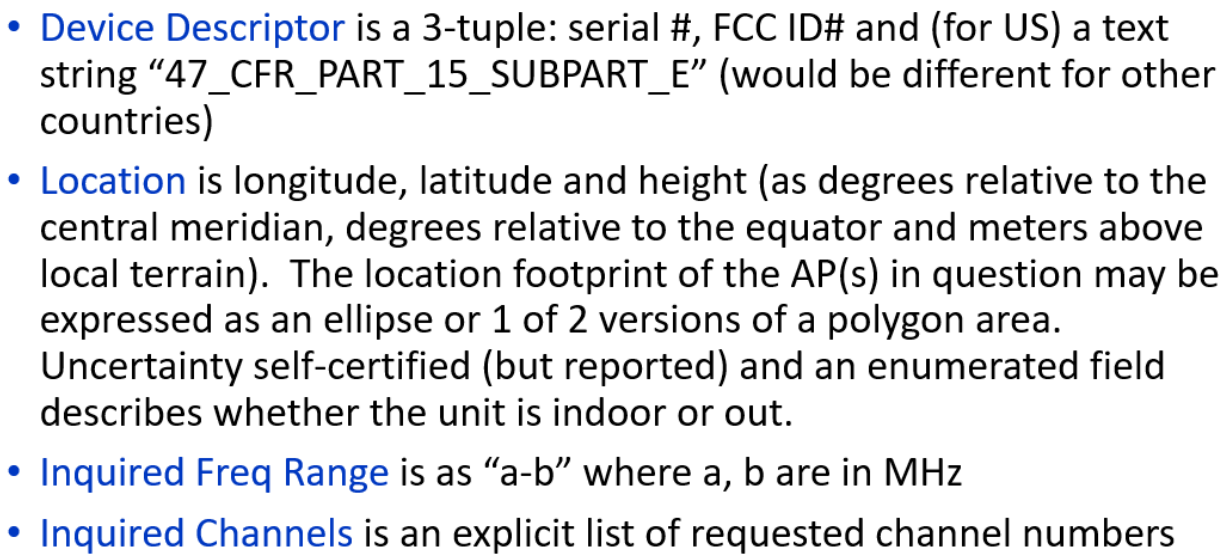


Figure 17 - Messaging Supplemental Definitions

Note that the exchanges listed may be executed by a proxy device (mobile, for example) on behalf of the AP. This would be accomplished with an NFC link between AP and proxy which would facilitate the cloud link connection and grant a pathway for user options and supplemental data entry on device location. The proxy shim would appear as follows:

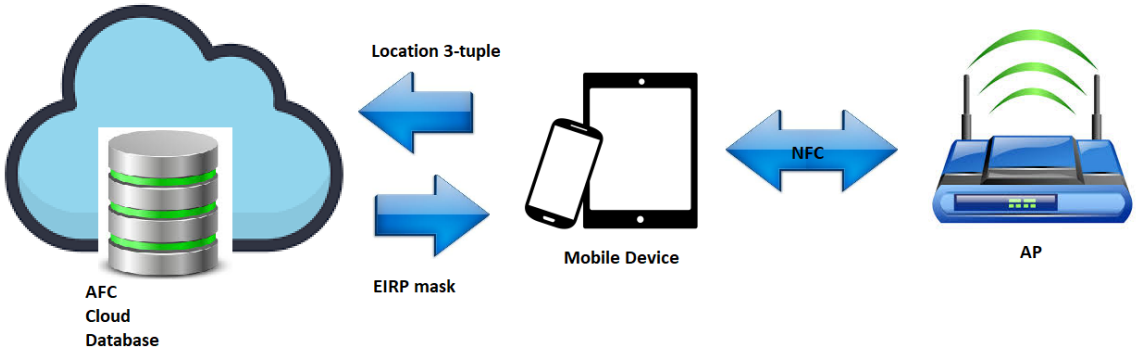


Figure 18 - AP Proxy Arrangement

The motivation for this added proxy lies in leveraging various methodologies to extend GPS position indoors, away from building apertures such as windows and to the immediate proximity of the AP (which, in a buried interior placement, may otherwise be blanked for GPS coverage). It also provides an application hosting environment where additional location data (height in floors or feet) may be inserted via the user (to improve on three-dimensional (3D) AP location uncertainty).

5.3 Nuances and Potential Optimizations

The operational dynamics of the AFC system suggest that, in the interest of speedier resolution of atomic requests from APs for access to standard power operation and a reduction in the query traffic impinging on the AFC from a nationwide population of devices, some amount of "query triage" might be considered. The driving consideration in this is that for a vast number of the requesting APs, lack of proximity to fixed wireless 6 GHz links guarantees deep marginalization of CCI possibilities which might otherwise negatively impact operation of the legacy fixed wireless links. In fact, CCI requires a multilateral conspiracy of overlaid/adjacent channels in the band, aligned FWA antenna aperture (as both azimuth and elevation above ground) and minimal geographical offset of the two contending endpoints in order for concern over potential CCI to be realized. These observations present the opportunity to streamline determination of exclusion zone limitations (reduce calculus overheads) and perhaps extend standard power AP leases past the 24-hour limit presently in the specification (in cases where the potential for interference is effectively and persistently nonexistent). Both efficiency gambits invoke some subtle tradeoff considerations, however.

5.3.1 Radius of Interference

First, as regards the reduction of calculation overhead, is the observation made above that the potential for CCI involves competitive channel access, unfortunate alignment of the FWA antenna with the source of possible interference and geographical proximity of the two potentially interfering endpoints (legacy FWA and Wi-Fi AP). Of these three parameters, inter-endpoint distance is easily the most significant contributor to CCI. It is only when this radius falls below some minimum that additional calculation need be performed to ascertain the possibility of CCI at a level which threatens operation of the FS or FSS link. For example, free-air LOS path loss over the radius involved (assuming optimal coupling of antenna apertures) and a conservative assignment of only a 10 dB penalty for dwelling egress (likely values will fall in a range of 10-30 dB at 6 GHz, depending upon exterior wall construction details and proximity of a

standard – non-e-glass -- window) generate the first cut at the potential for CCI based solely upon proximity of the FS receiver to the Wi-Fi AP.

Some example numbers would be helpful: at 6.5 GHz (mid-band 6 GHz) free-air path loss amounts to 113 dB at one mile. Coupling this with the maximum indoor AP power of 36 dBm and a building entry loss (BEL) of 10 dB suggests the propagated Wi-Fi signal amounts to no more than -83 dBm at a one-mile radius in any direction. The goal, however, is to get to I/N to -6 dB at the nearest FWA antenna (i.e., no closer than 6 dB below the noise floor at a particular BW). To be especially conservative on this gambit, we will neglect the interference moderating effects of clutter, scatter, terrain blanking (available in several standard outdoor propagation models) and narrow antenna aperture (assuming, in this case, no azimuth selectivity in the FWA antenna systems). For a 160 MHz overlaid band (between FS and AP) and a presumed 4 dB NF on the FS receiver, we get a maximum allowable interference footprint of -174 + 10*log(BW) + 4dB receiver NF – 6 dB minimum desired I/N ratio – 33 dB (FS boresight antenna gain for a small dish), or -127 dBm. Translating this to path loss of the AP signal (and accounting for benefitting only 10 dB on the BEL), the distance would have to amount to +36 – (-127) – 10, or 153 dB. At mid-band, this free-space loss comes at 102 miles displacement! Clearly, a single distance-to-FS receiver metric (under these stacked, overly conservative presumptions) does not represent much exclusion zone calculation triage benefit.

Perhaps a more realistic approach would be to find the distance at which a 4W AP presents no worse an interference candidate than an LPI device (which is permitted an indoor EIRP of 27 dBm @ 160 MHz BW *anywhere* in the US) – and then append a keep out area which, at its periphery, represents no greater a threat to FS operations than the energy footprint of an indoor LPI device exiting its enclosure (to a total throw of 300 feet or so) through an especially stout BEL. To set that energy threshold, one can conservatively set the BEL for the building housing the LPI AP to 30 dB (assumes stone exterior construction and e-glass windows), add 300 feet of arbitrary FSPL (as nearest proximity to an FS endpoint) and then use this abated field strength as the target value for a 4W AP to hit (under a less beneficial BEL of 10 dB, just to bake in margin). In round numbers, the target signal strength ends up being 27dBm – 88 dB (FSPL for 300 feet, mid-band) – 30 dB (BEL), or -91 dBm. This implies that one requires the FS spacing to the standard power AP to account for 117 dB worth of path loss (36 dBm – 10dB BEL – (-91 dBm target)). This arrives at a throw of 8500 feet (1.6 mi) at mid-band 6 GHz. This makes for a more manageable first cut.

So now the first triage step is to determine if the AP is further away from the nearest FS receiver by at least 1.6 miles; if this is the case, then the entirety of the U-NII-5 and U-NII-7 bands could be released for use by the petitioning AP – and the concern for interference potential relegated to being no worse (and this, actually by a fair margin) than that supplied by an equivalent indoor LPI device located in a stone home 100 yards from the FS receiver.

5.3.2 Antennae Alignment

Now the notion of antenna selectivity can be introduced. Note that with the 1.6-mile approximation above, all signal loss was put down to FSPL and BEL. With one additional step, APs which locate themselves less than 1.6 miles from a potential FWA site may be cleared to operate at standard power by evaluating whether the combination of FSPL, BEL and antenna pattern misalignment could defuse the possibility of interference. Note that FS endpoints have antenna radiation parameters logged into the Universal Licensing System (ULS, also referred to as national regulatory authority (NRA) in separate documentation) database referenced by the AFC. For the prevalent class A and B antennae, the pattern is

fairly narrow around the boresight direction of the element (for both azimuth and elevation). (In situations where sectorized coverages are combined – as would be the case for point-to-multipoint (P2MP) systems -- it is straightforward to select the most aligned element(s) of the array). An example of the most common apertures deployed in the field is captured below:

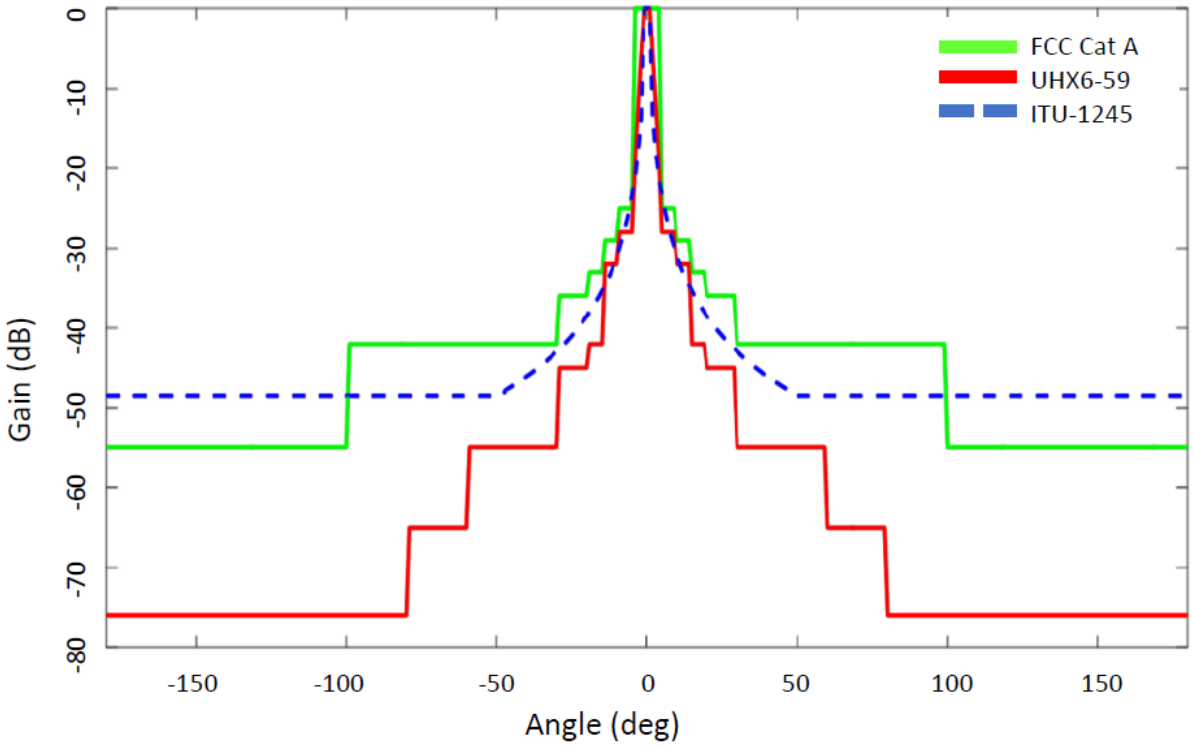


Figure 19 - FS Antenna Apertures (Azimuth Look Angle)

Zoomed in, the aperture looks like this:

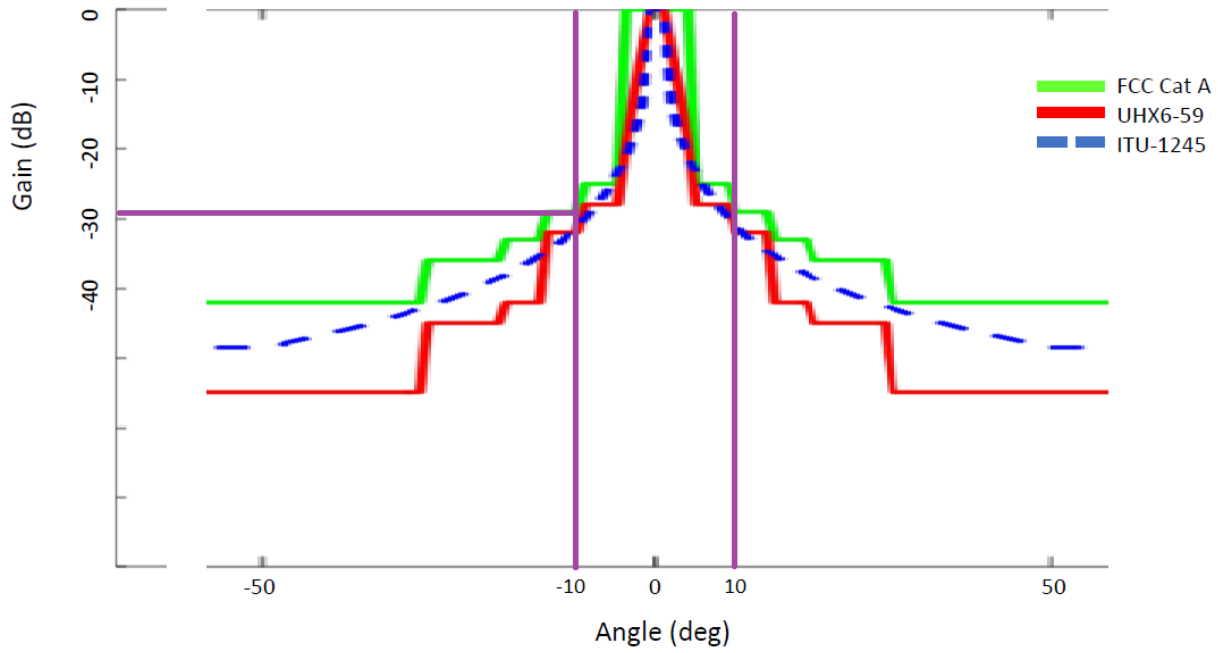


Figure 20 - Antenna aperture zoomed

Note that such performance is representative of 83% of the FS antenna population (per Commsearch research in 2011 – such not being recently re-evaluated, however). Even a cursory examination of the figure reveals that, once off-boresight by as little as +/- 10 degrees, any of the commonly used FWA antenna masks shown exhibit an azimuth selectivity which will produce > 30 dB loss relative to the aligned gain. Normally, one would consider boresight gain as the reference against which, would be applied the selectivity offset loss; however, recall that the distance calculation above applies to any angular displacement of AP-to-FS receiver, including exact boresight. Referencing our prior interference radius calculation, if we can count on an *additional* interfering signal loss due to the FS receive antenna's look angle to the candidate AP, our FSPL budget reduces to 117 – 30, or 87 dB. In mid-band, this would allow Wi-Fi AP's full rein on standard power as long as the AP distance from the FS endpoint is greater than 270 feet and the look angle exceeds 10 degrees on either side of target system antenna bore sighting. But (though in a significant minority) there are other antennae in use by legacy 6 GHz FWA infrastructure which feature a broader aperture and we will consider them, despite a low representation in the total number of target antenna systems, as the least common denominator default in calculating a conservative exposure profile for nearby FS receivers.

A representative sample of a popular (smaller) 0.9m diameter antennae, the Andrew VHLP3-6W, exhibits the following selectivity for horizontal and vertical, single polarizations:

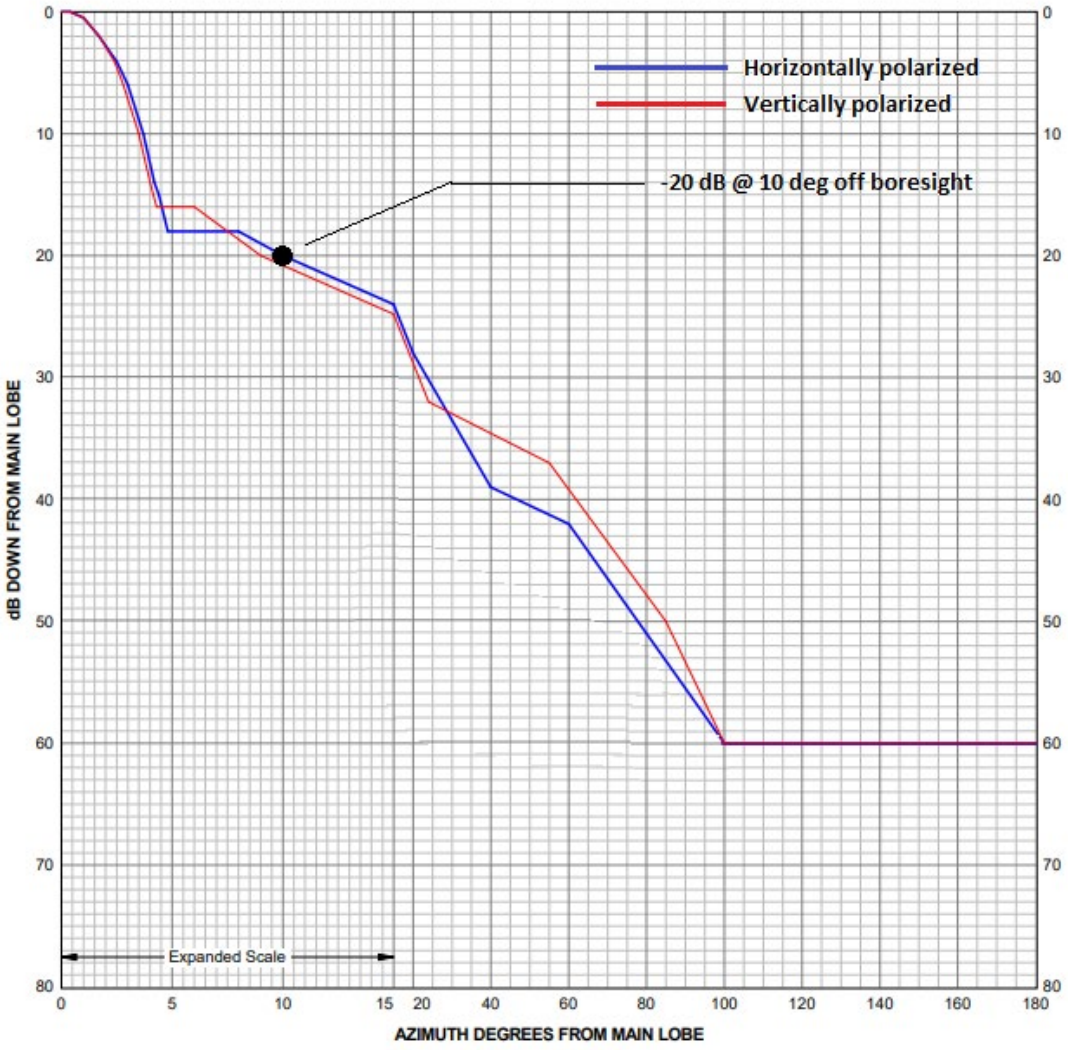


Figure 21 - Representative Small Diameter Antenna Single-Sided Selectivity @ 6 GHz

If we peg look angles > 10 degrees (between AP and FWA antenna) as a reference point, this smaller antenna exhibits worse selectivity than larger units (only ~ -20 dB or better, relative to the main lobe – versus the -30 dB exhibited by the majority of the FS antenna population). Using this as our generic interference possibility, our margin to interference collapses to 117 – 20, or 97 dB of required FSPL. The mid-band proximity associated with this number amounts to 850 feet (0.16 miles). The second "shortcut", then, for releasing the full extent of U-NII-5 and U-NII-7 spectrum for standard power use of APs has the dual qualifications that the distance to FS antenna must be > 0.16 miles *and* be off-boresight by at least 10 degrees.

Outside of these calculation shortcut cases, the AFC system would likely have to perform a full exclusion zone calculation (using distance, look angles, FS receiver link details and related) to evaluate whether Wi-Fi interference is a consideration.

5.3.3 The Argument for Variable Lease Duration

As regards the notion of extended leases, the FCC has decided that 24 hours (with an additional 24 hours grace period in the event of AFC system connection failure) will be the norm. This is based on the mere possibility of the ULS database being amended on 24-hour boundaries. Originally, the 6USC consortium had proposed a 30-day lease, projecting that the potential for interference nationwide would not significantly alter within a one-month timespan. It seems very likely that, in rural and suburban areas, the pace of 6 GHz infrastructure build-out will see a rate of database change very much slower than every 24 hours. It may be worthwhile for the FCC to re-examine standard power lease epochs – if for no other reason than this would greatly down-scale the communications traffic from APs to the AFC system (though, in fairness, this economy of communication does not appear to be an impediment for the system to operate, given the vast number of network paths in place to gather the AP solicitations and the ability to map AFC from the single cloud portal to multiple network edge locations). There certainly seems to be a case for applying some variation in lease lengths, depending upon location parametrics associated with the petitioning APs – especially if a case can be made to the AFC system that the device location is fixed (i.e., not a portable device).

Referring to our section on AFC calculation triage, a great many of the soliciting APs (perhaps as many as 90%), should require no full calculation of interference potential – which suggests longer leases may serve adequately.

5.3.4 Apportioning the Shared Spectrum

The final option we consider here is that, if the probability of the two wireless systems interfering with each other is not easily dismissible based upon proximity and alignment, we can examine the leverage of alternate in-band spectrum (non-overlaid assignment of spectral resources) to resolve the issue (and so grant some amount of leeway to the requesting AP(s), based upon available contingency spectrum). By appending guard bandwidth on either side of the incumbent system's operational link bandwidth, an exclusion portion of the spectrum for the soliciting Wi-Fi AP may be determined. (This would have to be done for every exposed FS node within a 1.6-mile radius -- depending upon boresight offset – each potentially with a differing link spectral profile.) Note that legacy FS links typically feature system BW << 100 MHz and there is 850 MHz of standard power 6 GHz spectrum across the two leverageable U-NII bands which can provide the frequency-division multiplexing (FDM) options desired – so the opportunity for mutual avoidance is quite high in all but the most crowded urban environments. Further, the Wi-Fi APs are already required to meet beyond-band-edge transmit power rolloffs which conform to the following out-of-band (OOB) energy mask:

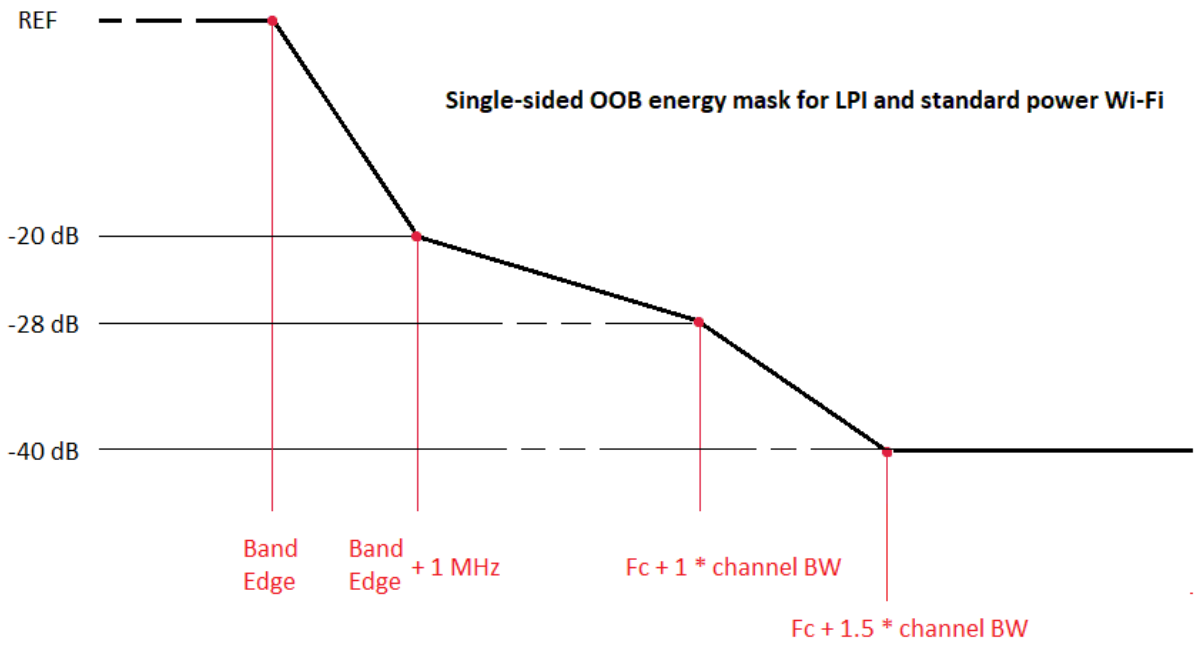


Figure 22 - Required OOB Energy Suppression by Wi-Fi APs

With 20 dB of specified OOB suppression at +/- one MHz from either band-edge of the AP's channel, and presuming 160 MHz of AP BW to set guard bands, one is tempted to represent the AP's spectral footprint as 162 MHz and have that augmented edge-of-band be no closer than the FS receiver's own band edge, to (conservatively) guarantee that the offset frequency AP's energy footprint within the FS link band is at least 11 dB lower (36dBm – 20 = 16 dBm) than that of a corresponding LPI footprint centered precisely on the FS link's center frequency. (As noted in prior sections, such is ubiquitously permitted without operational sanction anywhere in CONUS). So, a 1 MHz guard-band on either side of the FS channel to define a keep out region for overlaid AP energy is certainly sufficient. This has the takeaway specification that the nearest AP center frequency (Fc) for a 160 MHz channel can be no closer (on either band-edge side of the legacy FS link) than 81 MHz away to guarantee the necessary suppression. This can be shown pictorially as follows:

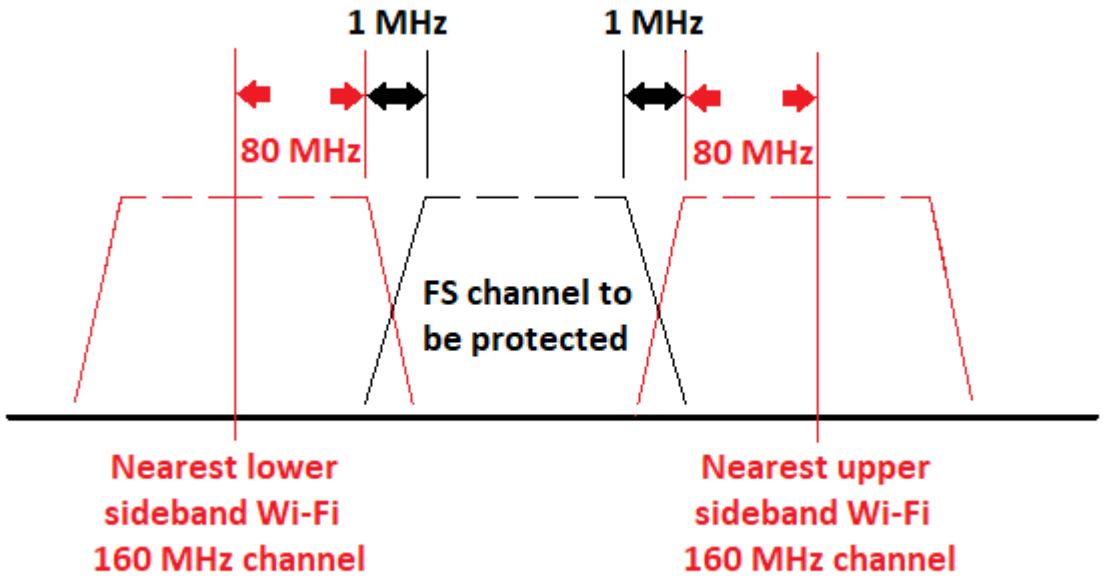


Figure 23 - Suggested FDM Keepout for Adjacent Wi-Fi Channels

In a spectral asset pinch (and as a last resort), Wi-Fi channel BWs less than the full 160 MHz may be considered for "spectrum puncturing" opportunity.

5.3.5 Ancillary AFC Services

The operational aspects of the AFC system in the prior sections have dealt exclusively with AP and cloud considerations for manufacturers and service providers. However, there is a consumer-facing aspect which needs to be acknowledged and dealt with: absent some type of *a priori* analysis, how could a consumer be assured that they would be able to leverage the standard power AP they are considering for purchase? The obvious answer seems to be that, like an AP, they should have the ability to query the AFC system (with appropriate geolocation detail) and have it deliver a verdict on the usability of a standard power device at the location they specify for use. Simply put, if they would not be permitted to exercise the standard power option where they reside, there would be no point in investing in a more powerful (and expensive) AP than one which exploits LPI.

The AFC specification also permits what amounts to a private data channel (vendor-specific extensions) which could be leveraged (along with a log of prior AFC communications) for client applications to provide some monitoring capability of the AFC interactions with the AP; such would be useful for determining the cloud-to-ground behavior of the AFC system and perhaps also assist with fault analysis.

5.4 AP Behavior Under AFC Management

The touchstone for managed AFC operation (and prerequisite for AP certification) is the compliance of the AP to the operating transmit power mask delivered to it from the cloud portal. That is, absent AFC connection and mask delivery, the AP cannot presume to operate at greater EIRP than that allowed for LPI operation. Further, it must default to no more than the LPI specification at lease timeout (currently specified to be 24 hours); finally, if granted a mask (as either available channels or frequency spans), it will not presume to a standard power EIRP in those spectral regions specifically masked off by the AFC query response and only leverage the spectrum allocated. In all cases, however, the AP will adopt a "listen before talk" spectrum deference posture (to mitigate interference from alternate, non-Wi-Fi, unlicensed links). This layered approach to spectrum exercise elicits several algorithmic methods to tune AP behavior for compliant operation and are examined below.

The following segments detail potential AP operating behaviors above the foundation presumptions just detailed above. Note that references to other bands besides 6 GHz are included only for completeness and acknowledge that 6 GHz band assignment for clients necessarily must account for the impact of mapping traffic to clients across *all* available Wi-Fi bands serviced by the AP.

5.4.1 Wi-Fi Airtime Engine (WAE)

The AFC-bound APs seeking to leverage standard power may comprise any of several multiband Wi-Fi options (up to, and including what some tag as "quad-band" APs – 2.4 GHz, 5 GHz low band, 5 GHz high band and 6 GHz). Assignment of multiband clients to 6 GHz (which define the magnitude of the exploit of 6E standard power) will require assessment by the AP regarding allocation of its band resources based upon projected bitrate consumption and the parametric contribution of several considerations. Among these are airtime budget/band available and in-use, desired margins to known airtime limits, band-use and bitrate-use profiles of the prospective client population, status of per-band data queues in the AP, limitations of the AFC-derived channel power mask, historical time-of-day prior behavior and detected desired-channel interference from other Wi-Fi APs or unlicensed 5G/LTE transmitters. The core determination effected is an assessment of airtime burn in each of the bands (especially 6E, given its higher degree of scheduling determinism and ultimate rate/latency performance). This evaluation we will refer to as the WAE.

As previously alluded to, consumed airtime per band is a result of supported bitrate at a particular set of MCS values; these latter are sustained by good C/N in the bandwidth being used – which itself is a function of radiated power and link length. The greater the EIRP and the shorter the path length (loss), the higher the potential MCS – and the less the airtime used to transmit the data packets. The reverse is obviously true: lowering the EIRP or extending the service radius to a far-flung client will lower the delivered signal C/N at the endpoint, reduce the operating MCS for the link and increase the airtime consumed to deliver the data packets (associated with that client, but nonetheless deducted from the overall airtime in the standard case of a single channel/band). Note that these aspects define the viewpoint from the AP to clients or peers (downstream traffic). Because the links support duplex operation, the corresponding upstream MCS and packet sizes must also be accounted for in the channel airtime usage. (And the upstream MCS will be lower than downstream in the 6 GHz band – given the Wi-Fi 6 FCC-mandated 6 dB backoff in client EIRP relative to the AP). The goal of the WAE is to canvas the available operating parameters and balance Wi-Fi band use to achieve best overall duplex data throughput (up to the SLA limits of the WAN).

The following diagram describes the algebra around estimations of band capacity (presume single channel/band) and band availability for future service mounts:

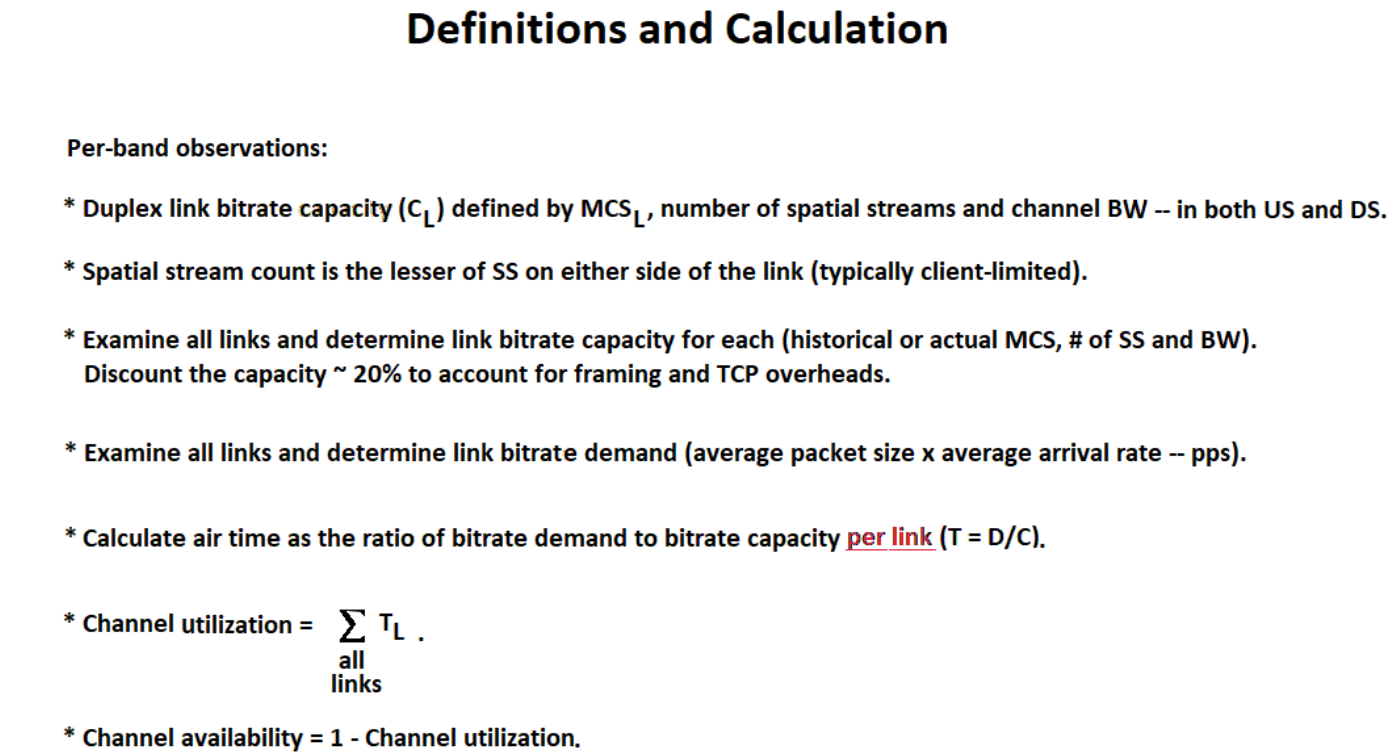


Figure 24 - WAE Term Definitions and Airtime Calculus

Note that the calculation involves acknowledgment of some amount of Wi-Fi framing overhead, coupled with TCP loop closure (as roughly 10% of available goodput airtime). This allows us to estimate goodput bitrate limits when calculating MCS-based link spectral density (as bps/Hz of channel BW) and then have the framing overhead count against that calculated link capacity. Using these calculations per band allows the AP to weigh service mounts based on actual available capacity (versus merely physical layer (PHY) expectations) per supported Wi-Fi band.

The general mechanism of consulting cached (historical) service mounts on a per-client basis would allow the WAE to do an initial assignment of client device to band based upon detection of the client's presence in the network, said client's own supported bands (and MACs) and a review of the client's recent past data consumption. The key historical factors are observed link upstream (US) and downstream (DS) MCS and data rate (calculated as a rolling average of packet size and arrival interval) historically associated with the client(s) in question. Actual MCS detail will be sounded out when the client attaches to its assigned band and may differ from historical indications for the case of a different applications environment, relocated mobile client devices or assignment to a different band (whether the device is mobile or fixed). By invoking historical record, the WAE can make a predictive estimate of the band availability impacts when a particular client lights up on the network.

Two diagrams immediately follow, one dealing with the static interplay of entities or parameters which influence the WAE and a second to demonstrate band assignment strategies for given mounts of clients/services. The third consideration (accessing 6E standard power against a stored mask of available power/channel in the 6 GHz band) will be detailed after these two.

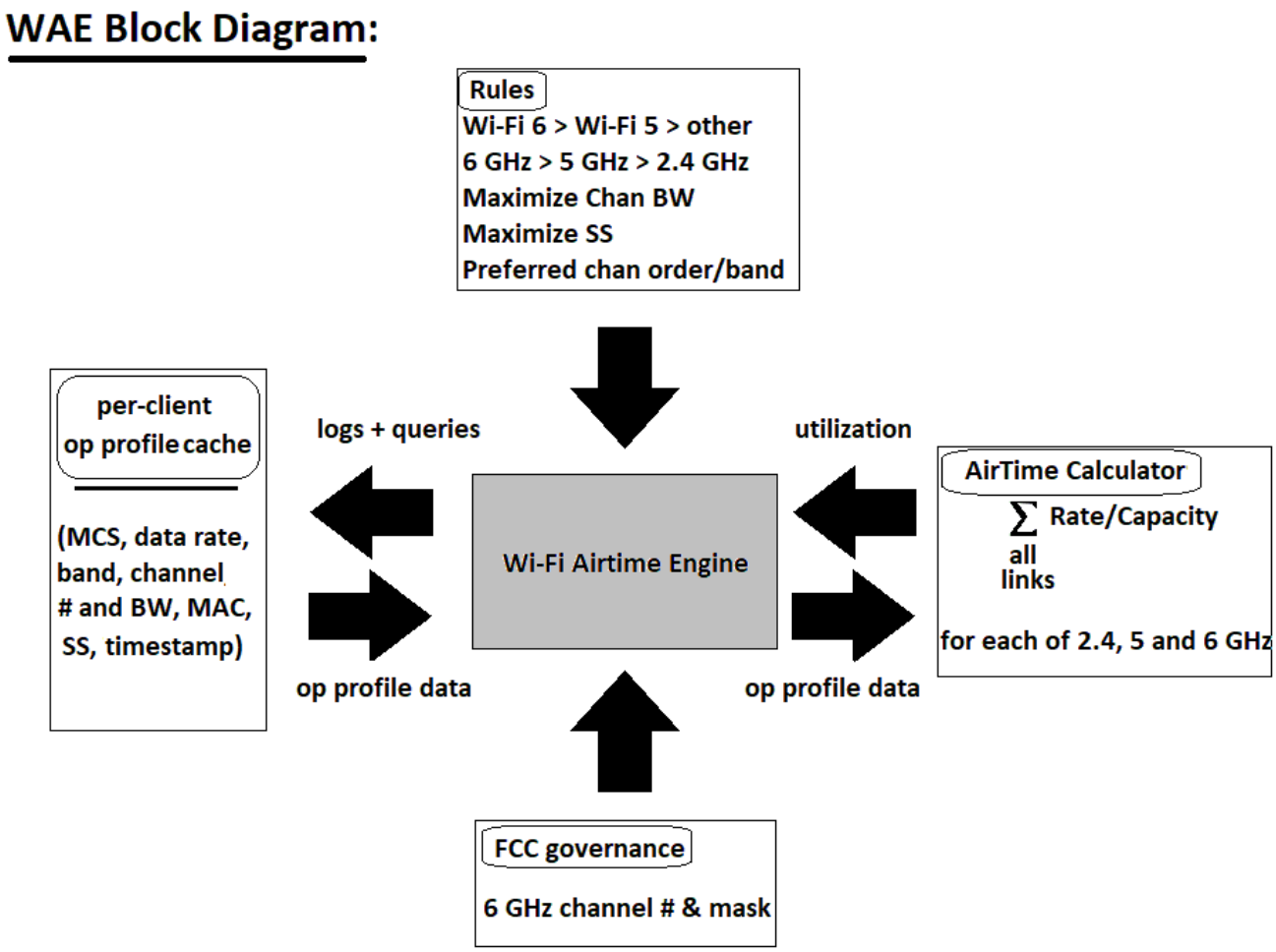


Figure 25 - WAE Interactions

The block diagram lays out the dependencies and interactions of the WAE and four operational data repositories: the profile cache, FCC governance, airtime calculator and client link parameter assignment rules. The rules block is essentially provisioned into the device in one of several ways: via cloud-based remote radio management (RRM), core software (and upgrades) or application-solicited user preferences and directives. The nature of the rules is to establish an order precedence to be used in the WAE's assignment of client operating parameters. Examples of these are shown in the diagram above; Wi-Fi 6 supported connectivity has a higher precedence than Wi-Fi 5 (especially at 6 GHz, where MAC homogeneity guarantees excellent, ungroomed, latency performance); which itself is the intermediate choice before indulgence of airtime-robbing older legacy Wi-Fi MACs. In similar fashion, where client band operation offers options, the preference would be to assign clients to 6 GHz before 5 GHz – with the 2.4 GHz band being reserved for devices which can only operate in that band. Generalized rules would be captured by notions such as maximizing channel BW and SS for any given link (to invoke the best bitrate – and hence, shortened airtime to deliver packets – for each client link). Other rules might establish channel number assignment in each band (based on prior attachment successes, say).

As regards the operating parameter cache, this defines a collection of most-recent client link parameter assignments over some finite rolling period with tagged timestamps. The cache reflects productive historical assignments which the WAE can reference in assigning initial operating points for a client link ahead of calculated airtime ratios based on the anticipated demands of any new service session. (This permits profiling of clients in terms of their data ingestion and production stochastics.) This should prove hugely beneficial for client devices which are single-purpose network mounts and unlikely to exhibit a wide spread of operating demands.

Finally, the airtime calculator maintains a monitor on the sum of all client link ratiometric (as rate/capacity) performance. Typical margin determination would assign a percentage (in the region of 20-25 %) of airtime excess in each band and the airtime calculator would establish band congestion based upon the available excess airtime available per band after the margin and all calculated airtime ratios are accounted for. This provides the WAE with an airtime budget per band which it utilizes in determining optimal client-to-band assignments (transferring bitrate loads as necessary to keep all bands at reasonable utilization levels).

5.4.2 Client Mounting Process

It follows then that these data repositories defined above can be leveraged in a controlled order to effect client band assignment for the AP WAE. An illustration of this follows:

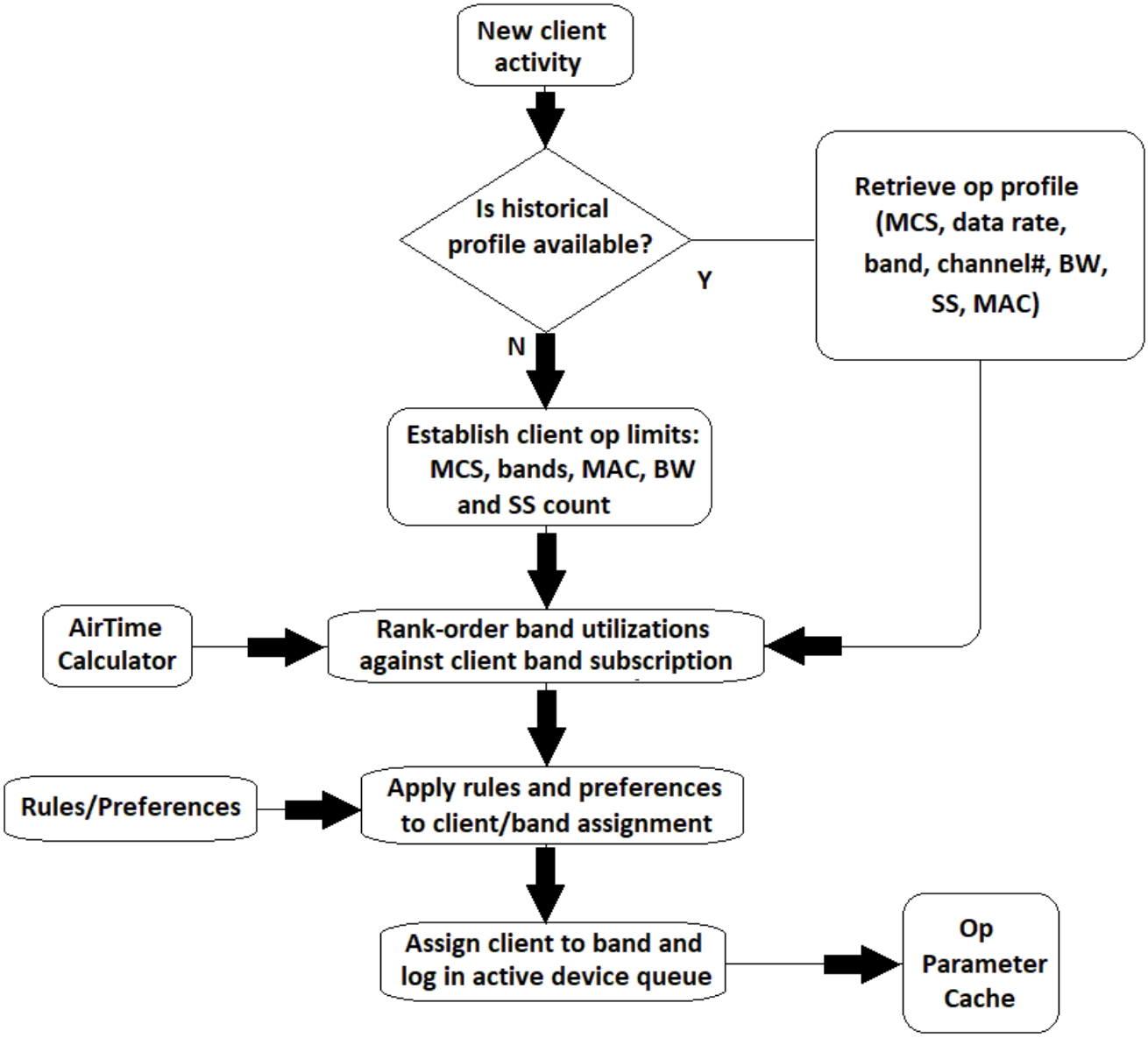


Figure 26 - Wireless Client Mounting Process

Note that housekeeping for the Op Parameter Cache is not included here but can be assumed to include retirement of latent client parameter detail for situations where the timestamps of the data exceed some "hold" timeout period.

The AP's decisions on band use and power/channel assignment are associated with time-of-day historical set points and available empty (or low CCI) channels. As regards the former, the WAE would make use of the Op Parameter logs and interpolate between those timestamped captures and the present time-of-day (TOD) to determine a matching set of parameters (modulo-24 hour).

5.4.3 The Importance of Channel Scanning

In the creation of an available channel map at 6 GHz the AFC-provided channel mask is the starting point of the process. However, responsible stewardship of the channels provided also requires the AP to develop a view of potential unlicensed competitors for the Wi-Fi channels granted it. (There is the unintended consequence, in multiple dwelling unit (MDU) situations, that nearby Wi-Fi customers may all receive the same AFC mask, which unfortunately serves to increase the risk of CCI on the available channels). RRM can assume the mantle of direction of the spectrum decisions; however, the process offered for consideration here is for the cases where RRM is either not available for the AP or there exist disparate, non-aligned RRM orchestrations proximate to the AP. The implication for the WAE is that it must maintain a background scanning function to determine a rank-order of acceptable channels to use (based upon detection of energy during the scan). The scan may be optimized for efficiency (time required to distill the exploitable spectrum) versus "freshness" of the available data. One such optimization would be a progressive triage of available channels based upon initial scanning at maximal bandwidth for the AFC-granted channels (from the collection of all but the current channel in-use) and then progressively narrower scans (by factors of 2, down to a limit of 20 MHz) of those channels which exhibit energy below some set threshold. The outcome could be used to create a candidate channel ranking down to an ultimate resolution of 20 MHz. (or to that bandwidth which has been determined – or commanded – to be the minimally acceptable one. It is certainly possible that, due to bitrate demands at 6 GHz, only options at 80 and 160 MHz may be considered). The algorithm (for determining the best option for a new channel in a single band) is logically illustrated as follows:

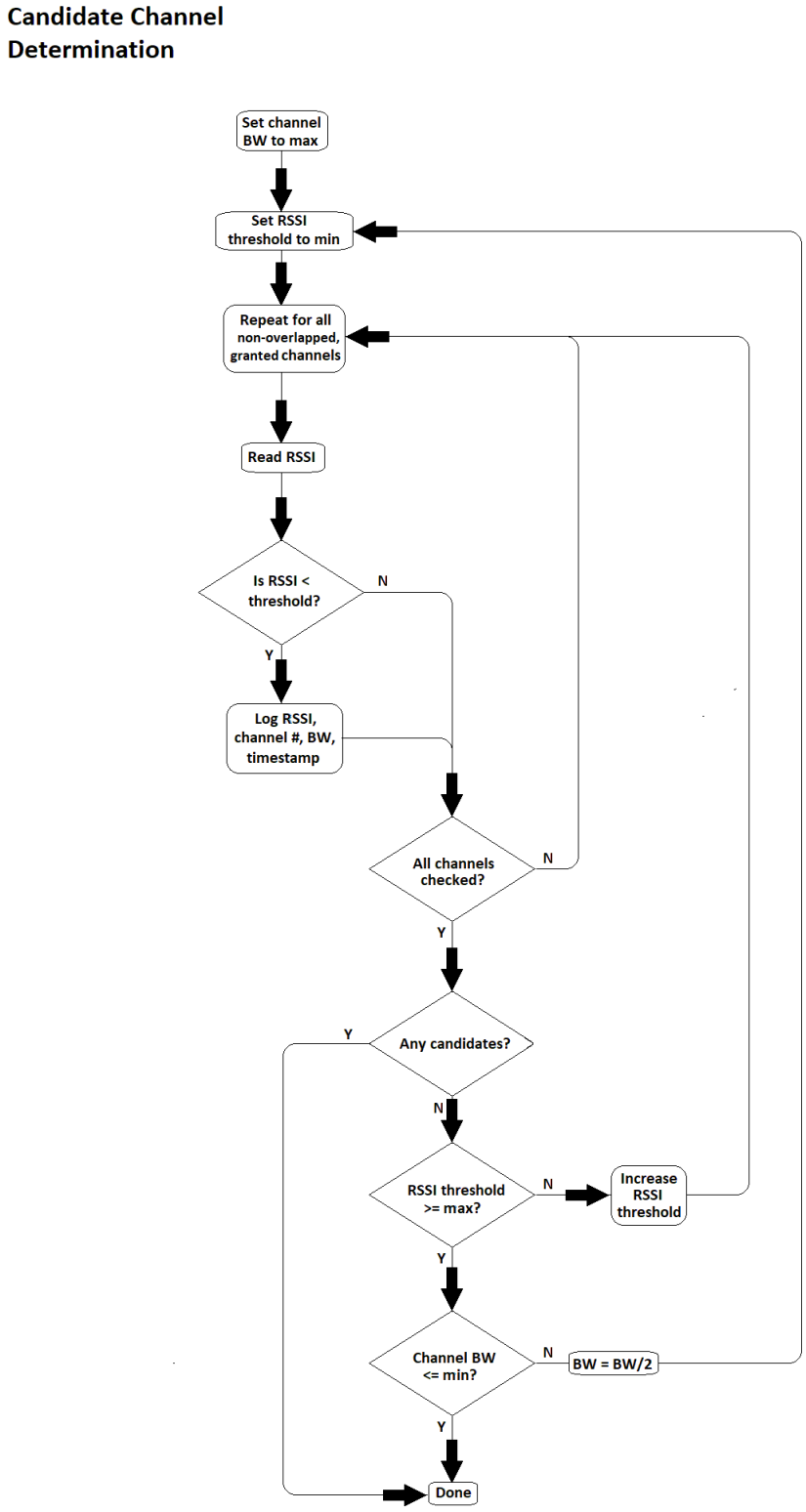


Figure 27 - Scanning Algorithm for Determination of Unconstested 6E Wi-Fi Channels

Given the still-pending maturity of the AFC specification, some of the details above are speculation – but in general, they represent the type of calculus and behavior a 4W AP would have to maintain to warrant its exploit of Wi-Fi 6E.

6 The Emerging Option of Indoor Adaptive Antenna Patterning

A brief aside on antenna developments appropriate to the various service mounts in Wi-Fi 6E is appropriate. Fundamentally, there are two classes of antenna arrangement which serve the interest of robust indoor Wi-Fi distribution: isotropic and narrow-beam patterns. As regards the former, it is expected that both client and AP groups will continue to support multi-direction antenna footprints for major slices of Wi-Fi airtime (to account for random AP-to-client spatial distribution and promote the least expensive method to implement fixed, full azimuth and elevation reach). Narrow-beam near-line-of-sight (nLOS) radiation patterns are, however, useful for P2P home inter-mesh trunk hauls, given that these endpoints are spatially fixed relative to each other. Their directional pattern distributes the allowed EIRP across a reduced spatial cone (improving link margin to one preferred, fixed receiver at the expense of sensitivity for multiple others) and thereby limits the scattering which promotes multipath-based symbol spreading. This unwanted aspect of isotropic radiation can dictate use of wider timing guard band intervals (for expansive floorplans) which can rob goodput throughput. It also inflates error vector magnitude (EVM) at a particular link operating point and hence, represents compromise of the delivered bitrate through potentially unwarranted MCS reduction.

In the scenario where bookended APs are used to establish a trunk haul between meshes (and involves at least a moderate concentration of 6E clients), it makes sense to consider smart antenna switching which, on a per client basis, alternates between isotropic radiation at some inclusive EIRP (which does not tax client receiver dynamic ranges for the respective -- likely 2x2 or even 1x1 -- clients of each AP endpoint), and a 4W, directionally specific, full spatial-stream-exploiting trunk link. The combined efficiency of sharing single-channel airtime between the two modes implies that a second channel (and radio) need not necessarily be resorted to; it is, at minimum, an option to consider for a high-end class of standard power AP.

7 Service Mount Considerations for Standard Power Operation

With the available recourse of higher EIRP, considerations of how best to integrate such a capability within a home Wi-Fi fabric can be evaluated. More transmit energy begets much-improved home wireless coverage, but there are architectural nuances to be evaluated. In addition to the increased bitrate comes the assurance of much better latency performance (between 10x and 100x, versus 5 GHz) for 6E links, primarily due to the single Wi-Fi media access control (MAC) to which every band participant must exclusively subscribe. At standard power, it should be possible to link large client populations throughout a spacious multi-story floorplan and still meet < 2 msec latency for all client links.

7.1 Client Backoff

By far the biggest impact on in-home Wi-Fi service throw at 6 GHz involves the FCC-mandated 6 dB backoff in client EIRP versus its servicing AP (whether this device operates at LPI or 4W). And it is at the point where client transmissions – even down to MCS0 – cannot be recovered by the AP that Wi-Fi framing collapses and the client is disenfranchised by the mesh. At this point in uplink communications,

the downlink is still capable of decent performance – so the argument goes that downlink-biased services (most are of this type; the exceptions being media backhaul for home security or work from home – WFH – scenarios) will not sputter and burn airtime on retries but more abruptly, simply (and prematurely, perhaps, from the transmitting AP's standpoint) fail.

The implication to whole home LPI wireless service is clear: once a client outreaches the service throw, the user has no other choice than to insert a wireless extender into the mix. But as pointed out earlier, at standard power operation, these service radii easily accommodate very large floorplans even when the single AP is not optimally located (i.e., in a lower corner of a multi-story home, for example). So perhaps the better way of viewing standard power is that it affords the opportunity (up to well above American average home floorplan size) to obtain the desired WAN-matching service-level agreement (SLA) with only a single AP whose location is determined solely by convenient access to its WAN. For larger homes, a summary takeaway might be that one could guarantee > 1 Gbps client service everywhere throughout the home, even with a corner basement mounted wide-area network (WAN) gateway.

7.2 Subordinate APs

It is noteworthy that the FCC has created an exception category to the backoff for what it deems "subordinate APs" – essentially peer nodes on trunk links who themselves have one or more clients to service. For these devices (think "extender" in the case of a large home) EIRP need not be reduced. This has the effect of providing for exceptionally robust connections between mesh nodes simultaneously serving proximate battery-constrained clients and implies that the insertion of an intermediate hub is almost always guaranteed to enfranchise clients whose path loss to a single, non-optimally placed, AP renders them otherwise "unreachable". In doing so, it also appreciably raises the average link MCS throughout the whole home, which pares the overall airtime spend and makes single-channel radio extension at least conceivable.

Note that this exception to EIRP backoff establishes standard power Wi-Fi as a less costly option for large homes. Simply put, even if the standard power AP is a bit more expensive than an LPI-constrained one, it costs less (in dollars and APs) to cover a whole home to a given aggregate bitrate service level than to attempt multiple LPI extensions to keep the wandering low-power clients happy. And not to put too fine a point on it, fewer (or no) linked extenders imply a much more robust network availability factor.

7.3 Standard Power Deployment Opportunities

So now we come to advised exploit of 4W indoors for Wi-Fi 6E APs. Certainly, the span of client capabilities (and service bitrate expectations) will define the most efficient leverage of multi-power Wi-Fi meshes for given service mounts. Using the testing at LPI and the projected improvement in reach afforded by standard power, we can imagine several architectures which would efficiently serve even the largest home layouts and unlock advantages of 6E. We already have data which indicates mobile client support anywhere in the Wi-Fi house based upon the servicing LPI AP being located at the home midpoint; when this AP is promoted to standard power, the rates cannot help but improve (up to the uplink starvation case previously mentioned). But this turns out to be a needless worry.

Some examples of the expectation for client devices in the Wi-Fi house if the central AP is bumped to 4W would clarify the monstrous benefit of a 4W midpoint AP in the 5300 square foot home. Referring to our data, a repainting of obtainable TCP bitrates for 2x2 mobile devices scattered throughout the Wi-Fi house

becomes a monotonous affair. No room on any floor is served with less than at least 1.4 Gbps downlink bitrate – and the entirety of the main floor (where resides the AP) sees the maximum 1780 Mbps. The tiling looks as follows:

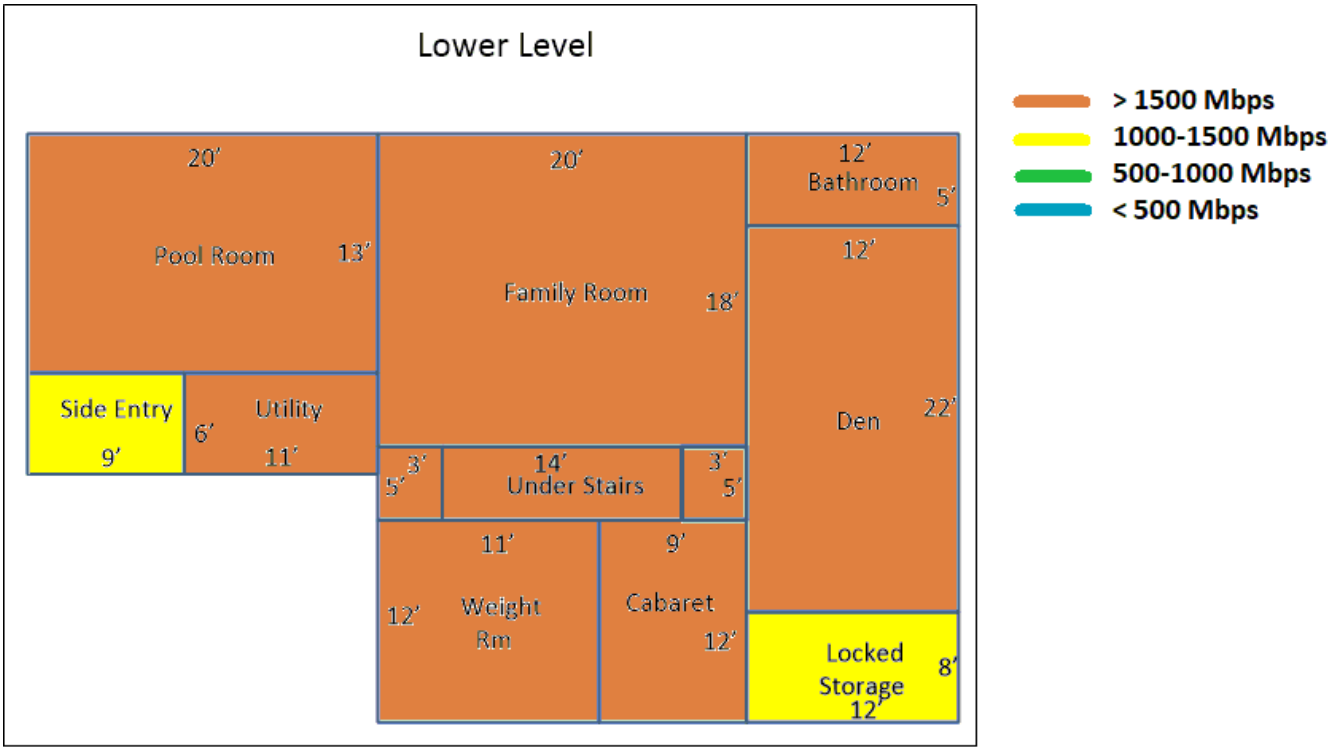


Figure 28 - Lower Level Service Coverage with 4W AP

In fact, though clients at the side entry and locked storage locations on the lower floor would be constrained to just under 1.5 Gpbs with an AP association from the home midpoint, in the case of the side entry location clients would be steered to a binding to the WAN gateway. Under this presumption the full 1780 Mbps becomes available there. Signal throw to the locked storage area, however, is actually worse to the WAN endpoint than the living room AP, so the modest attenuation of bitrate to 1.4 Gbps would remain.

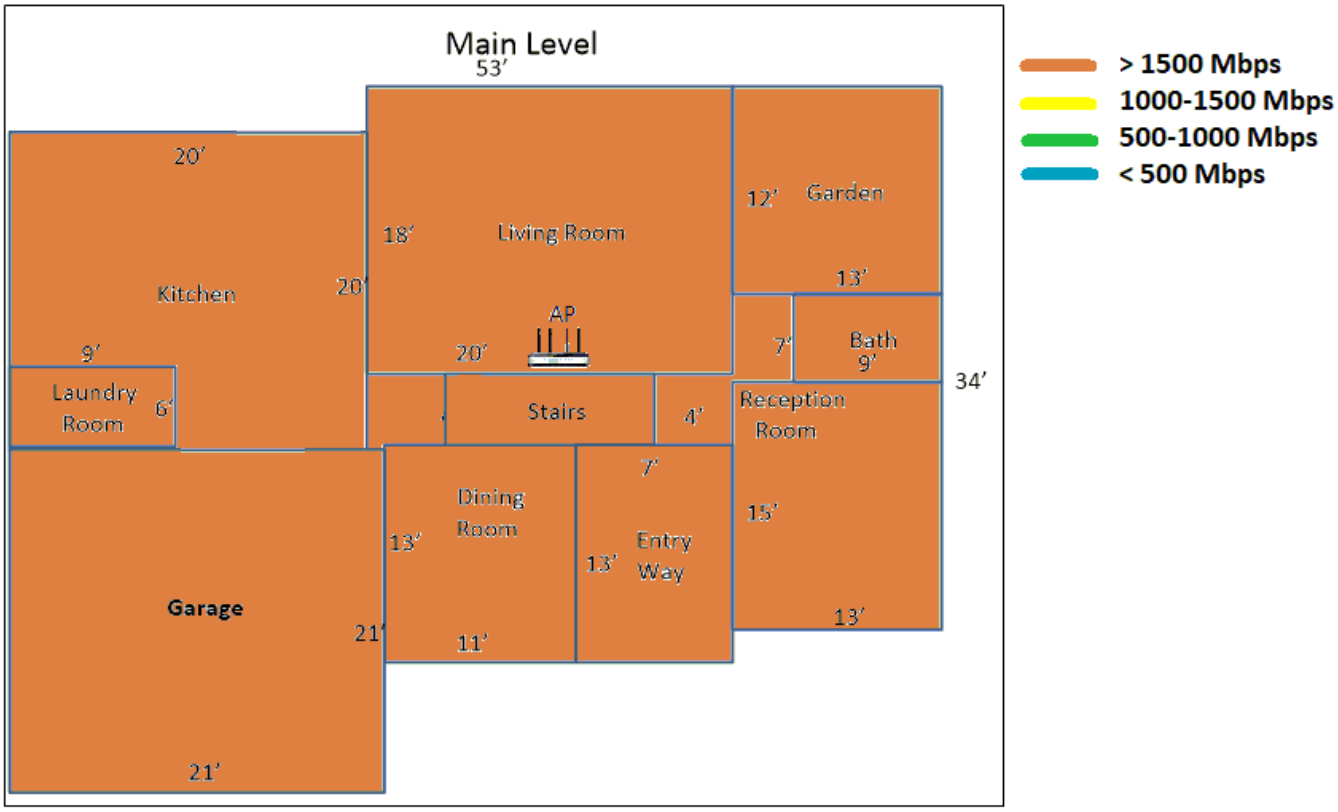


Figure 29 - Main Level Service Coverage with 4W AP

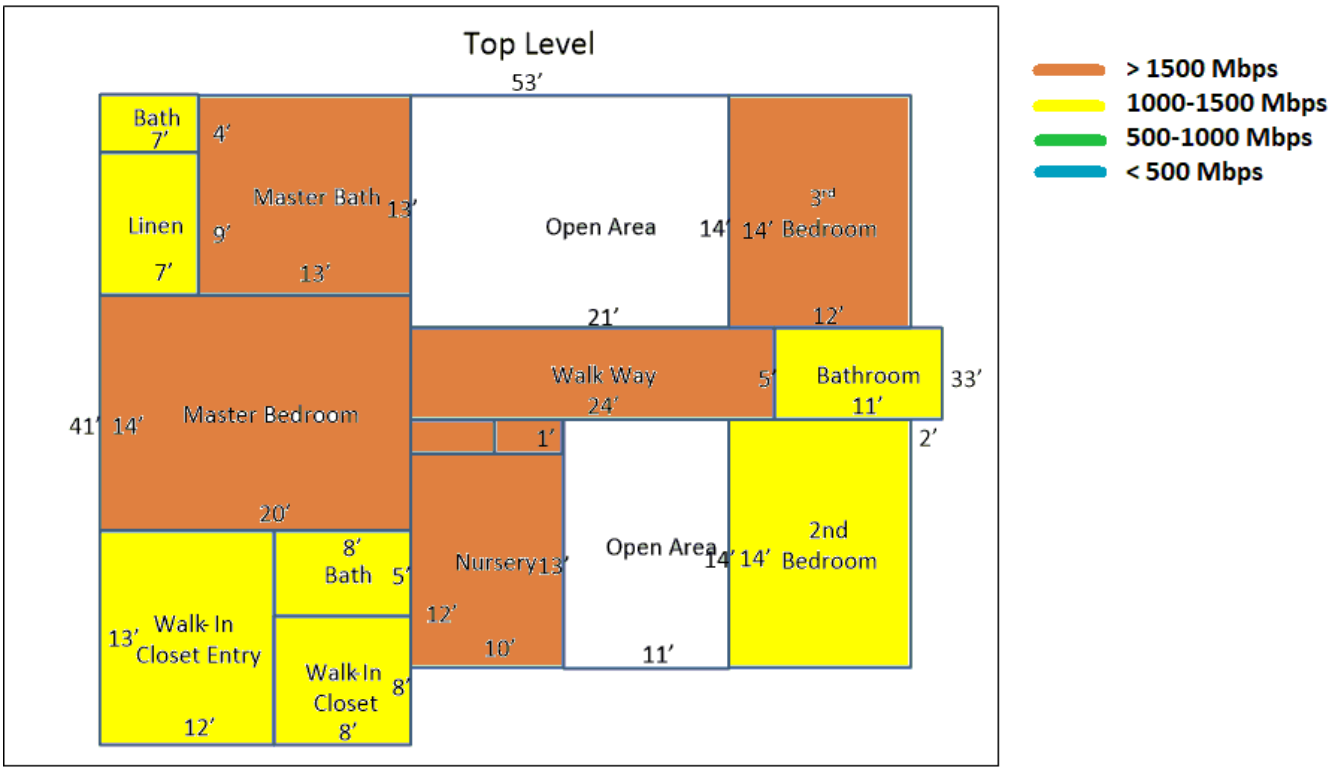


Figure 30 - Top Level Service Coverage with 4W AP

Noteworthy in all of this downlink behavior analysis is the comforting realization that, despite the rather onerous limitation of 20 dBm of transmitter power, the roaming mobile clients never experience uplink service anywhere in the house to < 550 Mbps (so the prior bookmark on potential uplink starvation may be retired).

But the midpoint approximation – not a bad alternative and easy enough to implement -- is likely not the best coverage option for extremely large floorplans. As the WAN attachment is in the basement pool room (a common enough home topology for connecting an outside WAN), it would make sense to move the extender all the way to the upper floor and closer to the two Jack/Jill bedrooms. This has the airtime benefit that more clients can be steered to the WAN-attached gateway, such steering subtracting airtime from the 4x4 trunk (and eliminating one link latency/client for those which can so attach). These latter outcomes are most welcome, in that the movement of the extender AP endpoint further away from the WAN will (slightly) penalize the trunk bitrate performance (so it having to carry less backhaul/fronthaul traffic tends to somewhat balance the capacity drop). The modified layout looks as follows:

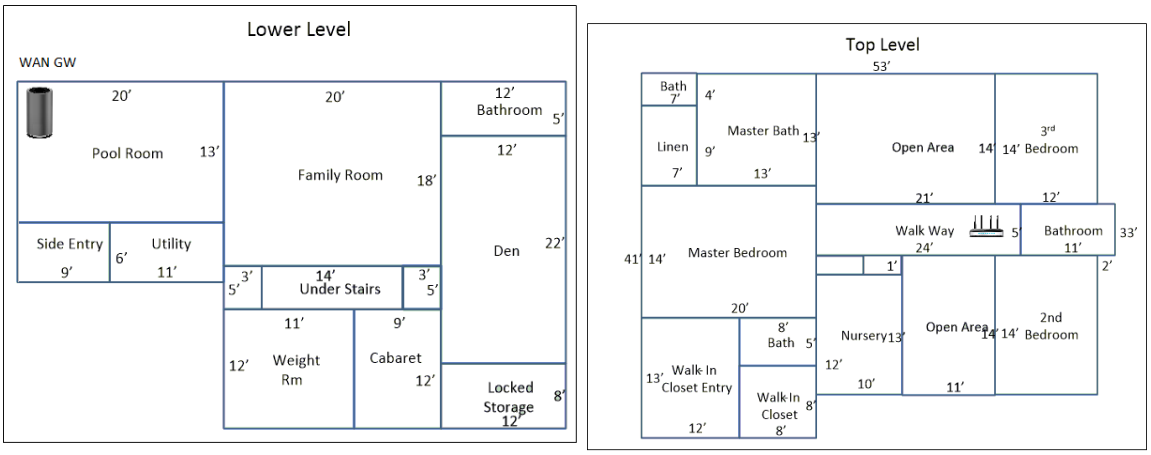


Figure 31 - Modified Trunk Topology to Improve Whole Home Coverage

In this scenario, trunk performance is somewhat degraded (recall the midpoint extender placement allowed for a 3.5 Gbps goodput carry) to something more like 2.5 Gbps. However, the client coverage everywhere in the Wi-Fi house (except for that poor locked storage location in the basement, which achieves "only" 1.25 Gbps now) easily averages 1.7 Gbps. For the case of low numbers of 6E clients in a home (as will be the norm for the next ~5 years), this ability to roam (or set up a fixed client endpoint) anywhere in the home and not experience rate throttling manifests obvious value. Note that, in terms of present-day services, there are precious few application bundles which, even in simultaneous aggregate on one device, approach a 100 Mbps/client downlink service requirement – so having 1.5-2 Gbps at your disposal everywhere in the home is a bona fide lottery win.

This client "overindulgence", then, begs the question of whether a single 4W WAN gateway in the low corner of the house could blanket it with sufficient signal throw to enable excellent (or even more-than-adequate) service bitrate in a home the size of the Wi-Fi house. Without performing a full volume analysis, it is instructive to examine what the worst-case client downlink/uplink support appears to be, given the geometries in play. Selecting the far corner of the second bedroom on the top level as a checkpoint, the service radius amounts to ~ 62 feet from the WAN gateway (with two floors and perhaps 3 walls to penetrate – though the two large open areas facilitate multipath reach to this client placement). In this far corner of the house, downlink would be expected to be 1030 Mbps and uplink, ~ 275 Mbps – a fair margin to uplink starvation which would otherwise defeat persistent client attachment. The clear implication is that one would have > 1 Gbps everywhere in such a large home, even with a single, inconveniently located, AP. And let's be clear – the Wi-Fi test house represents a floorplan which is 2x the US national average for size of new construction, single dwelling units (SDUs).

It appears, then, that the recourse of two 4W APs is not obviously necessary in up to at *least* 5300 square feet of living space -- which means we would reserve that strategy for extremely large floorplans (realtor estimation of the mansion footprint, for example, sets an 8000 square foot expectation on qualifying space to earn that moniker). And it is not clear that the recourse of 2 standard power APs to blanket such a homestead would represent much of an investment concern for users well-heeled enough to own it.

Statistically speaking, given the predominant 6E client target capability of 2SS/device (which bounds the 160 MHz BW delivery of services to around 1.8 Gbps TCP), the overwhelming majority of US homes would *at most* require a single standard power wireless gateway – no additional wireless meshing required.

8 Conclusion

There is great value to be had in policing and permitting standard power operation in the highly scheduling-deterministic Wi-Fi 6 GHz band – and so extend the promise of this vast spectral asset which the FCC released last year. Though limited-power mobile clients represent a challenge to the brute force expectation that standard power APs should always be considered, it is nonetheless true that the advantages of the expanded coverage footprint over LPI suggest that a standard power "lynchpin" AP find representation in all floorplans more than perhaps 1500 square feet, to properly anticipate and lever the multi-Gbps WANs which are in the offing. Furthermore, floorplans in the 6000-10000 square foot range can make use of a single, very robust 4x4, 4W trunk (with bookended -- full 4W EIRP at each endpoint -- APs) to seamlessly connect meshes with client devices up to the periphery of the expansive floorplans (and minimize the peppering of LPI extenders everywhere throughout the footprint). The FCC and the greater industry at large have taken great precaution in defining and (soon enough) implementing a 6 GHz band coexistence scheme which simultaneously protects legacy microwave infrastructure investments and liberates all the promise of Wi-Fi 6E. The in-home high speed wireless future has never looked so bright.

Abbreviations

3D	three-dimensional
5G	Fifth generation of mobile communication technology
AP	access point
AC	alternating current
AFC	automated frequency coordination
BEL	building entry loss
CCI	co-channel interference
CGI	computer-generated imagery
C/N	carrier-to-noise ratio
CPE	consumer premise equipment
CONUS	continental United States
dBm	decibel milliwatts
DS	downstream
DUT	device under test
EIRP	effective isotropic radiated power
EVM	error vector magnitude
FCC	federal communications commission
FDM	frequency-division multiplexing
FEM	front-end module
FS	fixed service
FSPL	free-space path loss
FSS	fixed satellite service
FWA	fixed wireless access
Gbps	gigabit per second
GEO	geostationary equatorial orbit
GHz	gigahertz

I/N	Interference-to-noise ratio
LOS	line-of-sight
LPI	low power indoor
LTE	long-term evolution
MAC	media access control (layer)
Mbps	megabits per second
MDU	multiple dwelling unit
MHz	megahertz
MS	mobile service
nLOS	near-line-of-sight
NRA	national regulatory authority
OOB	out-of-band
P2P	point-to-point
P2MP	point-to-multi-point
PHY	physical layer
PSD	power spectral density
SDU	single (family) dwelling unit
SLA	service-level agreement
SS	spatial stream(s)
STB	settop box
SUT	system under test
TCP	transmission control protocol
TOD	time-of-day
UDP	user datagram protocol
ULS	universal licensing system
U-NII	unlicensed national information infrastructure
US	upstream
WAE	Wi-Fi airtime engine
WAN	wide-area network

Bibliography & References

Frequency Sharing for Radio Local Area Networks in the 6 GHz Band, ver 3, RKF Engineering Solutions, LLC, January 2018; prepared for 6USC

ET Docket No. 18–295 and GN Docket No.17–183; FCC 20–51; FRS 16729, *Unlicensed Use of the 6 GHz Band,* FCC

COMPATIBILITY STUDIES IN THE BAND 5725 – 5875MHz BETWEEN FIXED WIRELESS ACCESS (FWA) SYSTEMS AND OTHER SYSTEMS*, 2005, ECC Report 68*

Why Scale Needs Unity:

One Operator's Journey

A Technical Paper prepared for SCTE by:

Elizabeth Riley-Wasserman, Ph.D.
VP/TPX Consumer & Strategic Programs
Comcast
Philadelphia, PA
215-313-7543
Elizabeth_RileyWasserman@cable.comcast.com

Shane Portfolio
SVP, Reliability Engineering
Comcast
Denver, CO
303-658-7993
s_portfolio@cable.comcast.com

1. Introduction

The “heavy lift,” when industries like ours consolidate though acquisitions and geographic clustering, is almost always carried by engineering. Why, because the biggest physical asset is the plant. And it wasn’t that long ago when dozens of different operators made hundreds of different decisions about capacity expansion methods, network topology, node sizes, even amplifier spacing and drop materials. Big decisions, that necessarily last for decades.

Then, one day, the convergence activities wind down, giving way to the undiluted pursuit of scale – scale of the network, and, by extension, to operations and “the field.” The pursuit of scale is a perpetual transformation, and involves aligning many of those previous decisions, made in previous times and under previous ownership. It’s a different kind of heavy lift because it’s as people-impacting as it is equipment-impacting. If you’ve ever tried to replace a favorite tool or dashboard with another one, for any reason, you know how what we’re talking about. Unity through scale is about aligning people with tools, processes, and standards.

Comcast’s unification blueprint focuses on three core areas: technology convergence, tools alignment, and process alignment. It’s more transformative process than organizational location, centered on a highly reliable, self-healing network. It uses real-time data to enable automation, and it’s all supported by the same tools, processes, and practices.

2. Compelling Case for Change, Vision and Guiding Principles

Background:

The Comcast network, like many other industry leaders, has grown through acquisition. While we have a single network, there is procedural variation in how work is executed. Individually, Divisions and Regions are highly successful. However, as we look to scale and optimize network performance, we are hampered by procedural and tool differences. If our tools and inputs into those tools are not the same, we are less able to take advantage of automation.

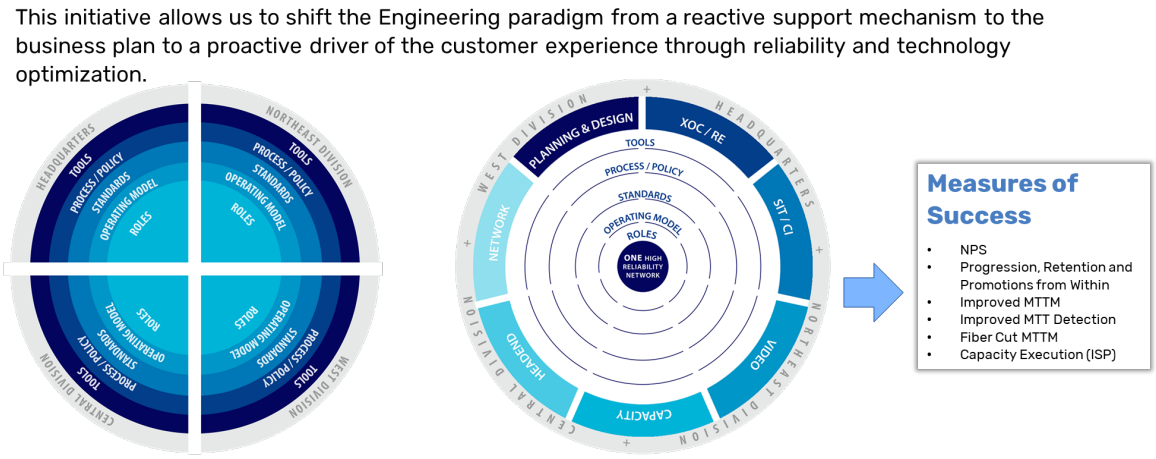


Figure 1 - Our Transformation Journey

Given the evolution of new technologies and tools, we now have the opportunity to shift the engineering paradigm from a reactive support mechanism to the business plan, to a proactive driver of the customer experience through reliability and technology optimization. Seeing the technology opportunity, the need to scale, and to evolve our tools, Tony Werner, President of Comcast Technology, and his engineering leadership across the network took two critical steps to prepare the organization to meet growing consumer demand on the network. In 2018, the company's three divisions began the process of developing division-specific standard operating models for Strategic Infrastructure/Critical Infrastructure (SIT/CI) Headend, Video, Network, XOC Planning and Design. In some locations, this also included shifting process ownership and accountability from regions to divisions. Simultaneously, Werner also asked his headquarters engineering teams to start to develop next generation tools that optimize cloud technology to better harness the power of automation and build a self-healing network.

- Our network operating **model** is fundamentally **changing**
- Moving **from 15 regionally diverse** networks **to a homogenous enterprise cloud-based** technology permitting unparalleled scalability
- Optimization of automation will **reduce escalations** and **trouble calls**
- **The pace** of the transition is **happening fast**

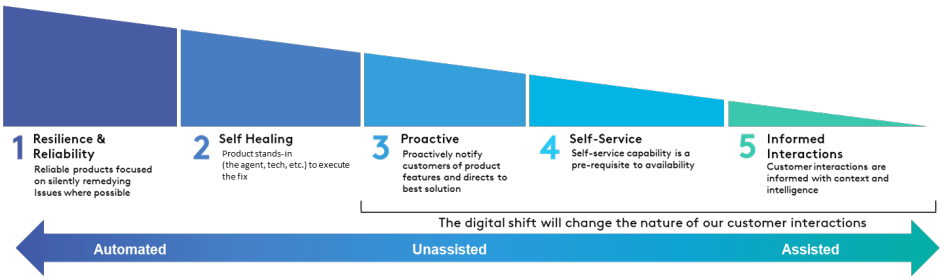
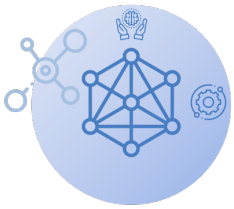


Figure 2 - Technology Transformation

In late 2019, it was increasingly clear that the development of technologies and tools was outpacing the development of common processes and the expressed elimination of redundant tools across the network. Divisional leaders were asking for tool customizations that would fundamentally undermine the potential power of the tools and network. This was the ideal time for the organization to launch an initiative that would ensure a clear understanding of our long-term vision and the pace with which we needed to realize our vision. To drive this change, Werner established an executive leadership coalition made up of headquarters engineering leadership and a representative of the division presidents to sponsor the change. They asked Shane Portfolio, SVP of Reliability Engineering, and his team to provide leadership for the change in partnership with his Engineering peers from headquarters and the 3 divisions (4 Engineering Leaders = E4). As VP of Consumer & Strategic Programs, I was asked to provide change management support for the initiative. The focus of this paper is on this second wave of change.

Common Vision and Guiding Principles: Based on leadership interviews and an organizational assessment, it became evident that each leader had a different vision and understanding of the guiding principles we would use to support decision-making and how we worked together. To that end, the E4 and Leadership Coalition developed the following vision and critical operating and design principles.

Vision: *We will design, build, and operate a highly reliable, self-healing network of unprecedented scale, that utilizes real-time data to enable automation supported by the same tools, processes, and practices. To realize this vision, we will move from regionally diverse networks to a homogeneous cloud-based technology providing scalability and efficiency to existing processes providing greater field focus on customer experience. The transition will align career and talent with our evolving technology. To be successful, our talent transition needs to leapfrog the technology transition. To fully capitalize on the technology available, in service of our customers, employees and shareholders, it is critical that we operate as one network to best position the company for the future.*



We will design, build and operate a highly reliable, self-healing network of unprecedented scale, that utilizes real-time data to enable automation supported by the same tools, processes and practices.

To realize this vision, we will move from regionally diverse networks to a homogeneous cloud-based technology providing scalability and efficiency to existing processes providing greater field focus on customer experience. The transition will align career and talent with our evolving technology. To be successful, our talent transition needs to leapfrog the technology transition.

To fully capitalize on the technology available, in service of our customers, employees and shareholders, it is critical that we operate as one network to best position the company for the future.

One Vision, One TEAM

Operating Principles

- Build and earn trust and respect
- Be system thinkers
- **Play our future into being by taking every opportunity to implement our vision**
- Be a coach for our teams and our peers
- Be authentic, honest and value constructive conflict
- Have a clear vision and strategy for the team
- Be transparent
- Get out of our own way
- Value collaboration, we will not act in silos

Design Principles

- **Same** Processes and Tools
- **Same** Architecture
- Data Inputs and outputs will be the **same** for each tool
- **Same** Deployment Practices
- **Same** Methods and Procedures

2020 Focus

- Business Case
- Executing on the tools strategy
- RE Realignment
- Skill and Career Development

Measures of Success

- eNPS
- tNPS
- rNPS
- Progression, Retention and Promotions from Within
- Improved MTTM
- Improved MTT Detection
- Fiber Cut MTTM
- Capacity Execution (ISP)
- TC Reduction
- Contact Rate Reduction
- Outages Due to Changes

Figure 3 - Our Vision

Operating Principles: Given that each engineering function rolled up under different P&L leaders, identifying and holding ourselves accountable to a set of operating principles was critical. Leaders quickly realized that trust, transparency, and a commitment to act together rather than in silos was essential.

Design Principles: In the organizational assessment, leadership realized that while their processes had become more aligned, they were still very different, and that in order to optimize new technology and tools, they needed sameness at each level of their work. So, the leadership agreed that all future work would focus on achieving sameness of architecture, process, tools, methods, procedures, and roles across the divisions and headquarters.

3. Campaign Approach to Change

Defining the critical path and the roles to achieve sameness was fundamental to getting started. We had 9 functions across headquarters and divisions with different levels of process alignment and tools. We also knew if we were going to have the leadership and talent to lead and manage the future network, that this transformation gave us the opportunity to grow our team capabilities rather than asking external consultants to "swoop in" to fix us. To transform the organization, we couldn't just focus on the operational processes, we needed to optimize enabling systems like goal setting, performance reporting, budgeting, talent management, training, etc. Fundamentally we needed to "play with other people cards" and enlist them to support and nudge our vision along. We couldn't be an island. The whole system needed to drive the change. In addition, we knew that we couldn't wait till the change was fully baked to

move forward we needed to "play our future into being," which meant every decision and action moving forward needed to be in support of sameness. We needed a multi-level approach.

We also understood that we needed to evolve into a culture of sameness. Culture is not something we or anyone could mandate. Culture is something we needed to grow over time, through the adoption of practices and new procedures that could be reinforced, supported, and adopted. We also knew that our future was not dependent on a single solution or action, but rather on the combination of elements that we would implement to create it.

Our transformation approach needed to focus on both building the organization's **capacity and capability to change.** To do this, we utilized a campaign model to design our change approach. To build our organizational capacity to change, we identified three levels of change: 1) organization design level - which focused on the redesign of processes, tools, and talent for each function; 2) macro change level - which focused on the optimization of the enabling functions and organization events to reinforce the change; and 3) micro change level - which focused on the individual workstream initiatives. As a technology organization we utilize agile design, however, to gain leadership support for major process changes we knew that they would need to see the end-to-end process design before we could implement. So, we agreed to utilize a fast-cycle redesign methodology that optimized a waterfall methodology.

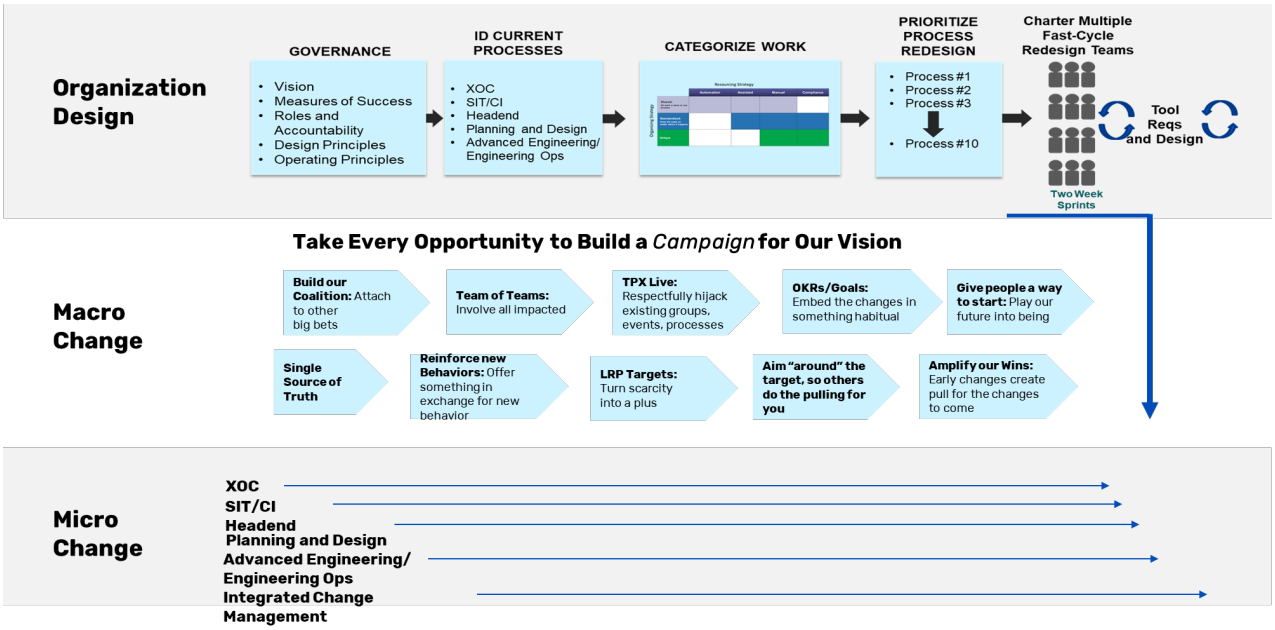


Figure 4 - Our Plan to Build Capacity

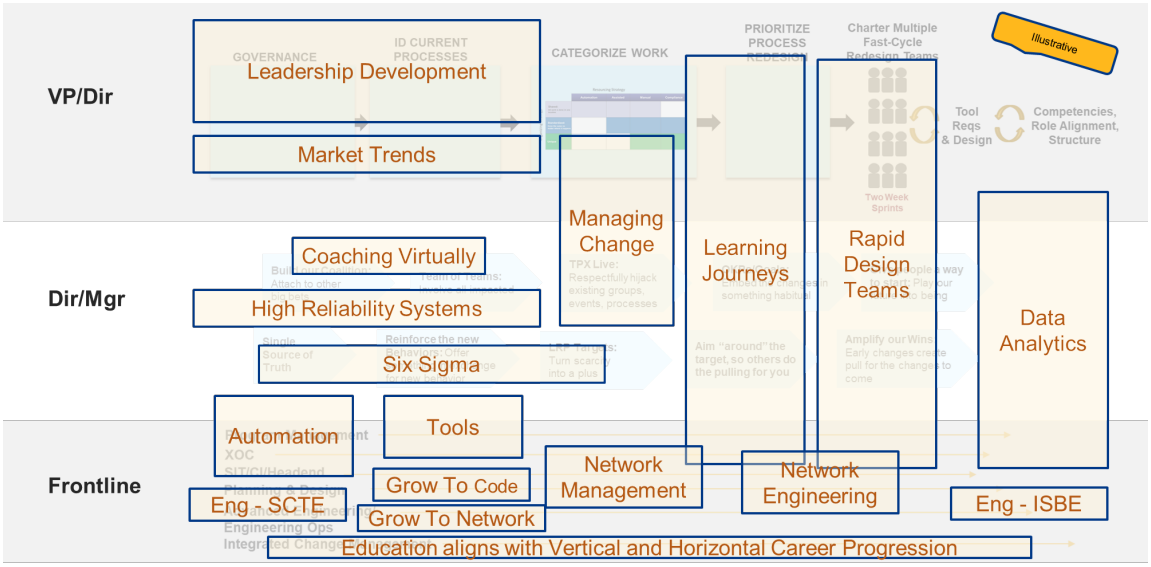


Figure 5 - Our Plan to Build Capability

To build capability, we agreed we would utilize an action learning approach and bolster our work with development at each level of the organization. From one vantage point, we were a transformation process. From another vantage point, this was a major leadership and skill development process.

The Executive Leadership Coalition, the E4 and the division Presidents gave their support for moving forward. The E4 as sponsors met with the leadership of each functional team to kick-off the work. The first step was for each functional leadership team to categorize the work, to identify what work needed to be the same, and whether the work should be done by a single team across the Divisions or needed to be done within the Divisions. In addition, we began to map current state processes to prioritize the work. Then, in late March 2020, COVID 19 hit.

4. A COVID Detour that Solidified Leadership Commitment and Approach

With COVID 19 came the retrenchment of consumers to their homes and significant dependency on networks for work, school and staying in touch with loved ones. Foreshadowing the increased need for scale, the Leadership Coalition asked leaders to explore options for accelerating the plan. To respond to the request, the team developed a business case for change and possible options. In July 2020, Leadership agreed to stay the course and because of a better understanding of the technology, tools, talent implications and transition timing, the division Presidents recommitted the strategy of sameness. The development of the business case surfaced several opportunities. First, we needed to solidify the return on investment (ROI) for the 10 most critical tools that our strategy was dependent on. Second, we needed to fortify the trust and transparency between headquarters and division partners. In particular, division leaders needed to feel confident they would achieve their targets even if they changed the processes that today were enabling them to be highly successful, or relinquished ownership of other processes to enable efficiencies and optimize automation. Finally, we needed to have headquarters technology and division presidents jointly signal their sponsorship.

5. Establishing a Center-Led Approach

To move forward, the E4 developed an approach which focused on accelerating decision-making, growing the partnership between headquarters and division teams and the operational discipline necessary to enable clear understanding of the changes as well as a highly reliable consumer experience. The approach had the following components:

Center-Led Leadership: In August 2020, The E4 agreed to implement a "Center-Led" operating model, to create a system of accountability to drive sameness. Being Center-Led meant that all changes moving forward within the 9 Pillars, nationally or within a division, needed to be agreed to by the three division and headquarters leaders. The objective was to focus our energy and resources to move towards sameness, not further apart.

pillar Role and Co-Leadership: The first step the E4 took was to identify the 9 functional Pillars, their role and strategy in the realization of the initiative vision. The E4 then identified a division and headquarters leader to co-lead each pillar on a yearly rotation. The objective was to create a single voice for each pillar to help drive requirements and prioritization. The role of the pillar co-leads is to create a "virtual cross functional leadership team" and facilitate the creation of process, tool, key performance indicators (KPI) and role sameness nationally within their pillar. A critical measure of success is to facilitate decision-making that reflects the operational discipline that drives transparency and builds a foundation of trust across the pillar. pillar co-leads were identified and announced in September 2020.

Three-Legged Support: To support the Pillars, each pillar was assigned 3 resources: Integration Lead, Organization/Change Management Lead, and Project Manager. The Integration Lead is a subject matter expert accountable for cross pillar integration. The Organization/Change Management Lead is accountable for ensuring a common operational problem-solving approach and real-time action learning. Project Manager is on point to support pillar Leadership. This 3-legged stool (Integration, Organization/Change and Project Management Leads) plus the Integrated Program Management (IPM) Lead are critical to success. All resources are internal.

Domain Integration: To ensure tight coupling across the 9 pillar teams the E4 decided that each Executive leader would provide sponsorship for 6 that are common cross the 9 pillar domains: Reliability, Talent, Partnership, Tools, Quality and Engineering Architecture.

Prioritized 90 Day Sprints: The E4 agreed to facilitate the transformation in 90-day sprints. Each quarter the pillar Co-leads were asked to prioritize their pillar's focus in partnership with their pillar leadership. During September 2020, the pillar Teams identified their priorities and developed a detailed charter for each priority that included: objective, scope, measures of success, design criteria, required resources and date driven workplan. In October 2020, each pillar team presented their recommended priorities and charter for review, feedback, and support to move forward. The E4 approved 24 Initiatives for the first 90 sprint (Q1 of 2021). In mid-November 2020, each pillar "kicked-off" their initiative workstreams. In Q2 2021 29 Initiatives were approved and in Q3 2021 17 Initiatives were approved.

Common Goal: In support of our collective success, the E4 agreed that each of their Senior Leadership Teams would have the following common goals for 2021.

1. Deliver the same highly reliable network experience across the enterprise by completing the 2021 pillar priorities.
2. Improved Partnership Scores

Linking Mechanisms: The E4 established regular weekly, bi-weekly, monthly, and quarterly connection points at the division President, E4, Integrated Program Management and pillar level to ensure transparency and thoughtful decision-making and change management. The Organization Development and Program Management team partnered to support each forum.

Sponsorship: Based on the E4's work and mutual commitment to success, on 11/30/20 the headquarters Leadership and division Presidents sent a joint letter to all leadership signaling their sponsorship that we will have common processes across the network:

> "The strategy represents a historic milestone within our company – an inflection point to change the way we manage our network that will have tremendous impact on our ability to continue providing the fastest broadband service to our customers and increase our agility to respond to the growing needs of the business. By developing strong partnerships among headquarters and division teams and implementing the exact same tools, technologies, and processes across all teams, we will continue to build the network of the future.
>
> This strategy hinges on partnership and 100% alignment. All of us are committed to the success of this initiative and dedicated to doing what it takes to make it happen. Now we are asking you to join us."

6. Disciplined Approach: Process, Tool, Function, Role Sameness and KPIs

As we began the initiative, one of the concerns expressed was regarding the pace of decision-making. Based on interviews with leaders at all levels, it became evident that when recommendations were presented, they lacked the back-up detail to allow for efficient decision-making. To help remediate this, we made a commitment to utilize process design best practices. To support each team's work and ultimately integration, the Organization Development/Change Management Team developed a process redesign toolkit to guide each team's work and to ensure we have the necessary documentation of decisions to develop the requisite business case and change management plans. One of the challenges we needed to navigate was helping teams to understand the type of work they were focused on and applying the right set of tools to help achieve their goal. For example, a sub process that needs to be redesigned requires a certain set of tools vs if the process needed to be completely redesigned. The need to apply these best practices has required "muscle building" and has met resistance. However, we have had several cases where the E4 has said, we like the vision, but we need the detail, or teams have spent months discussing strategy change and finally realized that they really didn't agree. The E4 has continuously shared that details matter and that "it is better to do it right the first time than having to do the work twice."

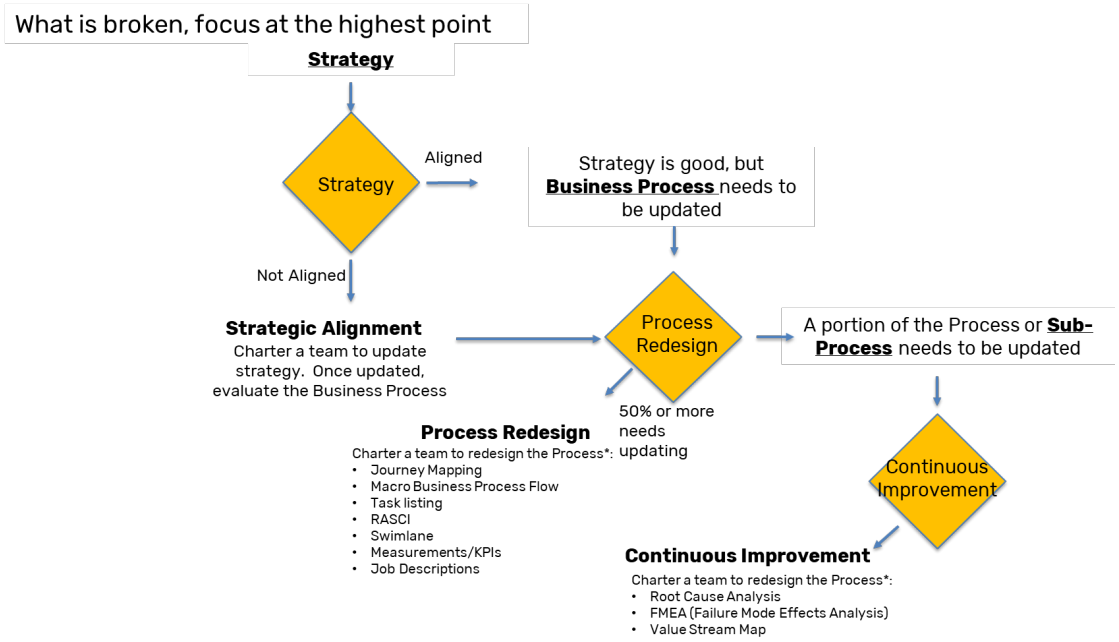


Figure 6 - What is the Right Approach

As a complement to each pillar's process work, they also worked on removing tool differences, ensuring that pillar KPIs are the same, developing talent strategies for each pillar and developing strategies to continuously improve employee satisfaction.

7. Building Trust and Partnership

To grow the headquarters and division pillar partnership, the E4 committed to regularly facilitating a partnership survey. The focus of the partnership survey is to ensure that headquarters and division Leaders are mutually living to our sameness principles and delivering what each other needs to be successful. So, we developed a short survey that as of the summer of 2021 has been administered twice to Director and above leaders. The survey results and the comments are reviewed openly with all partners and each pillar is asked to develop an action plan to improve results. The first survey surfaced several opportunities particularly around role and process. For example, we learned that division partners perceived that the headquarters agile sprint approach to development did not meet the planning and operational needs of their division partners. We also learned that a headquarters team's "big picture view" of customer outages did not feel customer-centric to their division partners, who are responding to the unsatisfied customer. Subsequently, we discovered that our division partner's need for data-driven specifics on what is changing was perceived as resistance to change. Virtually all organizations that have a division/region structure experience similar differences in experiences. What is different in this case is that we agreed to start myth busting to grow the partnership vs. ignoring it.

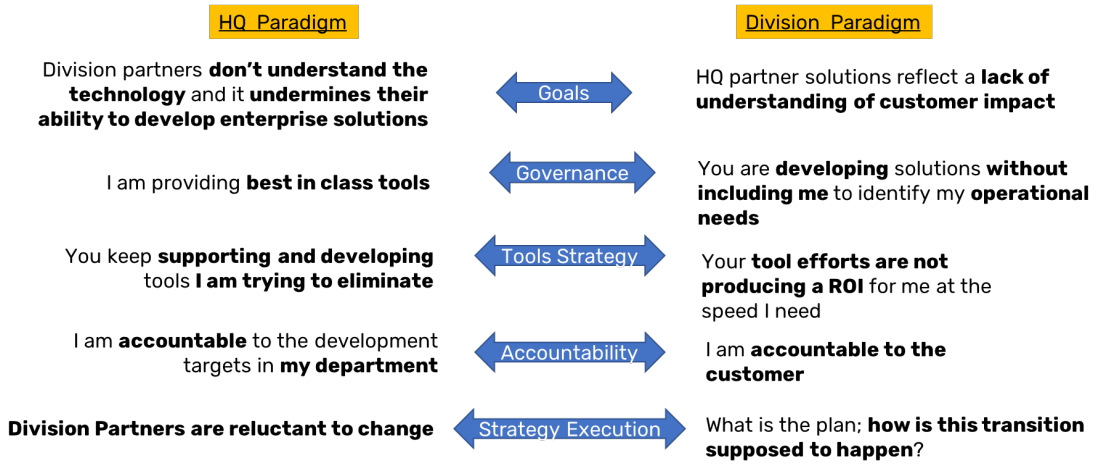


Figure 7 - Partnership Opportunities

8. Decision Protocol

In May of 2020, the Co-leads asked for clarity from the E4 about their decision-making protocol, the criteria necessary for the E4 to support a recommendation. Some of the co-leads expressed frustration with decision-making efficiency. In response the E4 developed a protocol to help facilitate decision-making. They perceived that there was an opportunity to help the teams better understand what they need, as leaders, to take the risk to change their processes. They felt that there were several case studies they could point to where decisions were expedited because the detail was fully developed and recommendations appropriately "stakeholdered." The E4 used these case studies to develop their protocol. Once they developed the protocol, they met with the Co-leads to share it with them. There were several themes in the protocol: 1) the foundational process, task and accountability level detail needs to be completed; 2) recommendations to the E4 need to be supported and presented by each of the four leads for the pillar (headquarters and 3 division Leads); 3) all recommendations should be stakeholdered with the E4 prior to the meeting; and 4) the E4 will not make decisions on the spot, they will confer and come back with recommended next steps.

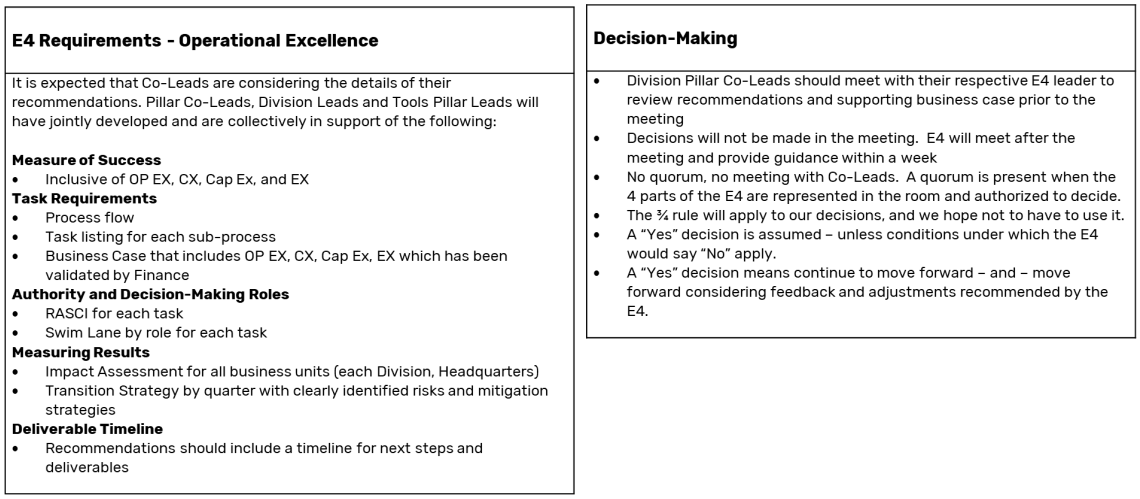


Figure 8 - Decision-Making Protocol

9. Creating our Future

To help build a high reliability culture that would propel our work, we agreed to create a "History of the Future." The state of any organization at some future time is a function of the interplay of three forces: its history, market events that are not in your control, and choices that a leadership team makes about its own future. A "History of the Future" exercise asks a leadership team to imagine their future success and identify the critical elements that enabled them to achieve success. (CFAR, 2003)

After understanding the partner survey feedback and meeting with each engineering leader, the E4 came together in March of 2021 to visualize their future and identify the critical elements of success. A primary theme that surfaced was an opportunity to increase VP ownership for the future. In addition, 6 future opportunities were revealed (below). To help build leadership and develop the future culture, pillar leaders were asked to develop recommended protocols that they could be implemented for each of the 6 future states below identified by leadership:

1. Create a protocol that would guide teams to design their vision, but incrementally implement with embedded opportunities for adjusting and readjusting to ensure reliability
2. Create a protocol to guide decision-making that enables an enterprise solution while caring for organization P&Ls
3. Design a feedback system that allows us to capture and respond quickly to front-line concerns and recommendations
4. Identify the operating practices we need to adopt, that will accelerate future development and allow us to pivot easily
5. Design a protocol for identifying talent within the divisions that we want to lift and shift and that provides relief for division P&Ls

6. Identify the practices we need to have in place to grow a culture that honors innovation without blame if it fails

In June 2021, the pillar leaders presented to the E4 recommended protocols for how the teams could move forward to create the culture and future the E4 visualized. In July 2021, the E4 Leadership supported all pillar lead recommendations and asked two Executive leaders to develop a strategy to implement the recommended practices as part of the reliability and quality domain.

10. Leapfrogging Our Talent in Front of the Technology Transformation

Early on, the E4 made a commitment to ensure we have the talent to manage our future network. In addition, they committed to ensure we grow our current talent into the future roles. To transition our talent, we developed a 4-phase process: 1) How We Operate Together; 2) Assess Change, Skill, Close Gaps and Align Talent; 3) Plan and Build Pipeline; and 4) Monitor Progress. Given that, the unification of tools and processes will drive enterprise-wide consistency. That automated and center-led systems will provide standardized workflows across the enterprise. We recognized that our engineers will become enterprise-wide subject matter experts who will problem solve across the system. In Phase 1 to support and proactively guide the talent needs across the Engineering organizations, we agreed to establish a single talent governance structure over the headquarters and division functions. To guide our work, we identified the following intention and design principles.

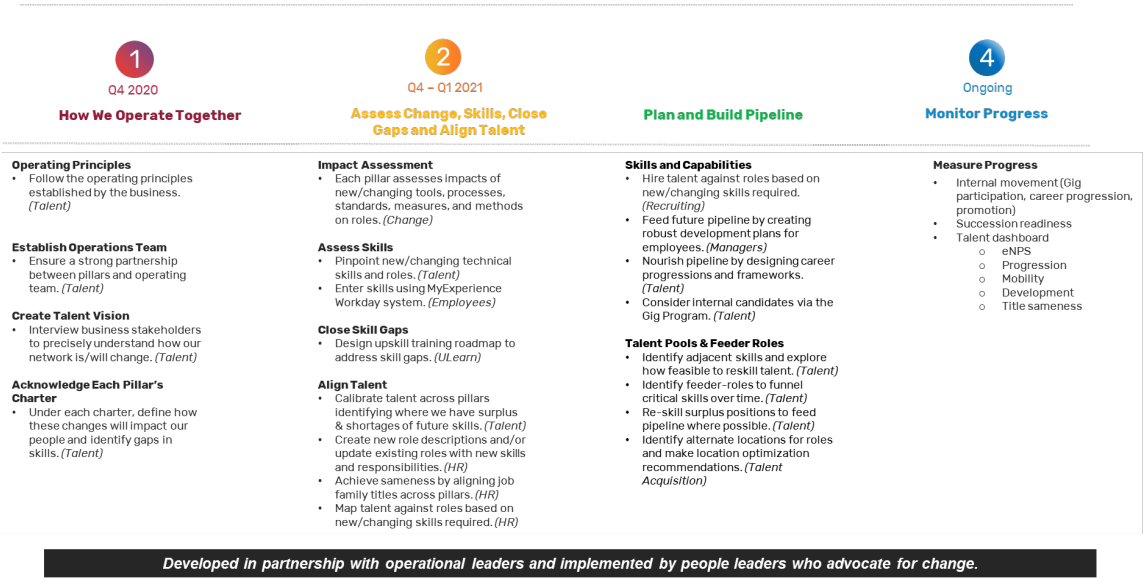


Figure 9 - Our Talent Approach

Our Intention:

- We are committed to enriching the careers of our people while elevating their confidence and motivation.

- We will achieve this by investing in career development opportunities while aligning role accountabilities across organization boundaries.
- This will allow us to mobilize our workforce with the right people in the right roles at the right time enabling organizational effectiveness and growth.

Our Design Principles:

- We align on the same titles, roles, responsibilities, & levels
- We design career movement the same and use the same standard approach
- We follow the same execution strategy (approvals, deployment practices, and communication methods)
- We don't act in silos
- We agree to use the same calculation to determine our Resource Needs

In Phase 2 we intricately are weaving in the updating of role profiles and career strategies with each process redesign. One of the first things we did was facilitate a title audit by function to understand how different we were across the organization. In Phase 3 we are developing pipeline strategies to ensure that as each new technology is rolled out, we have the talent to support it. We have completed the development of progression strategies for headend and engineering.

Our strategy is to resource new enterprise opportunities with current high performing talent from across the organization. Given that new technologies are still in their infancy and not widely understood by our front-line teams, we have developed a process for partnering with current managers to "tap high performers on the shoulder" to ask them if they are interested in new roles. This strategy was developed as part of our History of the Future work and has been highly successful. One of the bigger learnings in the development of this strategy was that we thought that normal recruitment processes would attract internal talent. We discovered that the information about new opportunities was not reaching the right talent and that there was not a clear understanding of the opportunities. Thus, the need to overtly reach out to the managers of high performers and mutually meet with them about future opportunities.

11. Evolving Operating Model

All of these elements combined are evolving into our new operating system. Fundamental to this new operating model has been establishing clear roles and a governance model for process changes.

In response to the Partner Survey, we focused on putting in place critical elements

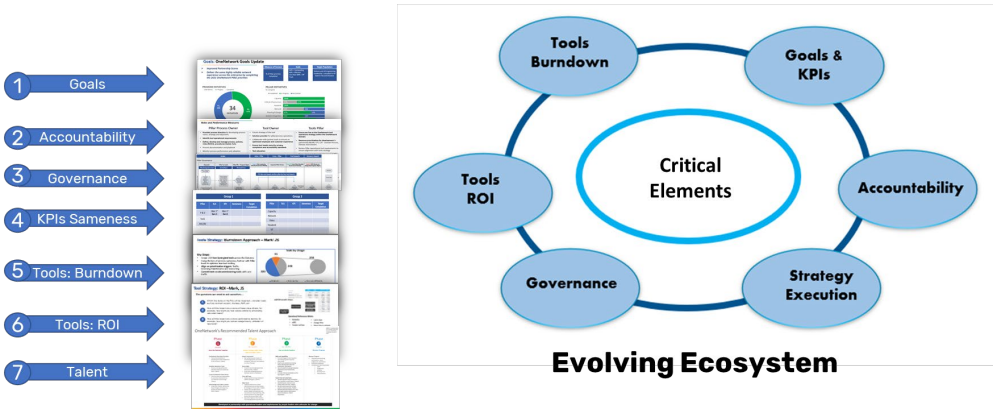


Figure 10 - Our Evolving Ecosystem

Accountabilities: To help clarify roles we adopted a business process management model. In this model **the Business Process Owners** come together to define the common workflow process, KPIs and creates a single set of requirements for tools. They are responsible for the decommissioning of legacy tools to ensure there is a single set of tools for each pillar. They are on point to monitor and continuously improve process performance.

The **Tools Owners** are solution providers. They develop, integrate, and operate the tools. They are on point to make sure our tools meet security, privacy, compliance, and accessibility standards.

The **Tools pillar Leads** coordinate and operationalize our tools strategy. They facilitate resource prioritization and manage intra tool dependencies. monitor tool decommissioning.

Roles and Performance Measures

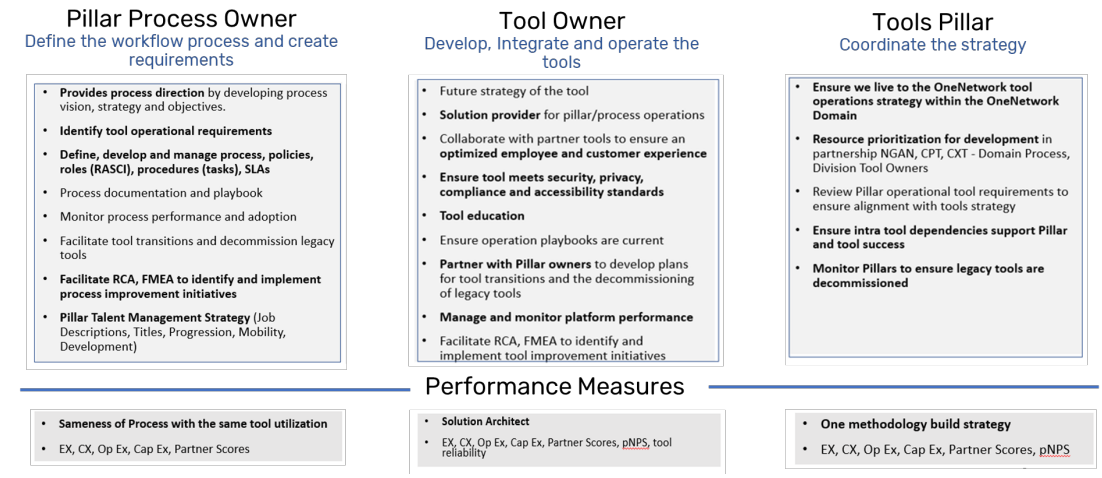


Figure 11 - Business Process Management

Clarifying roles and accountabilities has been critical. It has forced process owners to take up their responsibility to be clear about their process strategy. It has also been clear that it is their responsibility to live and enforce the tools strategy. Previously the tools owners were left to negotiate tool requirements across the multiple headquarters and divisional process owners. In addition, to support innovation, they were allowing independent tool development rather than coordinating pilots with tools partners.

KPI Reporting: We have established and resourced a **Reporting pillar** to ensure that we drive KPI sameness and primary source reporting. We have recently reached agreement that we will grow into a single reporting group.

Governance: To sustain our gains, we have established a **pillar governance model** that we are growing into. In this model we continue to position Process Owners as accountable for process design. However, to ensure integrated development across the functional pillars it forces process owners to gain support from other Pillars, and tools owners as well as the tools pillar support before making any changes to agreed-upon processes. Once any recommended change has support from all stakeholders it is presented to the E4 for final approval.

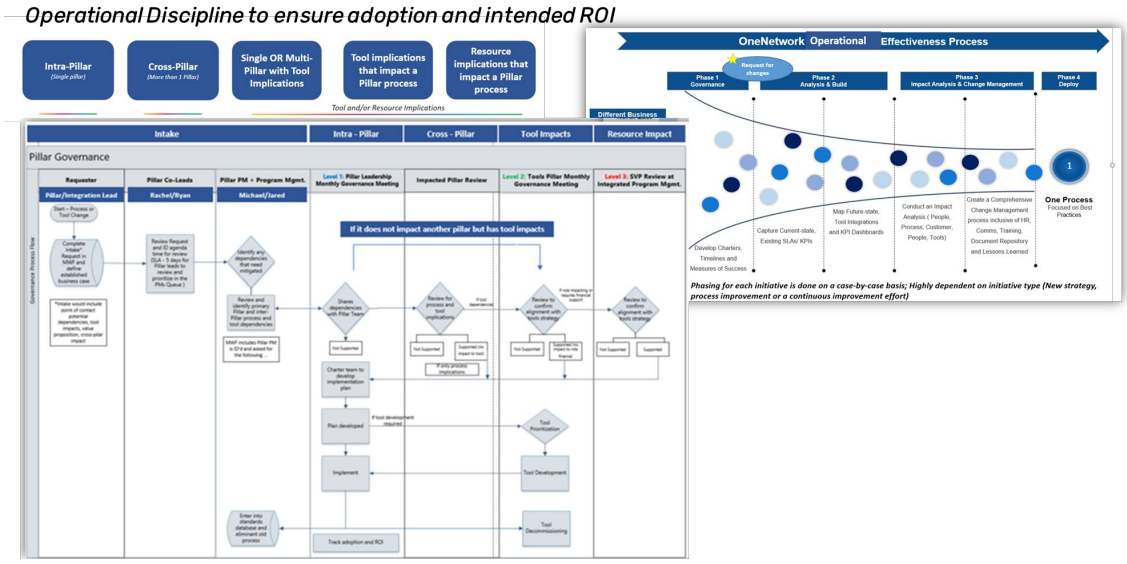


Figure 12 - pillar Governance Model

12. Program Maturity & Results

To date (summer 2021) this initiative has chartered and launched 66 initiatives in 2021. To date we have a 94% completion rate.

More importantly, we have a better understanding of where we are not the same, have clear accountabilities and agreement that sameness is our mutual goal, and have a status report of sameness maturity by pillar for Process, KPI, Tools and Talent.

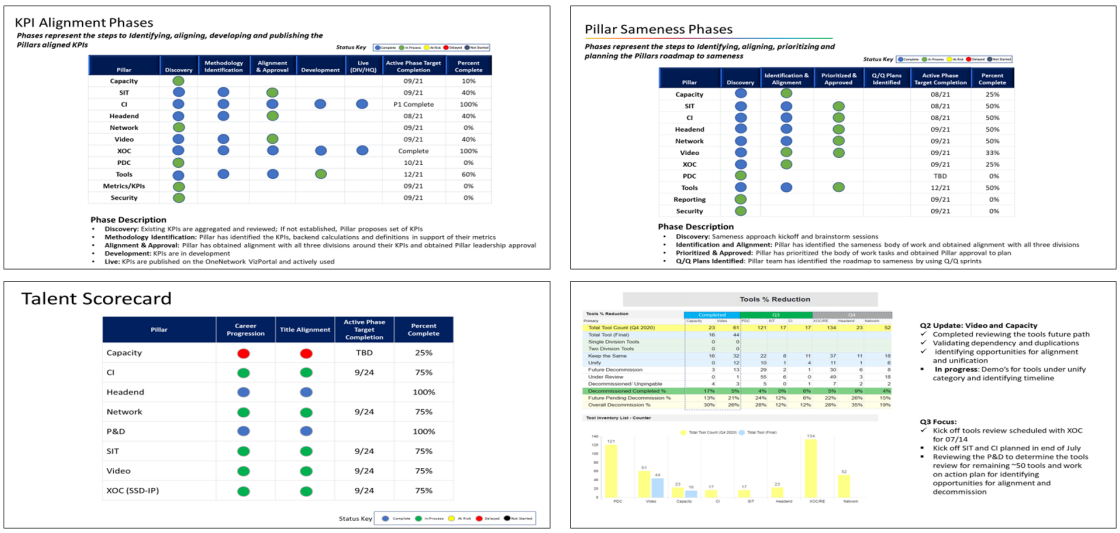


Figure 13 - Maturity Status

In addition, we are seeing improvement in our Partnership Survey scores.

13. Conclusion

Our experience over the past year has reinforced critical organization design practices that we learned we shouldn't take for granted. As a team we have reconfirmed the value of these practices and that we need to reinforce the need to apply them in everything we do.

Practice #1: Organize for the future you want. We reminded ourselves that if we wanted sameness, we needed to organize for sameness and have a set of principles that leadership was committed to.

Practice #2: Start with Sameness. We reminded ourselves that we needed to plan for strategy execution prior to design, that it could not be an afterthought. We needed to understand the variability of process prior to design and ensure all stakeholders included in the development of design requirements if we want to have a single playbook.

Practice #3: No process is an island. We reminded ourselves that Process design requires a cross pillar approach. Process design needs to consider the hand-offs between Pillars and tool integration points. For example, when redesigning our node split and headend processes, we need representation form multiple other Pillars at the table.

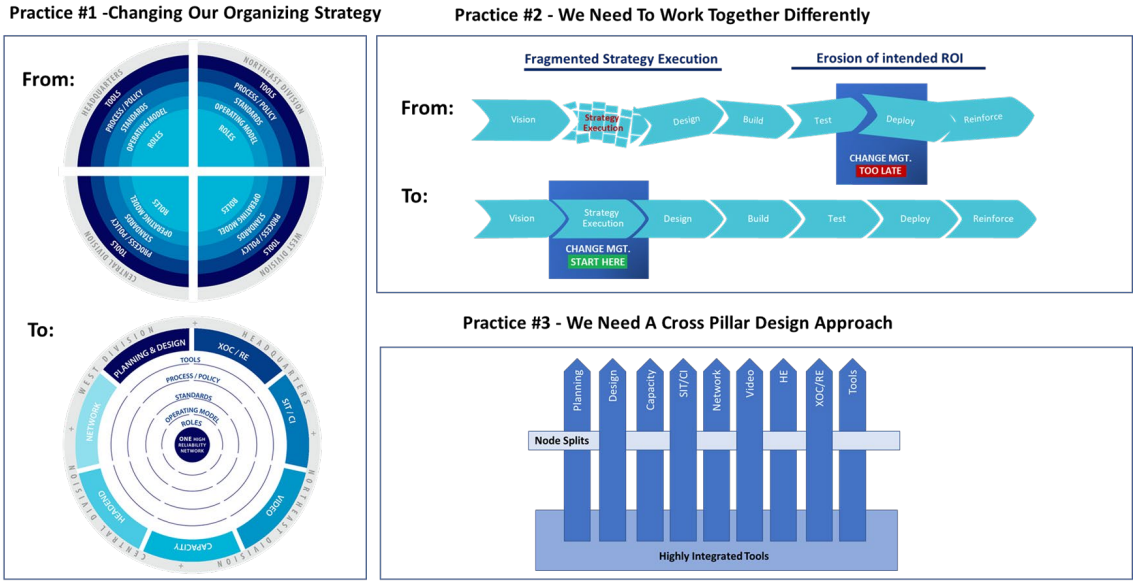
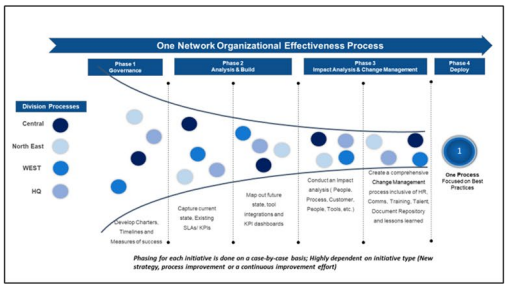


Figure 14 - Evolving Practices

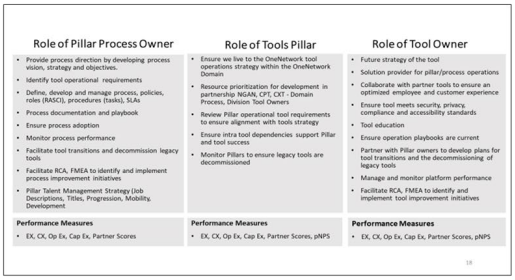
Practice #4: There is a science to organization design. We reminded ourselves that if we want a highly reliable system, we need to use the right tools. We needed to recommit to operational excellence and the tools that would enable us to ensure clarity of process, roles, and accountabilities. As we evolved as a leadership team, we began to realize that when we thought we had commitment, we really didn't because we were not discussing the process with enough detail on the table. As a team, we needed to learn that detail was our friend, not a burden that made the decision-making process longer. It reduced rework. We also began to realize that our leaders needed the detail to feel comfortable with the changes we were asking them to make. So now we do detailed macro process maps, a "swim lane" view of task listings and role and accountability charts for each task.

Practice #5: Role clarity is essential. We reminded ourselves that not everyone is clear about their roles. That we need to take the time to be overt about our expectations of each other and respective decision rights. This practice has been applied in several areas. We applied this practice when chartering each pillar, the role of pillar leaders, and roles within each process. But we also discovered that it was important to identify the roles of multiple levels of leadership. We clarified that executive leaders focus on a 3–5-year planning cycle. Senior Vice Presidents focus on the operationalization of strategy with a 2–3-year planning horizon and Vice Presidents focus on the near-term operationalization of strategy. Prior to this, our leaders were waiting for decisions to be made rather than feeling authorized to come together to develop and drive strategy.

Practice #4 – We Need to Adopt Operational Effectiveness Process/Tools



Practice #5 – We Need Role Clarity



Practice #6 – Tight Coupling is the Key to Reliability

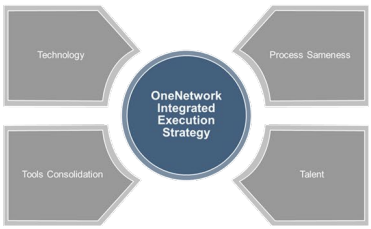


Figure 15 - Evolving Practices #2

Practice #6: Tight coupling and appropriate sequencing of process, tool, technology, and talent is critical to realizing consumer reliability. We reminded ourselves that we always need to start with the process.

We see ourselves as learning how to be a learning organization. We are actively developing the culture we need to grow into the phase of our future. We are looking forward to seeing how we can grow into our new practices.

Abbreviations

CI	critical infrastructure
E4	Engineering Leadership across 3 Divisions + Headquarters
HQ	headquarters
IPM	Integrated Program Management
KPI	key performance indicator
RASCI	Roles, Accountabilities, Support, Consult and Inform
RE	reliability engineering
ROI	Return on investment
SCTE	Society of Cable Telecommunications Engineers
SIT	systems integration
XOC	Excellence in Operations Centers

Bibliography & References

CFAR; History of the Future, 2003

Wireless IoT for Rural Use Cases

A Technical Paper Prepared for SCTE by

Joerg Ahrweiler
Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+15613065021
Joerg.Ahrweiler@charter.com

Mohamed Daoud
Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+13123639864
Mohamed.daoud@charter.com

Muhammad Khan
Senior Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+17205361578
Muhammad.J.Khan@charter.com

Hossam Hmimy
Senior Director Wireless R&D
Charter Communications
6360 S Fiddlers Green Cir, Greenwood Village, CO, 80111
+17204049716
Hossam.hmimy@charter.com

1. Abstract

Smart cities and communities is a concept that is gaining more attention because of its potential to enhance citizens' lives and make cities and communities more efficient. Rural America can also benefit from leveraging new technologies like IoT (Internet of Things) and Fixed Wireless Access (FWA) to achieve better operational efficiencies.

In this paper we present findings in a Proof of Concept (PoC) project that demonstrates how technology can help an equestrian farm transform into a smart farm. We discuss the details of connecting the farm with high-speed broadband using FWA in the Citizens Broadband Radio Service (CBRS) frequency band then meshing the barn with high-speed Wi-Fi using Point-to-Multipoint unlicensed band.

Within the project, we also deployed LoRaWAN technology to provide early-warning detection of potential illness in the horses on the farm through the use of a water intake monitoring solution in addition to LoRaWAN environmental sensors installed to monitor the horses' surroundings. Another use case is deploying a mix of LoRaWAN and video analytics to detect unauthorized access to facilities.

2. Introduction

The Wireless Technologies R&D team at Charter is executing a PoC trial at a local Colorado Equestrian Farm ("Farm") to ascertain the viability of IoT technology through a variety of Smart Farming use cases, using multi-access technologies, including access technologies such as LoRaWAN, FWA in CBRS, Wi-Fi, and application technologies like video surveillance and license plate recognition.

The primary goals of the Smart Farming use cases include:

- Providing early warning detection of potential illnesses in the horses on the farm through the use of a water intake monitoring solution.
- Improving security on the farm through LoRaWAN sensor monitoring and alerting and video surveillance.
- Providing Spectrum Internet ® Wi-Fi connectivity at the Farm

In this paper, we will describe the Farm layout, which is essential in the design of the system, followed by the logical network setup description. Both the CBRS main FWA connectivity to the Farm as well as the wireless mesh connectivity to the different barns are detailed in the next section along with the broadband use case of video surveillance. Next, we will describe the IoT network based on LoRaWAN and the associated use cases for the farm. Finally, we will conclude with the summary and next steps in the research and PoC.

2.1. PoC Venue – Equestrian Farm

The PoC venue is an Equestrian farm outside of the Denver area. Overview:

- 120 Acres
- 55 Total Horses
- 3 Barns – used to stable horses (includes both leased stalls and leased horses)

    - Main barn attached to indoor training arena
    - Upper barn located 200 feet southeast of the main barn
    - Lower barn located .25 miles south of main barn
- Large indoor training arena attached to main barn
- Outdoor training arena
- Outdoor horse walker
- 4 Pastures
- Security gate at entrance with minimal security
- Farm is set up on a hill with line of site to a CBRS FWA RAN (Radio Access Network) at Charter's Spectrum's CTEC office in Englewood

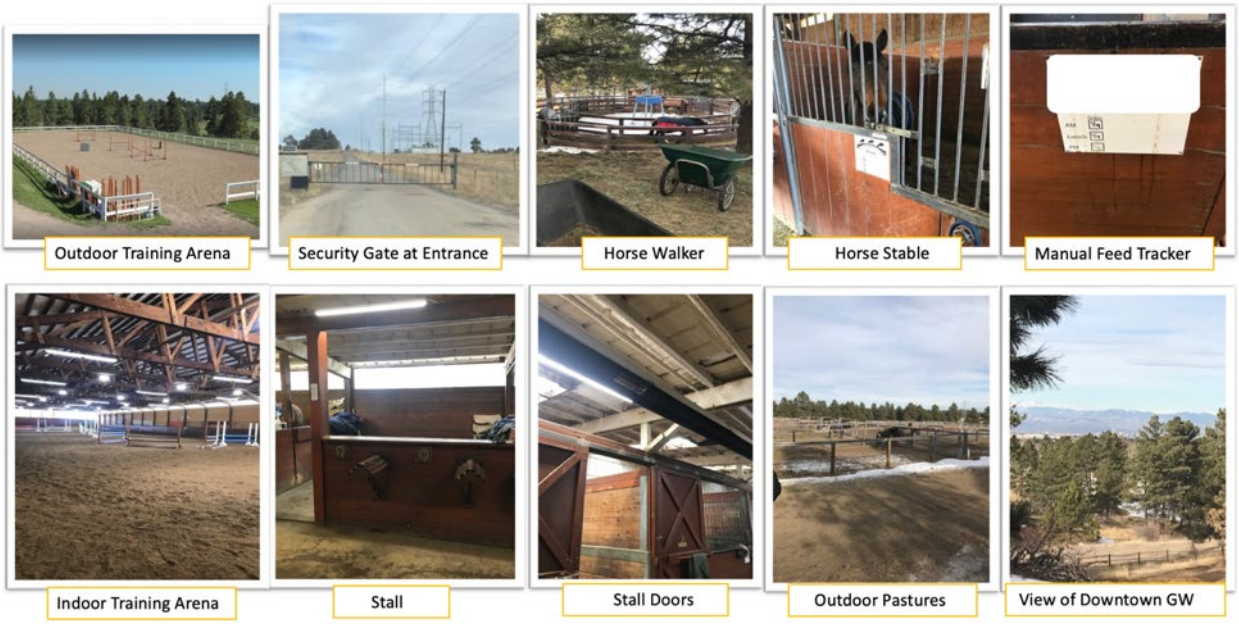


Figure 1 - Farm areas

3. Use Cases and Utilized Solution

The following figure depicts the logical network setup.

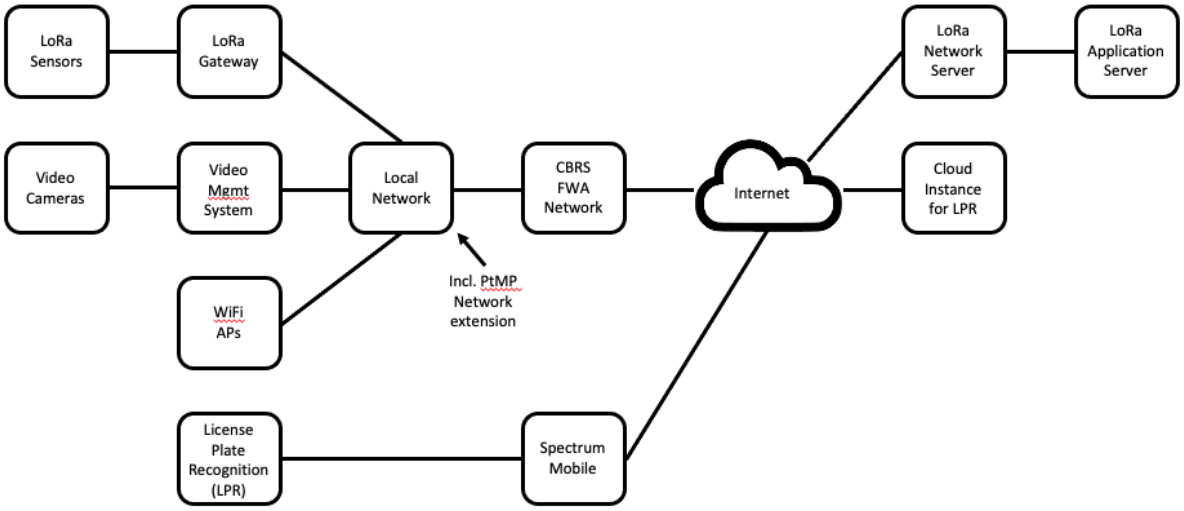


Figure 2 - Logical Network Setup

3.1. Internet Backhaul – CBRS Fixed Wireless Access

As a main 'anchor' to the internet, a Fixed Wireless Access link to the local Charter CBRS test network is utilized. The receiving eNodeB is located at Charter's CTEC1 building (5.5 miles away). Since line-of-sight conditions are available, the achievable throughput performance is close to ideal.

The 'FWA access link' is primarily used to transport traffic to the LoRaWAN network and application server, end-user internet backhaul via the Wi-Fi access point and other equipment cloud portals, but also provides the option to retrieve video streams from the local video management system. The License Plate Recognition system will have internet connectivity via Spectrum Mobile™ service.

Due to the superior Radio Frequency (RF) conditions and achievable performance results, the link to CTEC was selected.

Main Barn to CTEC Tower					
PCI	Distance	SINR	RSRP	Download	Upload
120	5.5 Miles	26 dB	-92 dBm	240 Mbps	11 Mbps

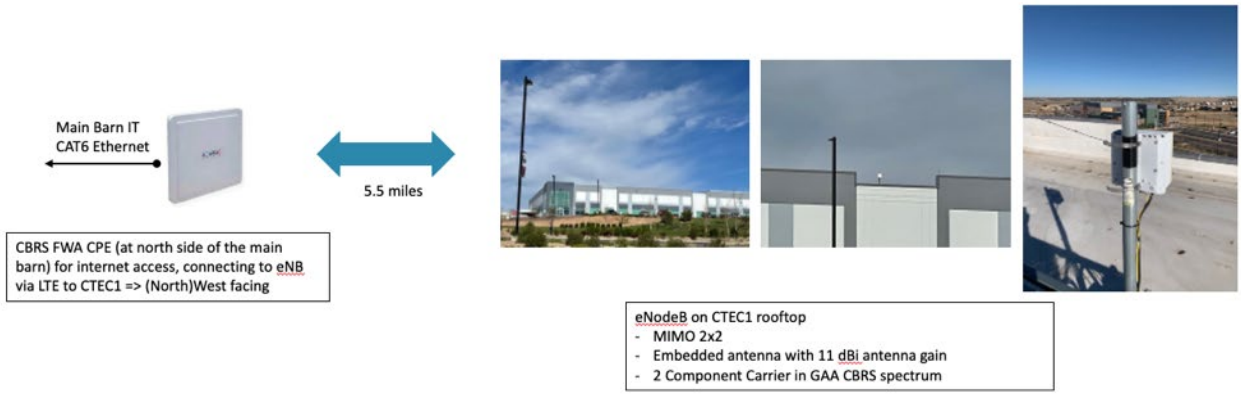


Figure 3 - eNodeB deployment at Charter CTEC1 building

3.2. Local Network Extension and Access for Farm Owner and Customers

Due to the distribution of Farm buildings – main/upper barns and owner's house – a wireless extension of the local network is required, mainly to achieve the following:

- Transport of the camera feeds from the upper barn stalls to the video management system (VMS) at the main barn
- Remote Wi-Fi access point at the farm owner's house to be able to access the system

For that purpose, a Point-To-Multipoint unlicensed 5GHz wireless system was chosen. The Base Station is mounted at the main barn and connected to the main distribution network switch. The subscriber units are installed at the upper barn and the farm owner's house.

In addition to the Wi-Fi AP at the farm owner's house, another unit is deployed inside the main barn for access to the network.

3.3. Video Surveilance

For the purpose of giving the option to the Farm owner and farm personnel to visually monitor designated areas in the farm complex, the following areas were equipped with cameras and connected to a video management system with integrated DVR:

- 4 horse stalls in main barn
- 5 horse stalls in upper barn
- Indoor training arena (4 cameras total) in the main barn
- Outdoor facing horse walker

3.3.1. Use Case 'video surveillance horse stalls':

Farm owners and personnel have the ability to remotely monitor the conditions of select horses for improved welfare and security. In the future, this could be extended to provide the same access to the horse owners. Additionally, processing the captured and stored video feeds allows for implementation of further analysis and detection of unusual horse behavior via AI (Artificial Intelligence) and ML (Machine Learning).

Figure 4 - Horse Video Surveillance

3.3.2. Use Case 'video surveillance indoor training facility':

Farm owners and personnel have the ability to detect abnormal behavior or 'not permitted' jumping based on time (after hours), in combination with installed LoRaWAN proximity sensors, and general remote monitoring of activities.

Figure 5 - Training Facility Video Surveillance

3.3.3. Use Case 'video surveillance outdoor horse walker':

Farm owners and personnel have the ability to remotely monitor activities at the outside horse walker, in combination with LoRaWAN sensors, in order to detect equipment malfunctions, horses escaping or other general remote monitoring of activities.

Figure 6 - Horse Walker Video Surveillance

3.4. LoRaWAN® Network and Sensors

For the purpose of monitoring conditions and environments in the Farm area, a LoRaWAN network setup consists of various sensors and an outdoor gateway, interconnecting via the internet to a network and application server. Sensors that are part of the solution include:

- Open/close door for monitoring the selected horse stalls
- Temperature and humidity in main, upper and lower barns and TAC room (contains equipment for the solution)
- Proximity sensors for training facility and outdoor horse walker (use cases described earlier)
- Water consumption monitoring. This sensor solution consists of a water meter, which is in line with the water supply and the water bucket for the horse (automatically refilled via a floating valve) and connected to a LoRaWAN pulse sensor.

Figure 7 - Water meter and LoRa pulse sensor solution

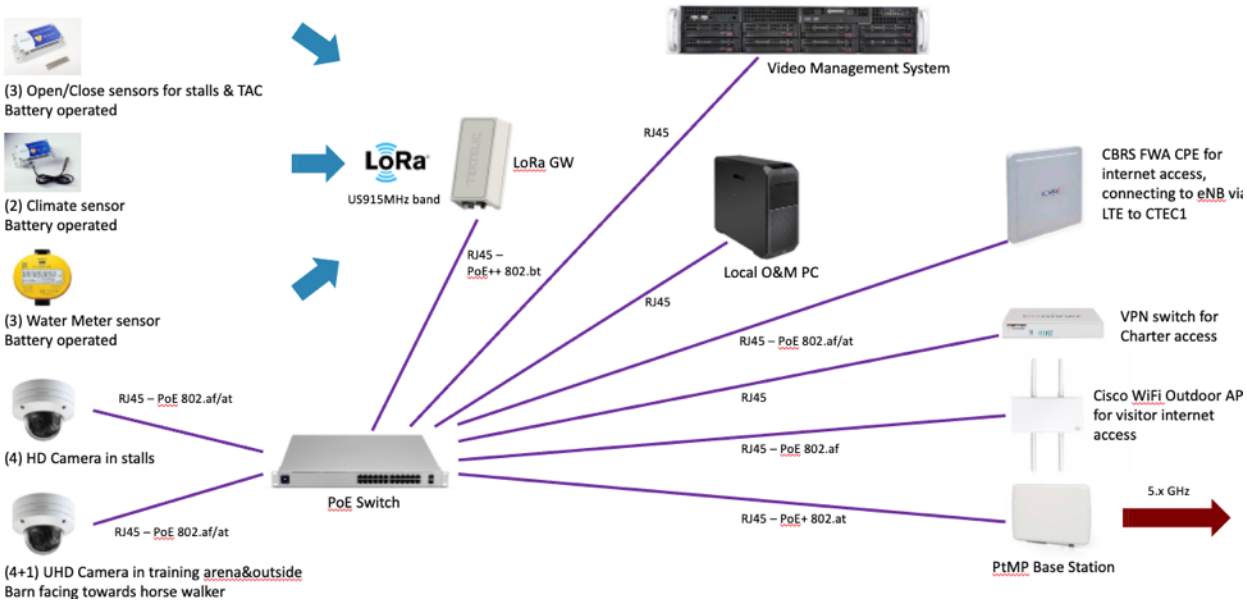


Figure 8 - Connectivity diagram example, here in the main barn

3.5. License Plate Recognition at Gate Entrance

The purpose of the license plate recognition (LPR) is to provide the Farm owner with the ability to monitor and track access, with dates and times, of visitors for potential later proof of entrance for those who entered the Farm area. As additional enhancement for the future, the LPR system could be linked to the gate control via a white/black list control.

Figure 9 - Example of outdoor LPR install

Figure 10 - Location of gate and existing poles

The LPR system is installed in close proximity to the gate, which is approximately 0.5 miles away from the Farm building complex.

The processing of the license plate recognition is done at the edge in a self-contained battery and solar powered system. The processed data is sent to a cloud portal via a cellular (using Spectrum Mobile service) link.

4. Conclusions and Next Steps

The proof of concept activities have already demonstrated that the use of hybrid wireless technologies combined with data processing at the edge and in the cloud is extremely helpful in a rural farm environment for its day-to-day activities. The trial will continue and learnings from the end users - farm owners and personnel - will be used to improve the overall solutions. Additionally, other use cases or enhancements to the existing use cases will be made as needed.

Abbreviations

BH	Back Haul
CBRS	Citizens Broadband Radio Service
CBSD	Citizen Broadband Radio Service Device
CPE	Customer Premisis Equipment
EPC	Evolved Packet Core
FWA	Fixed Wireless Access
GAA	General Authorized Access
IoT	Internet of Things
LoRaWAN	Long Range Wide Area Network
LOS	Line Of Sight
LPR	License Plate Recognition
LTE	Long Term Evolution
MIMO	Multiple Input Multiple Output
PCI	Physical Cell ID
PoC	Proof of Concept
PoE	Power over Ethernet
PtMP	Point to Multi Point

RAN	Radio Access Network
RF	Radio Frequency
RSRP	Reference Signal Received Power
SAS	Spectrum Access System
SINR	Signal to Interference plus Noise Ratio
UE	User Equipment
VMS	Video Management System
VPN	Virtual Private Network

Bibliography & References

• White Paper 'The CBRS Opportunity - The Wireless Infrastructure Association' by the Wireless Infrastructure Association

https://wia.org/wp-content/uploads/CBRS-Paper-3-20-20.pdf

• Technical Paper 'ON THE PERFORMANCE OF CBRS FIXED WIRELESS ACCESS: COVERAGE AND CAPACITY FIELD STUDY' prepared for SCTE•ISBE by Mohamed Daoud and others

https://www.nctatechnicalpapers.com/Paper/2019/2019-on-the-performance-of-cbrs-fixed-wireless-access

• White Paper 'THE FARMING OF TOMORROW IS ALREADY HERE HOW LoRaWAN® TECHNOLOGY SUPPORTS SMART AGRICULTURE & PRECISE ANIMAL PRODUCTION'

https://lora-alliance.org/resource_hub/the-farming-of-tomorrow-is-already-here-how-lorawan-technology-supports-smart-agriculture-precise-animal-production/

• 'Basics of LoRa Technology for Crop and Livestock Management' by John Nowatzki

https://www.ag.ndsu.edu/publications/crops/basics-of-lora-technology-for-crop-and-livestock-management

• 'Agriculture-Vertical-Market' – LoRa Alliance

http://pages.lora-alliance.org.pages.services/Agriculture-Vertical-Market/?ts=1601118451908

• 'What is a smart city' – Spectrum Enterprise

https://enterprise.spectrum.com/support/faq/smart-cities/what-is-a-smart-city.html6-
https://www.sqlite.org/index.html

xGitGuard: ML-based Secret Scanner for GitHub

A Technical Paper prepared for SCTE by

Bahman Rashidi
Senior Security Architect
Comcast Cable
Philadelphia PA
bahman_rashidi@comcast.com

1. Abstract

Misused leaked secrets on code sharing platforms such as GitHub (GH) have caused some of the data breaches of our time. Unfortunately, this kind of credential leak is quite common across the code sharing platforms. Developers and code contributors are required in many cases by organization's security policies to comply with security practices and remove sensitive information before they push their code to GitHub. However, sometimes inadvertently developers neglect to remove sensitive information, such as API tokens and user account credentials, from their code prior to posting it. Malicious attackers crawl through GitHub, hoping to find these secrets and thus grab foothold into an organization's territory. Companies have limited ability to address this risk as given the scale of GitHub it is difficult if not impossible to find leaked secrets before malicious attackers. Some companies leverage bug bounty programs as a way to incentivize third party agents to manually look for and report these secrets through responsible disclosure. Unsurprisingly, this process can create unnecessary exposure. Consequently, we at Comcast Cybersecurity Research designed and developed "xGitGuard,"a Machine Learning (ML)-based tool that uses advanced Natural Language Processing (NLP) to detect organizational secrets and user credentials at scale and with appropriate velocity in GitHub repositories. This paper begins with a description of the problem statement. Next, we discuss the design of xGitGuard and how it improves upon current solutions, and the solution space. Finally, we provide details about how xGitGuard can be deployed in different scenarios.

2. Introduction

GitHub is the biggest open-source community with more than 200 million repositories (and of those, 30+ million are public) and over 65 million users [1] [2]. Users on GitHub publish the code of their projects, collaboratively develop software, and use the code from the platform in development. Given the open nature of GitHub, there is an opportunity to publicly disclose otherwise confidential data. [3] [4]. For example, a project may disclose login credentials for remote servers and service accounts or API tokens if such information is embedded in code distributed on the platform.

Although the risk of disclosure on GitHub and Stackoverflow is known, it is not known the extent of confidential information disclosed, and how efficiently that information is identified by attackers.

In this paper, we describe xGitGuard, an ML-based tool that detects exposed organizational secrets on GH both the public and enterprise versions. xGitGuard is designed and developed with the goal of addressing the existing challenges associated with classic regex scanning approaches (solutions relying on finding regular expressions). xGitGuard takes advantage of new text processing algorithms that can find secrets within files with high level accuracy. xGitGuard scans the entire GH for secrets efficiently using an agile scanning search approach. The agility of xGitGuard helps incident response teams to take proper actions in timely manner. In next sections, we detail how it works and can be deployed.

The rest of this paper is organized as follows: we introduce an overview of xGitGuard in Section 3, alongside a workflow overview. We then further detail the components of the xGitGuard and how they work with each other. We then discuss the related work in Section 4. We finally discuss different deployment scenarios for xGitGuard in Section 5.

3. xGitGuard

xGitGuard is an ML-based tool that detects organizational secrets, including API tokens/keys and user credentials exposed on both public and enterprise GitHub. xGitGuard has two separate models: one for detecting credentials, such as password-like secrets, and one for detecting API tokens and keys. Table 1

shows examples of such secrets. It is designed to be both scalable and accurate in scanning and detecting secrets. The rest of this section details the architecture of xGitGuard and all of its components.

Table 1 – Examples of secrets for each category

Credentials	Tokens & Keys
Username & passwords	API tokens (AWS, Azure, Slack, etc.)
Server credentials	Encryption keys
Account credentials	Session tokens
DB access credentials	Session IDs

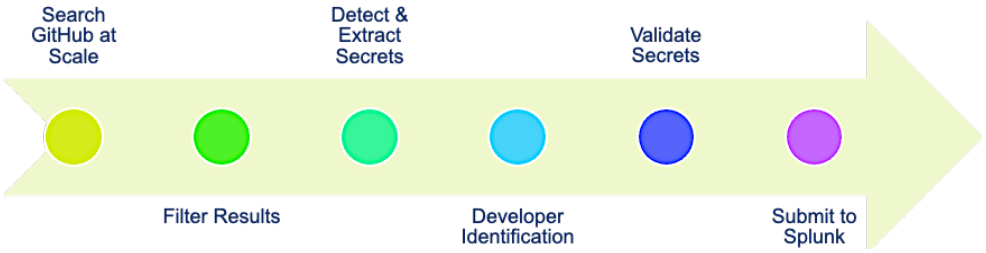


Figure 1 – xGitGuard workflow overview

3.1. Architecture

xGitGuard scans GH repos for documents potentially related to an organization and containing secrets. It also processes candidate documents for secrets, identifies developers using commit history, validates detected secrets using an ML component, and then calculates a confidence level for every detection to submit to a dashboarding tool. Figure 1 shows an overview of xGitGuard's overall workflow. The implementation of xGitGuard includes six main components. Figure 2 and Figure 3 show overview architectures of xGitGuard's workflow for credential and token detection models:

Search: xGitGuard has a unique approach to searching GitHub: it uses two types of keywords to craft GitHub queries, each for a different purpose. i) Primary keywords (PKEY): helps to search for documents that are related to the organization. ii) Secondary keys (SKEY): are then used to target documents that potentially contain secrets. xGitGuard uses two different lists of secondary keywords for credential and token detection. The two lists are deliberately different, as each will detect different types of secrets (credentials and tokens). However, the primary keywords are the same between both. With this unique approach, we scan the entire GitHub but only target documents that are relevant and sensitive.

Query Engine: This engine implements a multi-processing system that runs multiple queries in parallel to reduce the time to detect a secret. This helps with scalability. xGitGuard also maintains a hash list of documents that have been processed in the past and skips them in the future scans if they show up in query results. This way a file will be processed once, which reduces the processing time in total.

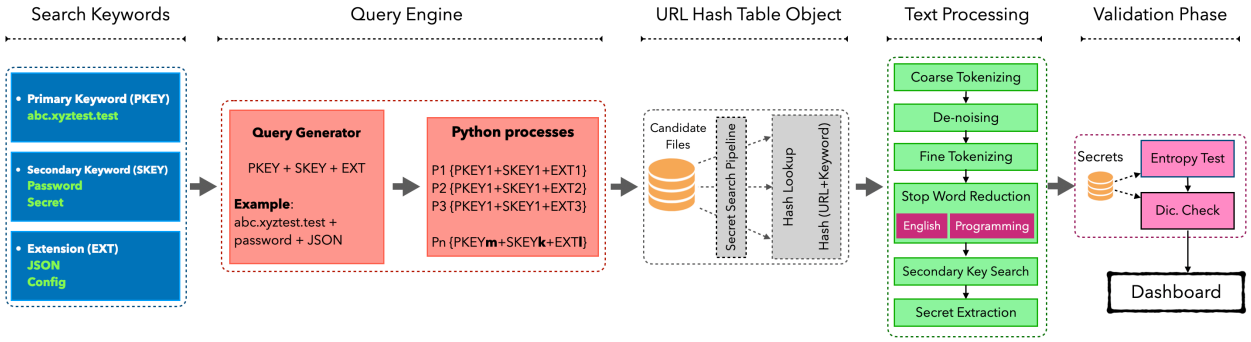


Figure 2 – Credential detection architectural overview

AI Model: This is the core of the xGitGuard, that processes the documents for secrets. This component is responsible for detecting the secrets within the documents. The core model of xGitGuard uses NLP and other in-house developed text processing algorithms to identify secrets. The model begins by breaking down documents to smaller tokens, then it removes the noise and stop words (meaning common English dictionary words, such as "the," "a" and "an," that are mutual terms/words among different languages, etc.), extracts the secrets and extra metadata (e.g., masked secret, line of code, etc.) around the detection. As Figure 3 shows, the only difference between credential and token detection models is that token detection model relies on using regexes (regular expressions) after de-noising the document.

Scoring: After detecting a secret in a document, this component calculates a confidence score for it. The higher the score, the more accurate the detection. This score is calculated based on several factors, such as the entropy of secret keys, n-gram similarity to English words, etc.

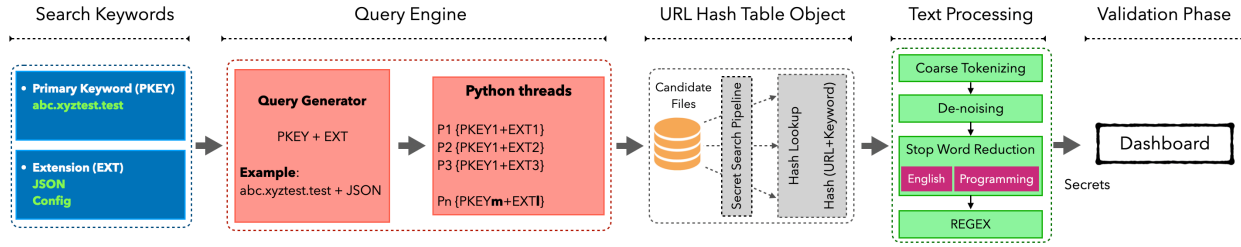


Figure 3 – Key & Token detection architectural overview

Validation Model: This ML-based model is responsible for validating detected secrets. The input to the model is a secret, and the output is whether the input is an actual secret or not. The model is built using a number of features, with the focus on the secret itself and the line of code where the secret was found. This model has an accuracy of above 90% in recognizing secrets from non-secrets text. The accuracy of the validator depends on how well we train the model and how comprehensive our training data is.

User Identification: xGitGuard is able to identify the developer who posted the code on GitHub, using commit logs related to the repos. This component extracts the email address and full name related to the developer.

3.1. Development and Deployment

The entire xGitGuard code base is developed in Python. This includes our algorithms, APIs, scanning and pre-processing documents, as well as submitting detections to dashboards. In order for xGitGuard to

perform better, all the text processing models and algorithms are persisted in the form of pickle files (serialized object models). This helps to avoid retraining our models every time a new cycle of xGitGuard starts. It also enhances agility and eases the process of deploying the final model. The modularity of pickle files helps to deploy the final model on many platforms.

4. Related Technologies

The problem of detecting leaked secrets within GH and other similar code-sharing platforms (e.g., Stackoverflow) is not a new problem in the information security community. In the past years, a number of tools from both the open source and commercial sectors tried to address such problems. However, existing solutions have some challenges, in particular scalability and accuracy.

There are three main limitations related to the existing tools. The first is that most of these tools, and especially the open source versions, exclusively focus on a single use case (they either focus on API tokens or passwords, not both). The second limitation is that almost all the solutions discussed here rely on pattern detection (regular expressions). This limits the types of secrets that are covered and only detects certain types of secrets. Finally, these approaches require users to point the tools to specific repos that they want scanned, because they lack scalability to scan the entire public or enterprise GH.

4.1. Truffle Hog

Truffle Hog is a free and open source tool that can help developers check for any hard-coded secrets. Truffle Hog is designed to scan for secrets through repositories (users need to point the tool to specific repos in order to scan them) and the entire commit history. The tool specifically searches for each diff (differences between file content) from each commit and evaluates them for secrets [5]. The Truffle Hog's core model relies on detecting high-entropy strings (average number of bits per symbol needed to encode a string) that could represent secrets, whether API tokens and keys or other types of credentials [6].

4.2. Nightfall for GitHub

Nightfall for GitHub is a commercial variant of similar tools that uses ML to detect secrets exposed on GitHub repos. Similar to Truffle Hog, this solution only works on manually-selected repos. Unlike Truffle Hog, this tool can be integrated with GH accounts in the form of an app and can automatically detect secrets as a user pushes code to repositories. This tool is an extension of Truffle Hog in terms of the approach to secret validation [7]. The main focus of Nightfall is to detect API keys and tokens and omits passwords and other credentials as a category of leak.

4.3. Gitguardian

Similar to Nightfall, Gitguardian is a commercial tool with a limited amount of publicly-available information. It describes the tool as ML-based and able to identify more than 200 types of API tokens. However, similar to Nightfall, Gitguardian exclusively targets API tokens and keys by relying on regular expression classifiers. Alongside the commercial variant that it offers, Gitguardian also provides a free service to scan one's repos on GitHub [8] [9].

4.4. EarlyBird

EarlyBird, developed by American Express, is another sensitive information detection tool relying on pattern detection (regular expression). EarlyBird exclusively scans repositories for clear text password violation, personally identifiable information (PII), sensitive file names and outdated cryptography

methods, etc. It functions on committed code (GH repos) and local files. Similar to Nightfall, it can also work as a pre-commit check. Besides the regular expressions that are used for secret detection, EarlyBird also uses entropy calculation for password detection [10].

5. Deployment

xGitGuard is a standalone and internally-developed and used application. Depending on the use case, it can be internally deployed and utilized in different ways. In this section, we discuss some of the ways that xGitGuard could be deployed.

5.1. Continuous Scanning

One way to deploy xGitGuard is to run it as a stand-alone application continuously running and scanning GitHub repos. This way, xGitGuard will continuously query GH in a cycle, receive documents and process them for secrets. This process continues until stopped manually. Using this approach, a longer time should be expected for the first cycle, as all the received files are new to xGitGuard and they all will be processed. In the second cycle and after, the previously processed files will not be processed (unless they have been modified) and it will take shorter time for xGitGuard to process the queries.

5.2. Git Hook

In this approach, xGitGuard is used as a hook integrated with GH. This approach will specifically be applicable on GitHub enterprise (GHE). The hook is responsible for intercepting new commits and processing them immediately for any hardcoded secrets. Depending on security operations center's (SOC) strategy, the commits containing secrets can be dropped or not.

5.3. BigQuery GH Datasets

GitHub has over 100 million repos with only ~30 million repos being public. The rest of repos on public GH are licensed. It is wise to scan not only the public GH repos, but also the licensed repos. GH provides weekly snapshots of open-sourced licensed repos that can be queried by Google BigQuery [11]. In this scenario, the scanned repos can be passed to xGitGuard for secret detection.

6. Conclusion

We developed an internal tool called xGitGuard to detect leaked secrets in open source code repositories, to keep our data safe from inadvertent leaks. It was designed to perform with a high level of agility and a low false positive rate. Its strength, compared to other methods using regular expression scanning, is its use of NLP and text processing to detect secrets, and machine learning/ML to validate the detections. Our in-house developed algorithms and technologies outperform existing solutions in this area. xGitGuard is a stand-alone model that can be used in different ways depending on the use case. If well-trained, the ML-based validator has a high level of accuracy. Such high accuracy reduces the rate of false positives.

Abbreviations

E	Extensions
GH	GitHub
GHE	GitHub Enterprise
ML	Machine Learning

NLP	Natural Language Processing
PII	Personally Identifiable Information
PKEY	Primary Keyword
SKEY	Secondary Keyword
SOC	Security Operations Center
xGG	xGitGuard

Bibliography & References

[1] GitHub, "GitGub About," [Online]. Available: https://github.com/about. [Accessed 8 July 2021].

[2] GitHub, "The 2020 State of the Octoverse," [Online]. Available: https://octoverse.github.com/. [Accessed 9 July 2021].

[3] Github Kills Search After Hundreds of Private Keys Exposed, [Online]. Available: https://it.slashdot.org/story/13/01/25/132203/github-kills-search-after-hundreds-of-private-keys-exposed. [Accessed 10 July 2021].

[4] M. Jackson, "Biggest Security Takeaway of 2020: Don't Leak Secrets on GitHub," DZone, [Online]. Available: https://dzone.com/articles/automating-secrets-detection-with-git-hooks. [Accessed 10 July 2021].

[5] T. Security, "Truffle Hot," 2021. [Online]. Available: https://github.com/trufflesecurity/truffleHog.

[6] T. Security, Truffle Security, [Online]. Available: https://github.com/trufflesecurity/truffleHog. [Accessed 2021].

[7] Nightfall, "Detect sensitive data in your GitHub repos," [Online]. Available: https://try.nightfall.ai/radar. [Accessed 8 July 2021].

[8] S. Lounici, "Optimizing Leak Detection in Open-source Platforms with Machine Learning Techniques," in *7th International Conference on Information Systems Security and Privacy*, Vienna, Austria, 2021.

[9] M. Mouw, "Profiling the abuse of exposed secrets in public repositories".

[10] American Express, "EarlyBird," 2021. [Online]. Available: https://github.com/americanexpress/earlybird. [Accessed 8 July 2021].

[11] Google BigQuery, "GitHub BigQuery Dataset," [Online]. Available: https://console.cloud.google.com/marketplace/details/github/github-repos. [Accessed 8 July 2021].

Index of Authors

Index of Authors

Index of Authors

Index of Authors

Index of Authors