THE COMPLETE
TECHNICAL PAPER PROCEEDINGS
FROM:

NCTA
Technical
Papers

A SIDE-BY-SIDE COMPARISON OF CENTRALIZED VS. DISTRIBUTED ACCESS ARCHITECTURES

Michael J. Emmendorfer, Thomas J. Cloonan, John Ulm, and Zoran Maricevic
All of the authors are from ARRIS

Abstract

This paper will define and extensively compare two (2) Classes of Access Architectures that will emerge this decade for Cable Networking. These two (2) Classes of Access Architecture may be referred to as Centralized Access Architecture (CAA) and Distributed Access Architecture (DAA). The use of Centralize Access Architecture (CAA) retains the MAC and PHY layer functions of the CMTS, Edge QAM, or CCAP in the headend or hub location. The use of Hybrid Fiber Coax (HFC), which utilizes Amplitude Modulation (AM) optical technology or analog optics, enables only Centralized Access Architecture.

However, a transition to digital optics for fiber to the node (FTTN) may enable either a Centralized Access Architecture (CAA) or a Distributed Access Architecture (DAA). In a Distributed Access Architecture (DAA) the MAC and PHY layers of the CMTS, Edge QAM, or CCAP may be split between headend and node devices or the MAC and PHY layer functions of the CMTS, Edge QAM, or CCAP may be placed entirely in the node, cabinet, or MDU location. As our industry considers digital optics between the headend and fiber node we need to understand the pros and cons of CAA and DAA.

Our industry has always placed the least amount of intelligence in the outside plant, thus keeping the intelligence together and only at the Headend and CPE locations (the bookends). The use of DAA fundamentally changes the style of access architecture cable has implemented since its inception.

Our industry is not aware of all these options, those that are aware from MSO to supplier are divided on which approach is best and why. We have compiled a complete evaluation criteria and side-by-side comparison of these six (6) different types Access Architecture, so that an MSO can make an informed decision.

Key Questions Examined in this Paper:

Some of the most often asked questions by cable industry forward-looking planners reflect the key challenges the industry is facing for this decade and beyond. Some of these challenges and questions include:

1. *Can Digital Fiber Coax (DFC) architectures maximize the coaxial segment revenue spectrum capacity?*
2. *Can Digital Fiber Coax (DFC) architectures maximize the optical segment wavelength capacity?*
3. *Can Digital Fiber Coax (DFC) architectures maximize facility space, power and cooling?*
4. *Can Digital Fiber Coax (DFC) architectures maximize long links and facility consolidation?*
5. *Can Digital Fiber Coax (DFC) architectures maximize the economics of OPEX and CAPEX?*

This paper will seek to provide some visibility and answers to these questions and key challenges.

REVIEW OF HYBRID FIBER COAX OPTICAL TECHNOLOGY

The Amplitude Modulation (AM) optical layer will be examined in this section and illustrated in figure 1. The paper will only examine the forward optical technologies and performance attributes. The optical transport return path technologies include: Amplitude Modulation (AM), commonly referred to as analog optics and Broadband Digital Return (BDR), which may be referred to as simply Digital Return.

This section will examine if the future capabilities of the cable access network will be limited by the fiber to the node (FTTN) optical technology. This section will examine the network capacity if we replaced the AM optics with digital optics, like those used for Broadband Digital Forward or Remote PHY or Remote CCAP is required.

This proved that AM optics used in today's HFC could support higher order modulations, such as those defined in DOCSIS 3.1. However, depending on spectrum, optical span, and optics type, use of the highest order modulations (yet to be defined) was not possible with current AM optics. There could be many other factors; the cable distribution network side, the size of the service group, the spectrum used, and it could be the optical technology.

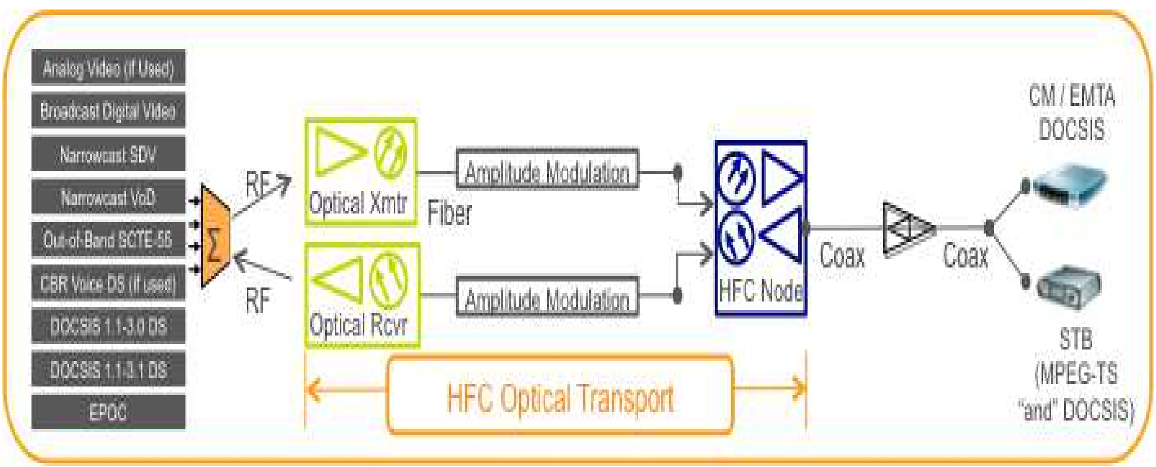


Figure 1: Overview of the Amplitude Modulation Optics

Overview of the "Current" FTTN Optical Technology

Amplitude Modulation (AM) optics when used in the return path had two types of lasers Fabry-Perot (FP) or Distributed Feedback (DFB) lasers. Though HFC Amplitude Modulation used DFB in the forward for many years. Analog return path transport is considered as a viable option for Mid-split and High-split returns; supporting short to moderate return path distances of 0-50 km. If the wavelength is changed to 1550 nm, with an EDFA, even greater distances are possible.

The analog optical return path transport presently supports up to 200 MHz loading; but typically only 5-42 MHz or 5-65 MHz is carried, depending on the distribution diplex filter split. The major benefit with analog optical return is its simplicity, lower cost, and flexibility, when compared with HFC style digital optical transmission. Distance is the chief challenge of analog optical transport and we will examine if support for very high order modulation, like that planned in DOCSIS 3.1, could be a factor.

Pros

The chief advantage of analog return is its cost effectiveness and flexibility. If analog return optics are in use in the field today, there is a good chance that they will perform adequately at 85 MHz; and even 200 MHz loading may be possible, if required in the future. This would allow an operator to fully amortize the investment made in this technology over the decade.

Important:

AM optics may support very high order modulation (4K & 16K QAM) though there are some restrictions mainly due to:

- Dependence on the type of optics in the forward and return
- Distance, spectral loading, spectral placement in the low frequency band to achieve the highest modulation order, and service group size (upstream)
- AM optics short distance or O-band optics will yield best performance
- Manufacturer consultation is needed to confirm performance thresholds

Cons

There are drawbacks to using analog optics. Analog DFB's have demanding setup procedures. RF levels at the optical receiver are dependent on optical modulation index and the received optical power level. This means that each link must be set up carefully to produce the desired RF output at the receiver (when the expected RF level is present at the input of the transmitter). Any change in the optical link budget will have a significant impact on the output RF level at the receiver, unless receivers with link gain control are used.

Also, as with any analog technology, the performance of the link is distance dependent. The longer the link, the lower is the optical input to the receiver, which delivers a lower RF output and lower C/N performance.

DIGITAL FIBER COAX (DFC) INTRODUCTION

Moving from AM Optics to Digital Optics for FTTN will force us to place PHY or MAC/PHY Access Layer Functions in the Node. What stays in the headend and what moves to the node? The industry will need to define a new access network architecture supporting digital connections between headend and fiber node. This new access network architecture will redefine the CCAP architecture and other headend platforms (e.g. Digital Optical Platforms) as well as the node platforms.

In this section the uses of Digital Optics is required and this will place new functions in the Node and add or remove functions from the Headend. It is of critical importance that we understand the functional layers and building blocks of MPEG-TS and DOCSIS MAC and PHY Functions as these functions may be split between the headend and node in the future. This section ends with several examples of Remote PHY layer or MAC and PHY functions in the node the node to support Digital Forward solutions.

As we examine the future to support higher data capacity in the optical and coax domain we may need to use digital optical technology for FTTN. We will examine this class of architecture we are calling Digital Fiber Coax (DFC). The DFC Architecture is a network class, which differs from HFC in that MAC/PHY or just PHY processing is distributed in the outside plant (node) or MDU. The DFC architecture also uses "purely digital" optical transport technologies such as standardized Ethernet, G.709, PON, or other transport methods providing optical capacity to and from the node. The industry may determine to call this class of architecture something else, but the functions, technology choices and architectures are different than HFC.

Digital Fiber Coax (DFC) is a "PHY or MAC/PHY Processing Architecture" in the node using Digital Optics to/from the node as seen in figure 2 and figure 3 is a side-by-side description of HFC and DFC. Thus Digital Fiber Coax (DFC) uses digital optical technology to and/or from the node as well as supports two (2) different Access Architecture options for FTTN as seen in figure 4 and figure 5 is a side-by-side description of Centralized Access Architecture (CAA) and Distributed Access Architecture (DAA). DFC uses digital optics for FTTN (to/from) in either a Centralized Access Architecture (CAA) "or" a Distributed Access Architecture (DAA). DFC in a Centralized Access Architecture (CAA) the CCAP MAC and PHY functions in Headend (HE) or Primary Hub (PH) only. DFC in Distributed Access Architecture (DAA) the CCAP MAC and PHY or PHY functions are placed in a node. As with Centralized Access Architectures there are several platform access architectures, this is even more the case with Distributed Access Architectures that will split up the MAC and PHY layers of CCAP between the headend and the node. In the full Remote CCAP option for DFC, the entire CCAP MAC and PHY layers are placed in the node or MDU location. This section will provide terms and definitions to the different Fiber to the Node Classes cable may select, like HFC or DFC as well as the two different Access Architecture classes options that may emerge this decade and beyond as seen in figure 6.

"Two (2) Different" Fiber to the Node (FTTN) Architecture Classes for Cable and Two (2) Different Access Architecture Classes

In this section, we describe the functions of several approaches for fiber to the node (FTTN). The following figures will aid in aligning the definitions with the list of functions; please refer to figures 2 through 6.

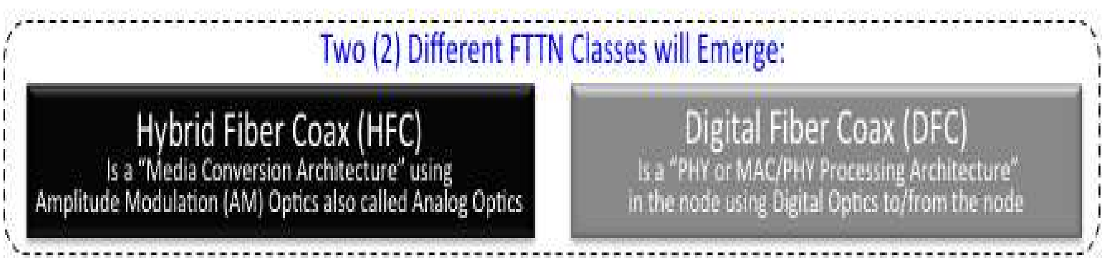


Figure 2 – Two (2) Different FTTN Classes for Cable will Emerge

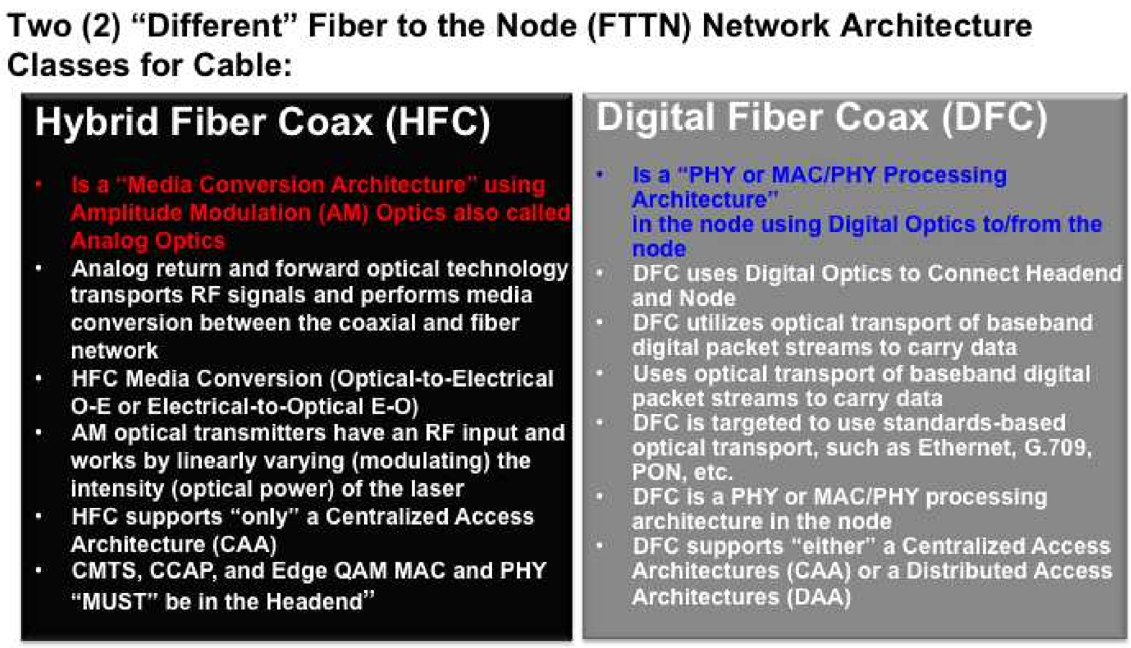


Figure 3 – Descriptions of the Two (2) Different FTTN Classes

Figure 4 – Two (2) Different Access Architecture Classes for Cable will Emerge

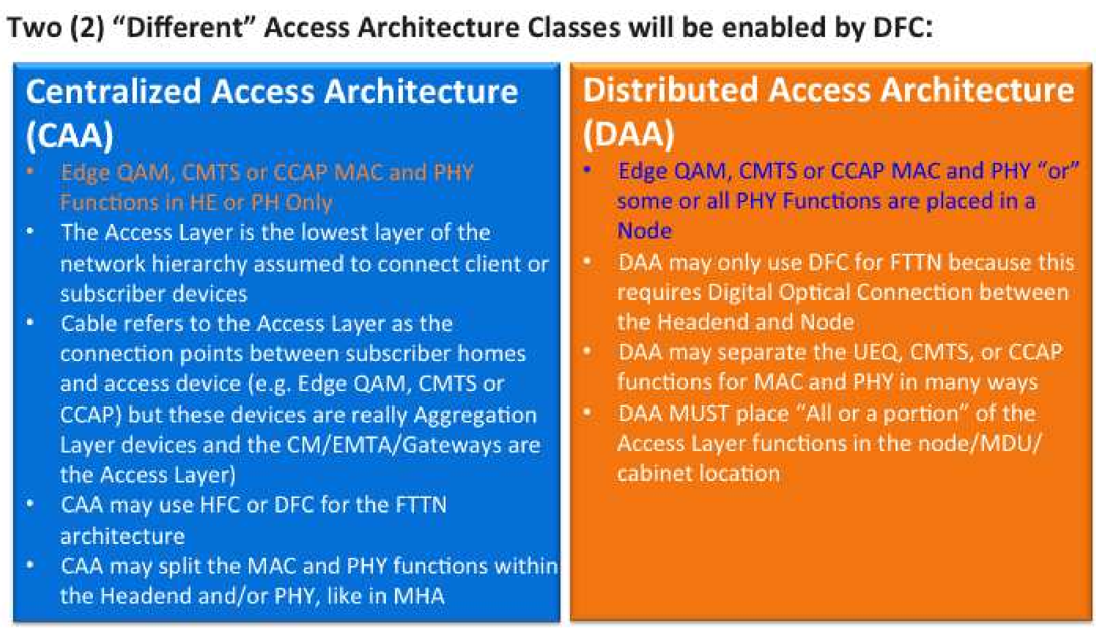


Figure 5 – Descriptions of the Two (2) Different Access Architecture Classes

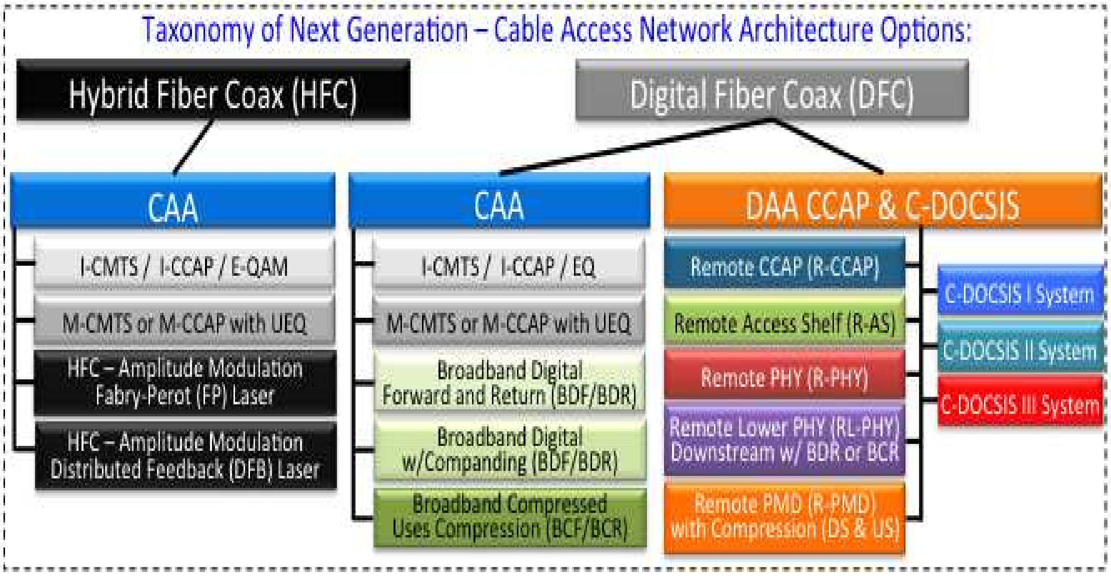


Figure 6 – Taxonomy of Next Generation – Cable Access Network Architecture Options

Overview of Current and Future FTTN Optical Technology

The optical layer and the relationship to the remote access layer architecture will be examined in this section.

Today, the two technologies used in optical transport for the return include Amplitude Modulation (AM) and Broadband Digital Return (BDR), as reviewed in the preceding section. The Broadband Digital term and current application is tied to the return path; **however, this could be used for the forward path as well.**
Broadband Digital Return places the lowest layer of the physical (PHY) layer called the PMD (Physical Medium Dependent) function in the Node. The PMD layer of the PHY is where the ADC/DAC (Analog-to-Digital or Digital-to-Analog) functions take place.

The FTTN technology and architecture for HFC has always retained one core function --- transparency of the underlying MAC/PHY technologies that travels through it. The transparency of the RF MAC/PHY technologies was possible because of the optical FTTN technology used to include either Amplitude Modulation optical technology or Broadband Digital.
In the future we need to consider the possibility of moving the IP/Ethernet transport past the HE/Hub locations to the node. We will examine what we are referring to as a new class of cable FTTN architecture called Digital Fiber Coax (DFC). The use of DFC may augment the existing HFC media conversion class of architecture that has been deployed for about two decades. We are suggesting that there are really two different Fiber to the Node (FTTN) architecture classes for Cable Networks. These will utilize FTTN and coaxial cable as the last mile media, but this is where the similarities will stop.

To simply summarize, the Two Different Cable FTTN network architecture classes are:

- HFC is a "Media Conversion Architecture"
- DFC is a "PHY or MAC/PHY Processing Architecture"

These new FTTN technologies and architectures have or will emerge, that if implemented "may" remove this transparency.

Should the cable industry change the definition of HFC to mean multiple functions, "or" define a new term(s) for this fundamentally different Class of FTTN Network Architecture that uses Digital Optics to/from the node as illustrated in figure 9 which keeps the MAC and PHY functions of the CMTS, EdgeQAM and CCAP in the headend and enable Digital transport through a separate optical transport shelf using a Digital Fiber Coax architecture called Broadband Digital. If it is decided to break up the CCAP and place digital optics directly on the CCAP, a Remote PHY CCAP Class of DFC is possible as seen in figure 10 or MAC and PHY CCAP Class of DFC functions in the node as seen in figure 11 is possible.

The figures in the sections represent the high-level functions and technology placement in the headend and node.

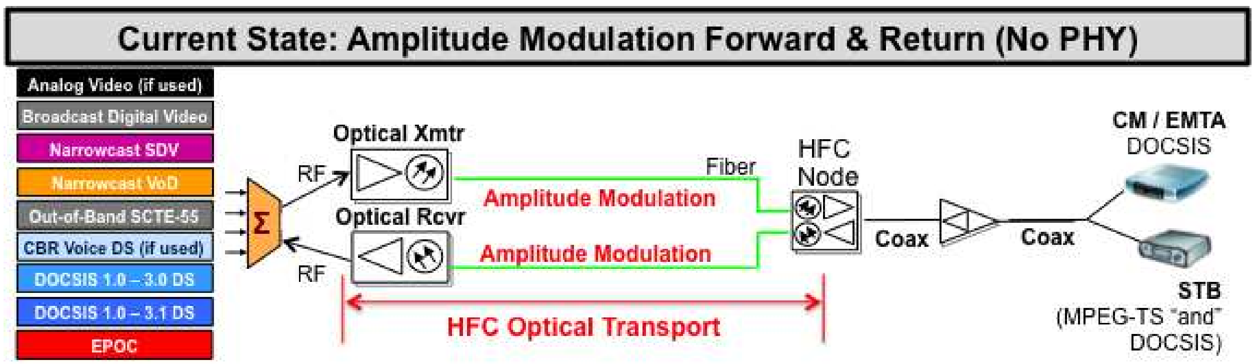


Figure 7: HFC Amplitude Modulation Forward and Return

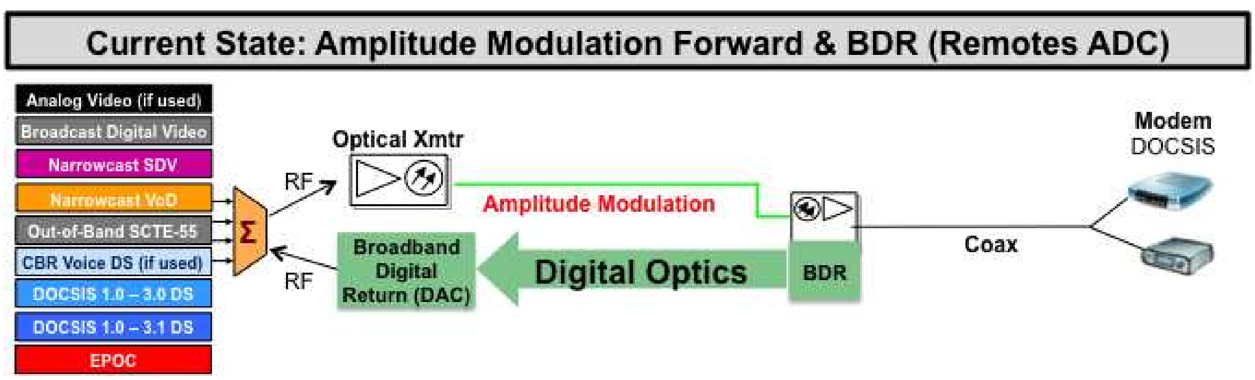


Figure 8: HFC Amplitude Modulation Forward and DFC Broadband Digital Return (BDR)

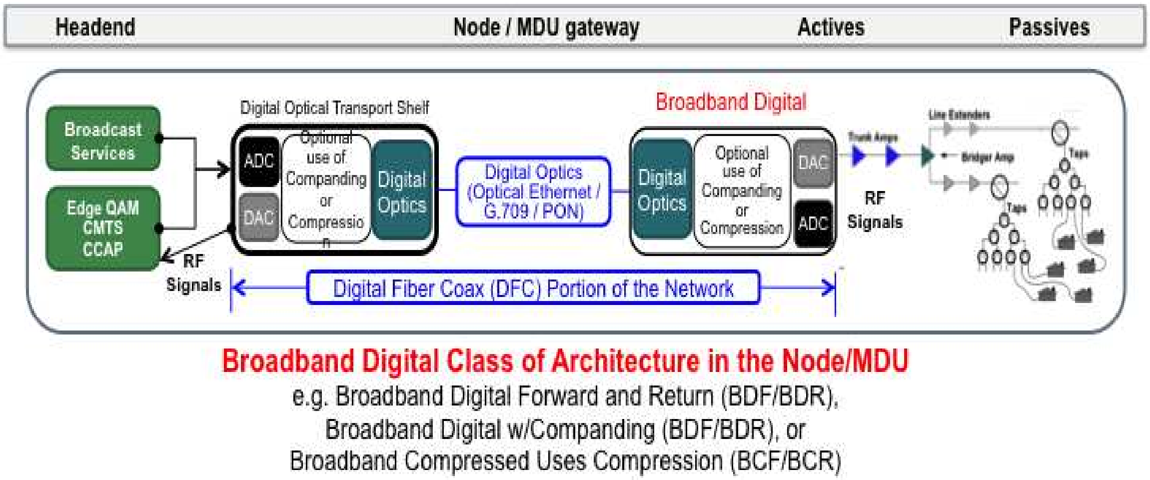


Figure 9: Digital Fiber Coax – Broadband Digital Class

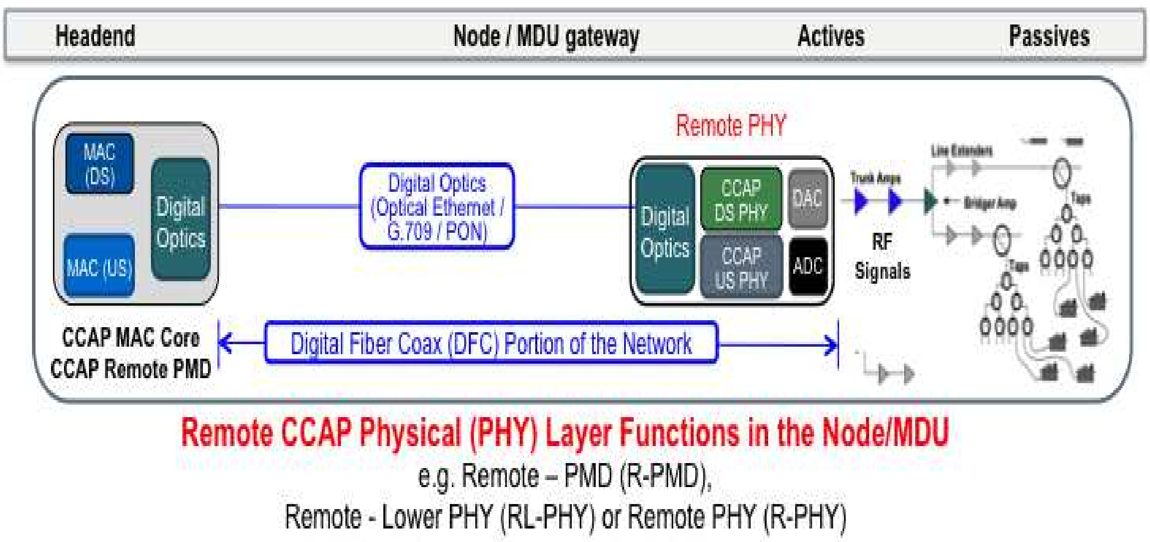


Figure 10: Digital Fiber Coax – Remote PHY CCAP Class

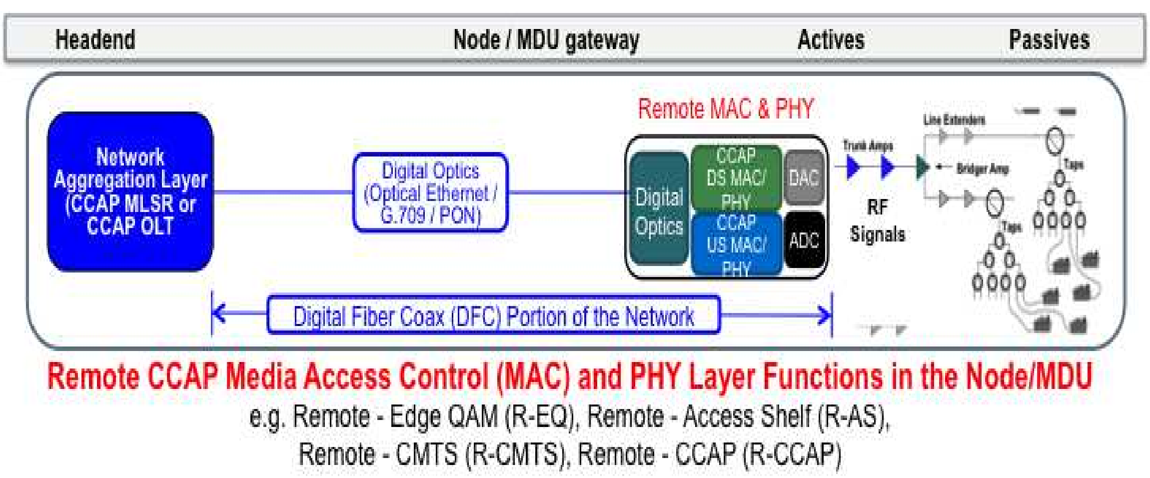


Figure 11: Digital Fiber Coax – Remote MAC and PHY Class

Downstream DOCSIS and Edge QAM Functional Alignment to Headend and Node Platforms

This section and associated figures are meant to align cable technologies to the OSI reference model. The technologies examined include DOCSIS 3.0 and Edge QAM functions to the left which both use Recommendation ITU-T J.83 as the Physical Layer. The right side of the figure 12 is an attempt to define the “possible” framework for DOCSIS 3.1 currently in development. This figure is based on the DOCSIS specifications, ITU-T J.83-B, and DVB-C2. This is aimed to help show the functions of the Remote Access Layer Architecture that may remain in the headend and that which is placed in the node.

The figure below captures the downstream DOCSIS and Edge QAM functions. The figure is intended to show the relationship with headend functions defined today and functions that will change in the headend CCAP and the node to support Remote Access Layer Architectures. The red boxes represent node functions and all align with the functions defined on the left of the figure. Please note that the figure above places the Edge QAM MAC functions partially in the PHY layer and this is because all edge QAMs products contain the Edge QAM MAC and the J.83 PHY used for video and DOCSIS. The figure below remove the Edge QAM MAC functions from the PHY and places this alongside the DOCSIS MAC functions see figures 10 and 11.

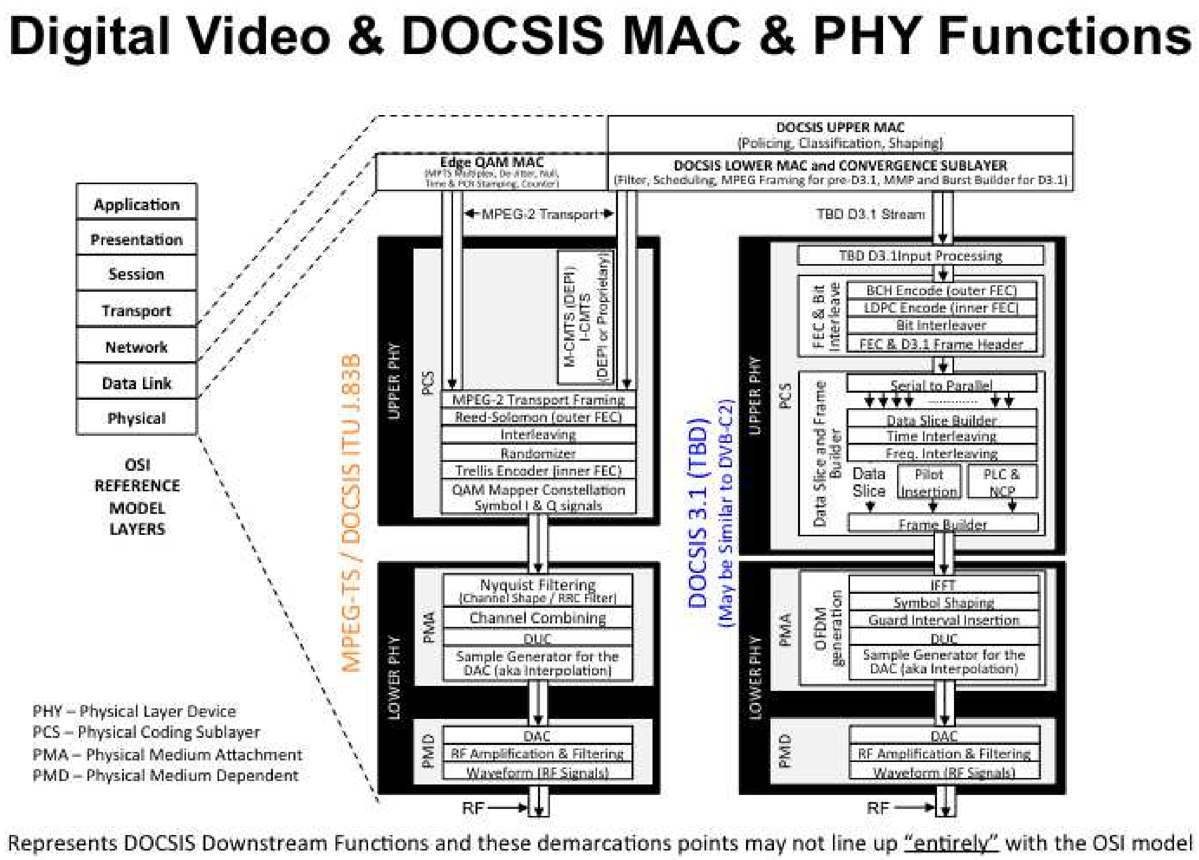


Figure 12: Detailed Digital Video and DOCSIS MAC and PHY Functions for the Downstream

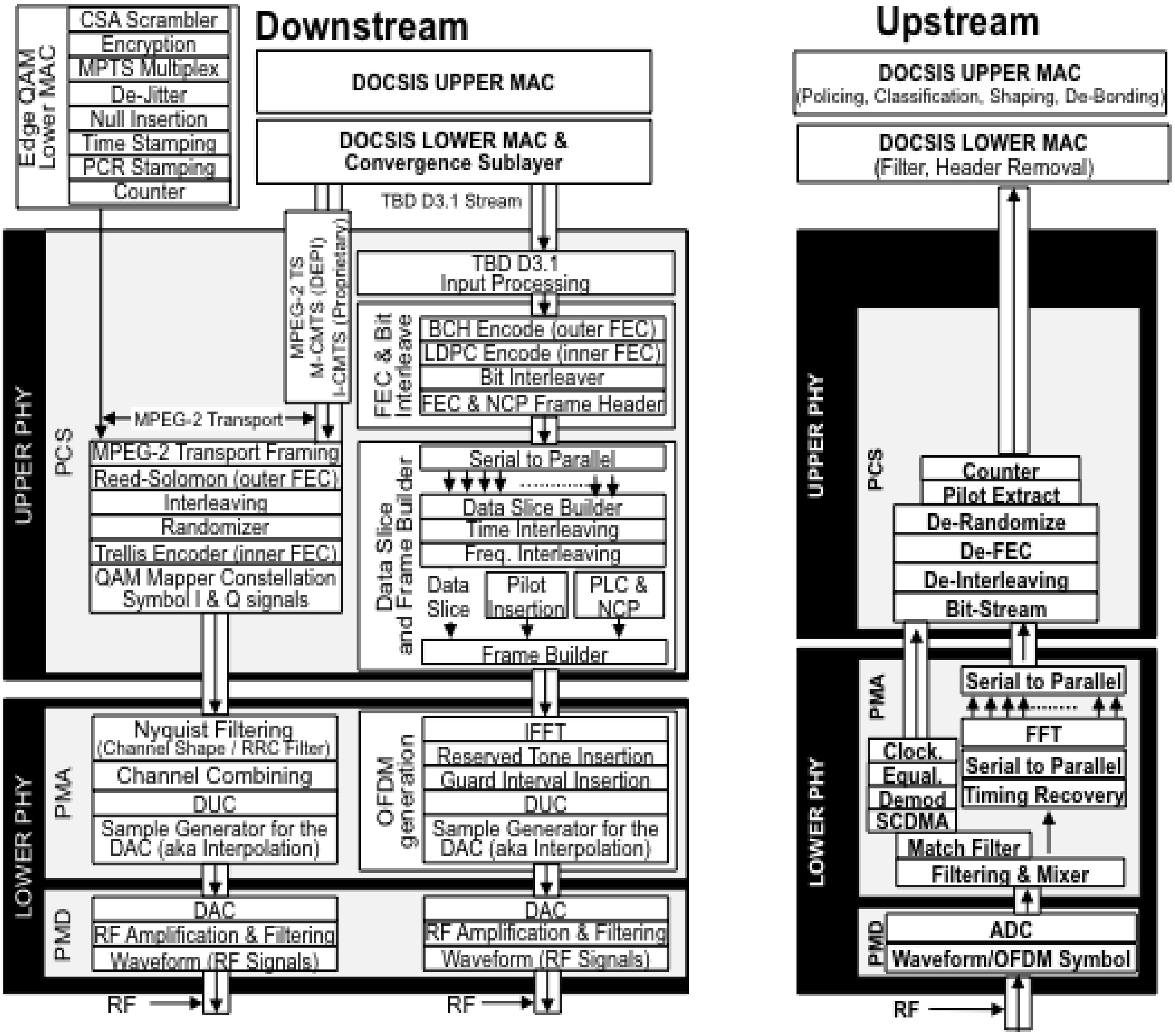


Figure 13: Functional Review of the RF MAC/PHY Layers Downstream and Upstream

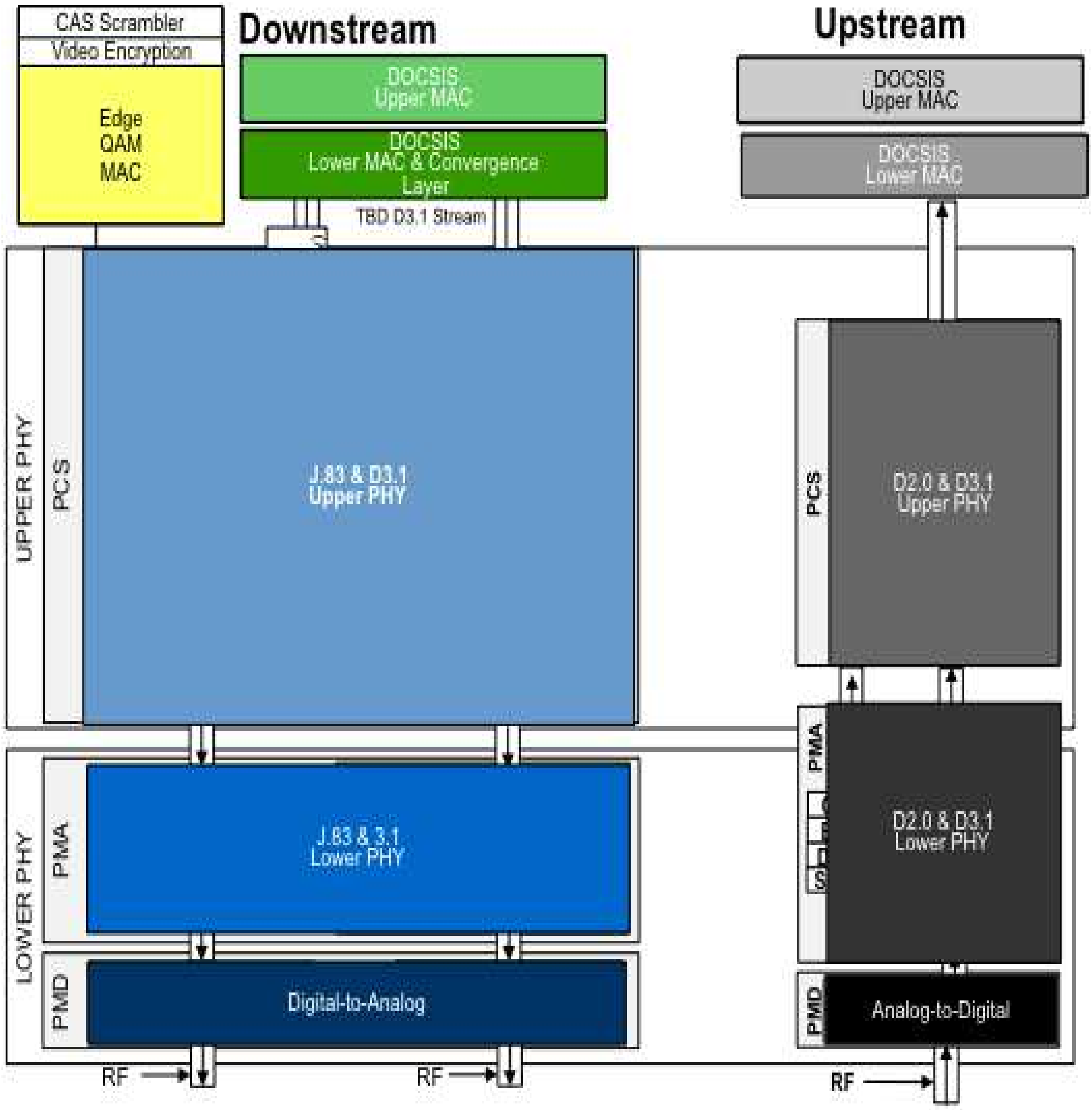


Figure 14: Simplified View of the RF MAC/PHY Layers Downstream and Upstream

These fundamental building blocks as illustrated above may serve as demarcation for functions that may be kept in headend platforms and those placed in Node or MDU locations. The next section takes these blocks and moves them to headend or node locations to illustrate the different architectures that may exist in the future to enable Digital Fiber Coax (DFC).

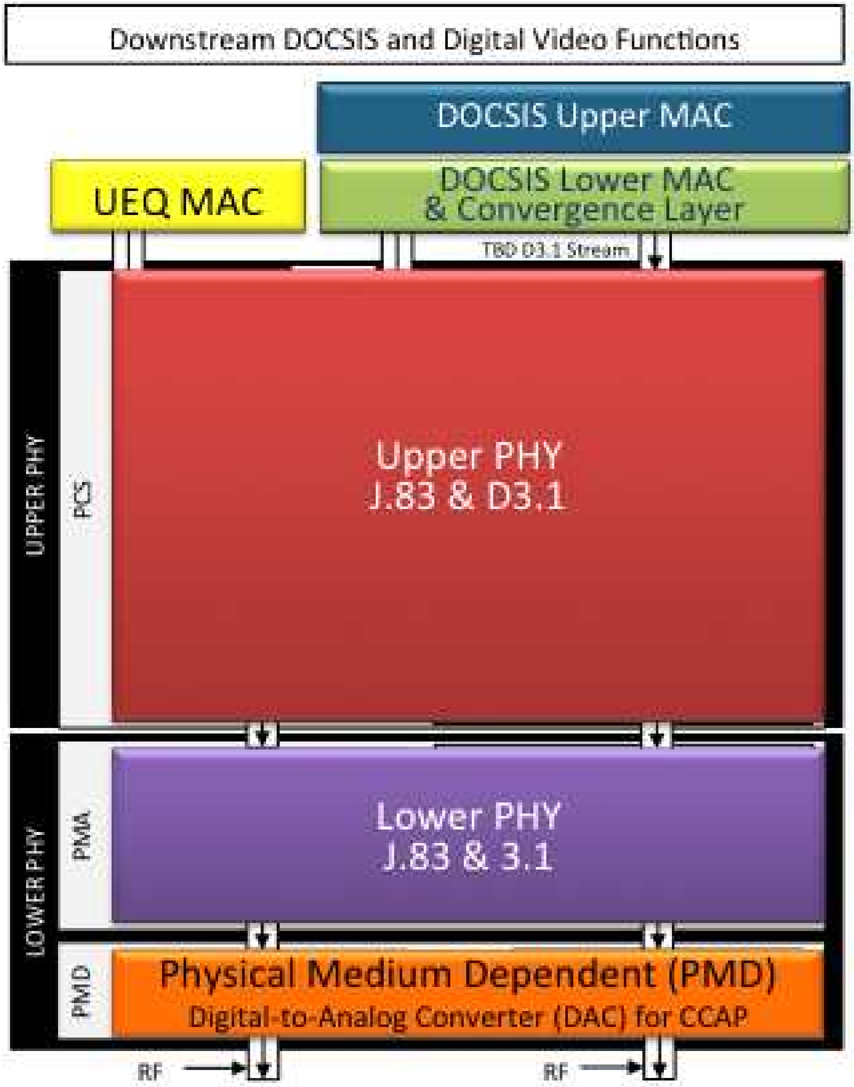


Figure 15: Summary Digital Video and DOCSIS MAC and PHY Functions for the Downstream

Hybrid Fiber Coax (HFC) Class of FTTN

1. Optical Amplitude Modulation uses Media Conversion (Optical-to-Electrical or Electrical-to-Optical) allowing for transparency of the RF MAC/PHY technologies. This is what we have used for decades. **Please refer to figure 16.**

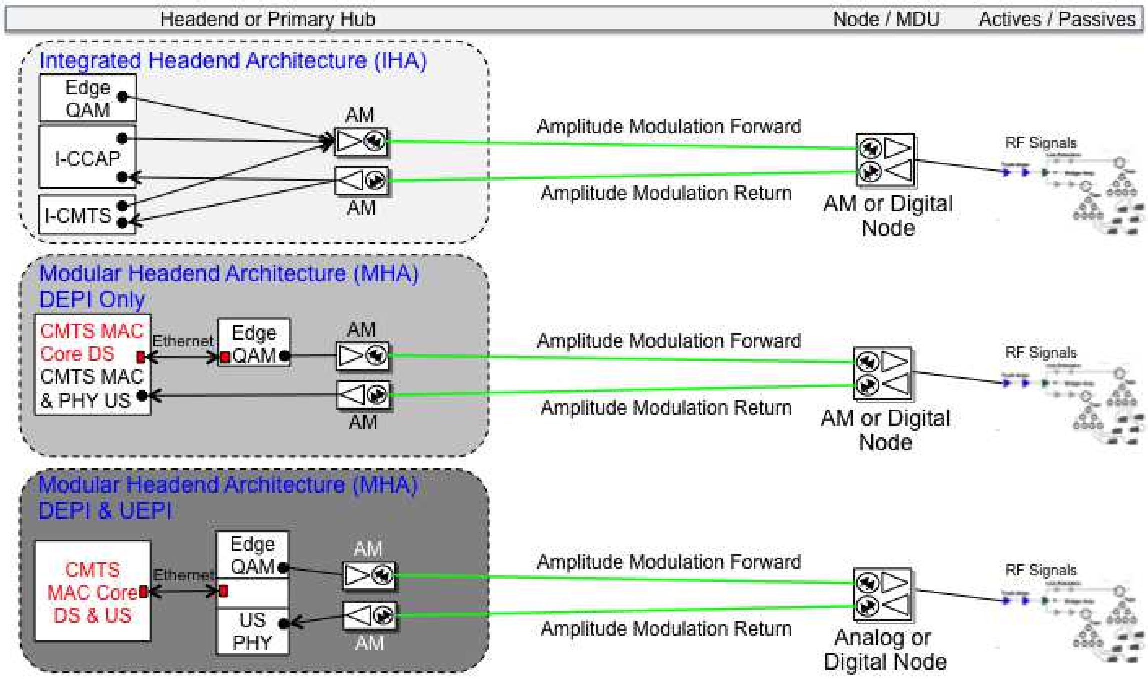


Figure 16: Summary Digital Video and DOCSIS MAC and PHY Functions for the Downstream

Digital Fiber Coax (DFC) Class of FTTN

The Digital Fiber Coax (DFC) Class of Fiber to the Node (FTTN) can be separated into six (6) different types of architectures. In the figure below the type of CAA and DAA are associated with types of DFC architecture. Todays, CAA can be carried over HFC optics as well as Digital Optics like Broadband Digital Forward (BDF) or Broadband Compressed Forward (BCF). In figure 17, the left side of the figure summarizes the functional layers for downstream DOCSIS and digital video. The right side of the image captures the platform or system architectures, or the network elements and what functions each contain. For example, I-CMTS or I-CCAP has a bar spanning from the top to the bottom of the functional diagram, thus all the functions are in those platforms. Likewise the far right bar, called Remote CCAP (R-CCAP) contains all the function as well, but this has a red highlight around it meaning that this is all in a node housing. The color codes represent the highest layer of function in the node and any of the gray bars represent functions that will remain in the headend. At the top of the bar charts these are group by Centralized Access Architectures (CAA) and Distributed Access Architectures (DAA). Please note that the two Centralized Access Architectures have RF outputs in the headend or primary hub but these may be part of Digital Fiber Coax when a separate optical shelf is used in the headend to enable digital communications to and from the fiber node. In the CAA all of the Edge QAM, CMTS, or CCAP functions and network elements remain intact maintaining the MAC and PHY layers in the headend. The DAA distributes the entire MAC and PHY functions “or” may distribute portions of the CCAP to the node keeping the remainder in the headend, thus in DAA there is no CCAP RF ports in the headend.

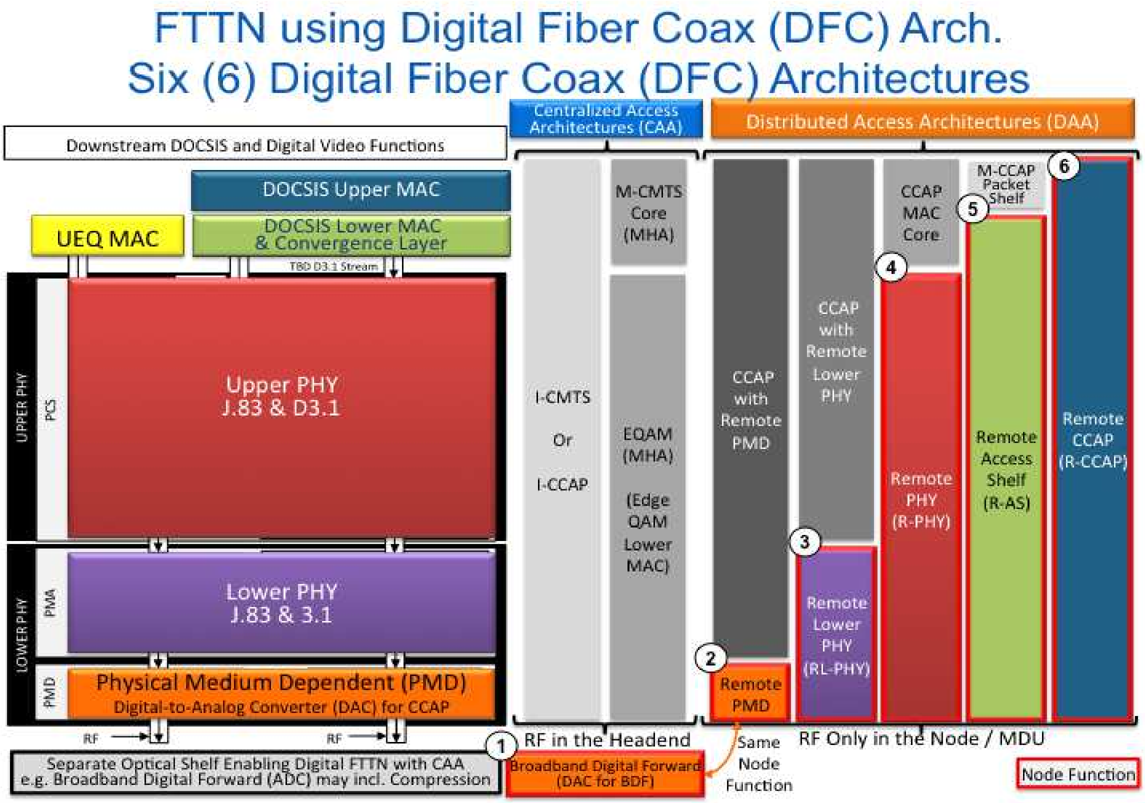


Figure 17: Platform / System Architectures (Headend + Node)
MPEG TS & DOCSIS Downstream

1. **Broadband Digital:** Assumes a separate optical shelf receiving RF sources from analog video, Edge QAM, CMTS, CCAP, RF Out-of Band, and RF Test equipment. The Broadband Digital equipment receives RF and digitizes the spectrum transported to or from the node. Key components of this process are the Analog-to-Digital Converter (ADC) and Digital-to-Analog Converter (DAC). This approach allows transparency of the RF MAC/PHY technologies in the outside plant. This is in use today for the upstream called Broadband Digital Return (BDR) and this type of approach may be used in the downstream direction as well called Broadband Digital Forward (BDF). Suppliers may add innovations to reduce the capacity requirements imposed when the analog signal and spectrum is digitized. These are proprietary solutions today but could easily be standardized. This approach is the "only" Remote PHY architecture that maintains the transparency of the underlining MAC/PHY technologies that travels through it and uses digital optics. **Please refer to figure 18 through 21.**
2. **Remote PMD (R-PMD):** The term PMD refers to the Physical Medium Dependent sub-layer of the PHY that contains the ADC/DAC (Analog-to-Digital or Digital-to-Analog). The PMD layer is part of the CMTS, Edge QAM or CCAP platforms. This is similar to Broadband Digital, however this just removes the PMD layer in the CMTS, Edge QAM or CCAP platform and places this function in the node or MDU location. This type of architecture has not been done in the cable space, but if desired could be called Remote Physical Medium Dependent (R-PMD). We are suggesting the term Remote PMD because this better defines the remote

PHY layer that is placed in the node. The cable industry could define a standards based Remote PMD Architecture for the return and forward path similar to that, which was done when the PHY layer was removed from the CMTS in the Modular Headend Architecture (MHA). As in the case with Broadband Digital suppliers may add innovations to reduce the capacity requirements imposed when the analog signal and spectrum is digitized and this could also become standardized. **Please refer to figures 22 and 23**

3. **Remote Lower PHY (RL-PHY):** Remote Lower PHY is placed in the node where constellation symbols or groups of constellation symbols are received from the headend to the node lower PHY for modulation. This represents the modulation functions and is sometimes called Remote Mod. Remote Lower PHY is only an option for the downstream and not the upstream. **Please refer to figures 24 and 25**
4. **Remote PHY (R-PHY):** This places the full PHY layer including the FEC, symbol generation, modulation, and DAC/ADC processing in the node. This is analogues to the Modular Headend Architecture (MHA), but of course different in may ways, such as timing and support for extreme separation of the MAC and PHY layers as well as support for DOCSIS 3.1 would have to be written. This approach could be called Remote PHY Architecture (RPA). **Please refer to figure 26**
5. **Remote - Access Shelf (R-AS):** Places the entire "Edge QAM" MAC and PHY layer functions in the node. Video security and encryption may or may "not" be placed in the node. The Lower "DOCSIS" MAC functions for scheduling and the entire PHY functions are placed in the node. This could be referred to as the Remote Access Shelf. The M-CCAP Packet Shelf remains in the headend and performs the DOCSIS upper MAC functions while the M-CCAP Remote Access Shelf performs Edge QAM MAC and Lower DOCSIS MAC functions. **Please refer to figure 27**
6. **Remote CCAP (R-CCAP):** Places the entire upper and lower MAC and PHY layer functions in the node. This places the CMTS, Edge QAM and CCAP functions into the node. **Please refer to figure 28**

<u>Example Platform and Network Architectures</u>

Broadband Digital Return and Forward Architecture

In figure 18 please refer to the definition above called "Broadband Digital" above but as far as a brief description Broadband Digital Return and Forward will be a separate optical shelf that interfaces with devices with RF ports and digitizes the signals between the headend and node. Today, Broadband Digital is used only for the return path. In the figure the I-CCAP has all functions for video, DOCSIS J.83, DOCSIS 2.0, and DOCSIS 3.1 all in a single platform, however a MHA could have been used with RF outputs in the headend. The key point for Broadband Digital is that RF interfaces remain in the headend and these devices interface with an optical shelf that enables a digital connection between headend and node. Like Amplitude Modulated (AM) optics used in HFC, Broadband Digital is completely transparent.

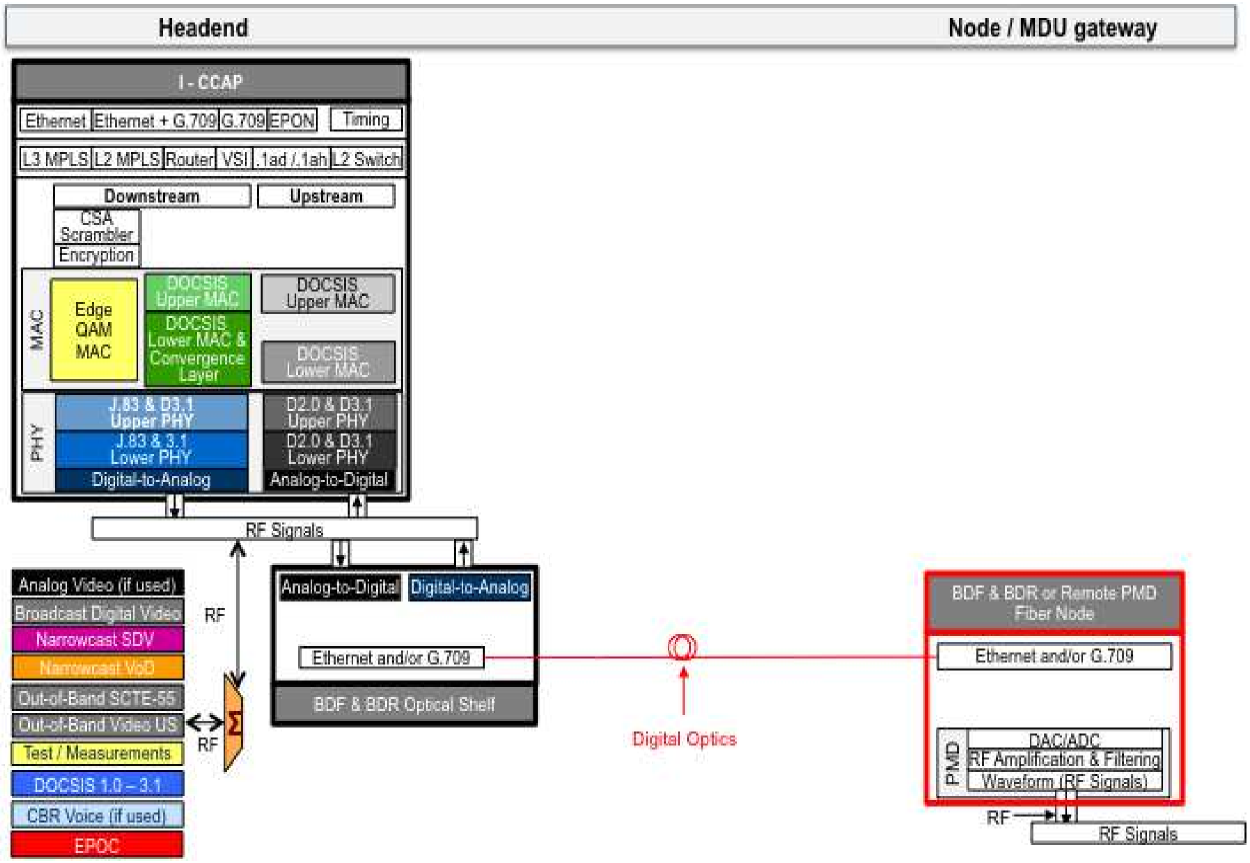


Figure 18: Broadband Digital Return and Forward Architecture

__Broadband Digital Forward or Broadband Compressed Forward__

The use of Broadband Digital has been used in the return patch for many years now. The use of Broadband Digital in the forward direction has not been a viable option because of the large optical link to carry full spectrum, perhaps as high as 25 Gbps optical link to carry 1 GHz worth of spectrum. The introduction of the next evolution Broadband digital use compression technologies to reduce the overhead typically required to transport RF spectrum such perhaps approximately 800 MHz of spectrum may be carried in a 10 Gbps optical class link. The use of Broadband Compressed Forward (BCF) could be used as part of Full Spectrum architecture or as part of the Broadcast Cast and Narrow Cast architecture. The use of Broadband Compressed Froward will enable high order modulation over long optical spans or distance, using of lots of optical wavelengths, and even in Narrow Cast architectures.

In the diagram below the use of amplitude modulation optics is used and with analog optics the distance and number of wavelength trade-off is illustrated.

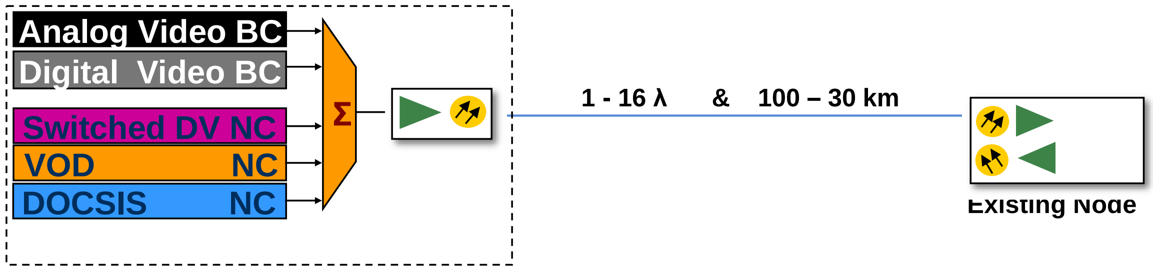


Figure 19: Amplitude Modulation Optics with Wavelength and Distance

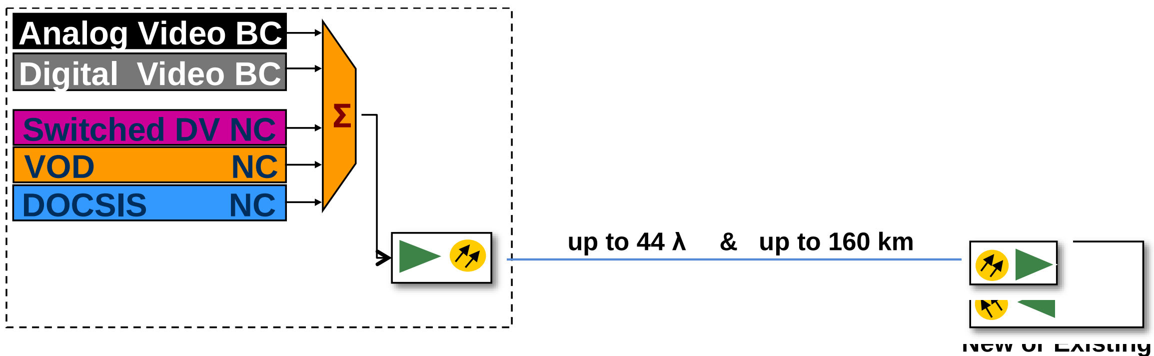


Figure 20: Broadband Compressed Forward (BCF) integrates the entire Legacy

The use of Amplitude Modulation (AM) Compared with Broadband Compressed Forward is that the services transported and the equipment in the headend used for HFC AM optical transport can be used for DFC use of BCF. The use of BCF will transport any MAC and PHY technology transparently. The uses of BCF can carrier analog services and out-of-band (OOB) signals. BCF carries the services, leverages the existing headend data and video equipment, and the transparency of AM optics, but BCF has the benefits of digital transport. Similar to Remote PHY and Remote MAC and PHY discussed later in this paper, the used of BCF enables the high performance of any of the other DFC architecture such as: distance / number of wavelengths, pluggable optics modules with ease of setting, deploying and sparing.

The figure below illustrates that BCF will enable I-CCAP or MHA architectures to remain in place and even to continue to be deployed while enabling a digital option for these platform architecture should DFC be desired. The RF combined is combined with analog and digital video broadcast and CCAP narrowcast services are shown using the same digital optical link. The use of BCF and CCAP enable a digital option where and if needed at all, while keeping the CCAP functions only in the headend and very little intelligence in the outside plant node.

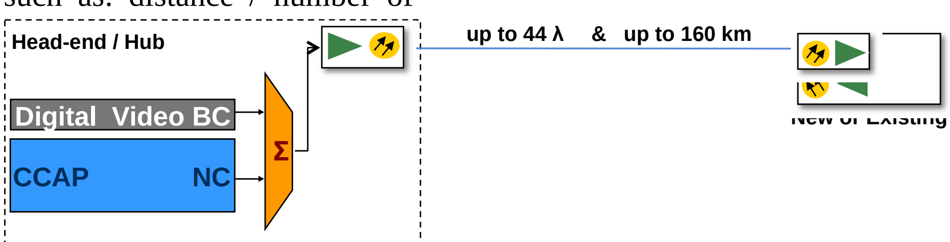


Figure 21: BCF Enables a Digital Transport for I-CCAP

Remote PMD Architecture

In figure 22 please refer to the definition above called "Remote PMD (R-PMD)". In this architecture the term PMD refers to the Physical Medium Dependent sub-layer of the PHY that contains the ADC/DAC (Analog-to-Digital or Digital-to-Analog). The PMD layer is part of the CMTS, Edge QAM or CCAP platforms. This is similar to Broadband Digital, however this just removes the PMD layer in the CMTS, Edge QAM or CCAP platform and places this function in the node or MDU location.

Remote PMD Architecture the Broadband Compressed Forward (BCF) gets subsumed into the CCAP shelf in this example. The CCAP incorporate the BCF processing for part or the entire spectrum. The use of RF outputs in the CCAP is removed and replaced with standards based 10G optical technology and if an MSO deployed BCF using a separate headend optical shelf to a BCF node, the existing BCF node could be used.

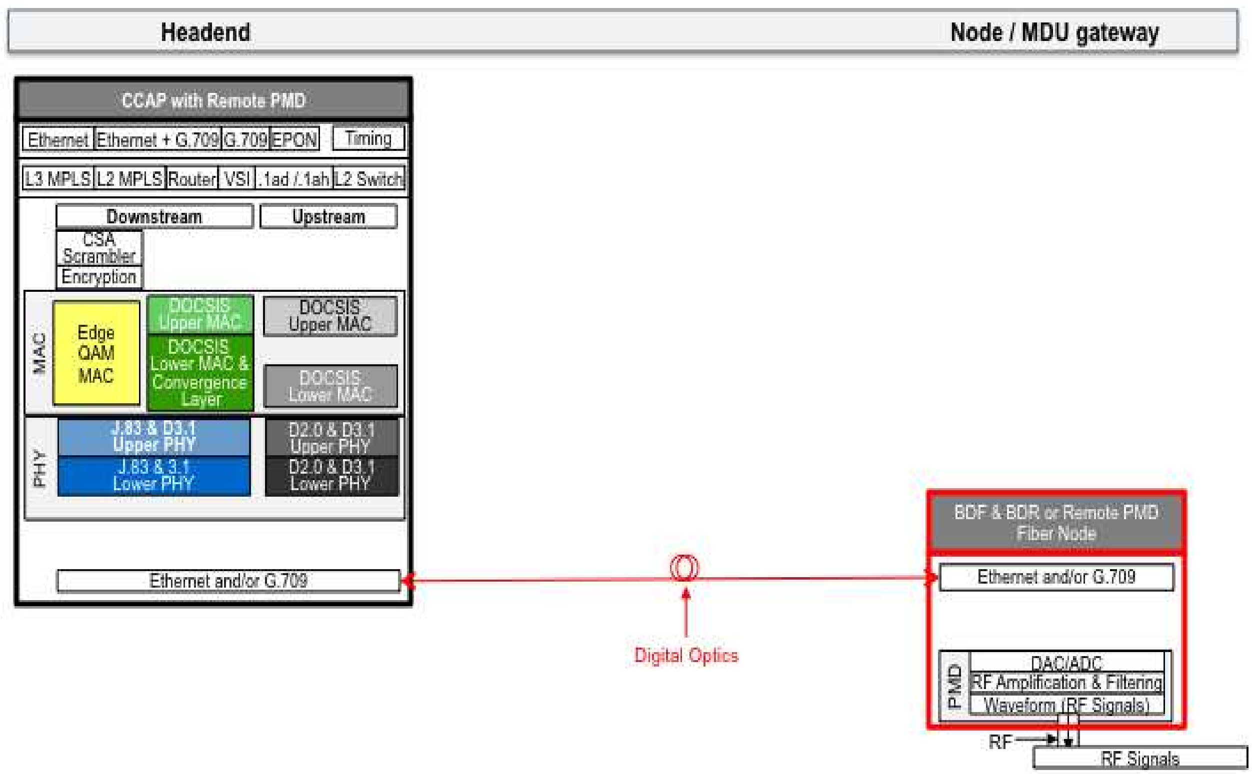


Figure 22: Remote PMD Architecture

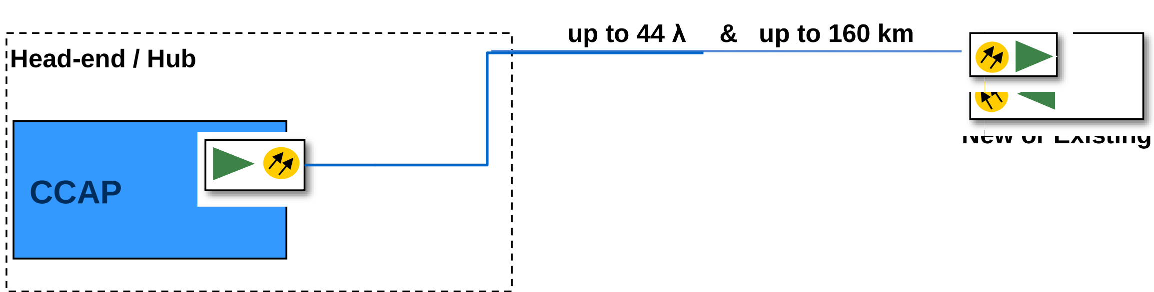


Figure 23: Remote PMD Architecture

Remote Lower PHY (RL-PHY) Architecture

In figures 24 and 25 please refer to the definition above called "Remote Lower PHY (RL-PHY)". These represent two different architectures to implement Remote Lower PHY. As with Remote PMD a portion of the PHY is removed from the headend and placed in the node location.

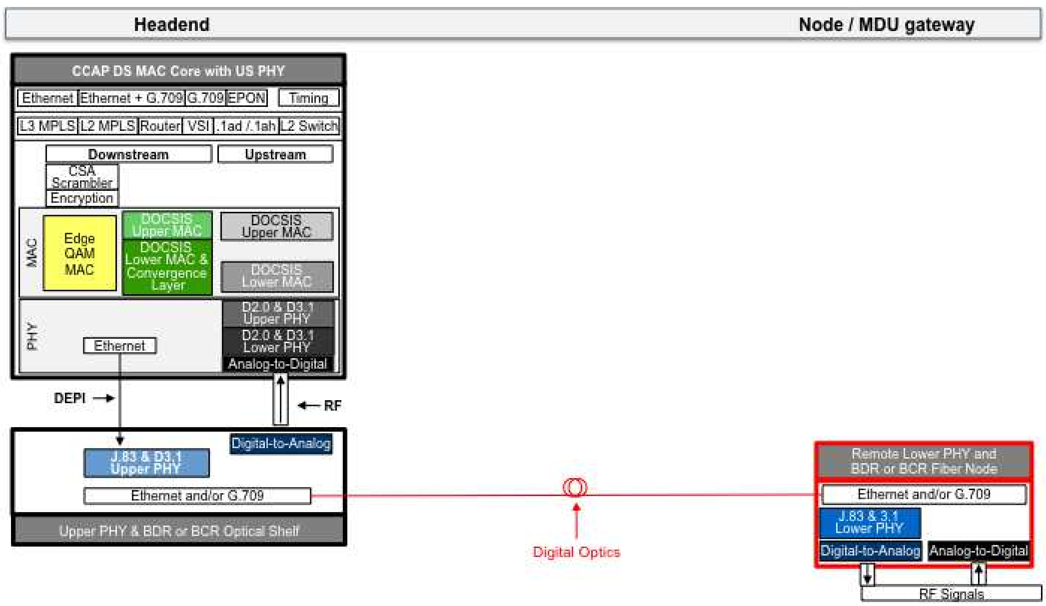


Figure 24: Remote Lower PHY and BDR Separate Headend Optical Shelf Architecture

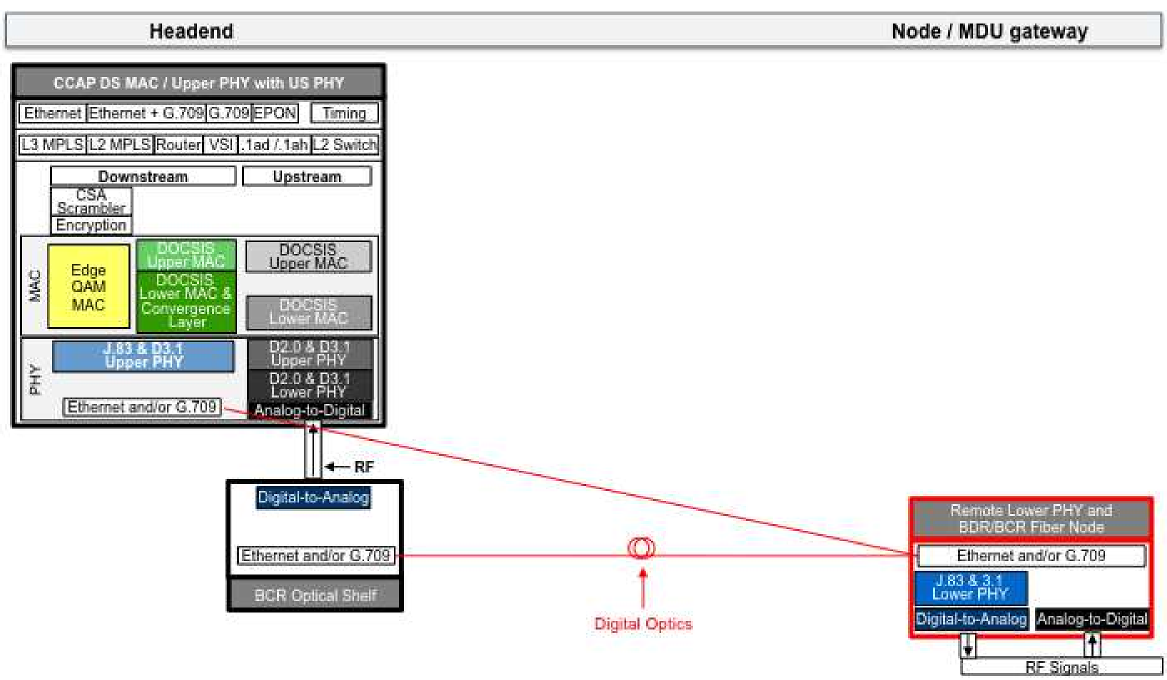


Figure 25: Remote Lower PHY CCAP and BDR Separate Headend Optical Shelf Architecture

Remote PHY (R-PHY) Architecture

In figure 26 please refer to the definition above called "Remote PHY (R-PHY)". The architecture of using a CCAP MAC Shelf with a Remote PHY could be called Remote PHY Architecture (RPA), as this resembles in some ways the Modular Headend Architecture (MHA) defined by CableLabs.

Figure 26: Remote PHY Architecture (RPA)

Remote Access Shelf Architecture

In figure 27 please refer to the definition above called "Remote Access Shelf (R-AS)". This is very similar to the Modular CCAP architecture that defined a Packet Shelf containing the DOCSIS Upper MAC functions and the Access Shelf (AS) containing the DOCSIS Lower MAC and full PHY functions.

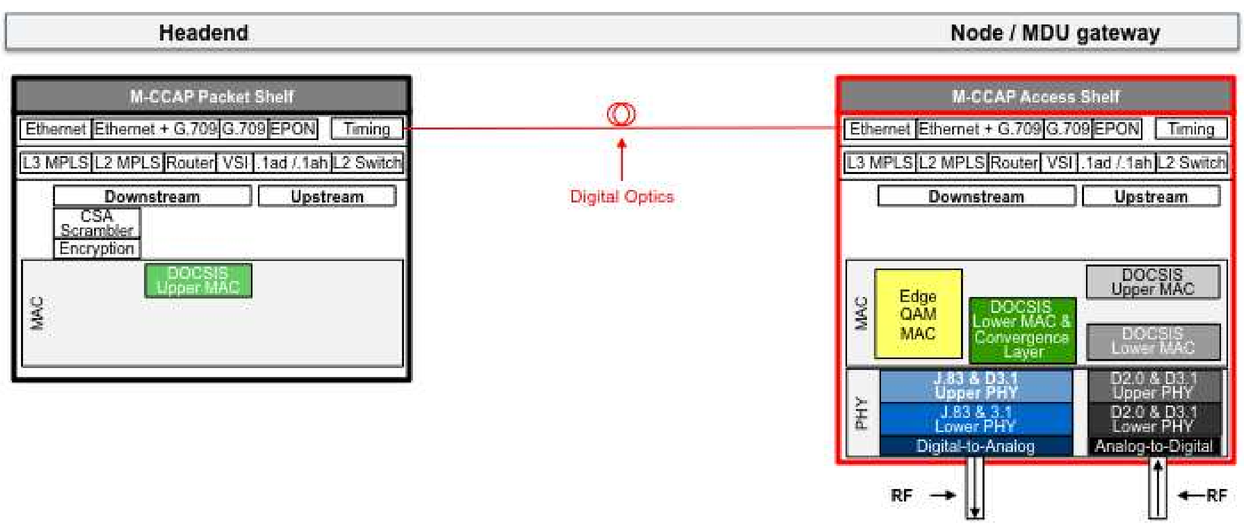


Figure 27: Remote Access Shelf (R-AS) Architecture

Remote CCAP Architecture

In figure 28 please refer to the definition above called “Remote CCAP (R-CCAP)”. This is the entire CCAP in the node minus the CSA Scrambler and Video Encryption.

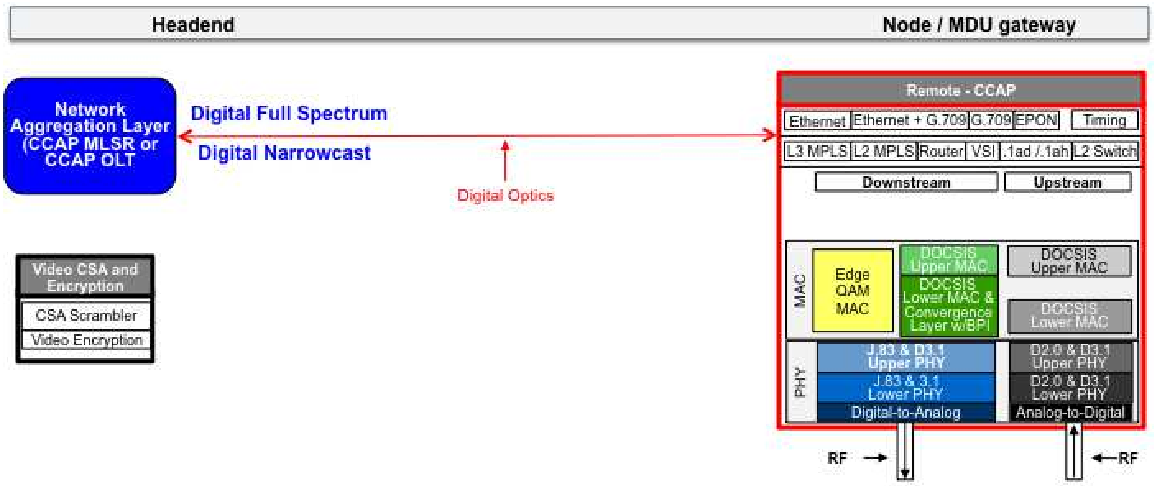


Figure 28: Remote CCAP (R-CCAP) Architecture

THE COAXIAL SEGMENT REVENUE SPECTRUM CAPACITY CONSIDERATION DETAILS

Comparison of HFC AM Optics and DFC Architectures

This is likely one of the most critical analysis of the paper because there have been concerns in the industry of the ability for amplitude modulation optics to support that modulation formats that are defined in DOCSIS 3.1. The three figures below represent an analysis comparing Hybrid Fiber Coax (HFC) optics using amplitude modulation (AM) and two types of Digital Fiber Coax (DFC) CAA using Broadband Compressed Forward (BCF) and also DFC using Remote Gadget, and this refers to either Remote PHY CCAP or Remote CCAP. It is assumed in the model will estimate the End of Line (EoL) performance and will then align that to the highest order modulation possible to support full spectrum using a given modulation order. Please note the assumptions below for each of the model comparisons.

Assumptions:

- Capacity Comparisons of HFC AM Optics vs. DFC BCF VS. DFC Remote Gadget
- Assumes: Full Spectrum DOCSIS 3.1 Gen2 CCAP, up to 8 Lambdas, up to 1 GHz N+2, and 1.2 GHz N+0
- Capacity estimates based on modeled End of Line (EoL) performance with alignment to the highest order modulation possible
- Please note though EoL estimated may support given Modulation Order no DOCSIS 3.1 systems are available to confirm estimates.

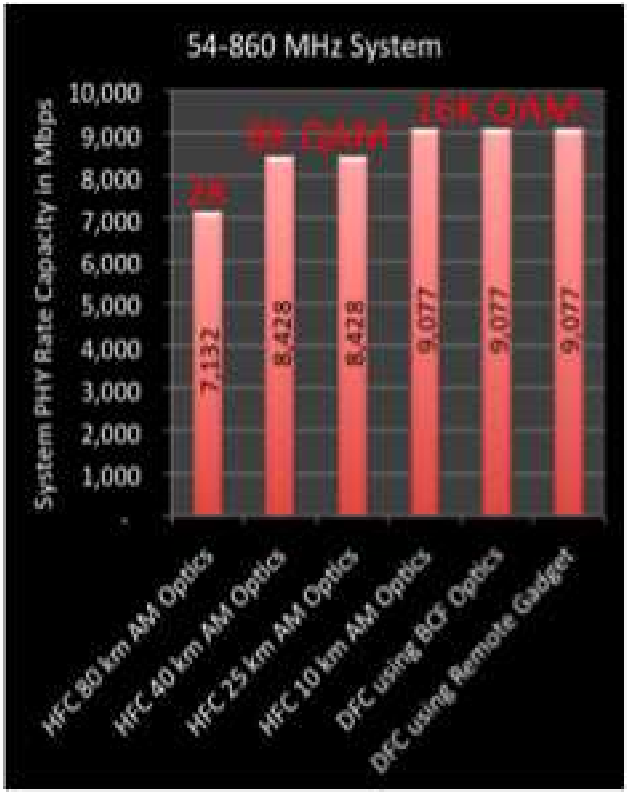


Figure 29: End-of-Line (EoL) Estimates and DOCSIS 3.1 Modulations in a 860 MHz Systems

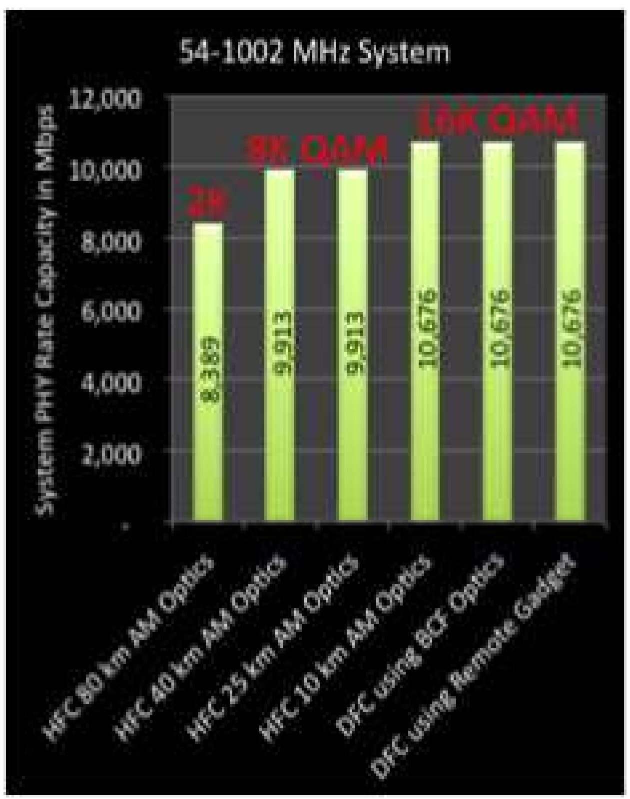


Figure 30: End-of-Line (EoL) Estimates and DOCSIS 3.1 Modulations in a 1 GHz Systems

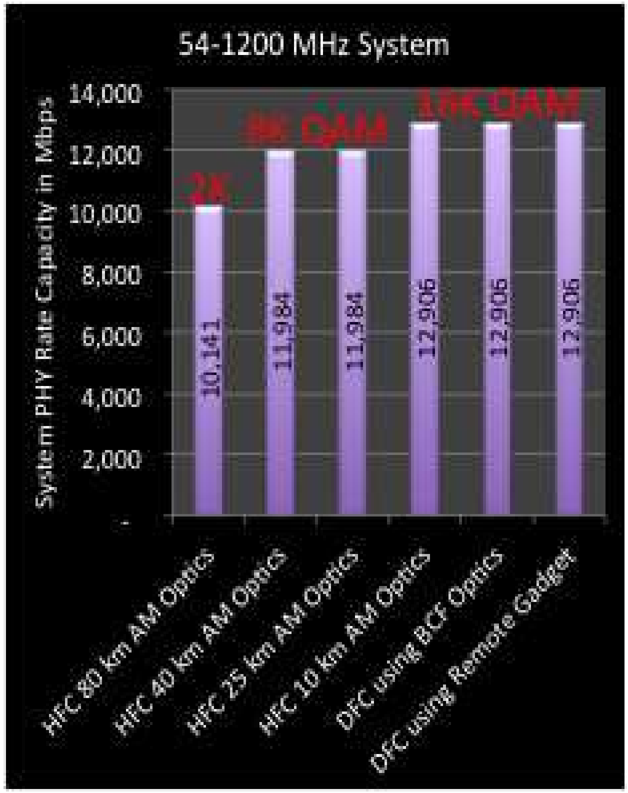


Figure 31: End-of-Line (EoL) Estimates and DOCSIS 3.1 Modulations in a 1.2 GHz Systems

A major takeaway from this section is that Hybrid Fiber Coax (HFC) use of amplitude modulation (AM) optics will enable the high order modulation defined in DOCSIS 3.1. It is unclear if the end-to-end systems (CCAP and CM) or if the operation can support such high order modulation in real-world, but it should be pleasing that the simple, transparent and flexible HFC Fiber to the Node (FTTN) class of architecture will support the future needs of the MSO. This will allow the MSOs keep intelligence out of the outside plant (OSP) until such point where performance, distance, or wavelengths demand the use of Digital Fiber Coax. The use of digital optics does not require the breakup and dismantling of the CCAP, an MSO if desired could keep the CCAP MAC and PHY functions in the headend and/or primary hub site and use Broadband Digital Optics like Broadband Compressed Forward (BCF) to meet the performance of Remote Gadget.

THE OPTICAL SEGMENT WAVELENGTH CAPACITY CONSIDERATION DETAILS

The use of Amplitude Modulated (AM) optics can take advantage of CWDM and DWDM optical transmission techniques to maximize the optical segment of the network. In a narrowcast overlay architecture, we assume as many as 40 wavelengths / lambdas per fiber, 80 QAMs of narrowcast spectrum, and a reach of approximately 100 km to the node. HFC optical distance will vary based on many factors, including narrowcast channel loading, the number of analog video channels, and many other factors. However, in the example above the use of AM optics for full spectrum and the desire that entire spectrum enable the highest order modulation possible this will reduce the number of wavelengths. In the model we assumed that 8 wavelengths were possible for AM optics to enable the modulation orders and capacities of each system bandwidth. The us of digital optics will maximizes the optical segment while enabling the high modulation orders, the wavelength capacity for digital optics may be between 3 to 4 times AM optics wavelength capacity. If digital optics uses CWDM supporting 16 lambdas and using DWDM supporting 44 lambda this is a combined up to 58 lambdas and at this level this is 7 times that AM optical wavelength capacity and enable the high order modulation at nearly any distance.

In some cases, fiber count is insufficient, regardless of the distance. Therefore, to avoid over lashing new fiber to service groups, separate wavelengths are placed on the fiber. The use of HFC analog optics today supports far fewer optical wavelengths than that which is supported using digital optical technology. This may be a challenge for HFC AM optics.

THE HEAD END SPACE & POWER CONSIDERATION DETAILS

Scaling Centralized Access Architectures

With the continued 50% growth rates in bandwidth capacities, some operators are concerned that they may need to continue to split nodes and need a dozen or more times the number of Service Groups (SG) than they have today. There is a fear among some that traditional CCAP boxes will not keep pace with this growth in SG. This could result in operators running out of both space and power in their existing Head End facilities. [ULM] studies this in detail.

While there are significant variations between head ends, that paper focused on a conservative "normalized" head end footprint that is represented in the first row of Table 1. The model head end requires 10 racks of space today to support about 200 Service Groups (SG) using existing CMTS, EQAM, RF Combining and Optic shelves. That's an average of 20 SG per rack. This is the baseline for the analysis.

The next step is the migration to today's CCAP platforms and existing optic shelf technology (row 2). This could squeeze 200 SG into 3 total racks. This uses 2013 technology of 56 SG per CCAP and optic rack density of 60 SG per 12RU. That results in an average of 70 SG per rack, which results in a 3.5X improvement over our baseline configuration. This means that we might fit 700 SG into the existing 10-rack footprint.

A 2nd generation traditional CCAP (row 3) might achieve at least 25% increase in SG density. This pushes the CCAP up to ~70 SG per chassis. At the same time, optic shelf rack density increase in the last year from 60 SG per 12RU up to 80 SG per 12RU. Using these two inputs, the next step in the CCAP evolution should get us down to 2 racks to support 200 SG. That's an average of 100 SG per rack for a 5X increase over our baseline of today's CMTS/EQAM based head ends.

Configuration	Space Needed For ~200 SG	SG per 1 Rack	Relative Scale
2012 Head End – CMTS, EQAM, RF Combining, Optics	~10 Racks	~20 SG	1X

2013 Traditional CCAP (56 SG) + **Optics Shelf** (60 SG per 12RU)	~3 Racks	~70 SG	3.5X
2nd Gen CCAP (~70 SG) + **2014 Optics Shelf** (80 SG per 12RU)	~2 Racks	~100 SG	5X
Future 2020 CCAP (~200 SG) + **Optics Shelf** (120 SG per 12RU)	~1 Rack	~200 SG	10X

Table 1 – CCAP Space Savings Example, CAA

The analysis in [ULM] shows that traditional CAA-based CCAP systems can just about quadruple densities of today's CCAP by the year 2020. That would put them around 200 SG per CCAP chassis. This combines with the expected continued advances in optical shelf rack densities to 80 SG per 8RU. The result (row 4) achieves 200 SG in a single rack within this decade. That provides a 10X increase in SG growth within today's existing head end footprint.

In addition to this SG growth, D3.1 will also give a giant boost to the SG capacity. Starting from today's CCAP system that provides about 1 Gbps (i.e. 32 DOCSIS® channels), D3.1 can provide more than 10Gbps per SG. So, the bottom line is that traditional CAA-based head end systems can leverage CCAP + AM optic advances to get both a 10X increase in SG counts in conjunction with 10X increase in capacity per SG before the end of this decade.

Head End Space for DAA

As was shown in [ULM], the primary limitation for SG density in CCAP is the RF connectors. For a DAA based head end system, the RF in the head end goes away. The DAA impact is shown in the expanded Table 2.

Our analysis shows that both BCF and Remote PHY can achieve roughly twice the SG density in the CCAP once the RF connector restriction is removed. This could allow SG densities of 400 SG per rack as shown in rows 5 and 6.

Configuration	Space Needed For ~200 SG	SG per 1 Rack	Relative Scale
2012 Head End – CMTS, EQAM, RF Combining, Optics	~10 Racks	~20 SG	1X
2013 Traditional CCAP (56 SG) + **Optics Shelf** (60 SG per 12RU)	~3 Racks	~70 SG	3.5X
2nd Gen CCAP (~70 SG) + **2014 Optics Shelf** (80 SG per 12RU)	~2 Racks	~100 SG	5X
Future 2020 CCAP (~200 SG) + **Optics Shelf** (120 SG per 12RU)	~1 Rack	~200 SG	10X

Future CCAP (~400 SG) + BCF/BCR	**~0.5 Rack**	**~400 SG**	**20X**
Future CCAP (~400 SG) + R-PHY	**~0.5 Rack**	**~400 SG**	**20X**
Future CCAP (~800 SG) + R-CCAP aggregation	**~0.25 Rack**	**~800 SG**	**40X**

Table 2 – CCAP Space Savings Example, CAA + DAA

For R-CCAP, the entire MAC and PHY layers have been moved out of the existing head ends. What's left in the head end or "cloud" from the data path perspective is primarily Ethernet aggregation. This provides further consolidation in head end space and roughly doubles the SG density compared to BCF/BCR and Remote PHY. This could push SG density up to 800 SG per rack for a 40x SG increase within today's existing space and power footprint.

SUMMARY OF THE SIDE-BY-SIDE COMPARISON HFC VS. DFC AND CAA VS. DAA

Some of the most often asked questions by cable industry forward-looking planners reflect the key challenges the industry is facing for this decade and beyond. Some of these challenges and questions include:

1. **Can Digital Fiber Coax (DFC) architectures maximize the coaxial segment revenue spectrum capacity?**

   Answer:
   - Yes and No (yes, assuming extremely high modulations are possible in the real world)
   - Yes, DFC may increase spectral capacity by 27% in 80 km spans and by 8% 25-40 km spans
   - No, DFC and HFC AM optics up to 10 km may support the same modulation order

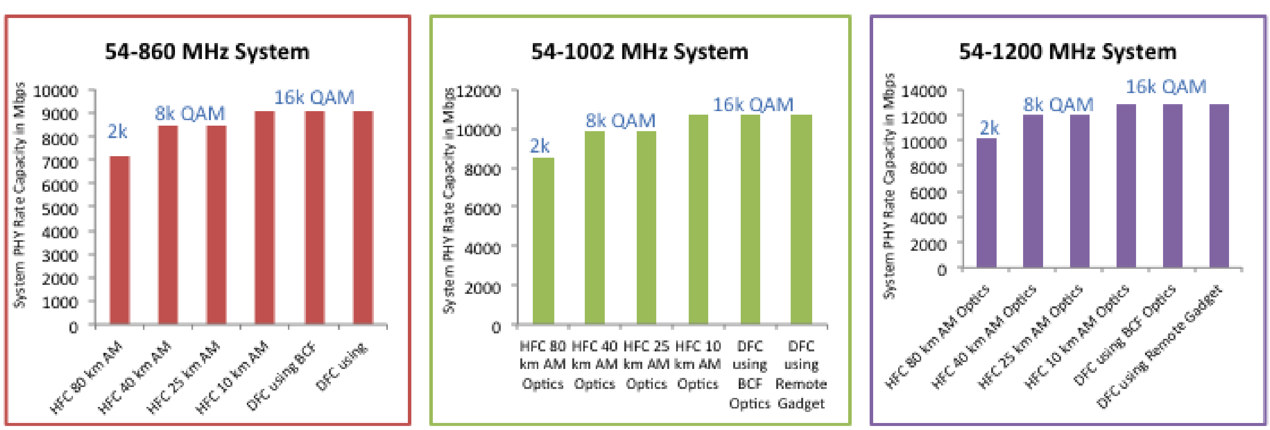


Figure 32: End-of-Line (EoL) Estimates and DOCSIS 3.1 Modulations Several Cable Systems

Assumptions:
- Capacity Comparisons of HFC AM Optics vs. DFC BCF VS. DFC Remote Gadget

- Assumes: Full Spectrum DOCSIS 3.1 Gen2 CCAP, up to 8 Lambdas, up to 1 GHz N+2, and 1.2 GHz N+0
- Capacity estimates based on modeled End of Line (EoL) performance with alignment to the highest order modulation possible
- Please note though EoL estimated may support given Modulation Order no DOCSIS 3.1 systems are available to confirm estimates.

2. **Can Digital Fiber Coax (DFC) architectures maximize the optical segment wavelength capacity?**

Answer: Yes

- Digital Optics Maximizes Optical Segment Wavelength Capacity between 3 to 4 times AM Optics Wavelength capacity

- Digital Optics using CWDM supports 16 lambdas and using DWDM supports 44 lambda with a combined up to 58 lambdas

3. **Can Digital Fiber Coax (DFC) architectures maximize facility space, power and cooling?**

Answer:

- It Depends this may reduce the headend requirements & increases the OSP
- HFC or DFC with CAA will meet or exceed the service group growth projections in most cases
- DAA reduces space, power & cooling in the headend but increase these factors in the OSP.

Access Architecture (HFC or DFC and CAA or DAA) Headend Equipment Types	Space For ~200 SG	SG per 1 Racks	Advance Scale
HFC using AM Optics and **CAA (CMTS)** Year 2012	~10 Racks	~20 SG	1X
HFC using AM Optics and **CAA (CCAP Gen1)** Year 2013	~3 Racks	~70 SG	3.5X
DFC using Broadband Digital and **CAA (CCAP Gen2)** ~2016	~2 Racks	~100 SG	5X
DFC using Broadband Digital Tx/Rx and **CAA (CCAP)** ~2020	~1 Rack	~200 SG	10X
DFC using **DAA with CCAP Remote PMD**	~0.5 Rack	~400 SG	20X
DFC using **DAA with CCAP MAC Core and Remote PHY**	~0.5 Rack	~400 SG	20X
DFC using **DAA with Remote CCAP Aggregation**	~0.25 Rack	~800 SG	40X

Table 3 – CCAP Space Savings Example, CAA + DAA

It is assumes ~32RU per rack are available after power supplies and the table does not show continued improvements in HFC AM Optical Headend Densities (contact ARRIS)

4. **<u>Can Digital Fiber Coax (DFC) architectures maximize long links and facility consolidation?</u>**

CAA with AM Optics limits long links & headend consolidation

- 10 km End of Line support estimates of 16K QAM in full spectrum
- 40 km End of Line support estimates of 8K QAM in full spectrum
- 80 km End of Line support estimates of 2K QAM in full spectrum

DOCSIS MAC in the headend limits long links & headend consolidation

- 160 km limit when DOCSIS MAC is in the headend (CAA BCF or DAA R-PHY)
- The DOCSIS MAC in the HE/Hub “plus” Digital Optics exceeds AM optics without reducing use of high order modulation but the DOCSIS MAC to CPE separation shall not exceed 160 km, thus DOCSIS is the limiting factor

DAA with Remote Access Shelf (R-AS) or Remote CCAP (R-CCAP)

- Over 160 km & virtually no headend-to-customer limit (Remote Access Shelf R-AS or Remote CCAP R-CCAP)
- Placing the DOCSIS “MAC” in the node / MDU removes the DOCSIS 160 km distance limitation, so now digital optics performance vs. cost vs. facility consolidation will determine how far beyond 160km is practical

5. **<u>Can Digital Fiber Coax (DFC) architectures maximize the economics of OPEX and CAPEX?</u>**

ANSWER: It is too early to tell for sure when weighing all factors

Identified previous benefits of DFC Drivers will improve the Economics (OPEX and/or CAPEX)

- Coax Segment: capacity (b/s/Hz) range from zero to small (depends)
- Optical Segment: wavelength are maximized with DFC
- Space: I-CCAP with AM optics densities will exceed SG growth rates
- Long Links and Headend Consolidation: is expanded with DFC without reducing coaxial capacities

End-to-End Solution OPEX and CAPEX

- Increase in service group OSP plant power and battery suppliers
- Overall failure rates could increase with intelligence in the OSP
- MTTR/MTTD may increase with more intelligence in the OSP
- Benefits of sharing optical transport link to carry other technologies
- End-to-End cost could improve with standard digital optics.
- True costs comparisons are unknown and validated

CONCLUSIONS

The use of Digital Forward and Return may place the lowest layer of the PHY in the node, like the ADC and DAC to the entire PHY and may also place the entire MAC and PHY in the node. It is too early to tell which Remote Access Layer architecture is best to enable digital optics. It should be noted that AM optics will support high order modulations in the majority of MSO FTTN applications today, but there are limitations. The use of digital forward and return independent of which architecture may not be desired or used by all MSOs and even within an MSO. Further industry research is needed to determine the best DFC architecture.

Conclusion Summaries

Digital Fiber Coax (DFC) just an additional tool in the MSO tool bag

- CAA using Broadband Compressed Forward or
- DAA with Remote Gadget (Remote PHY or Remote CCAP)
- The DFC solutions could just be used where / when / if needed

Keeping the CCAP intelligence together and in the headend

- Keeps the OSP simple and transparent
- Enable by the HFC optics and CAA option
- Enable by the DFC (Broadband Compressed Forward) and CAA option

DFC may be used as a tool in extreme cases:

- Where there is a need for Massive Service Group (SG) expansion
- Example: one (1) node per 500 HHP moves to 20 nodes per 500 HHP SG
- If locations are fiber starved and/or headend space constraints exist

Where there is a need for extremely long distance between facility and fiber node

- Broadband Compressed Forward or Remote PHY extends reach to 160 km
- Remote CCAP (MAC/PHY in the Node) extends reach beyond 160 km to enable massive headend consolidation virtually without limits

A major takeaway from this section is that Hybrid Fiber Coax (HFC) use of amplitude modulation (AM) optics will enable the high order modulation defined in DOCSIS 3.1. It is unclear if the end-to-end systems (CCAP and CM) or if the operation can support such high order modulation in real-world, but it should be pleasing that the simple, transparent and flexible HFC Fiber to the Node (FTTN) class of architecture will support the future needs of the MSO. This will allow the MSOs keep intelligence out of the outside plant (OSP) until such point where performance, distance, or wavelengths demand the use of Digital Fiber Coax. The use of digital optics does not require the breakup and dismantling of the CCAP, an MSO if desired could keep the CCAP MAC and PHY functions in the headend and/or primary hub site and use Broadband Digital Optics like Broadband Compressed Forward (BCF) to meet the performance of Remote Gadget. Perhaps conside use of Remote CCAP in the extreme cases where AM optics and BCF and Remote PHY cannot meet the needs of extreamly long fiber spans for headend consolidation and if there are serious space concerns.

REFERENCES

(1) [ULM] – “Scaling Traditional CCAP To Meet The Capacity Needs Of The Next Decade”; J. Ulm, J. Finkelstein, S. Rahman, J. Salinger, The Cable Show Spring Technical Forum 2014

AN IMPROVEMENT PROPOSAL FOR THE TIMING AND SCALING OF DOCSIS IP MULTICAST SERVICES

William T. Hanks, Tom Cloonan, Amit Eshet, Jeff Howe,
John Ulm, Ian Wheelock
ARRIS

Abstract

This paper will present a proposal for an extension to the DOCSIS protocol to facilitate the use of a "DOCSIS Multicast Carousel" table which periodically transmits DOCSIS-specific IP multicast forwarding information for each of the currently active multicast group addresses within a MAC Domain Downstream Service Group (MD-DS-SG). The CM may utilize this table of information to rapidly change its own multicast filters (and, if necessary, quickly retune one or more pre-designated downstream channel receivers) in response to subscriber IP video channel change requests without requiring additional control signaling to or from the CMTS. When the subscriber has settled on a program selection for a provisioned timeframe, the final selection and a list of interim selections may be signaled to the CMTS for metrics collection. This proposal is being submitted in an effort to mitigate several issues that have been identified as MSOs and vendors continue to explore IP Video Multicast deployment scenarios. These issues include the large bandwidth requirements that may be needed to support unicast bursts for fast-channel change assistance. The issues also include the high processing and message exchange rates (between CMs and CMTSs) that may be required when large numbers of subscribers initiate simultaneous channel change requests,

PROBLEM BACKGROUND

Viewer Behavior

Television program selection menu systems have been around for many years - allowing a direct selection of a single channel. In spite of this, many television viewers have the habit of using a television or set-top box remote control to rapidly switch between multiple channels – using either the channel up/down buttons or a "next favorite channel" button - before settling on a program. This behavior tends to occur en-mass at half-hour intervals as a previous program ends and a new one begins. This behavior can also occur at other moments during the hour when different network broadcasters tend to have un-planned synchronization between their respective commercial breaks.

IP Multicast Transport Technique

Modern Internet Protocol (IP) Video technology utilizes the concept of IP multicast to distribute a linear (continuous, sequential programming) video stream transmission from a single video IP source to a potentially large number of subscribers over a wide area in an efficient manner. The use of IP multicast allows a transmission to only be carried to the subscribers which have requested (and are entitled to receive) the video program using only the transmission links which are necessary to reach them. Furthermore, if multiple endpoint clients have access to the same transmission link (as they do in a shared medium technology such as Ethernet or DOCSIS), then these endpoint clients can each receive (and share) the same transmission simultaneously. This sharing of multicast transmissions can result in great bandwidth savings within the DOCSIS network, because it eliminates the wasted bandwidth that would otherwise be created by the many simultaneous unicast transmissions of identical video content to the multiple endpoint clients within a common Service Group. In some cases, up to 80% bandwidth savings can be realized in IP Video distribution networks through the use of IP multicast transmissions (in place of the less-efficient IP unicast transmissions). These dramatic improvements in bandwidth efficiency have made IP multicast a very important technology within most of the MSO IP Video distribution networks that are currently being planned for the future.

DOCSIS IP Multicast

The majority of modern cable high speed data transmission systems follow the Data Over Cable System Interface Specification (DOCSIS) protocol [1] [2] [3]. The cable modems (CMs) that predate the release of version 3.0 of DOCSIS (pre-3.0 DOCSIS) relied on the snooping of IGMPv2 messaging by the CM. The DOCSIS 3.0 protocol, in an attempt to make the cable modem be multicast protocol agnostic, moved the multicast control plane hooks from the CM device to the Cable Modem Termination System (CMTS). This approach was taken in an attempt to simplify the cable modem operation and to reduce the overall cost of deploying multicast solutions.

This newer DOCSIS 3.0 multicast architecture requires that the CMTS label all packets which are part of a particular multicast group transmission with a 20-bit Downstream Service ID (DSID). The CM uses this DSID to identify the packets which must be forwarded as part of this multicast transmission. The architecture relies on the successful runtime execution of a heavyweight, fault-tolerant, three-way Dynamic Bonding Change (DBC) signaling message transaction between the CMTS and the CM for each new multicast group service flow that the individual subscriber device wishes to receive. This exchange communicates the DSID, the downstream channel set over which the multicast transmission will be sent, and optionally, some packet resequencing parameters and forwarding directives to the CM.

The Perfect Storm

The scenario intersection of the en-mass rapid channel-changing behavior of video viewers and the overhead associated with the DOCSIS 3.0 multicast DSID forwarding architecture may cause problems both in viewer expectations and system capabilities. The signaling associated with a sequence of rapid channel changes may cause previously unexperienced latency for each single subscriber viewer. The signaling overhead for a large number of subscriber viewers performing rapid channel changes at the same time may also cause scaling issues for the DOCSIS CMTS system.

In addition, many MSOs are proposing to use unicast bursts of Adaptive Bit-Rate (ABR) video traffic to rapidly fill the video buffers within their video client devices to provide Fast Channel Change operations. The simultaneous transmission of these unicast bursts of ABR video traffic when many clients request channel changes at the same time can aggregate together to create very high transient bandwidth requirements within the DOCSIS network. Architecting the network to support the bandwidth needs of these transient bursts may yield higher-cost, higher-bandwidth systems with very low average utilization levels- which may be viewed as cost-prohibitive. Architecting the DOCSIS networks with lower bandwidth levels that ignore these unicast ABR bandwidth bursts may, however, lead to periodic reductions on subscriber Quality of Experience levels for all services, as the simultaneous unicast ABR bandwidth bursts will undoubtedly lead to periodic congestion within the DOCSIS network.

In a recent study of channel change behavior of more than 100,000 tuners (active video clients + digital video recorders) over several weeks with an available catalog of between 300 and 600 video programs across multiple operators and multiple video technologies, we observed that the worst case service group with 500 tuners had as many as 104 channel changes over an 8 second window, 95 channel changes over a 4-second window, and a peak of 50 channel changes within a one second window. While channel change statistics for larger populations of tuners (multiple service groups serviced by the same device) do not scale linearly, a population of almost 30,000 tuners can be expected to require a maximum of about 700 channel change operations in a one second interval.

If the servicing device has DOCSIS 3.0 CMTS functionality, then each of these channel change operations may result in a multicast group membership LEAVE operation (to leave the channel previously viewed) followed by a multicast group membership JOIN operation. In turn, each LEAVE or JOIN operation requires a successful three-way DBC signaling transaction between the CMTS device and the requester's CM. Thus, each single channel change event within a subscriber's home produces a total of eight protocol exchanges between the CM and the CMTS, so the 50 channel changes per second that are expected for a small 500-tuner Service Group can result in a total of 400 protocol exchanges per second between the CM and the CMTS (which equates to a protocol exchange every 2.5 milliseconds on average). A CMTS with many Service Groups would therefore experience much higher protocol exchange rates. The CMTSs must also perform multicast routing protocol exchanges with their Northbound network to initiate the IP multicast flows.

Shell Game for Protocol-agnostic CMs

As mentioned previously, the DOCSIS 3.0 MAC protocol moved the multicast control plane hooks from the CM device to the Cable Modem Termination System (CMTS) device in an effort to simplify the cable modem operation and to reduce the overall cost of deploying multicast solutions. These are very reasonable goals to strive for. However, the architectural solution adopted by the DOCSIS 3.0 protocol moved the requirement for multicast protocol snooping (IGMPv2 at the time but now extending to IGMPv3 and MLD v1/v2) from the CM device to the CMTS. In doing so, the CMTS must not only look for every possible IP multicast signaling packet in the hardware-based data stream and send them all to a control-plane processor (an activity which is not at all desirable when vendors are being asked to lower infrastructure costs on a multiple-gigabit router such as the CMTS device), but the processor must then also ensure that the requested multicast group media stream is available at the CMTS NSI (possibly issuing a multicast routing protocol message to the northbound router cloud) and then execute a heavyweight, fault-tolerant (possibly requiring message retransmissions), three-way Dynamic Bonding Change (DBC) signaling message transaction over a cable upstream transmission medium which is prone to noise.

What this multicast processing shell game has done is concentrate the vast majority of the multicast processing requirements onto a single device- the CMTS. Furthermore, the MSO demands for increased scaling on this device means that this single CMTS device must be engineered to address the peak scaling requirements of all multicast requestors across the entire CMTS; not just the much lower average processing requirements. This need for engineering to peak rates seems to be in conflict with demands for decreased infrastructure costs.

While IGMP (for IPv4 networks) and MLD (for IPv6 networks) are the protocols typically used for signaling multicast group membership requests, these are not the only ways that the CMTS might be signaled to add a particular device to an IP multicast group. IP Multicast signaling extensions have been added to the PacketCable Multimedia protocol [4]. Additionally, vendor-proprietary extensions may exist whereby a CMTS or other device might snoop a different type of media request (perhaps HTTP-Get request from a unicast adaptive bitrate (ABR) request?) from a customer video device and then signal the CMTS in another way to initiate a multicast flow.

Regardless of the signaling mechanism utilized to initiate the multicast stream transmission to each multicast group client, the same heavyweight DBC transaction must be used in order to communicate the DOCSIS MAC-level details for the IP multicast stream from the CMTS to the CM.

In addition, the simultaneous transmission of unicast ABR bursts for Fast Channel Change operations can lead to increased bandwidth requirements within the DOCSIS network.

A single, simple technique for mitigating these two fundamental IP Video problems (heavy CMTS processing loads and high DOCSIS bandwidth requirements) may be beneficial for the cable industry to consider at this point in time.

THE PROPOSED SOLUTION

SDV Inspiration

The authors propose borrowing a solution from the Switched Digital Video (SDV) solution space by implementing a DOCSIS Multicast Carousel (DMC) which is periodically transmitted and is received by all cable modems within a service group. In keeping with a protocol layering architecture, the DMC may well be published in concert with and in support of any application-layer IPTV signaling carousel. The contents of each Carousel type-length-value structure (described later) consists of the DOCSIS data which is necessary for the cable modem to receive and forward the contents of exactly one IP multicast group.

Architecture Assumptions

The authors have made the following assumptions about the DOCSIS IP Video system:

1) In order to eliminate the need for extra group replications and the need for channel reassignment for group reception, all Multicast Group sessions will be carried together on as few DOCSIS downstream channels as possible. For this discussion, we will call the number of downstream channels N.

2) All embedded or non-embedded cable modem devices that are to carry multicast IP Video are compatible with the GMAC Promiscuous DSID mode of Multicast DSID forwarding of the DOCSIS specifications and are capable of receiving at least N downstream channels. Ideally, these devices are capable of bonding all multicast group flows over the N channels as well.

3) The IP video application layer signaling will somehow cause the subscriber's CM device (either standalone CM or embedded within a home gateway) to attempt to join a particular IP multicast group session. This application layer signaling may be triggered by the IP Video player device, by a video application running on a home gateway,

some application sending directives from the network cloud, or any other possible method.

4) The IP video application layer signaling is NOT commingled with the contents of the DOCSIS Multicast Carousel. This way the DOCSIS Multicast Carousel can be legitimately defined as a DOCSIS MAC layer extension.

5) The Receive Channel Configuration (RCC) mechanism of DOCSIS by which a CMTS assigns channels to CMs can and will be augmented in a manner to designate a number of receivers to be under the control of the CM (i.e. not directly assigned by the CMTS) for the purposes of tuning to IP Multicast groups in a manner to be described later.

MAC Subinterface Access Control List

The Multiple Services Operator (MSO) may choose to make different IP Video multicast group sessions available to different sets of CMs based upon cable-side topology. Specifically, the CMs that are associated with one MAC Domain Downstream Service Group (MD-DS-SG) might be allowed to receive a slightly different set of multicast sessions from the CMs associated with a different MD-CM-SG. An Access Control List (ACL) with permit/deny directives per group IP address might be assigned to each MD-DS-SG. Ideally, this ACL might be assigned with directives per source IP, Group IP (S, G) tuple.

In order to facilitate this ability, a set of configuration directives may be required within the CMTS to map an MD-CM-SG to a MAC Domain subinterface. Furthermore, each MAC Domain subinterface might present a different cable-helper address to the Dynamic Host Configuration Protocol (DHCP) server for the CMs of a MD-DS-SG once the CMTS determines (or is told by the CM as a result of CM topology resolution) the MD-DS-SG identifier of the CMs.

The DOCSIS Multicast Carousel

In order to define the parameters necessary for a DOCSIS Multicast Carousel, we begin by studying the parameters of the communications that take place when a television subscriber tunes an IP Video Player device to receive program content data over an IP multicast group using a DOCSIS 3.0 CM device.

Table 1 contains the structure of the payload of a DOCSIS Dynamic Bonding Change Request (DBC-REQ) message which was used by a CMTS to successfully install a resequencing DSID for a DSID-forwarded multicast group. The payload is in the Type-Length-Value (TLV) format that is specified in [1]. Items which are not necessary for the DMC are stricken with the explanation to follow.

Type	Len	Type	Len	Type	Len	Type	Len	Value
50	N							DSID Encodings
		1	3					DSID value (20 bits)
		~~2~~	~~1~~					~~Downstream Service Identifier Action: Add~~
		~~3~~	~~N~~					~~Downstream Reseq. Encodings~~
				~~1~~	~~1~~			~~Reseq. DSID flag (1=Resequencing DSID)~~
				2	n			DS Channel ID (byte) array [DCID1, DCID2, …, DCIDn]
				~~3~~	~~1~~			~~DSID reseq. wait time (1-180) x 100 µs~~
				~~4~~	~~1~~			~~DSID Reseq. Warning Threshold (1-179) x 100µs~~

Type	Len	Type	Len	Type	Len	Type	Len	Value
				~~5~~	~~2~~			~~CM-STATUS Hold-off Timer for Out-of-range events (in 20 ms units)~~
		4	~~N~~					~~Multicast Encodings~~
				~~1~~	~~N~~			~~Client MAC Addr Encodings~~
						~~1~~	~~1~~	~~Action (0=add; 1=delete)~~
						~~2~~	~~6~~	~~Client MAC Address joining or leaving~~
				~~2~~	4			~~Mcast CM Interface Mask bitmap~~
				3	~~n~~			~~Mcast Group MAC Address array [GMAC0, GMAC1,…,GMACn]~~
31	1							Key Sequence Number
~~27~~	~~20~~							~~HMAC-Digest~~

Table 1: Contents of a DBC-REQ message to install a multicast DSID

Note: A typical shorthand used when discussing TLV subtypes is to list the cascading types (ignoring the intervening lengths) with dotted notation beginning with the main type. For example, the DSID Reseq. Warning Threshold (`Table 1`) would be known as TLV type 50.3.4

No Resequencing

If we assume that all multicast flows are UDP–based and are therefore resequenced by an upper layer application protocol then TLV type 50.3.1 becomes unnecessary. Optional parameters (TLV Types 50.3.3 and 50.3.4 in light grey) are also resequencing parameters and will not be included in the DMC.

No Individual CM Directives

Individual CM directives (TLV Types 50.4.2 and 50.4.3 in medium grey) are sent from the CMTS to the individual CM to tell the CM upon which interface to forward the multicast group packets. The information in these TLVs can be worked around by the CM noting upon which CM interface the video client is requesting the JOIN operation so these messages are not appropriate for the DOCSIS Multicast Carousel.

No Action Control

If we assume that the carousel message would include information for multicast groups which are available and would not include information for multicast groups which are not, then the "action" parameters of TLV Type 50.2 and 50.4.1.1 become unnecessary.

Since the goal of this feature is to have the CM receive the multicast request and, provided that the multicast group stream is available, to have enough information to process the JOIN; the information about the Client MAC can be determined from the interface upon which the request itself was received and the entire 50.4 TLV branch becomes unnecessary.

Since the CMTS is the only device which can issue a DMC message on the downstream, authentication of the sender via the HMAC–Digest is not necessary. Remember, IP video content itself is expected to be protected via end-to-end encryption.

Finally, since the TLV 50.3.2 is the only piece of information remaining under 50.3, TLV 50.3.2 can be promoted one level for the DOCSIS Multicast Carousel and 50.3 is not needed.

The remaining fields (in **shaded** rows) are necessary information that will form the basis for the information carried within an entry of the Multicast Carousel.

An IGMP request will have a Group IP address with an optional Group Source IP address (for Source-Specific Multicast or SSM). If the multicast stream is not source specific then the Group Source IP will be 0.

These are the database keys (not to be confused with the security key) that the CM will use to find the correct DSID and Key Sequence Number in the Carousel.

In the end, the DOCSIS Multicast Carousel message will consist of multiple TLV Type 1 messages as shown in Table 2.

Type	Len	Type	Len	Type	Len	Value
1	N					Multicast Group Carousel Encoding
		1				Group ID (GIP)
		2				Group Source IP (optional for SSM)
		3	N			DSID Encodings
				1	3	DSID value (20 bits)
				2	n	DS Channel ID (byte) array [DCID1, DCID2, …, DCIDn]
		4	1			Key Sequence Number

Table 2: Proposed contents of a Multicast Group Carousel Encoding entry for one multicast group session

The Multicast Group Carousel Encoding TLV Type 1 will be repeated within the carousel messages for each multicast group to be described. It is anticipated that all multicast groups which are active (or statically provisioned) within the CMTS and available (per ACL configuration) to CMs within the MD-DS-SG will be reflected in the DOCSIS Multicast Group Carousel message.

The multi-part Multicast Group Carousel (MGC) message will consist of one or more numbered fragments consisting of DOCSIS frames; each containing a number of Multicast Group Carousel Encoding TLVs. The MGC message fragments will include a configuration change count so that CMs monitoring the stream know when the stream has been updated. All MGC fragments from the same MAC Domain IP Address containing the same Configuration Change Count and MD-DS-SG-ID belong to the same MGC message. The CM MUST successfully receive all fragments of an MGC message before using the contents of the message. Fragment Sequence Numbers begin with fragment number zero and increase by one for each successive fragment in the message.

Priority of CMTS-Assigned DSIDs

Use of Multicast DSID Forwarding (MDF) for multicast IP Video requires the use of processor and/or memory resources on the CM that are identified with a Downstream Service IDentifier (DSID). The number of DSIDs (and their accompanying processor and/or memory resources) are limited by the CM’s design and the maximum number of DSIDs is communicated from the CM to the CMTS as part of the modem registration process. Per the DOCSIS protocol, the CMTS assumes that each DSID resource is managed and can be assigned by the CMTS at any time of its choosing.

For this reason, the CM may only use DSIDs and associated resources which have not been explicitly assigned by the CMTS. Furthermore, if all such resources are in use (explicitly assigned + Multicast Group Carousel applications) and the CMTS explicitly assigns a new DSID to the CM, the CM MUST immediately cease using one of the DSIDs for Multicast Group Carousel use (thereby losing its subscription to one of the Carousel-defined multicast groups) and immediately assign the DSID and resources as directed by the CMTS.

Keeping the CMTS Apprised of Multicast Group Subscription

The CM MUST inform the CMTS of its use of multicast DSIDs as part of Multicast Group Carousel processing. The CMTS MAY retain this information for accounting or other use but there will no longer be an intent to track all multicast membership in real time. In fact, periodic multicast polling of general and group-specific queries will be disabled to prevent further CM query response message avalanches at the CMTS. A new unicast MAC Management Messaging mechanism should be created to allow the CMTS to more slowly poll each single CM for its current multicast membership and recent multicast membership changes. The CM would respond in a unicast message with all of its current multicast DSID resource usage and a listing of (S,G) membership changes with an associated timestamp for each. The CMTS might use the returned information to update standard multicast management information as well as to determine when a dynamically added multicast group stream might be pruned from the MD-CM-SG. However, once the CMTS believes that a multicast stream may be pruned, the CMTS should then send a multicast group-specific membership query to make sure that no CM has dynamically JOINed since its last unicast poll.

Adding a New Multicast Group to the Carousel

A multicast group session may be created by any of the following ways:

1. Multicast group protocol membership request (JOIN)
2. CMTS static multicast session CM configuration file encoding (TLV Type 64, see [1] – Section C.1.1.27)
3. Statically-provisioned multicast group sessions
4. Other methods? (PCMM, etc.)

The Multicast group protocol membership request is the most common, and is typically the method used today on IP Video networks. This method uses a multicast protocol such as IGMP (version 2 or version 3) for devices running over IPv4 and the MLD protocol (usually version 2) of IPv6. This processing could proceed exactly as it does with the DOCSIS 3.0 protocol (using DBC messages to install DSIDs in the requesting CM(s)) and then the CMTS could add the group session information to the Multicast Group Carousel message for any MD-DS-SGs that might be allowed (via ACL provisioning) to carry the session. Care must be taken to handle race conditions between multiple requestors for the same session and the resultant Carousel update.

The static multicast session CM configuration file encoding is used by the MSO to provide the CMTS with the static ASM or SSM multicast sessions to which the CM should be configured to forward multicast traffic at registration time. The CMTS Static Multicast Session Encoding contains the Static Multicast Group Encoding and, if SSM, Static Multicast Source Encoding.

Statically provisioned multicast sessions might be created via the CMTS command line interface (CLI). The CMTS might be responsible for initiating a network-side JOIN using a multicast protocol (PIM SSM or similar) and then the session information might be added to the Multicast Group Carousel message for the assigned MD-DS-SGs.

Backward Compatibility with non-Carousel-Capable CMs

In order for this DOCSIS Multicast Carousel mechanism to work harmoniously with the general population of CMs, the CMTS must understand which CMs must use DBC messaging for each MDF multicast session and which ones can use the Carousel. This requirement implies the need for a CM capability exchange similar to the ones that are used for other DOCSIS features.

A USE CASE

First Client Joins to Initiate Multicast Stream

To illustrate how the DOCSIS Multicast Carousel might be used, different variations of a possible IP Video use case will be

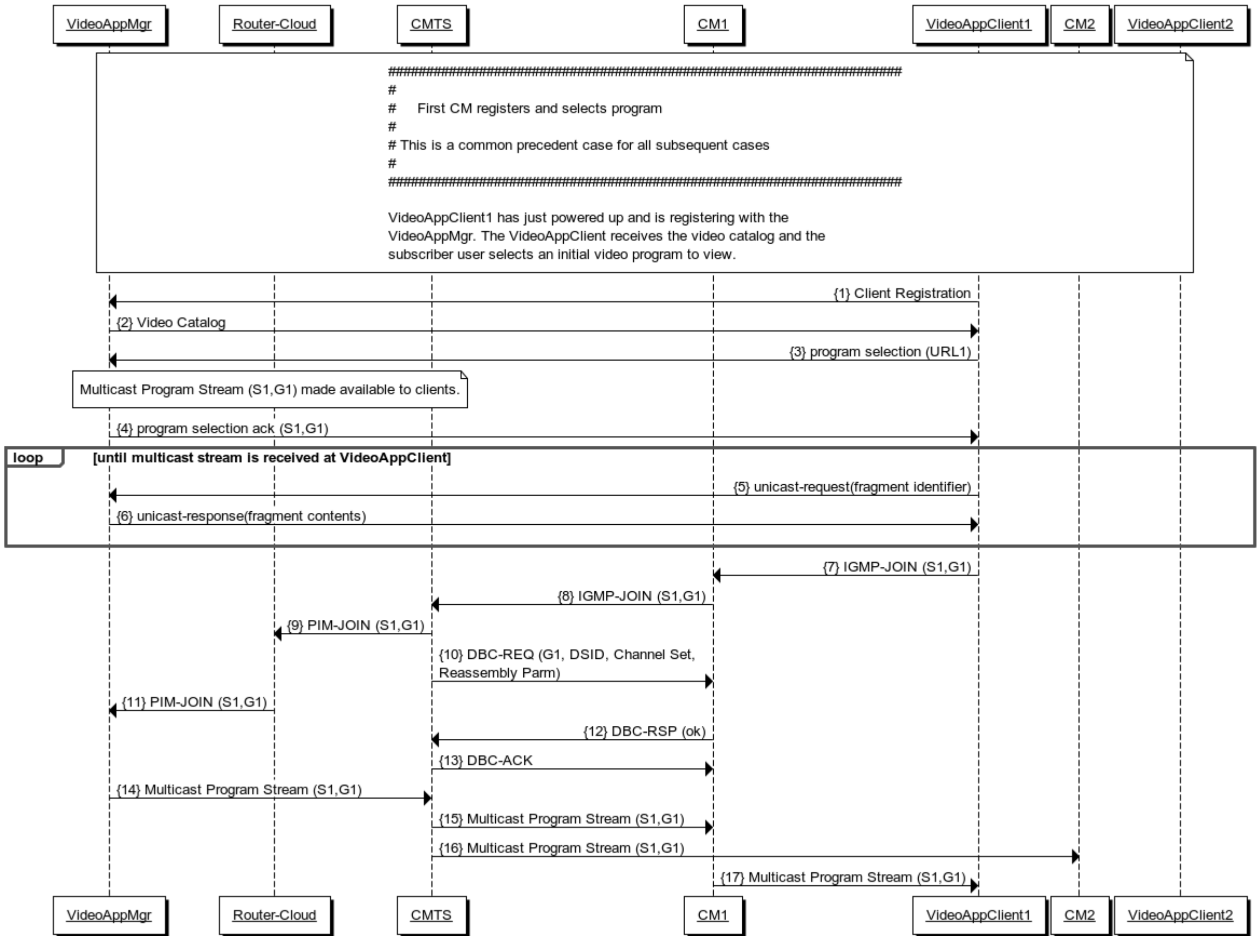


Figure 1: First client joins the IP Multicast group (common to all scenarios)

presented, beginning with the case of the very first client selecting a program which corresponds to a multicast group stream. Then the case of a second viewer of the same program within the same MD-CM-SG will be presented – using only the toolset as published in [1]. Finally, the case of a second viewer of the same program within the same MD-CM-SG will be presented with the use of the DOCSIS Multicast Carousel. This example presents just one of many possible scenarios that involve multiple CMs within the same MD-CM-SG becoming involved in the same IP multicast group session. Similar efficiency benefits can be gained through the use of the DMC in nearly all of the alternate scenarios.

In the following descriptions, each message number in braces {} refers to the numbered message in Figure 1.

A VideoAppMgr (which may include both video control and video media in this example) resides somewhere within an MSO network and a VideoAppClient1 behind a CM powers up and registers with the VideoAppMgr {1}. The VideoAppMgr provides a Video Catalog {2} (perhaps in the form of a browser menu screen) and the user makes a program selection. In this scenario, the VideoAppClient1 communicates this selection via the use of a URL {3}. The VideoAppMgr acknowledges the selection and responds that the video program can be found on a multicast video stream denoted by network source IP address S1 and multicast group IP address G1 {4}.

As is common in IP Video systems, the VideoAppClient1 uses unicast requests to the server {5, 6} to request video to fill its video jitter buffer before playing. VideoAppClient1 also initiates an IGMP-JOIN request for the multicast stream denoted by (S1, G1) {7}. The VideoAppClient1 will continue to request unicast video fragments until it begins to receive the multicast stream packets.

The CMTS, upon receiving the JOIN request{8}, must request the proper multicast stream from the router cloud using a protocol such as PIM-SSM {9}. The CMTS must also initiate a DBC transaction {10,12,13} to cause the CM1 serving VideoAppClient1 to begin forwarding multicast packets tagged with the DSID on the specified downstream channel set to the VideoAppClient1.

At some point the packets of the Multicast Program Stream identified by (S1, G1) begin to flow through the network, to the CMTS {14} and over the downstream channel set indicated in the DBC-REQ message {15,16}. The CM, having successfully completed its three-way DBC transaction, strips the DSID and forwards the multicast packets to VideoAppClient1 {17}. At this point, VideoAppClient1 ceases its unicast requesting cycle and switches to play the packets received via the multicast stream.

Subsequent Client(s) Join the Multicast Stream

At this point, the packets of the multicast program stream are being tagged with the DSID by the CMTS and are being sent down a downstream channel set that is within CM1's Receive Channel Set. At this point, if other subscribers (single instance using VideoAppClient2 behind CM2 is shown but the example can be repeated for multiple instances of JOINers at the same service group) wishes to join the same multicast group service flow denoted by (S1, G1), then a comparison can be made between the message flow using the DOCSIS 3.0 protocol as defined in [1] and the message flow as it might appear using the proposed DOCSIS Multicast Carousel.

NOTE: The message flows of the following use case diagrams are intentionally not monotonically sequential. Instead, corresponding messages are numbered the same on both diagrams and unique flows are labeled with either a "-a" or "-b", depending on which case to which they belong.

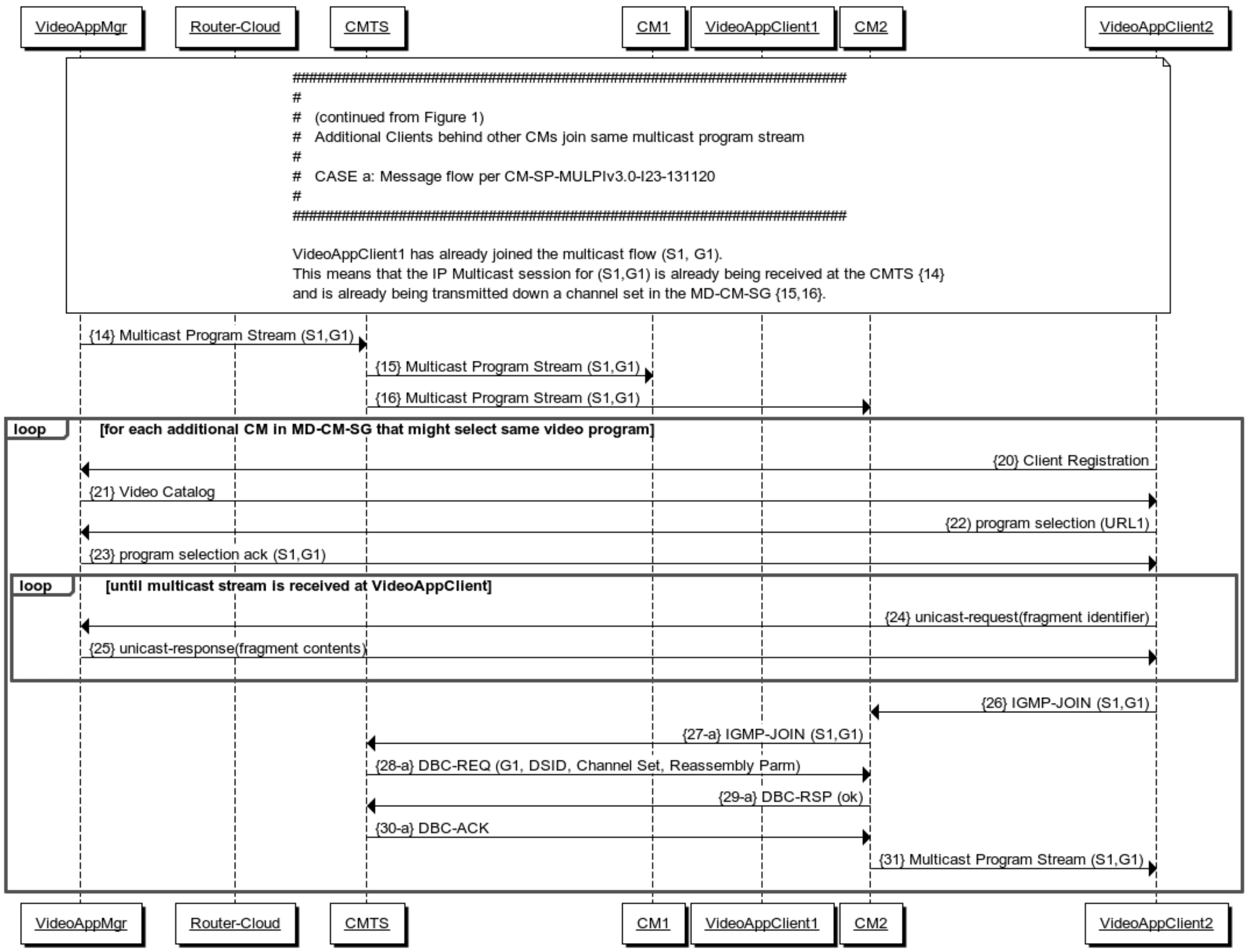


Figure 2: CASE a: Additional JOIN using DOCSIS 3.0 CMTS Control

<u>CASE a: Subsequent Clients using original DOCSIS 3.0 signaling</u>

Using the DOCSIS 3.0 protocol as defined in [1], the message flow would look something like Figure 2. Being a new VideoAppClient, the Client would need to register {20}, receive the video catalog {21}, select a program {22}, and receive a program selection acknowledgement {23}. Also, in order to fill the video jitter buffer, the Client would issue unicast requests {24} and receive fragment contents {25}. Upon receiving the IGMP-JOIN request {26} from the Client, CM2 blindly forwards the request to the CMTS {27-a}. Since the CMTS is already receiving the multicast program stream {14}, it does not need to issue any northbound requests to the Router-Cloud to get it. The CMTS does still need to initiate the three-way DBC transaction {28-a, 29-a, 30-a} to install the DSID onto CM2. Once the CM installs the DSID, it can begin forwarding the Multicast Program Stream to the VideoAppClient {31}.

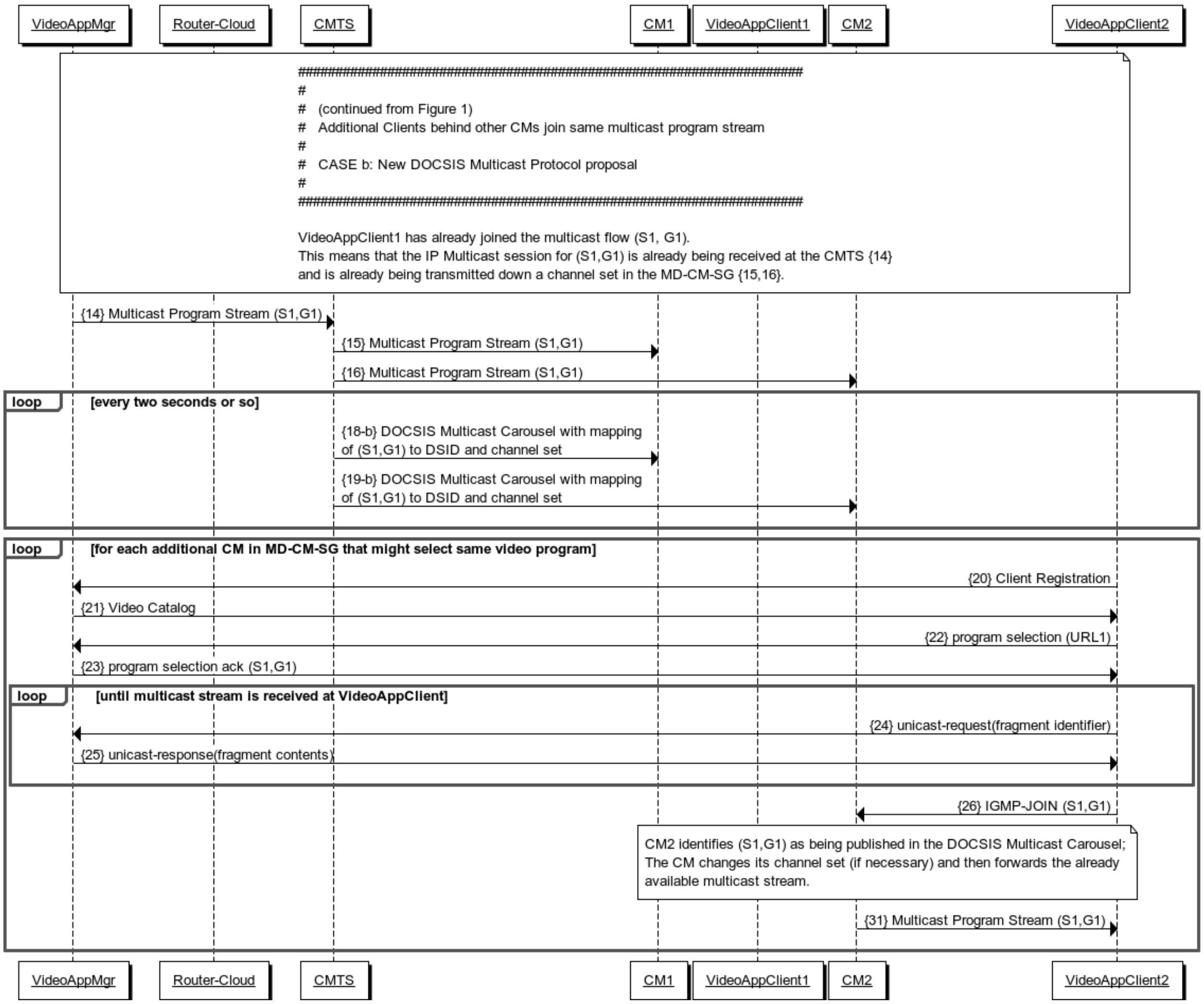


Figure 3: CASE b: Additional JOIN using DOCSIS Multicast Carousel

CASE b: Subsequent Clients using proposed DOCSIS Multicast Carousel

Using the proposed DOCSIS Multicast Carousel protocol, the message flow would look something like Figure 3. With the DOCSIS Multicast Carousel, the CMTS would periodically broadcast the mapping of the multicast stream corresponding to (S1, G1) to the DSID, channel set, and security information within the Multicast Carousel {18-b & 19-b; although this is really only one message that gets to all clients simultaneously}. Once again, being a new VideoAppClient, the Client would need to register {20}, receive the video catalog {21}, select a program {22}, and receive a program selection acknowledgement {23}. Also, in order to fill the video jitter buffer, the Client would issue unicast requests {24} and receive fragment contents {25}.

The interesting operation occurs when the Client sends the IGMP-JOIN request to the CM {26}. Since the CM is already receiving the (DSID, channel set, security info) mapping for the requested multicast group, it can retune one or more pre-designated (perhaps via a new parameter in the DOCSIS RCC encodings?) CM multicast IPTV receivers if necessary and install the required DSID and security information to immediately begin forwarding the Multicast Program Stream to VideoAppClient2 {31}.

Silent Multicast Group Leaves

With the DMC approach, multicast session LEAVEs can be processed in a similar manor as the JOINs without informing the CMTS about the LEAVE. As mentioned previously, a new unicast polling mechanism is being proposed to allow a CMTS-controlled reporting of channel changes. This mechanism can be used to determine when a multicast session might be a candidate to be pruned from the MD-CM-SG. Again, a group-specific query should be issued before the session is actually pruned to catch any CM which have silently JOINed since they were last polled.

Statistics Collection for Subscriber Viewing Activities

In many set-top box environments and/or media gateway environments today, MSOs have the ability to monitor and observe the fine-grained viewing activity of every subscriber, knowing when and how every channel change event was initiated. The use of the DMC approach eliminates (by desire) most of the transmission of multicast JOINs and LEAVEs between the CMs and the CMTS, so the CMTS is no longer cognizant of every channel change event. At first glance, this may appear to reduce the level of observability that the MSO may have into the channel change activities of their subscribers. This does not have to be the case.

Slight augmentations to the DMC proposal can restore the fine-grained observability of subscriber channel change activity to the MSO. In particular, the IP Video client application could be required to archive every channel change event, which could then be periodically collected by a Subscriber Management system to provide the MSO with a detailed view of each subscriber's channel change activities. Alternatively, the CM could be required to archive every channel change event, which could then be periodically collected by a Subscriber Management system or could be periodically collected by the CMTS to provide the MSO with a detailed view of each subscriber's channel change activities.

Benefits of DOCSIS Multicast Carousel

Looking at Figure 1, Figure 2, and Figure 3; the following desirable benefits result:

- Only the first JOINer to new multicast group will require direct CMTS action {8, 9, 10, 12, 13, 15, 16} as part of the channel change; steps {27-a, 28-a, 29-a, and 30-a} can all be avoided for every subsequent JOINer. In addition, even the delay for some of the first JOINers to multicast groups might be avoided if the most popular Multicast Program Streams are statically forwarded in each MD-CM-SG at system startup time.

- Every time that the CM can process the request on its own, significant time and message processing is saved in the channel changing process. In turn, this means that the unicast load imposed by fast-channel-change algorithms on the entire system will be lower than when using a CMTS-controlled multicast model.

- Since the DOCSIS Multicast Carousel is broadcast to all CMs at once, rather than being unicast to each CM like the three-way DBC transaction is, the CMTS can adjust the channel set of one or more multicast streams dynamically and all CMs can retune very quickly if one of the channels experiences a failure.

- Since the DOCSIS Multicast Carousel permits CMs to begin passing any accessible IP Video multicast packets directly through to the IP Video clients without incurring the delays of the IGMP/PCMM and DBC protocol exchanges, it is feasible that (with appropriate GOP selections) the IP Video multicast streams may be able to be rendered quite rapidly, resulting in relatively Fast Channel Change times without a unicast ABR burst. (Note: Only the first JOINers to an IP Video multicast stream would typically be forced to make use of unicast ABR bursts to expedite their channel change times). If this is the case, then the DOCSIS Multicast Carousel may greatly reduce the need for transient unicast ABR bursts to provide for Fast Channel Changes. As a result, the DOCSIS Multicast Carousel may also help to solve the aforementioned DOCSIS bandwidth problems associated with the aggregated bandwidth requirements of many simultaneously-transmitted unicast ABR bursts.

CONCLUSION

The effects of rapid changes amongst multicast program streams may pose a challenge for many DOCSIS-based IPTV systems in terms of multicast processing power, membership messaging latency, and unicast bandwidth needs for fast-channel-change techniques. This challenge can be overcome by using the DOCSIS Multicast Carousel - a technique borrowed from SDV technology - to transmit the information necessary for a CM to make its own tuning and multicast forwarding decisions, thereby relocating most of the time-critical multicast processing from the CMTS back amongst the distributed processors of the CM population at the very edge of the DOCSIS network.

Author Contacts

Tom Cloonan
Tom.Cloonan@arrisi.com

Amit Eshet
Amit.Eshet@arrisi.com

William T. Hanks
Bill.Hanks@arrisi.com

Jeff Howe
Jeff.Howe@arrisi.com

John Ulm
John.Ulm@arrisi.com

Ian Wheelock
Ian.Wheelock@arrisi.com

AN OVERVIEW OF HTTP ADAPTIVE STREAMING PROTOCOLS FOR TV EVERYWHRE DELIVERY

Yuval Fisher
RGB Networks

Abstract

In this paper we review the advantages of adaptive HTTP streaming and detail its usefulness in delivering content on both managed and unmanaged networks. We review the differences between these protocols and discuss their strengths and weaknesses. Additionally, we'll give a detailed look at the packaging and delivery mechanisms for TV Everywhere in which we explain the implications for storage, CDN distribution, ad insertion and overall architecture.

INTRODUCTION

The traditional television viewing experience has clearly changed as viewing video online, on a tablet or smartphone, or on the living room TV thanks to Internet-delivered content is increasingly commonplace. With considerable speed, consumers have passed the early-adopter phase of TV Everywhere and today an ever increasing number expect any program to be immediately available on any viewing device and over any network connection. What's more, they expect this content to be of the same high quality they experience with traditional television services. Regardless of whether this explosion of multiscreen IP video is a threat or an opportunity for cable operators and other traditional video service providers (VSPs), it's clear that it's here to stay.

Despite advancements in core and last mile bandwidth in the past few years, the bandwidth requirement of video traffic is prodigious. Combining this with the reality that the Internet as a whole is not a managed quality-of-service (QoS) environment, means new ways to transport video must be employed to ensure that the high quality of experience (QoE) consumers enjoy with their managed TV delivery networks is the same as what they experience across all their devices and network connections.

To address the conundrum of ensuring optimal quality despite the bandwidth-hungry nature of video and the lack of QoS controls in unmanaged networks, Apple, Microsoft, Adobe and MPEG have developed adaptive delivery protocols. These have been broadly adopted by cable operators and other VSPs, including traditional players and new entrants. The result is that networks are now equipped with servers that can 'package' high-quality video content from live streams or file sources for transport to devices that support these new delivery protocols.

This paper reviews the four primary HTTP adaptive streaming technologies: Apple's HTTP Live Streaming (HLS), Microsoft Silverlight Smooth Streaming (MSS), Adobe's HTTP Dynamic Streaming (HDS) and MPEG's Dynamic Adaptive Streaming over HTTP (DASH).

The paper is divided into three key sections. The first is an overview of adaptive HTTP streaming which reviews delivery architectures, describes its strengths and weaknesses, and examines live and video-on-demand (VOD) delivery. The second section looks at each technology, details how they work and notes how each technology differs from the others. The third and final section looks at specific features and describes how they are implemented or deployed. In particular, this last section focuses on:

- Delivery of multiple audio channels
- Encryption and DRM
- Closed captions / subtitling
- Ability to insert ads
- Custom VOD playlists
- Trick modes (fast-forward/rewind, pause)
- Fast channel change
- Failover due to upstream issues
- Stream latency
- Ability to send other data to the client, including manifest compression

ADAPTIVE BITRATE HTTP VIDEO DELIVERY

In Adaptive Bitrate (ABR) HTTP streaming, the source video – whether a live stream or a file – is encoded into discrete file segments known as 'fragments' or 'chunks.' The contents of these fragment files can include video data, audio data or such other data such as subtitles, program information or other metadata. These data may be multiplexed in the file fragment or can be separated into distinct fragment files. The fragments are hosted on an HTTP server from which they are served to clients. A sequence of fragments is called a 'profile.' The same content may be represented by different profiles that may differ in bitrate, resolution, codecs or codec profile/level.

Clients play the stream by requesting fragments from the HTTP server. The client then plays the fragments contiguously as they are downloaded. If the audio, video or other data are stored in separate fragment sequences, all are downloaded to form the video playback.

The video in each fragment file is typically encoded as H.264, though HEVC or other video codecs are possible. AAC is generally used to encode the audio data, but again, other formats are also used. Fragments typically represent 2 to 10 seconds of video. The stream is broken into fragments at video Group of Pictures (GOP) boundaries that begin with an IDR frame, which is a frame that can be independently decoded without dependencies on other frames. In this way, a client can play a fragment from the beginning without any dependence on previous or following fragments.

As the name suggests, adaptive delivery enables a client to 'adapt' to varying network conditions by selecting video fragments from profiles that are better suited to the conditions at that moment. Computing the available network bandwidth is easily accomplished by the client, which compares the download time of a fragment with its size. Using a list of available profile bitrates (or resolutions or codecs), the client can determine if the bandwidth available is sufficient for it to download fragments from a higher bitrate/resolution profile or if it needs to change to a lower bitrate/resolution profile. This list of available profiles is called a 'manifest' or 'playlist.' The client can quickly adapt to fluctuating bandwidth – or other network conditions – every few seconds as it continually performs bandwidth calculations at each fragment download. Local CPU load and the client's ability to play back a specific codec or resolution are factors – in addition to available bandwidth – that may affect the client's choice of profile. For example, a manifest file may reference a broad collection of profiles with a wide selection of codecs and resolutions, but the client will know that it can only play back certain profiles and therefore only request fragments from those profiles.

This model enables delivery of both live and on-demand-based sources. A manifest file is provided to the client in both scenarios. The manifest lists the bitrates (and potentially other data) of the available stream profiles so the client can determine how to download the chunks; that is, the manifests tells the client what URL to use to fetch fragments from specific profiles. In the case of an on-demand file request, the manifest contains information on every fragment in the content. This is not possible when it comes to live streaming. DASH, HLS and HDS deliver 'rolling window' manifest data that contains references to the last few available fragments, as shown in Figure 1. Continual updating of the client's manifest is necessary to know about the most recently available chunks. For MSS, a rolling window-type manifest isn't necessary because MSS delivers information in each fragment that lets the client access subsequent fragments.

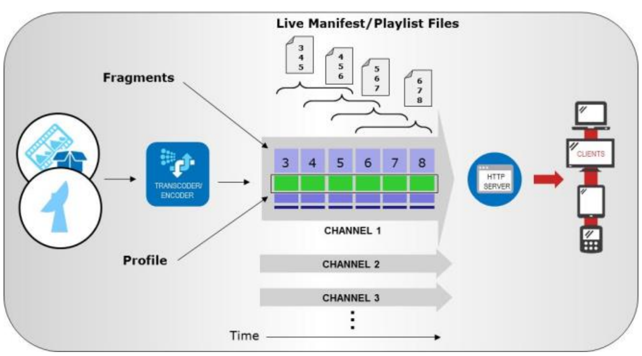


Figure 1. The content delivery chain for a live adaptive HTTP stream. The client downloads the 'rolling window' manifest files which references the latest available chunks. The client uses these references to download chunks for sequential play back. In the figure, the first manifest refers to chunks 3, 4 and 5, which are available in multiple bitrates. The playlist is updated as new chunks become available to reference the latest available chunks.

The advantages of adaptive HTTP streaming include:

- Being able to utilize generic HTTP caches, proxies and content delivery networks, as are used for web traffic;
- Content delivery is dynamically adapted to the weakest link in the end-to-end delivery chain, including highly varying last mile conditions;
- Playback control functions can be initiated using the lowest bitrate fragments and then transitioned to higher bitrates, enabling viewers to enjoy fast start-up and seek times;
- The client can control bitrate switching taking into account CPU load, available bandwidth, resolution, codec and other local conditions;
- Firewalls and NAT do not hinder HTTP delivery.

There are also a few issues with adaptive HTTP streaming:

- The end-to-end latency of live streams is increased as clients must buffer a few fragments to ensure that their input buffers aren't starved;
- HTTP is based on TCP and TCP recovers well when packet loss is low, meaning that video playback has no artifacts caused by missing data. However, TCP can completely fail when packet loss rises. Because of this clients usually experience good quality playback or the playback stops entirely – there is no middle ground. This is as opposed to quality that degrades proportionally to the amount of packet loss in the delivery network. Typically the Internet is sufficiently reliable that

the benefit of completely clean video at low packet drop rates outweighs the value of some poor quality video at high packet drop rates (e.g. with UDP-based streaming).

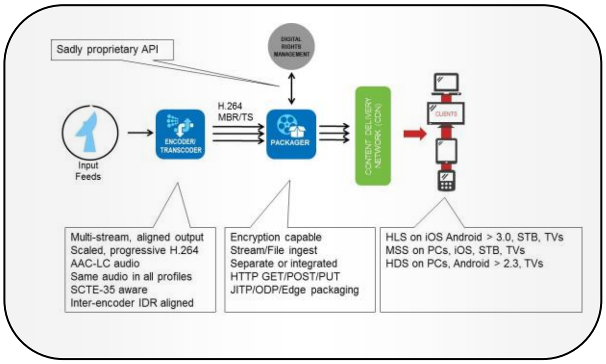


Figure 2. The components of an HTTP streaming system.

Components of Video Delivery

The key components in an adaptive HTTP streaming data flow consist of an encoder or transcoder, a packager (sometimes called a 'segmenter' or a 'fragmenter') and a CDN. In this section we will discuss the features of these components (see Figure 2) as related to adaptive streaming.

The Encoder/Transcoder

The ingestion and preparation of content for segmentation is the responsibility of the transcoder – or encoder, if the input is not already encoded in another format. The following features require support by the transcoder:

- The transcoder must de-interlace the input as the output video must be in progressive format.
- The video must be encoded into H.264 (or HEVC).
- The output video must be scaled to resolutions suitable for the client devices.
- The different output profiles must be IDR-aligned so that client playback of the chunks created from each profile is continuous and smooth.
- Audio must be transcoded into the AAC codec most often used by DASH, HLS, HDS and MSS.
- The same encoded audio stream generally must be streamed on all the output video profiles; to avoid clicking artifacts during client-side profile changes.
- If SCTE-35 is used for ad insertion, it is recommended that the transcoder add IDR frames at the ad insertion points to prepare the video for ad insertion. Fragment boundaries can then be aligned with the ad insertion points so that ad insertion is accomplished by the substitution of fragments, as compared to traditional stream splicing.
- An excellent fault tolerance mechanism allows two transcoders ingesting the same input to create identically IDR-aligned output. This allows the creation of a redundant backup of encoded content so that the secondary transcoder can seamlessly backup the primary transcoder should it fail for any reason.

Because consumers' QoE requires multiple different profiles for the client to choose from, it is best that the encoder be able to output a large number of different profiles for each input. Deployments may use anywhere from 4 to 16 different output profiles for each input. Naturally, the more profiles there are means the operator can support more devices and deliver a better user experience. The table below shows a typical use case for the different output profiles:

Width	Height	Video Bitrate
1920	1080	5 Mbps
1280	720	3 Mbps
960	540	1.5 Mbps
864	486	1.25 Mbps
640	360	1.0 Mbps
640	360	750 kbps
416	240	500 kbps
320	180	350 kbps
320	180	150 kbps

The Packager

As its name implies, the packager takes the transcoder's output and packages the video for a specific delivery protocol. Ideally, a packager will have the following features and capabilities:

- Encryption – the packager should be able to encrypt the outgoing chunks in a format compatible with the delivery protocol.
- Integration with third party key management systems – the packager should be able to receive encryption keys from a third party key management server that is also used to manage and distribute the keys to the clients.
- Live and on-demand ingest – the packager should be capable of ingesting both live streams and files, depending on whether the workflow is live or on-demand.
- Multiple delivery methods – the packager should support multiple ways to deliver the chunks, either via HTTP pull or by pushing the chunks using a network share or HTTP PUT/POST.

The CDN

CDNs do not need to be specialized for HTTP streaming nor are any special streaming servers required. For live delivery, it is ideal to set the CDN to rapidly age out older chunks as there is no need to keep them around long. A minute is usually enough, but the actual duration is dependent upon on the duration of the chunks and latency in the client.

Note that the number of chunks can be very large. For example, a day's worth of 2-second chunks delivered in 10 different profiles for 100 different channels creates 43 million files! Clearly the CDN must also be capable of handling a large number of files.

The Client

Obviously HLS is available on all iOS devices, but its availability on Windows devices is only thanks to third party products that are not always complete or sufficiently robust. Android include HLS natively, though not all features are well supported, for example stream discontinuity indications. MSS on a PC requires installation of a Silverlight runtime client. However, there are native Smooth Streaming clients for multiple devices, including iOS devices. HDS is native to Flash 10.1 and later releases and is also supported on some smart TVs and STBs. Adobe and Microsoft have announced that they will support DASH.

Despite being supported by a number of clients on various platforms, DASH is not yet experiencing the widespread adoption that the other protocols are. Though, it should be noted that some deployments of HbbTV utilize DASH.

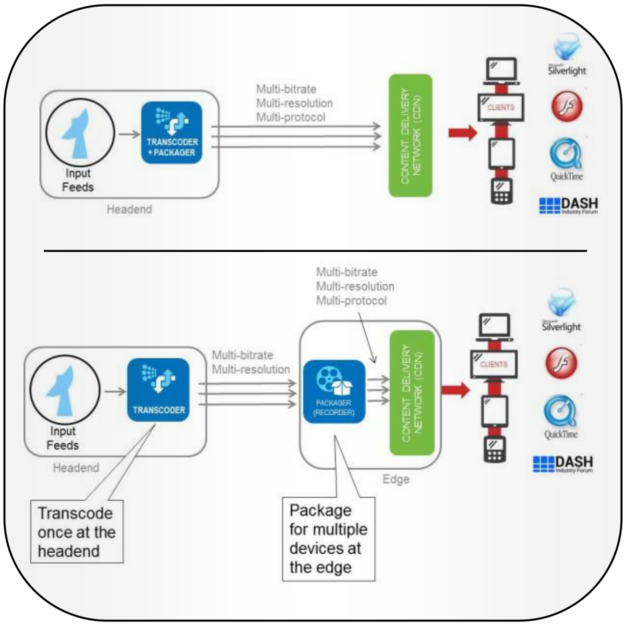


Figure 3. Integrated and remote segmentation of streams: When multiple formats are used, segmenting closer to the edge of the network (top) saves core bandwidth, as streams can be delivered once and packaged at the edge into multiple delivery formats. However, if the core network is susceptible to packet loss, segmenting at the core ensures that segments will always be delivered to the CDN (bottom).

Workflow Architecture

A flexible architecture in which the transcoder and packager can be separate is ideal. The key benefit of this approach is that the input video only needs to be transcoded once at the core, then delivered to the network edge where it's packaged into multiple formats. Without this separation, all the final delivery formats must be delivered over the core network, unnecessarily increasing its bandwidth utilization. This is shown in Figure 3.

Additional Adaptive HTTP Streaming Technologies

In addition to the "Big 4" there are several other adaptive streaming technologies, most notably:

- EchoStar acquired Move Networks and has integrated Move's technology into their home devices. Move played a major role in popularizing adaptive HTTP streaming and has multiple patents on the technology (though chunked streaming was used before Move popularized it).
- 3GPP's Adaptation HTTP Streaming (AHS) is part of 3GPPs rel 9 specification (see [AHS]). And 3GPP rel 10 is working on a specification called DASH as well.
- The Open TV Forum has developed its own HTTP Adaptive Streaming (HAS) specification (see [HAS]).
- MPEG Dynamic Adaptive Streaming over HTTP (DASH) is based on 3GPP's AHS and the Open TV Forum's HAS and is completed. It specifies use of either fMP4 or transport stream (TS) chunks and an XML manifest (called the media presentation description or MPD) that can behave similarly to both MSS or HLS. As a kind of combination of MSS and HLS, DASH allows for multiple scenarios, including separate or joined streaming of audio, video and data, as well as encryption. However, it makes clients complex to implement due to its generality, which is as much a drawback as an advantage. To address this issue, the DASH Industry Forum (see [DASHIF]) created the DASH-264 (and DASH-265)

specifications which make specific profile suggestions to create a more readily implementable and interoperable specification, focused on the Base Media File Format (BMFF) using H.264 (or H.265) and AAC audio. Questions about Intellectual Property rights may be slowing its adoption, but DASH-264/265 has the potential to become the format of choice in the future.

- Though some DRM vendors still have their own variation of these schemes, none appear to have any significant traction.

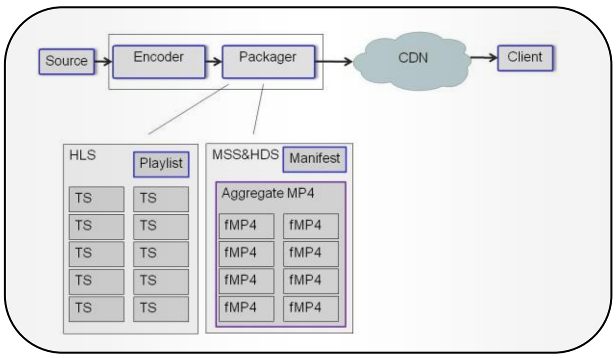


Figure 4. A comparison of HLS and MSS/HDS: The latter can create aggregate formats that can be distributed on a CDN for VOD, whereas for live video, all distribute chunks on the CDN. MSS allows audio and video to be aggregated and delivered separately, but HDS and HLS deliver these together.

THE INTERNALS OF ADAPTIVE STREAMING

Let's now review some of the details of HLS, HDS, MSS and DASH. Each protocol has its own unique strengths and weaknesses, which will be reviewed in the following sections.

Apple HTTP Live Streaming (HLS)

Interestingly, Apple chose not to use the ISO MPEG file format (which is based on its own MOV file format) in its adaptive streaming technology, unlike Adobe and Microsoft. Instead, HLS takes an MPEG-2 transport stream and segments it to a sequence of MPEG-2 TS files which encapsulate the audio and video. These segments are placed on any HTTP server along with the playlist files. The playlist (or index) manifest file is a text file (based on Winamp's original m3u file format) with an m3u8 extension. Full details can be found in [HLS].

HLS defines two types of playlist files: normal and variant. The normal playlist file lists URLs that point to chunks that should be played sequentially. The variant playlist files points to a collection of different normal playlist files, one for each output profile.

Metadata is carried in the playlist files as comments – lines preceded by '#'. In the case of normal playlist files, this metadata includes a sequence number that associate chunks from different profiles, chunk duration information, a directive signaling whether chunks can be cached, the location of decryption keys, the type of stream and time information. In the case of a variant playlist the metadata includes the bitrate of the profile, its resolution, its codec and an ID that can associate different encodings of the same content.

Figure 5 and Figure 6 show a sample HLS variant playlist file and normal playlist file. For an HLS client to know the URLs of the most recently available chunks, it's necessary for a playlist file corresponding to a live stream to be repeatedly downloaded. The playlist is downloaded every time a chunk is played, and thus, in order to minimize the number of these requests, Apple recommends a duration of 10 seconds, which is relatively long. However, the size of the playlist file is small compared with any video content, and the client maintains an open TCP connection to the server, so that this network load is not significant. Shorter chunk durations can thus be used, so the client can more quickly adapt to bitrates. VOD playlists are distinguished from live playlists by the `#EXT-X-PLAYLIST-TYPE` and `#EXT-X-ENDLIST` tags.

```
#EXTM3U
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=531475
mystic_S1/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=381481
mystic_S2/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=531461
mystic_S3/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=781452
mystic_S4/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=1031452
mystic_S5/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=1281452
mystic_S6/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=1531464
mystic_S7/mnf.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=3031464
mystic_S8/mnf.m3u8
```

Figure 5. An HLS variant playlist file showing eight output profiles with different bitrates. The URLs for the m3u8 files are relative, but could include the leading 'http://…'. In this example, each profile's playlist is in a separate path component of the URL**.**

Only the HLS protocol does not require chunks to start with IDR frames. It can download chunks from two profiles and switch the decoder between profiles on an IDR frame that occurs in the middle of a chunk. The downside to this is the requirement for extra bandwidth as two chunks corresponding to the same portion of video are downloaded simultaneously.

<u>HLS Considerations</u>

The advantages of HLS include:

- It is a simple protocol and is easily modified. The playlists can be easily accessed and their text format makes modification for applications such a re-broadcast or ad insertion simple.
- The use of TS files means that there is a rich ecosystem for testing and verifying file conformance.
- TS files can carry SCTE 35 cues, ID3 tags (see [HLSID3]) or other such metadata.
- Monetizing HLS is more easily accomplished as it's native to popular iOS devices, the users of which are accustomed to paying for apps and other services.

```
#EXTM3U
#EXT-X-KEY:METHOD=NONE
#EXT-X-TARGETDURATION:10
#EXT-X-MEDIA-SEQUENCE:494

#EXT-X-KEY:METHOD=NONE
#EXTINF:10,505.ts
505.ts
#EXTINF:10,506.ts
506.ts
#EXTINF:10,507.ts
507.ts
```

Figure 6. An HLS playlist file from a live stream showing the three latest available TS chunks. The #EXT-X-MEDIA-SEQUENCE:494 is used by the client to keep track of where it is in the linear playback. The fragment name carries no streaming-specific information. The #EXT-X-TARGETDURATION:10 tag is the expected duration (10 seconds) of the chunks, though durations can vary. The #EXT-X-KEY:METHOD=NONE tag shows that no encryption was used in this sequence. The #EXTINF:10 tags show the duration of each segment. As in the variant playlist file, the URLs are relative to the base URL used to fetch the playlist.

The disadvantages of HLS include:

- HLS is not supported natively on Windows OS platforms.
- Apple's aggregate format stores all the fragments in one TS file and uses byte-range URL requests to pull out the fragment data. Unfortunately CDNs are sometimes unable to cache based on such requests, which limits the usefulness of this aggregation format. Without an aggregation format, HLS must store each fragment as a separate file, so that many files must be created. For example, a day's worth of programming for a single channel requires almost 70,000 files,

assuming eight profiles with 10-second chunk duration. Clearly the managing of such a large collection of files is not convenient.

- HLS is an ecosystem in which different iOS clients have different capabilities, and this limits the value of the later improvements to HLS. This because HLS has evolved from requiring fragments that mux audio and video to allowing separate fragments (as well as other developing features).

Microsoft's Silverlight Smooth Streaming (MSS)

Silverlight Smooth Streaming delivers streams as a sequence of ISO MPEG-4 files (see [MSS] and [MP4]). Usually these are pushed by an encoder to a Microsoft IIS server (using HTTP POST), which aggregates them for each profile into an 'ismv' file for video and an 'isma' file for audio. The IIS server also creates an XML manifest file that contains information about the bitrates and resolutions of the available profiles (see Figure 7). When the request for the manifest comes from a Microsoft IIS server, it has a specific format:

```
http://{serverName}/{PublishingPointPath}/{Pub
        lishingPointName}.isml/manifest
```

The `PublishingPointPath` and `PublishingPointName` are derived from the IIS configuration.

In MSS, the manifest files contain information that allows the client to create a RESTful URL request based on timing information in the stream, which differs from HLS in which URLs are given explicitly in the playlist. For live streaming, the client computes the URLs for the chunks in each profile directly, rather than repeatedly downloading a manifest. The segments are extracted from the ismv and isma files and served as 'fragmented' ISO MPEG-4 (fMP4) files. MSS (optionally) separates the audio and video into separate chunks and combines them in the player.

```
<SmoothStreamingMedia MajorVersion="2"
MinorVersion="0" TimeScale="10000000" Duration="0"
LookAheadFragmentCount="2" IsLive="TRUE"
DVRWindowLength="300000000">
    <StreamIndex Type="video" QualityLevels="3"
        TimeScale="10000000" Name="video" Chunks="14"
        Url="QualityLevels({bitrate})/Fragments(vide
        o={start time})" MaxWidth="1280"
        MaxHeight="720" DisplayWidth="1280"
        DisplayHeight="720">
        <QualityLevel Index="0" Bitrate="350000"
        CodecPrivateData="00000001274D401F9A6282833F
        3E022000007D20001D4C12800000000128EE3880"
        MaxWidth="320" MaxHeight="180" FourCC="H264"
        NALUnitLengthField="4"/>
        <QualityLevel Index="1" Bitrate="500000"
        CodecPrivateData="00000001274D401F9A628343F6
        022000007D20001D4C12800000000128EE3880"
        MaxWidth="416" MaxHeight="240" FourCC="H264"
        NALUnitLengthField="4"/>
        <QualityLevel Index="2" Bitrate="750000"
        CodecPrivateData="00000001274D401F9A6281405F
        F2E022000007D20001D4C12800000000128EE3880"
        MaxWidth="640" MaxHeight="360" FourCC="H264"
        NALUnitLengthField="4"/>
        <c t="2489409751000"/>
        <c t="2489431105667"/>
        <c t="2489452460333"/>
        <c t="2489473815000"/>
        <c t="2489495169667"/>
        <c t="2489516524333"/>
        <c t="2489537879000"/>
        <c t="2489559233667"/>
        <c t="2489580588333" d="21354667"/>
    </StreamIndex>
    <StreamIndex Type="audio" QualityLevels="2"
        TimeScale="10000000" Language="eng"
        Name="audio_eng" Chunks="14"
        Url="QualityLevels({bitrate})/Fragments(audi
        o_eng={start time})">
        <QualityLevel Index="0" Bitrate="31466"
        CodecPrivateData="1190" SamplingRate="48000"
        Channels="2" BitsPerSample="16"
        PacketSize="4" AudioTag="255" FourCC="AACL"/>
        <QualityLevel Index="1" Bitrate="31469"
        CodecPrivateData="1190" SamplingRate="48000"
        Channels="2" BitsPerSample="16"
        PacketSize="4" AudioTag="255" FourCC="AACL"/>
        <c t="2489408295778"/>
        <c t="2489429629111"/>
        <c t="2489450962444"/>
        <c t="2489472295778"/>
        <c t="2489493629111"/>
        <c t="2489514962444"/>
        <c t="2489536295778"/>
        <c t="2489557629111"/>
        <c t="2489578962444" d="21333334"/>
        </StreamIndex>
</SmoothStreamingMedia>
```

Figure 7. A sample MSS manifest file. The elements with 't="248…"' specify the time stamps of chunks that the server has and is ready to deliver. These are converted to Fragment timestamps in the URL requesting an fMP4 chunk. The returned chunk holds time stamps of the next chunk or two (in its UUID box), so that the client can continue fetching chunks without requesting a new manifest.

The URLs below show typical requests

for video and audio. The `QualityLevel` indicates the profile and the `video=` and `audio-eng=` indicate the specific chunk requested. The `Fragments` portion of the request is given using a time stamp (in hundred nanosecond units) that the IIS server uses to extract the correct chunk from the aggregate MP4 audio and/or video files.

```
    http://sourcehost/local/2/mysticSmooth.isml/QualityLevels(350000)/Fragments(video=2489452460333)
    http://sourcehost/local/2/mysticSmooth.isml/QualityLevels(31466)/Fragments(audio-eng=2489450962444)
```

In the VOD case, the manifest files contain timing and sequence information for all the chunks in the content. The player uses this information to create the URL requests for the audio and video chunks.

It is important to recognize that the use of IIS as the source of the manifest and fMP4 files doesn't prohibit using standard HTTP servers in the CDN. The CDN can still cache and deliver the manifest and chunks as it would any other files. More information about MSS can be found at Microsoft (see [SSTO]) and various excellent blogs of the developers of the technology (see [SSBLOG]).

<u>MSS Considerations</u>

The advantages of MSS include:

- IIS creates an aggregate format for the stream, so that a small number of files can hold all the information for the complete smooth stream.
- IIS has useful analysis and logging tools, as well as the ability to deliver more MSS and HLS content directly from the IIS server.
- Rapid adaptation during HTTP stream playback is achieved when the recommended use of a small chunk size is followed. The delivery of different audio tracks requires only a manifest file change due to video and audio files being segregated.
- The aggregate file format supports multiple data tracks that can be used to store metadata about ad insertion, subtitling, etc.

The disadvantages of MSS include:

- The network is data flow is slightly more complex and an extra point of failure is added due to need to place an IIS server in the flow.
- On PCs, MSS requires installation of a separate Silverlight plug-in.

```
<manifest>
    <id>USP</id>
    <startTime>2006-07-24T07:15:00+01:00</startTime>
    <duration>0</duration>
    <mimeType>video/mp4</mimeType>
    <streamType>live</streamType>
    <deliveryType>streaming</deliveryType>
    <bootstrapInfo profile="named" url="mysticHDS.bootstrap"/>
    <media url="mysticHDS-audio_eng=31492-video=3000000-" bitrate="3031" width="1280" height="720"/>
    <media url="mysticHDS-audio_eng=31492-video=1500000-" bitrate="1531" width="960" height="540"/>
    <media url="mysticHDS-audio_eng=31492-video=1250000-" bitrate="1281" width="864" height="486"/>
    <media url="mysticHDS-audio_eng=31492-video=1000000-" bitrate="1031" width="640" height="360"/>
    <media url="mysticHDS-audio_eng=31492-video=750000-" bitrate="781" width="640" height="360"/>
    <media url="mysticHDS-audio_eng=31492-video=500000-" bitrate="531" width="416" height="240"/>
    <media url="mysticHDS-audio_eng=31469-video=350000-" bitrate="381" width="320" height="180"/>
</manifest>
```

Figure 8. A sample HDS manifest file.

<u>Adobe HTTP Dynamic Streaming (HDS)</u>

Adobe HDS uses elements of both HLS and MSS as it was defined after them (see [HDS]). In HDS, an XML manifest file (of file type f4m) contains information about the available profiles (see Figure 8 and [F4M]). Like HLS, the HDS client repeatedly downloads data that allows it to know the URLs of the available chunks -- in HDS this is called the bootstrap information. This bootstrap information isn't human-readable as it is in a binary format. As in MSS, segments are encoded as fragmented MP4 files containing audio and video information in a single file. HDS chunk requests have the form:

```
http://server_and_path/QualityModifierSeg'segm
   ent_number'-Frag'fragment_number'
```

where the segment and fragment number together define a specific chunk. As in MSS, an aggregate (f4f) file format is used to store all the chunks and extract them when requested.

HDS Considerations

The advantages of HDS include:

- The Flash client is available on multiple devices and is installed on almost every PC worldwide.
- HDS is a part of Flash and can make use of Flash's environment and readily available developer base.

The disadvantages of HDS are:

- HDS is relatively new and could suffer more from stability issues than its more mature counterparts.
- Deploying a stable ecosystem is difficult due to Adobe's Flash rapidly changing access roadmap.
- The ecosystem of partners offering compatible solutions is limited due to the binary format of the bootstrap file.

Dynamic Adaptive HTTP Streaming (DASH)

DASH is the most feature complete and complex of all the protocols, as it incorporates many features similar to those in HLS and MSS. It can make use of both TS fragments and fMP4 fragments, and it supports both repeated manifest downloads or URLs derivable from a template. Like MSS and HDS, the DASH manifest, or MPD, uses an XML format. Figure 9 shows a sample DASH MPD.

```
<?xml version="1.0" encoding="utf-8"?>
<MPD
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:DASH:schema:MPD:2011"
xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:2011"
type="static"
mediaPresentationDuration="PT12M34.041388S"
minBufferTime="PT10S"
profiles="urn:mpeg:dash:profile:isoff-live:2011">
<Period>
<AdaptationSet mimeType="audio/mp4"
 segmentAlignment="0" lang="eng">
 <SegmentTemplate timescale="10000000"
     media="audio_eng=$Bandwidth$-$Time$.dash"
     initialisation=" audio_eng=$Bandwidth$.dash">
 <SegmentTimeline>
   <S t="667333" d="39473889" />
   <S t="40141222" d="40170555" />
   ...
   <S t="7527647777" d="12766111" />
  </SegmentTimeline>
 </SegmentTemplate>
<Representation id="audio_eng=96000" bandwidth="96000"
     codecs="mp4a.40.2" audioSamplingRate="44100" />
</AdaptationSet>
  <AdaptationSet mimeType="video/mp4"
     segmentAlignment="true" startWithSAP="1"
     lang="eng">
   <SegmentTemplate timescale="10000000"
      media="video=$Bandwidth$-$Time$.dash"
      initialisation="video=$Bandwidth$.dash">
   <SegmentTimeline>
      <S t="0" d="40040000" r="187" />
      <S t="7527520000" d="11678333" />
   </SegmentTimeline>
  </SegmentTemplate>
<Representation id="video=299000" bandwidth="299000"
     codecs="avc1.42C00D" width="320" height="180" />
<Representation id="video=480000" bandwidth="480000"
     codecs="avc1.4D401F" width="512" height="288" />
<Representation id="video=4300000" bandwidth="4300000"
     codecs="avc1.640028" width="1280" height="720" />
</AdaptationSet>
</Period>
</MPD>
```

Figure 9. A sample DASH MPD.

DASH Considerations

The advantages of DASH include:

- It is based on an open standard.
- The DASH-264 specification has strong industry support which bodes well for solid interoperability.

The disadvantages of DASH include:

- It's all-encompassing, so different DASH implementations will most likely be incompatible unless they focus on exactly the same specification profiles.
- The Intellectual Property (IP) rights associated with DASH are not completely clear. Normally, the holders of essential patents have (for the most part) participated in the MPEG process for other standards – leading to so-called reasonable and non-discriminatory (RAND) licensing terms for the IP associated with the specification. In the case of DASH, the picture is more murky.

COMPARING FEATURES

Let's now compare the usability of DASH, HLS, HDS and MSS usability in several common usage scenarios.

Delivering Multiple Audio Channels

HLS is now equal to MSS in its ability to deliver audio and video separately and easily, which is thanks to the release of iOS5 which added the ability to deliver audio separately. Delivering audio separately is, of course, important in locales where multiple languages are used and only one is to be consumed. HDS is still primarily used with audio and video muxed together in the fragments. DASH can flexibly separate video and audio.

Encryption and DRM

HLS supports encryption of each TS file, meaning that all of the data contained in the TS file is encrypted and cannot be extracted without the decryption keys. All metadata related to the stream (e.g. the location of the decryption keys) must be included in the playlist file. While functioning well, HLS does not specify a mechanism for authenticating clients to receive the decryption keys. This is considered a deployment issue. Several vendors offer HLS-type encryption, generally with their own twist which makes the final deployment incompatible with other implementations.

Microsoft's PlayReady is used by MSS to provide a complete framework for encrypting content, managing keys and delivering them to clients. Because PlayReady only encrypts the payload of the fragment file, the chunk can carry other metadata. Microsoft makes PlayReady code available to multiple vendors that productize it, and so a number of vendors offer PlayReady capability (in a relatively undifferentiated way).

HDS uses Adobe's Access, which has an interesting twist that simplifies interaction between the key management server and the scrambler that does the encryption. Typically, keys must be exchanged between these two components, and this exchange interface is not standardized. Each pair of DRM and scrambler vendors must implement this pair-wise proprietary API. With Adobe Access however, key exchanges are not necessary as the decryption keys are sent along with the content and are encrypted themselves. Access to those keys is granted at run time, but no interaction between the key management system and scrambler is needed.

DASH allows DRM systems to share keys, encryption algorithm and other parameters via its support for the so-called common encryption format. This enables the same content to be managed by different clients and DRM systems which implement key distribution and other rights management. This represents a major achievement in the breaking up the traditional DRM vendor model, which locks-in users by making content playable only by clients associated with a specific vendor's encryption.

Closed Captions / Subtitling

As of iOS 4.3, HLS can decode and display closed captions (using ATSC Digital Television Standard Part 4 – MPEG-2 Video System Characteristics - A/53, Part 4:2007, see [ATCA]) included in the TS chunks. For DVB teletext, packagers must convert the subtitle data into ATSC format or wait for clients to support teletext data. HLS also supports WebVTT, which holds subtitles in a separate file.

MSS supports data tracks that hold Time Text Markup Language (TTML), a way to specify a separate data stream with subtitle, timing and placement information (see [TTML]). For MSS, packagers need to extract subtitle information from their input and convert it into a TTML track. Microsoft's implementation of MSS client currently offers support for W3C TTML, but not for SMPTE TTML (see [SMPTE-TT]), which adds support for bitmapped subtitles, commonly used in Europe.

HDS supports data tracks that hold subtitles as DFXP file data or as TTML. In a manner similar to MSS, HDS clients can selectively download this data.

Though the DASH specification supports various forms of subtitling, client support is inconsistent.

Of the many things the formats have in common, it's worth noting that none support DVB image subtitles particularly well. HLS can support these using ID3 signaling and proprietary Javascript wrapped around the client. The other formats can also "support" this using proprietary clients as well.

Targeted Ad insertion

HLS is the simplest protocol for chunk-substitution-based ad insertion. With HLS, the playlist file can be modified to deliver different ad chunks to different clients (see Figure 10). The `EXT-X-DISCONTINUITY` tag can tell the decoder to reset (e.g. because subsequent chunks may have different PID values), and only the sequence ID must be managed carefully, so that the IDs line up when returning to the main program. HDS also supports the repeated downloading of bootstrap data used to specify chunks, and this can be modified to create references to ad chunks. However, because the bootstrap data format is binary, and the URLs are RESTful with references to chunk indexes, the process is complex.

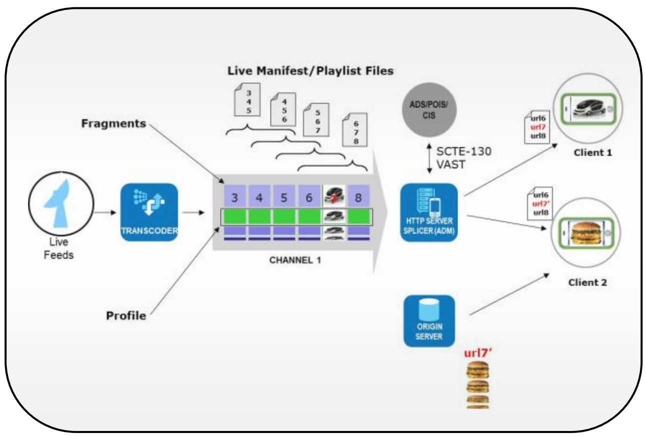


Figure 10. HLS ad insertion in which changes to the playlist file delivered to each client cause each client to make different URL requests for ads and thus receive targeted ad content.

Chunk-based ad insertion for live streams is more complicated with MSS. Because fragments contain timing information used to request the next fragment, all ad fragments must have identical timing to the main content chunks. Regardless, a proxy can redirect RESTful URL fragment requests and serve different fragments to different clients.

MSS and HDS can both deliver control events in separate data tracks. These can trigger client behaviors using the Silverlight and Flash environments, including ad insertion behavior. However, this is beyond the scope of this paper, which is focused on 'in stream' insertion.

A particularly nice feature in DASH allows it to switch from template-based URLs to repeated download of the manifest. The advantage is that the network program can be delivered using templated URLs which stop and switch to downloaded manifests for the targeted ads.

nDVR Mezzanine Format

HLS is clearly the winner when it comes to the stream-to-file capture of adaptive bitrate formats. It captures and recombines fragments much more easily than possible with MSS or HDS. DASH can do the same, in its TS file profile, but not so readily in DASH-264.

Trick Modes (Fast-forward / Rewind)

None of the protocols implement VOD trick modes, such as fast-forward or rewind, well at all. HLS, DASH and HDS offer support for fast-forward and rewind in the protocols, but actual client implementations are non-existent. MSS defines zoetrope images that can be embedded in a separate track. These can be used to show still images from the video sequence and allow viewers to seek a specific location in the video.

Custom VOD Playlists

Being able to take content from multiple different assets and stitch them together to form one asset is particularly convenient. This is readily done in HLS and DASH, where the playlist can contain URLs that reference chunks from different encodings and locations. Unfortunately for MSS and HDS, constructing such playlists is basically impossible because of the RESTful URL name spaces and the references to chunks via time stamp or sequence number.

Fast Channel Change

Adaptive HTTP streaming can download low bitrate fragments initially, making channel 'tune in' times fast. The duration of the fragment directly affects how fast the channel adapts to a higher bandwidth (and higher quality video). This is an advantage for DASH, MSS and HDS, which are tuned to work with smaller fragments, and which tend to work a bit better than HLS.

Failover Due to Upstream Issues

To counter situations in which content is not available, HLS manifests can list failover URLs. The mechanism used in the variant playlist file to specify different profiles can specify failover servers, since the client (starting with iOS 3.1 and later) will attempt to change to the next profile when a profile chunk request returns an HTTP 404 'file not found' code. This is a convenient, distributed redundancy mode.

MSS utilizes a fully programmable run-time client. Similarly, HDS works within the Flash run-time environment. This means that the same failover capabilities can be built into the client. Despite this, neither protocol has a built-in mechanism supporting a generic failover capability.

Each protocol will failover to a different profile if chunks/playlists in a given profile are not available. Potentially, this could enable any of the protocols to be used in a "striping" scenario in which alternate profiles come from different encoders (as long as the encoders output IDR aligned streams), so that an encoder failure causes the client to adapt to a different, available profile.

Stream Latency

Adaptive HTTP clients buffer a number of segments. Typically one segment is currently playing, one is cached and a third is being downloaded. This is done to ensure that the end-to-end latency is minimally about three segment durations long. With HLS recommended to run with 10-second chunks (though this isn't necessary), this latency can be quite long.

MSS is the sole protocol with a low latency mode in which sub-chunks are delivered to the client as soon as they are available. The client doesn't need to wait for a whole chunk's worth of stream to be available at the packager before requesting it, reducing its overall end-to-end latency.

Sending Other Data to the Client (Including Manifest Compression)

HLS, DASH and HDS use playlist and manifest files to send metadata to their clients. MSS and HDS allow data tracks which can trigger client events and contain almost any kind of data. HLS allows a separate ID3 data track to be muxed into the TS chunks. This can be used to trigger client-side events.

MSS and HLS also allow manifest files to be compressed using gzip (as well as internal run-length-type compression constructs) for faster delivery.

CONCLUSION

A snapshot of how these four formats compare is shown in the table below.

Feature	HLS	MSS	HDS	DASH
Multiple audio channels	☺	☺	☺	☺
Encryption		☺	☺	☺
Closed captions / subtitling	😐	😐	😐	☺
Custom VoD playlists	☺			☺
Ability to insert ads network-side	☺			☺
Value as a stream-to-file format	☺			😐
Trick modes (fast forward / rewind)	😐	😐	😐	😐
Fast channel change		☺	☺	☺
Client failover	☺			☺
Stream latency		☺	☺	
Metadata	☺	☺	☺	☺

Though DASH appears strong, in reality its dearth of widely adopted and feature-rich clients makes it look better on paper than reality – notwithstanding the use of DASH by Netflix. Quite understandably, HLS benefits greatly from the ubiquity of iPads and iPhones, and it is the protocol most commonly used by video service providers. Smooth Streaming benefits from the success of the Xbox and the strong brand awareness around Microsoft's PlayReady DRM, which content owners are comfortable with. Regionally, it has slightly stronger support in Europe, versus the Americas. HDS benefits from the ubiquity of Flash on PCs and laptops, though this does make it the most susceptible to declining usage trends. However, HDS benefits from Adobe Access's respected DRM, and Adobe's commitment to focus on DRM and clients across multiple platforms. The conclusion is that at this time, no single format is poised to strongly dominate (and no format shows signs of near-term death). An operator's final decision must be based on the client devices served, DRM requirements inherited from content owners, and lastly the underlying services delivered to customers. For services that include ad insertion, nDVR and linear TV to iPads, for example, the choice is easy. For other combinations, the decisions are more subtle.

REFERENCES

[HLS] HTTP Live Streaming, R. Pantos, http://tools.ietf.org/html/draft-pantos-http-live-streaming-06

[HLS1] HTTP Live Streaming, http://developer.apple.com/library/ios/#documentation/NetworkingInternet/Conceptual/HTTPLiveStreaming/_index.html

[MP4] International Organization for Standardization (2003). "MPEG-4 Part 14: MP4 file format; ISO/IEC 14496-14:2003"

[SSBLOG] http://blog.johndeutscher.com/, http://blogs.iis.net/samzhang/ , http://alexzambelli.com/blog/, http://blogs.iis.net/jboch/

[SSTO] Smooth Streaming Technical Overview, http://learn.iis.net/page.aspx/626/smooth-streaming-technical-overview/

[ATCA] ATSC Digital Television Standard Part 4 – MPEG-2 Video System Characteristics (A/53, Part 4:2007), http://www.atsc.org/cms/standards/a53/a_53-Part-4-2009.pdf

[TTML] Timed Text Markup Language, W3C Recommendation 18 November 2010, http://www.w3.org/TR/ttaf1-dfxp/

[SMPTE-TT] SMPTE Time Text format, https://www.smpte.org/sites/default/files/st2052-1-2010.pdf

[MSS] IIS Smooth Streaming Transport Protocol, http://www.iis.net/community/files/media/smoothspecs/%5BMS-SMTH%5D.pdf

[F4M] Flash Media Manifest File Format Specification, http://osmf.org/dev/osmf/specpdfs/FlashMediaManifestFileFormatSpecification.pdf

[HDS] HTTP Dynamic Streaming on the Adobe Flash Platform, http://www.adobe.com/products/httpdynamicstreaming/pdfs/httpdynamicstreaming_wp_ue.pdf

[AHS] 3GPP TS 26.234: "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs".

[HAS] OIPF Release 2 Specification - HTTP Adaptive Streaming http://www.openiptvforum.org/docs/Release2/OIPF-T1-R2-Specification-Volume-2a-HTTP-Adaptive-Streaming-V2_0-2010-09-07.pdf

[HLSID3] Timed Metadata for HTTP Live Streaming, http://developer.apple.com/library/ios/#documentation/AudioVideo/Conceptual/HTTP_Live_Streaming_Metadata_Spec/Introduction/Introduction.html

[DVB-BITMAPS] Digital Video Broadcasting (DVB); Subtitling systems, ETSI EN 300 743 V1.3.1 (2006-11), http://www.etsi.org/deliver/etsi_en/300700_300799/300743/01.03.01_60/en_300743v010301p.pdf

[DASHIF] DASH Industry Forum, http://dashif.org/

AN SDN BASED APPROACH TO MEASURING AND OPTIMIZING ABR VIDEO QUALITY OF EXPERIENCE

Sangeeta Ramakrishnan, Xiaoqing Zhu
Cisco Systems

Abstract

The volume of video traffic is continuing to grow rapidly over cable networks. While a majority of IP video on DOCSIS networks today is over-the-top video from third parties, operators are increasing the amount of programming their subscribers can access on IP video-capable devices. Additionally many operators are pursuing IP video deployments that are at various stages from planning, testing, trialing to deploying. Most OTT deployments use Adaptive Bit Rate (ABR) technology (also known as HTTP Adaptive Streaming). And ABR has also been adopted by cable operators as the technology of choice as they begin their migration to IP video.

ABR video was primarily developed to work as well as possible despite the network. This approach made sense for over-the-top content providers who have no control or influence of the network. Such a server-client method has so far worked well for OTT providers. Some of the challenges with such an ABR delivery method are masked because of end-users' acceptance of a poorer user-experience since OTT providers are perceived to be lower cost options. Also the amount of ABR video on the cable network today, while significant, will still be dwarfed by the amount of ABR video that will come on to the network when cable operators migrate to a ABR-based IP video delivery method. Hence the combination of more ABR traffic along with higher user-experience expectations of a cable subscriber, may pose challenges to cable operators as they deploy ABR-based IP video.

Another concern with ABR video is that different segments of the network can have similar network utilization levels yet there could be a large difference in end user Quality-of-Experience (QoE) between these segments of the network. Currently operators do not have good visibility to the actual QoE of subscribers; instead they primarily monitor network utilization levels to identify those segments that need to be upgraded. Operators need better tools to identify where QoE is below their service objectives so that they can target their network investments accordingly.

In this paper, we present a novel SDN-based solution to solve some of the challenges that we anticipate will occur with heavy ABR usage. Our proposal will help operators to improve their visibility of QoE and optimize their network investments. We also present a technique to help operators improve aggregate QoE of all users on their network. Alternatively this method can be used to pack significantly more streams in the given bandwidth with comparable quality to conventional ABR. Our studies indicate that with this approach bandwidth requirements can be reduced by a third or more, thereby saving DOCSIS channels and HFC spectrum, and reducing the cost of the overall solution.

INTRODUCTION

Over the past few years, video streaming traffic has been growing at a rapid pace, and is anticipated to dominate next generation networks. According to [1] global consumption of Internet video viewed through a TV doubled in 2012; video-on-demand traffic is projected to nearly triple by 2017. Cable operators, are seeking novel solutions to fend off the impending bandwidth crunch introduced by video traffic on their existing network infrastructure.

Additionally, the cable industry is on the cusp of a migration to IP video. Various factors are contributing to the desire to migrate. One of the biggest drivers is the desire to deliver any content to subscribers, anywhere and anytime they want it. This includes the ability to deliver content to consumer owned devices such as tablets, smartphones, PCs, laptops, game consoles and Internet enabled TVs. Delivering video to such consumer owned devices not only helps cable operators to meet subscribers' needs and expectations but has the added benefit that operators do not have to incur capital expenditure to deploy and maintain these devices. STBs deployed in subscribers' homes are a significant portion of the capex budget in offering video services. Reducing that expenditure by either leveraging consumer owned devices or lower cost IP STBs could help operators to improve their bottom-line significantly. Additionally, operators may be able to charge fees for making services available on new outlets such as tablets. So overall IP video provides an opportunity to both improve top-line and bottom-line for operators.

Additionally with IP video, operators may be able to leverage the rapid innovation in the Internet space to offer newer services, and a better user experience. It also enables operators to consolidate the video services infrastructure.

One of the major challenges faced by operators as they make this migration to IP video is the bandwidth required to deliver this service and the capex requirements associated with it. As shown in [2] the bandwidth needs for IP video are highly dependent on the type of service offered and can range typically from 20-40 downstream channels to serve a typical Service Group of 250-300 subscribers. The increasing focus on Cloud DVR like technology will in fact lead to higher capacity requirements due to the unicast nature of the delivery method. Hence there is likely to be significant increasing amounts of Adaptive Bit Rate video on the cable network in the future and operators will need better ways to manage it.

SOFTWARE DEFINED NETWORKING

Software Defined Networking (SDN) is a term that began being used widely in the industry around 2009. SDN is defined as the separation of Control and Data planes using an open standard protocol to communicate between them. This differs from a traditional network device (such as a Router, Switch, or CMTS) in which the data and control planes are vertically integrated. SDN promises flexibility and rapid innovation by virtue of the fact that the control software would be removed from the relatively constrained network device to a generic server that can be easily scaled to have more processing and memory capabilities.

SDN is typically implemented with a Controller based architecture as shown in

Figure 1, where the Controller interfaces with network devices via standard interfaces (for example Openflow [7], PCMM). Therefore, the controller presents a level of abstraction of the underlying network. Applications communicate with the controller using "controller APIs" and the controller in turn interacts with the network. In other words the controller acts as middleware that provides a higher lever of abstraction to network application developers.

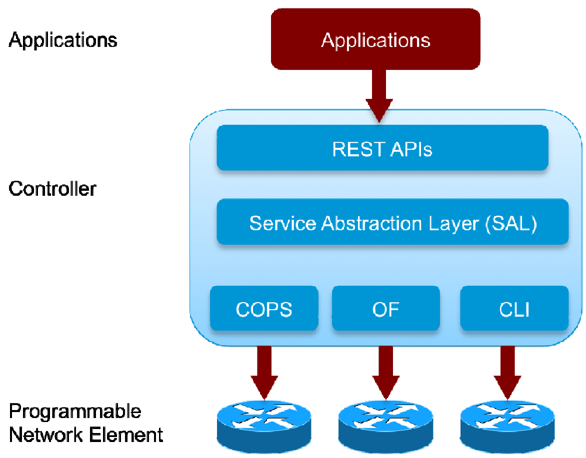


Figure 1 SDN Architecture

Another key attribute of SDN is the two-way communication with the network devices. The network can be thought of as a large distributed database of flows and states. In an SDN architecture, applications can learn from the Controller the state of various devices and can also program the devices via the Controller. An SDN-based architecture enables Applications to be developed in a relatively device-agnostic way.

ADAPTIVE BITRATE STREAMING

Recent years have seen a major technology shift as Internet video delivery solutions are converging towards the adaptive bit rate (ABR) streaming paradigm. Since its inception in 2007 by Move Networks [3], ABR has been quickly adopted by major vendors and service providers, including YouTube, Netflix, Hulu, Akamai, Microsoft Smooth Streaming [4], Apple HTTP Live Streaming (HLS) [5], and Adobe HTTP Dynamic Streaming (HDS).

Figure 1 shows the architecture overview of an adaptive bitrate streaming (ABR) system. The media contents are either pre-stored or captured live at the source. Multiple quality versions of the same video content are generated via transcoding. Moreover, each media file is broken down into many small fragments. The origin HTTP server keeps track of these fragments either as a large collection of separate physical files (e.g., in Apple HLS), or as logical separations via indexing (e.g., in Microsoft Smooth Streaming). Additional content delivery

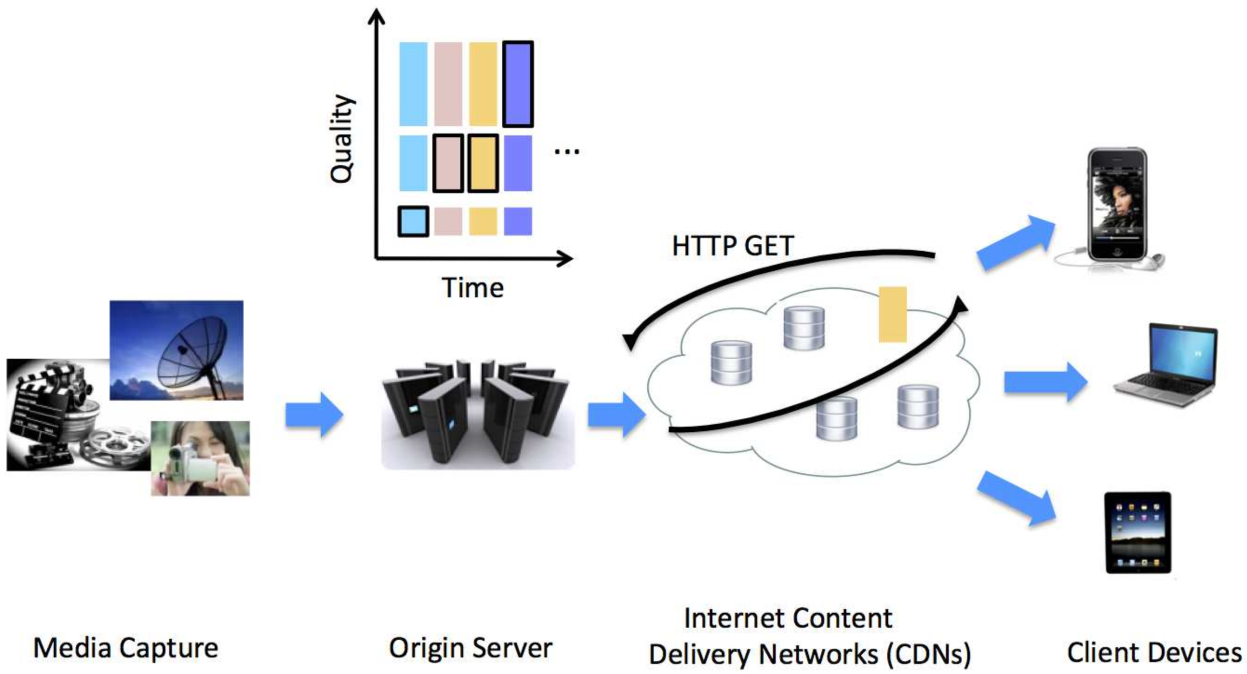


Figure 2 Architecture overview of an adaptive bitrate streaming (ABR) system. Multiple quality versions of the same video content are generated via transcoding or parallel live encoding. Each media file is broken down into many physical or logical fragments. The client can adaptively request different quality versions of the media fragments based on its own estimate of available network bandwidth

networks (CDNs) may also be leveraged at the edge of the network, so as to assist in disseminating video contents to a wide range of end users.

In ABR, the client can dynamically change its video rate and quality on a fragment by fragment basis. In face of temporary network congestion, the client can switch to a lower rate (and hence quality) version of the video to avoid buffer underflow; when connection speed recovers, the client can switch back to higher quality. Such flexibility in rate adaptation can be advantageous in the presence of dynamic network conditions, especially in mobile environments. In 2012, the Motion Picture Experts Group (MPEG) has joined forces with 3GPP (3rd Generation Partnership Project) in defining the recently standardized Dynamic Adaptive Streaming over HTTP (DASH) specifications [6]. The MPEG-DASH standard has intentionally left out of its scope the definition of client behavior for content fetching, rate adaptation heuristics, and video playout, thereby allowing plenty of space for innovation-based competition in industry.

QoE MEASUREMENTS

Today most operators measure the quality-of-experience (QoE) of their subscribers in terms of network utilization levels. Network utilization is typically averaged across over few minutes (typically 5~15 mins). Such average utilization is measured throughout the day. Most operators consider their networks to be congested if utilization exceeds a certain threshold such as 70-80% during peak hours.

If many such measured samples of utilization level exceed their pre-determined threshold, operators typically declare those interfaces to be congested and plan upgrades of their network to address the congestion. The advantage of such a QoE measurement is it is simplicity – easy to measure and easy to monitor.

Challenges with ABR QoE Measurements

As the industry moves toward ABR-based video delivery, this new paradigm poses some challenges for the aforementioned QoE measurement mechanism. This is because with ABR, clients up-shift to higher rate profiles when bandwidth is available. This is in contrast to standard Internet browsing, where higher bandwidth availability simply causes the file download to finish faster. For example, downloading a file when plenty of bandwidth is available does not cause file size to grow! TCP simply takes advantage of the bandwidth available, and file download completes faster.

With ABR, however, clients download larger and larger files along with growing available bandwidth, by upshifting to higher profiles. So bandwidth demand from ABR delivered video is elastic in nature (up to a certain point) and can easily use up all available bandwidth.

As a result, the measured network utilization level can always look high. But in some cases, end users may be quite happy because they are all receiving highest quality streams. Or, in other cases, end users can be quite unhappy because they are all putting up with lowest-profile streams. An operator can no longer tell whether end users are having a good quality-of-experience or not by simply looking at network utilization measurements.

In fact, high utilization on multiple segments of the network may cause operators to assume all such network segments need to be upgraded whereas if they had a better view of the users' QoE they may be able to better determine which segments need to be upgraded immediately and which can be deferred. Having better visibility on actual QoE would enable operators to spend capital on upgrading segments of the network that would yield the most improvement in subscriber QoE.

We ran a simple test to illustrate this point and the results of that test are included in figures that follow. Two interfaces were configured with bandwidth just under 20Mbps, however the number of ABR clients on each interface was varied, 4 in one case and 3 in the other. As seen in `Figure 3` and `Figure 5` below, the utilization levels of the two interfaces are comparable. However as seen in `Figure 4` and `Figure 6` the actual rate profiles selected by the clients in the two cases are quite different.

When 4 clients are sharing an interface we see rate profile selections generally in the range of 3~5 Mbps. However when 3 clients are sharing an interface we see rate profile selections in the range of 4~7Mbps, which typically result in better QoE than when rate profiles selected are in the range of 3~5 Mbps.

The above example is obviously overly simplistic but it does illustrate the point that higher order data such as interface utilization may not give a good indication of QoE due to ABR's behavior of stepping up or down in profiles to match bandwidth available.

Additionally the mix of client devices, content mix, etc., may also be quite different between two network segments. For example the network segment where rates were lower, could in fact be the segment with a larger mix of premium content/subscribers and/or big screen devices. Whereas, the network segment receiving higher rates might have actually a larger mix of handheld devices. So a better measure of QoE should also take into account other characteristics such as device type, content type, and subscriber type

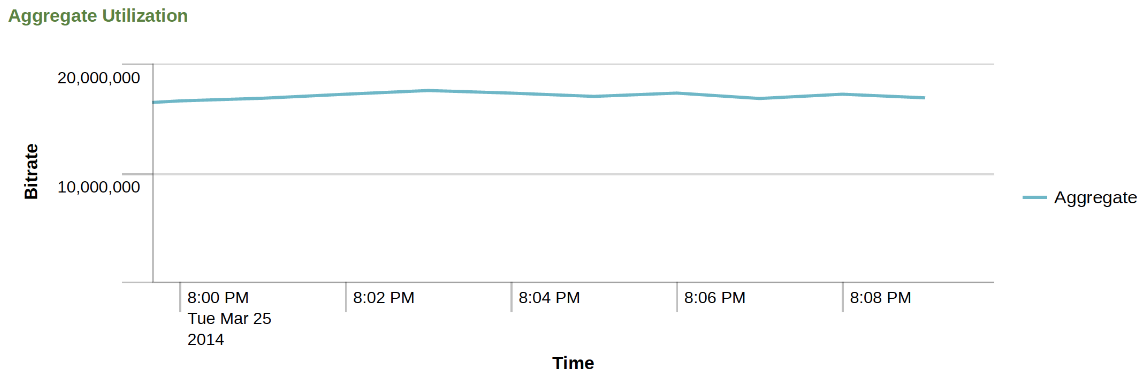


Figure 3 Aggregate utilization of an interface with 4 ABR clients

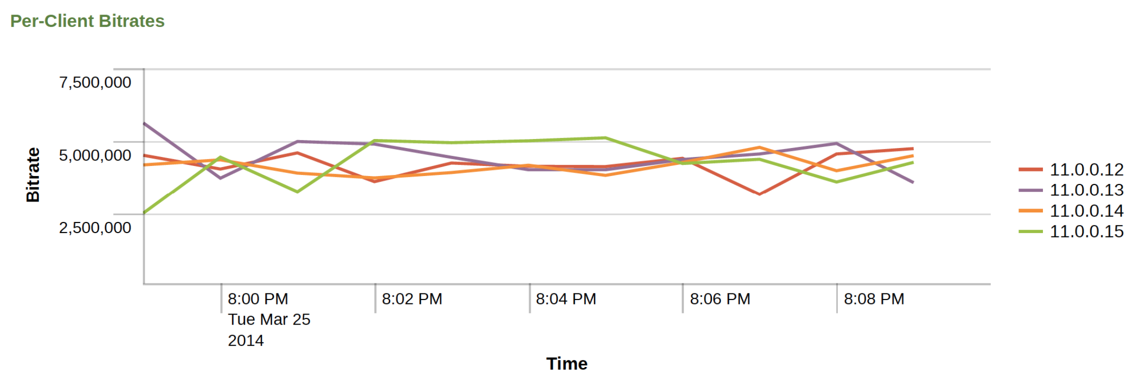


Figure 4 Bitrate per client when 4 ABR clients are sharing an interface

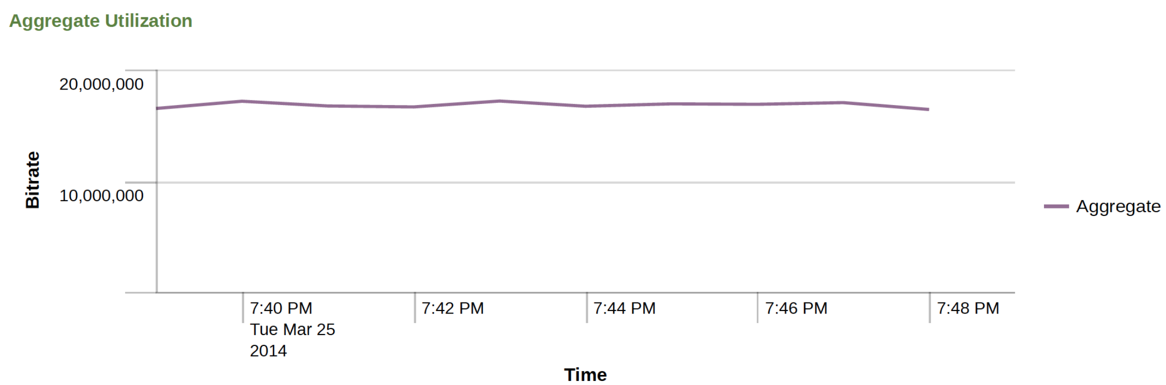


Figure 5 Aggregate utilization of an interface with 3 ABR clients

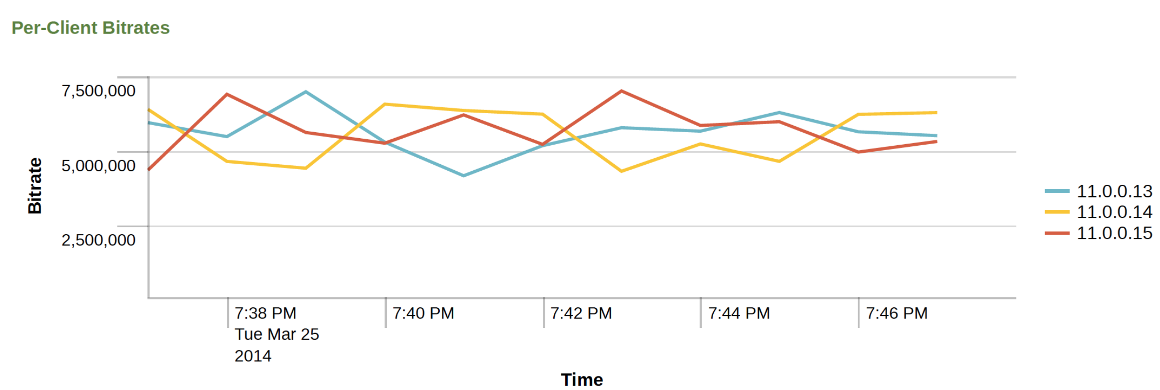


Figure 6 Bitrate per client when 3 clients are sharing an interfa

NETWORK AGNOSTIC ABR DELIVERY

Today Adaptive Bit Rate delivery of video works despite the network. This was fine for over-the-top (OTT) content since OTT providers anyway has no control over the network. Additionally OTT originally started off being a very small portion of the network bandwidth. So in fact each client being greedy and maximizing its own experience was a reasonable approach for a OTT provider solely focused on maximizing their own subscribers' experience when competing for bandwidth with other applications.

However for MSO managed content, the landscape will change in two ways – one, a much larger proportion of the traffic on the network will be ABR video and secondly cable subscribers will be competing with each other for bandwidth. Therefore, operators may want to ensure fairness across a pool of users, to maximize quality of experience across their set of users rather than be satisfied with a greedy approach where each client is operating solely to maximize its own benefit at the cost of other users' experience.

PROPOSED SOLUTION

We hereby propose an alternate architecture --- built upon SDN --- that is better suited for monitoring and delivery of managed video. In this our SDN-based solution, streamers, networks and clients *work together* to both report on and improve users' quality-of-experience (QoE) in the video network.

The SDN architecture aids in this solution by providing a common framework for collecting information from various network entities. By design, different applications can interface with the SDN Controller to extract the particular information they need for their application-specific purposes. For example, there may be a number of applications interested in topology information. Instead of each application collecting such topology information it is more efficient and scalable to have the Controller be the single entity that collects such information from the CMTS and then have the applications interface with the Controller.

For ABR video streaming, we envision a Video QoE Application (VQA) that collects information from various points in the network and analyzes it to provide a more accurate estimate of end users' QoE. For example, as shown in Figure 7, the VQA may collect information about the content from the video steamer/CDN, client metadata (such as device type) from clients and network information from the CMTS. By combining information about the content, client, and network, the VQA can generate analytics to aid the operator in better understanding end users' QoE in different segments of the network.

The Controller can be used to collect network topology information from the network devices (such as CMTS), and end device type (smartphone, STBs etc.). The VQA can then query the Controller for required information. Video QoE metrics can be generated per Fiber Node, per CMTS, or per region. QoE could initially be measured as simply the bitrate achieved by video flows. It could be further enhanced to reflect the actual quality of the flows, where quality is measured by some objective metric such as PSNR (Peak Signal to Noise Ratio), VQM (Video Quality Metric), or SSIM. There are proposals in DASH industry forum on enhancing the Media Presentation Description (MPD) files to include quality information of video fragments. By collecting such rate and quality information from MPD files, HFC topology information from CMTS, and client stream and device information from either the client or the server, the VQA can pool together meaningful QoE analytics for the Operator.

Cable operators can use the QoE analytics for a number of purposes. For instance, they can more clearly identify which segments of their network are most congested and need to be upgraded. This will allow them to spend their upgrade budget more intelligently, rather than simply relying on the interface utilization as a poor measure for QoE.

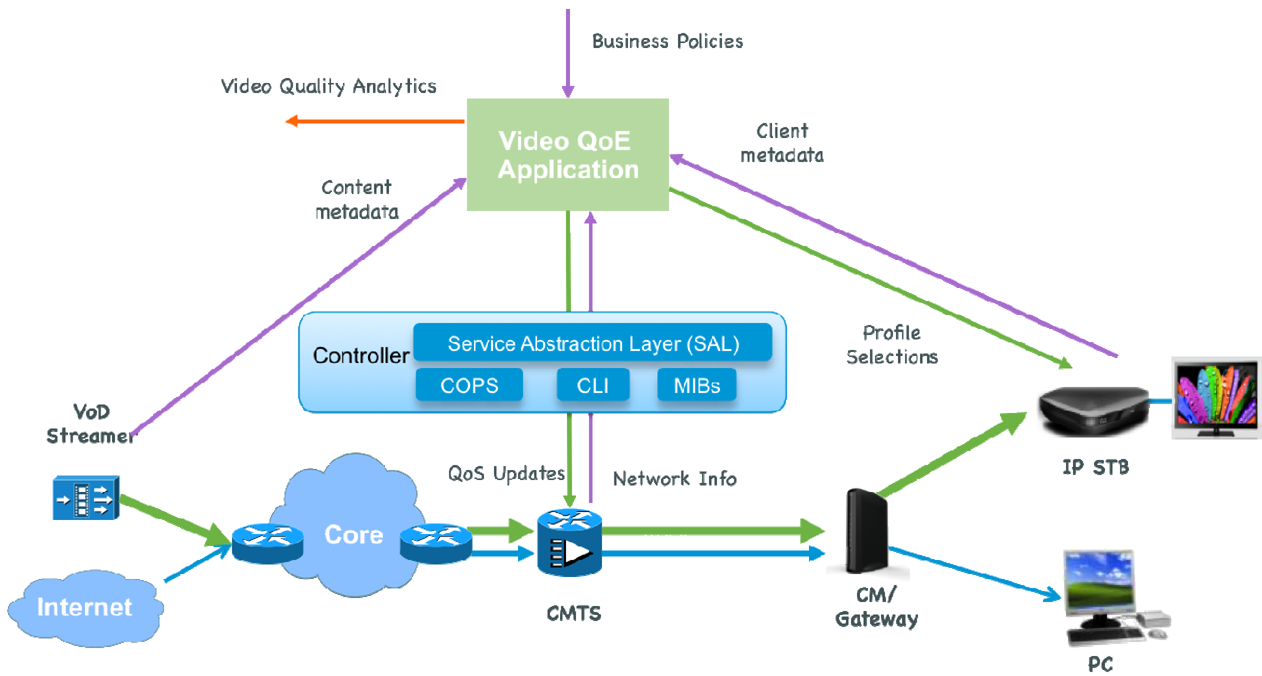


Figure 7 SDN based architecture for QoE Analytics and QoE Optimization

In addition, such analytics can be used to debug and prevent issues in the field. Availability of the QoE analytics makes it much easier for operators to identify where QoE is poor, so they can address the problems in a proactive manner rather than waiting for subscribers to complain about problems.

The proposed solution not only presents a platform for aggregating and visualizing QoE analytics, it further enables the Operator to optimize its network resource for improved end-user quality-of-experience (QoE). For this second use case, the VQA, as shown in Fig. 8, collects a similar set of data as in the Analytics use case. In addition, it can also modify the QoS on the CMTS to optimize users' QoE as needed. One example of how this can be done is via a SDN Controller that has a COPS plug-in that is able to leverage the PCMM support built into CMTSs to apply QoS changes to flows that need it.

The VQA may decide on the bandwidth allocated to each flow based on device type (e.g., STB vs. tablets vs. smartphones), on codec in use (HEVC vs. MPEG4), or on video content complexity (e.g., sports vs. talk shows). Business policies such as premium subscriber or premium content may also influence the QoS that the flows receive. The VQA may in fact use any other information --- e.g., bitrate and video quality scores carried within the Media Presentation Description (MPD) or as companion metadata --- to optimize the QoE for individual flows, so as to achieve certain performance objectives (e.g., maximizing the aggregate quality across all flows, equalizing the quality across all flows in a given user group).

Such an approach will ensure that the bandwidth is available for clients that need it the most. While any approach that strives for aggregate fairness across a group of users will inevitably allocate less bandwidth to some clients than what they would have obtained in the absence of such a mechanism, the intent is that the loss of bandwidth doesn't lead to a noticeable reduction in QoE for that user, whereas the increased bandwidth for some other users contribute to their significant QoE improvement.

Benefits to the Operators are two-fold with this approach: improved aggregate QoE for subscribers and the ability to support more video streams that meet a certain QoE criteria at a given network bandwidth. Indeed, increased packing efficiency can provide significant bandwidth savings for Operators, thereby providing capex savings.

To evaluate the proposed approach we ran simulations with varying number of clients sharing a 100Mbps link. Figure 8 shows the average quality across all the flows (measured in terms of PSNR) for each test.

As can be observed from the figure, the proposed quality-optimized approach yields a significant improvement in average quality across clients. Another way to view the results is that the same average PSNR (of ~42 dB in this particular example) can be attained with our quality-optimized approach while supporting 25 streams at the same bandwidth as opposed to only 20 clients via existing methods. This improvement translates to a 25% increase in packing efficiency.

In fact, the packing efficiency improvement would be even greater if the quality metric in use is the minimum quality across clients. This is because the proposed quality-optimized approach tends to reduce the variation in quality across clients and over time, thereby boosting the minimum quality across all video segments. Using minimum quality as a metric is analogous to a Service Level Agreement (SLA) where an operator would like for the video to stay above a certain threshold most of the time.

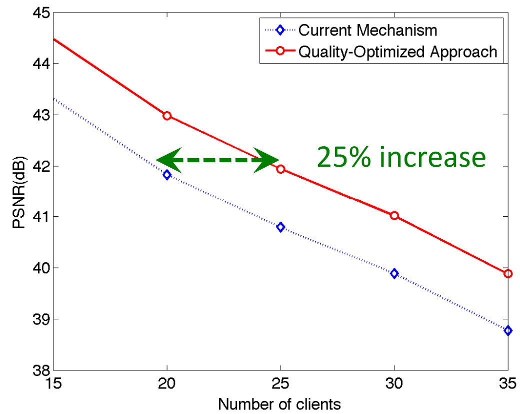


Figure 8 Comparison of the proposed quality-optimized approach against current QoE-oblivious approach. Performance results are plotted in terms of average video quality (PSNR). In this experiment, the number of competing clients vary from 15 to 35 over a 100 Mbps bottleneck link.

SUMMARY

In this paper, we have identified the challenges posed by ABR video for operators in measuring and managing their Access networks. We have also proposed a unique solution to the problem of assessing subscriber QoE and optimizing it. The proposed QoE optimization techniques can be used by operators to either improve QoE of groups of users or to pack more users in a given bandwidth while maintaining QoE. Simulation results have shown that significant improvements in packing efficiency can be achieved thereby providing significant CAPEX savings to cable operators. The proposed solution is based on SDN, and hence can leverage the flexibility of an SDN infrastructure to build a more efficient and intelligent video delivery network.

REFERENCES

1. Cisco White Paper, “Cisco visual networking index - forecast and methodology, 2011-2016,” http://newsroom.cisco.com/dlls/index.html.
2. “HFC capacity planning for IP Video”, Sangeeta Ramakrishnan, Cisco Systems, SCTE 2011.
3. History of Move Networks, Available Online: http://www.movenetworks.com/history.html
4. A. Zambelli, “IIS Smooth Streaming Technical Overview,” http://www.microsoft.com.
5. Apple Inc., “HTTP Live Streaming Overview,” https://developer.apple.com/library/mac/documentation/networkinginternet/conceptual/streamingmediaguide/Introduction/Introduction.html
6. MPEG, “ISO/IEC 23009-1:2012 Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats,” 2012.
7. https://www.opennetworking.org/sdn-resources/onf-specifications/openflow

APPLICATION OF POLICY BASED INDEXES AND UNIFIED CACHING FOR CONTENT DELIVERY

Andrey Kisel
Alcatel-Lucent

Abstract

Caching technologies are used to improve quality of experience while controlling transit and transport costs. Two common caching technologies are considered transparent caching (TC) and on-net content delivery network (CDN). We illustrate cache avoidance, analyze and model two common approaches - content hash indexes and policy based indexes - to minimize the impact. We show how the use of policy based indexes and applying advanced features of IP routing can be used to merge on-net CDNs and TC together into a single unified caching system with superior and quantifiable improvement in cache efficiency, while providing resilience and simplifying operations.

INTRODUCTION

The content industry is being reshaped by exponential video traffic growth, rapid proliferation of connected devices and delivery formats, and surging consumer demand for video from many different online sources. To stay competitive, MSOs need solutions that can cost effectively deliver pay TV content and Internet traffic with high quality of experience while optimizing transit and peering costs.

The majority of the growing Internet traffic originates from outside of MSO's networks and from Content Providers with whom they do not have business agreements to deliver content. Internet traffic is diverse by nature, and while video is the largest type by volume, software updates and Web traffic constitute substantial proportions.

MSOs are using caching technologies to improve quality of experience while controlling delivery costs for own video services and Internet traffic. They use transparent caching for unmanaged over-the-top Internet content and on-net content delivery network for managed or premium content they own or distribute for partners. Currently both solutions operate independently from each other.

There are several fundamental challenges that MSOs are facing with unmanaged OTT content: cache avoidance, a technique used by content providers to prevent MSOs from caching their content, different traffic types that needs to be cached, and maintaining high network resilience where caching is deployed. In this paper, we illustrate and model the impact of cache avoidance on TC and introduce two common approaches - content hash indexes and policy based indexes - to minimize this impact. Using traffic data from field trials and through numeric modeling we compare the two approaches and show how the use of policy based indexes improves cache efficiency and reduces latency for any traffic type. We then discuss application of advanced features of IP routing to resiliently deploy TC into greenfield and existing on-net CDNs creating a single unified caching system with improved resilience and simplified operations.

IMPROVING CACHE EFFICIENCY

Cache Avoidance

Caching is intended to provide distributed and scalable content delivery from inside MSO networks, optimizing network savings while maintaining a consistently high QoE. Cache efficiency is typically measured in cache-hit rate, the ratio of Bytes delivered form cache storage B_C versus total Bytes delivered B_C.

$$R_{CH} = {}^{B_C}/_{B_T} \quad (1)$$

HTTP cache control headers such as “Cache-Control", "Etag", "Date", "Last-Modified" are commonly used to improve cache efficiency and are well defined by RFC 2616. However, some content providers do not make their content objects easy to cache by either obfuscating object URLs, using unique object URLs for the same objects, setting incorrect values for cache control headers such as “Cache-Control", "Etag", "Date", "Last-Modified" or a combination of these methods, even if objects are cacheable by nature. While there may be valid reasons for using such approach, such as load balancing, often there are other considerations such as promoting deployment of content provider’s own caching inside MSOs networks. Regardless of the reason, this approach is known as cache avoidance, and there are two broad types of cache avoidance based on using semi-dynamic or fully dynamic URLs.

Semi-dynamic URLs are content URLs containing different object requests for the same content object. Different hostnames may also be used, for example, for load-balancing. A common feature of semi-dynamic URLs is extensive use of dynamic query string parameters and random parameters attached to object URLs to make it unique, even if the object name is the same. Semi-dynamic URLs make traditional on-net CDN caching un-efficient because multiple URLs point to the same content. An example of a semi-dynamic URL from NetflixTM is presented below:

http://31.55.163.113/737473630.ismv/range/34270098-34927096?c=gb&n=25127&v=3&e=1391197153&t=2EWCIyTIS201F4kw3AAyYWfUhb8&d=silverlight&p=5.rjKPKTOOlAKs2xPbICBh007uSde_EIHN1XbhqYSxeAs&random=976648079 (2)

The URL includes ‘random’ query string parameter making content request unique among different users requesting the same video object. In other approaches object name, ‘737473630.ismv‘ in our example, may be unique for the same video objects, and the video object may be only identifiable by static query string parameters inside semi-dynamic URL.

Fully dynamic URLs completely randomize object name and remove query string parameters, a combination of which may uniquely identify video object, form object URLs. For example, two URLs below requested the same object:

http://www846.megavideo.com/files/d008f8c759a4f4b3f07ccef7ea7588a4/
http://www763.megavideo.com/files/f66f936f7fc39f1bb9524f49c5f54184/ (3)

<u>Caching Policy and Content Hash Indexes</u>

In order to improve TC R_{CH} two main approaches have been developed: content policy index (content policies) and content hash index (content hashing). Content policy is a set of TC instructions how to group similar hostnames and identify the same content objects being delivered from each hostname using different semi-dynamic URLs. Essentially the policy defines parts of a URL that can be omitted and parts that shall be used for unique content identification - cache index. Content policy relies on presence of unique static parts or parameters inside each URL identifying content object. For example, a policy to cache semi-dynamic URLs (2) is illustrated below:

[policy-netflix]
match url regex http://([0-9]{1,3}\.){3}[0-9]{1,3}/(?<chunk>\d+\.ismv)\?c=\w+&n=\d+&v=\d+&e=\d+&t=(?<t>[\w|-]+)&d=(?<device>\w+)&p=5\.(?<p>[\w|-]+)
cache_index = netflix-$e-$chunk-$device (4)

The policy effectively defines that TC can use a combination of chunk name, device name and ‘e’ parameter to uniquely identify content, omitting the remaining of the URL. Content policy can identify content objects without waiting for any response from the Origin server, and therefore content delivery can start faster improve time to first byte T_{FB}, a time between sending request and receiving the first byte of response. Content policies can efficiently cache objects identified by semi-dynamic URLs.

Alternative approach uses content hash indexes. Content hash is a hash computed form the first Bytes of object, and used to uniquely identify the complete object. In order to compute the hash, TC passes original object URL to the Origin, waits for the object delivery to start, computes hash of the first Bytes, and checks whether the hash matches hashes for content objects already stored in the in the cache. If the match is found, TC takes over object delivery and disconnects the Origin. Typically the size of data to compute hash is measured in the number of IP packets, for example, often 10 IP packets are used to identify video objects. Multiple hashing algorithms can be utilized, for example, MD5 (RFC1321) or SHA-1 (RFC3174) is often used.

Content hash indexes can efficiently cache content identified by both semi-dynamic and fully-dynamic URL. However, the proportion of fully dynamic URLs in the total traffic volume observed in a field trial is negligible, as discussed below, and they are considered instead an alternative to content policies when caching content identified by semi-dynamic URLs, rather then a supplement. Content hash indexes cannot improve T_{FB} because TC has to wait for the Origin's response, and are less efficient in caching smaller object where a larger proportion of the object needs to be used for hash computation and therefore received from the Origin. We will compare efficiency of both approaches in the next section using object distribution inside traffic volume obtained from field trial.

<u>Cache Efficiency</u>

We will use two characteristics when analyzing cache efficiency: R_{CH} (1) and P_{CT} – is the number of bytes required to compute hash index and B_T is defined in (1).

$$P_{CT} = \left(1 - {B_H}/{B_T}\right) \quad (5)$$

Figure 1 illustrates percentage of video traffic excluded form caching if content hash indexes are used for $B_H = 15000$ Bytes, (10 IP packet, 1500 B each), 2s Smooth streaming and 10 sec HLS segments.

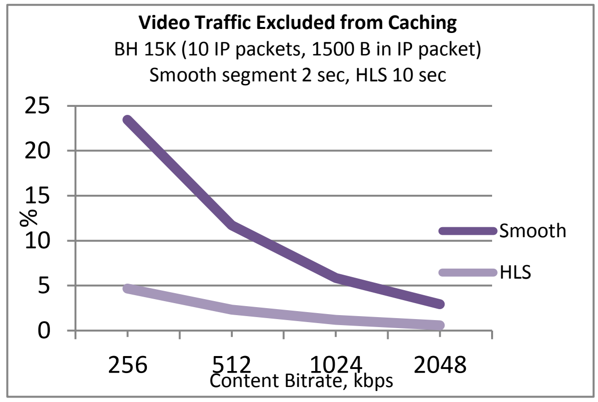


Figure 1. Video traffic excluded from caching.

As illustrated on Figure 1 more traffic is excluded from caching for lower video bitrates, decreasing cache efficiency for networks with such traffic profile, for example, serving large number of mobile clients.

While Figure 1 illustrates theoretical impact of content hashing on HTTP Adaptive streaming, to consider practical impact on all Internet traffic requires introduction of object distribution. For simplicity of practical simulations we split object sizes into bands, and define object distribution as

$$P_{ONi} = \frac{N_{O\,i}}{\sum_{k=1}^{M} N_{Ok}} \quad (6) \text{ and}$$

$$P_{OVi} = \frac{B_{T\,i}}{\sum_{k=1}^{M} B_{Tk}} \quad (7),$$

where P_{ONi} is a percentage of objects inside *i-th* band in relation to total number of objects, P_{OVi} is a percentage of traffic inside *i-th* band in relation to total traffic volume N_{Oi}- number of objects inside *i-th* band, B_{Ti}-size of all objects inside *i-th* band and M- the number of bands. Figure 2 illustrates observed object distribution for a large NAR ISP.

On the next step let's consider P_{CT} gains - ΔP_{CTi} for object distribution presented in Figure 2. Figure 3 illustrates gains in the percentage of cacheable traffic for content

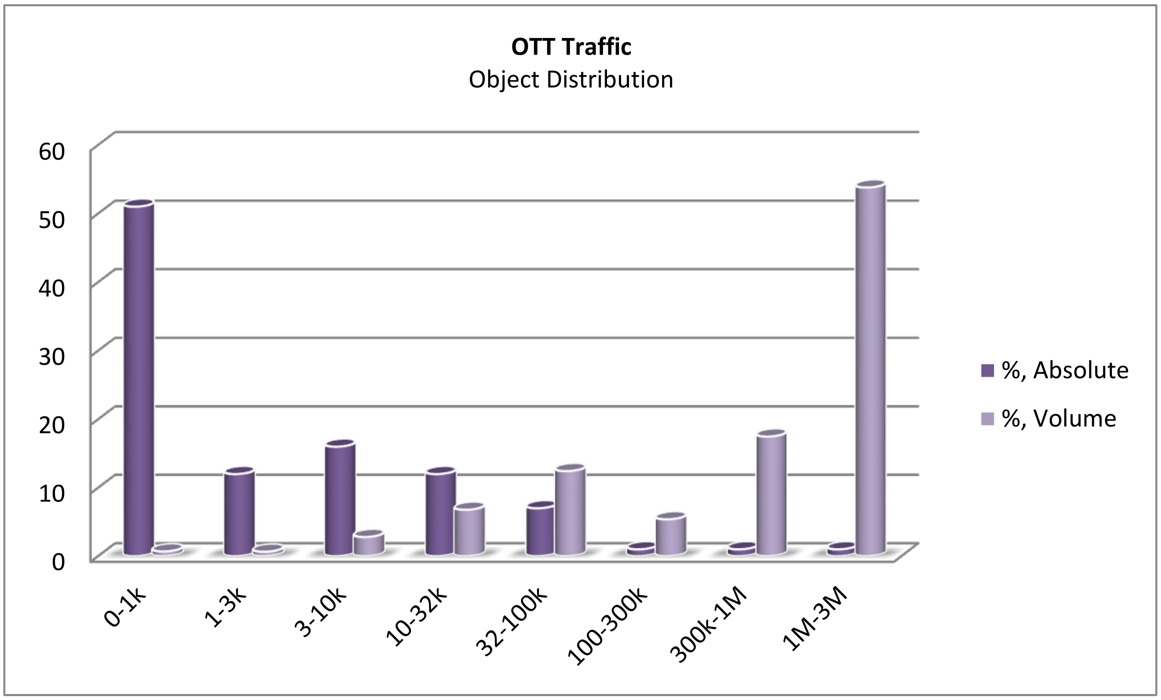


Figure 2. Object distribution for Internet (port:80) traffic.

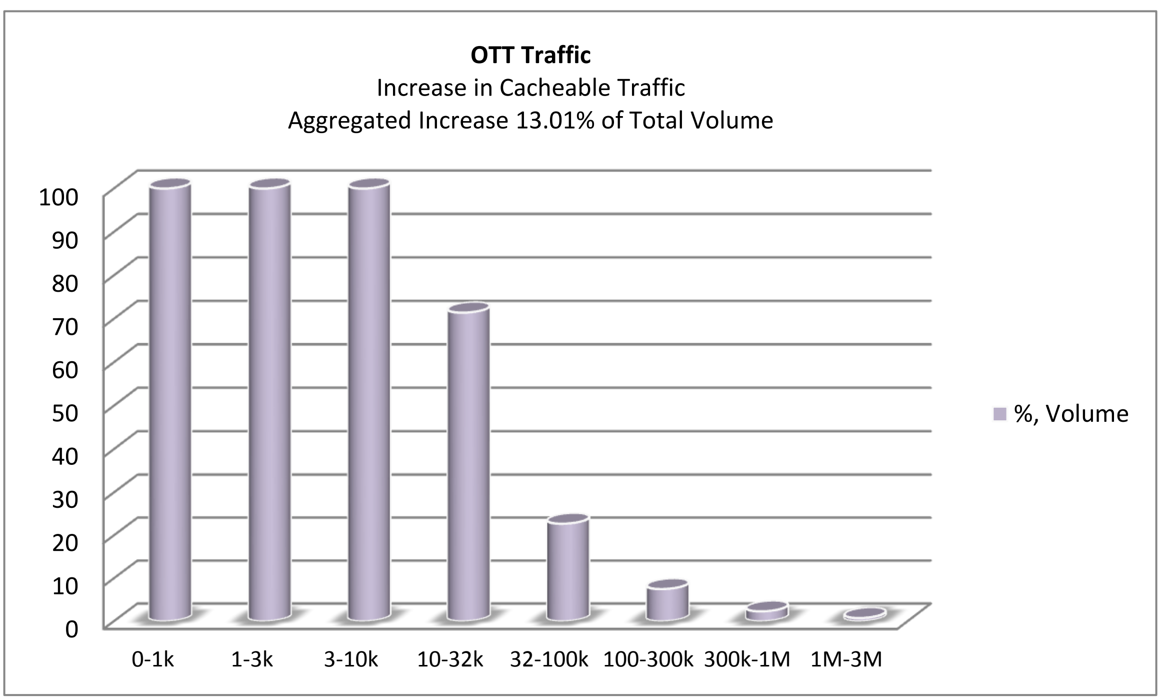


Figure 3. Increase in cacheable traffic.

policies for each of object bands, and aggregated gain P_Σ across all bands factoring into account relative contribution of traffic inside each band into total the volume (illustrated on Figure 2).

When calculating aggregated gain we applied unified object distribution inside each band for simplicity.

$$\Delta P_{CTi} = \begin{cases} 100\%, & B_H \geq \overline{B_{Oi}} \\ (B_H/\overline{B_{Oi}}) \times 100, & B_H < \overline{B_{Oi}} \end{cases} \quad (8),$$

where $\overline{B_{Oi}}$ is the average object size for *i-th* band, B_H is introduced in (5) $P_\Sigma = (\frac{1}{100}) \sum_{i=1..M} P_{OVi} \times \Delta P_{CTi}$ (9),

where P_{OVi} is introduced in (7) and ΔP_{CTi} is introduced in (8).

As illustrated by Figure 3 overall increase in cacheable traffic for object distribution in Figure 2 is 13%. The increase depends on object distribution and will be larger for bigger proportion of smaller objects. Figures 2 and 3 show two interesting trends. First, in terms of object numbers, the majority of objects are small, therefore any caching solution that can deliver small objects with better T_{FB} and lesser latency would improve QoE. Policy based indexes reduce T_{FB} for over 80% of traffic objects more then content hashing. Second, in terms of traffic volume, larger objects dominate traffic and make main contribution to R_{CH}, however, the proportion of smaller objects (e.g. below 100 kB) is non-negligible.

So far, we considered the gains in cacheable traffic by using policy indexes instead of content hash indexes. Let's consider impact of using content hashing on cache hit rate R_{CH}. R_{CH} is a linear function of B_C, and if portion of the traffic is excluded from caching, B_C and R_{CH} will linearly decrease as illustrated on Figure 4. The rate of decrease will be higher for larger potentially achievable cache R_{CH}, therefore content hash indexes impact most the better performing caching systems.

$$R_{CH} = R_{CH\,MAX}\left(1 - {P_{EXC}}/{100}\right) \quad (10),$$

where $R_{CH\,MAX}$ cache hit rate when policy indexes are used, P_{EXC} percentage of traffic excluded from caching.

Content policies offer noticeable

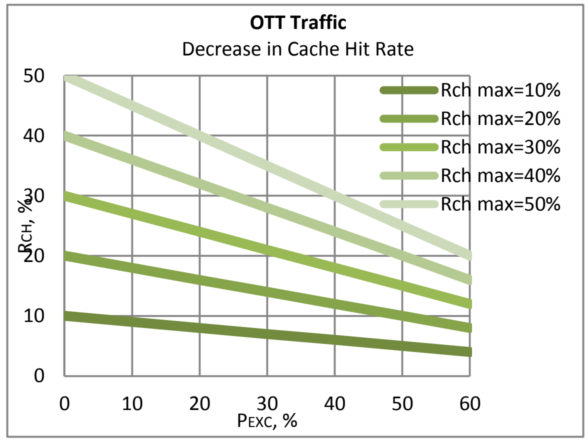


Figure 4. Cache hit rate

decrease.improvement both in increased volume of cacheable traffic and in cache-hit rates. The improvement is dependent on the traffic profile inside ISP's network. Traffic profile from a large NAR ISP yield 13% improvement in cacheable traffic and derived 6% gains in cache hit rate. Another important characteristic, often overlooked when analyzing transparent cache efficiency, is caching of small objects and improved T_{FB}. Content policies enable to improve T_{FB} for 80% of more objects, which otherwise would have slipped through caching using content hash indexes. While an argument can be made that content hashing can cache objects identifiable by fully dynamic URL, we have not observed traffic with fully dynamic URL among top 10 sites contributing over 80% of traffic volume in the field trial, therefore on their own content hashing is a weaker alternative to content policies.

Content policies can be configured to initially contact Origin, similar to content hashing, if, for example, preferred for operational reasons. However, content policies allow initial contact with the Origin to be made without initiating content delivery by using 'If-Modified-Since' HTTP header, preserving cache efficiency. While another argument can be made about management of content policies, we observed that content policies are relatively static, and none of them needed updates after initial tuning likely due to fairly static format of content URLs used by Content Providers. Moreover, policy updates could be fully automated. Therefore both arguments do not disprove superior cache efficiency of content policies, or introduce noticeable barriers against applying content policies.

RESILIENT DEPLOYMENTS

Application of Policy Based Routing

Traditional deployment of transparent caching relies on using policy based routing (PBR) to divert all or HTTP (port:80) traffic to the cache. This approach places the cache on data path, and therefore resilience is one of the main considerations for MSOs. Load-balancers and *N+1* or *N+N* redundancy are typically used for resilience, although at extra costs and complexity. We will consider alternative approaches using advanced functions already available in IP routing. The approaches rely on the ability of IP routing to reroute traffic from failed caches using conditional redirects, or ability to duplicate traffic, letting the Origin deliver objects instead of a failed cache. The approaches do not require any modifications to the cache, for example, adding additional protocols or signaling.

First approach is based on using conditional redirects applied to policy based routing. Figure 5 illustrates this approach together with logical modifications of a routing plain.

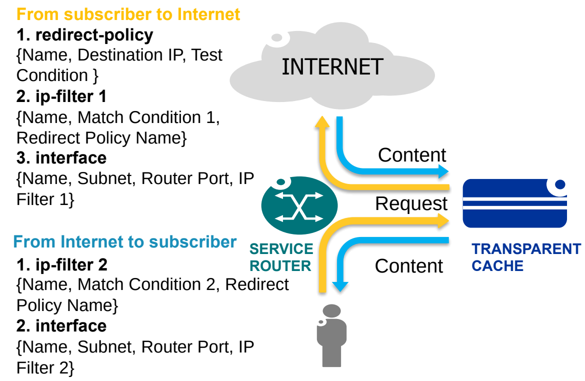


Figure 5. Resilience based on PBR with conditional redirects.

Let's consider logical modification to the routing plain for resilient handling of subscriber traffic first. A redirect policy is added to route traffic to the *Destination IP* of the cache if *Test Condition* is passed. *Test Condition* verifies cache availability, and most of the routers can test ICMP ping, HTTP GET or SNMP messages. Second, *ip-filter1* is created incorporating redirect policy and filtering traffic passing *Match Condition 1,* for example, match condition '*protocol:TCP, destination port:80'* would effectively diverting to the cache only HTTP traffic that it can potentially cache. Removing other traffic types from the cache frees cache's resources for caching instead of analyzing and returning without caching other traffic types. Next, the *ip-filter 1* needs to be applied to the router port handling traffic from subscribers.

Resilient handling of return traffic from the Internet requires similar logical steps (Figure 5) with main differences that *Match Condition 2 = protocol:TCP, source port:*80 selects HTTP traffic based on source port and the new filter needs to be applied to the router port handling return traffic from the Internet.

In case of transparent cache failure, the traffic is routed to the upper caching layers or to the Origin; therefore they would deliver more content until failed cache is restored. Field simulations in a large NAR ISP network showed that conditional redirects applied to policy based routing did not cause any noticeable network outage following simulated cache failure, and from the user experience, there were no disruptions to unicast streaming services from Amazon and YouTube, no disruption in constant ping, or in ability to surf Web. Therefore, policy based routing with conditional redirects offered a robust and cost efficient deployment for transparent caching without need to introduce extra network functions, for example, load-balancers or *N+1*/*N+N* redundancy.

Application of Traffic Mirroring

Let us consider alternative approach to achieving resilience where conditional redirects are not available. The approach is based on feature of IP routing to duplicate traffic, or 'port mirroring', and requires the cache to intercept delivery for the content objects it decides to serve. The 'port mirror' approach is illustrated on Figure 6.

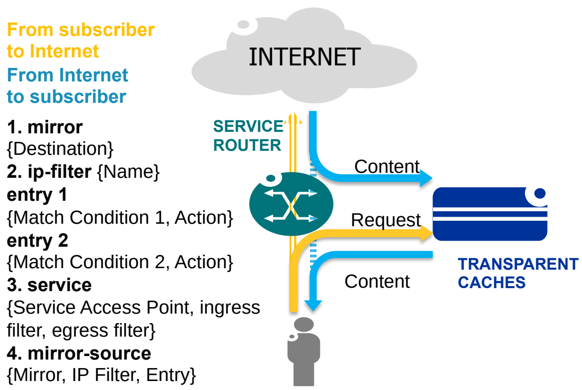


Figure 6. Resilience based on traffic mirroring.

Logical modifications of a routing plain for 'port mirror' based resilience requires creation of a mirror destination point connected to TC, a filter forwarding (*Action=forward*) HTTP traffic to and from the Internet, for example where '*Match Condition 1 = protocol:TCP, destination port:80*' and '*Match Condition 2 = protocol:TCP, source port:80*', a service applying the filter to the ingress and egress Internet traffic, and defining filtered traffic as a source for previously defined mirror destination point.

With 'port mirror' approach both TC and Content Origin can see object requests, and TC needs to take over delivery for objects it has in the cache. Figure 7 illustrates flow diagram how TC can intercept delivery of content objects.

As shown on Fig. 7, initial request is sent to both Origin and TC, and there is a race between the Origin and the cache when replying to the request. The cache needs to win the race for successful delivery, which can be achieved by inserting HTTP 302 Redirect, message spoofing the Origin and instructing
the client to reconnect directly to the cache (or it's peer), and followed by instructions to the client to close current TCP connection to the Origin, e.g. by sending TCP FIN.

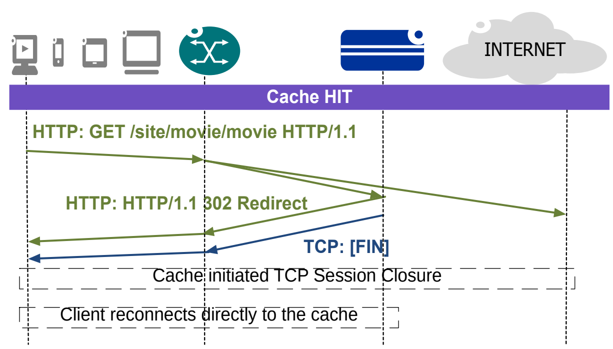


Figure 7. Flow for TC intercept of object delivery.

A cache failure in 'port mirror' approach would not affect requests delivered to the Origin because the cache is no longer on the path and resilience is achieved natively. Recovery mechanism for clients with in-progress object or video delivery during the cache failure is the same as in PBR approach, the client needs to re-establish TCP connection and re-request object being requested but not received when the cache failed, therefore practical testing of the user experience obtained for PBR mode equally apply. Similarly to PBR mode, in case of transparent cache failure, the traffic is routed to the upper caching layers or to the Origin.

However, port mirror approach has drawbacks. First, some client may either not support HTTP 302 or being prevented by local security setting from following 302 off domain for the initial object requests. Second, the race condition means that if there is small latency between the client and the Origin the cache may not win all races reducing P_{CT} and overall R_{CH}. Impact of P_{EXC} on R_{CH} follows the same linear function as for applying content hashing and as illustrated on Figure 4. The main difference is that in this case traffic would be excluded form caching due to lost races with the Origin rather then due to the need to hash first bytes of each object.

Unified Caching

The PBR with conditional redirect or port mirror based resilience enable MSOs to build robust distributed unified caching architecture where transparent caches are deployed close to network edges caching both OTT and CDN content, while CDN appliances are deployed closer to the network core. Both architectures provide caching resilience without relying on lab-balancers allowing upper-layer CDN or the Origin serve content requests instead of failed transparent caches.

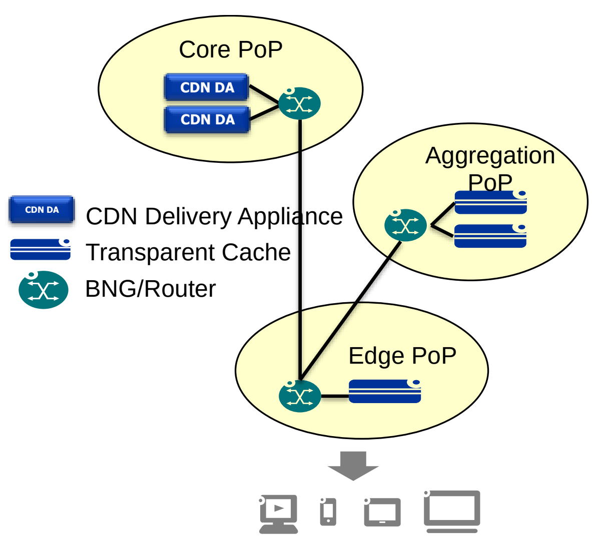


Fig. 8 Resilient unified caching.

Center the captions under each illustration and make the text large enough so that captions are easy to read.

CONCLUSIONS

We considered two fundamental challenges for MSOs when caching unmanaged OTT content: cache avoidance and maintaining network resilience, and introduced two approached to deal with cache avoidance: content policy indexes and content hash indexes. Content policies show improvement in volume of cacheable traffic and in cache-hit rates and are a stronger alternative to content hashing. The improvement is dependent on traffic profile in an ISP network. For traffic profile observed in one of the NAR ISPs policy indexes yield 13% improvement in cacheable traffic and 6% gains in cache hit rate. Other parameters affecting QoE, for example, caching of small objects and time to first byte are also discussed. Content policies enable to improve T_FB for 80% more of objects compared to content hash indexes for the traffic profile from the same ISP. That traffic would otherwise would have slipped through caching, and although 80% of objects do not translate in equal volume of cacheable traffic in Byte terms, it contributes to QoE improvement.

Further we introduced two approaches to maintaining network resilience without need for deploying extra functions in networks like load-balancers. The approaches used advanced features readily available from network routing of underlying networks: PBR with conditional redirect and port mirror. Both approaches enable MSOs to build robust distributed unified caching architecture where transparent caches are deployed in distributed network locations. However, port mirror approach is restricted to client that can support HTTP 302 redirect or are allowed to follow off domain redirects, and the race condition may reduce P_CT and overall R_CH if the latency in the network is small.

In conclusion, using PBR with conditional redirect and policy based indexes enables MSO to benefit from a more robust unified caching solution for premium and OTT content resulting in improved cache efficiency and resilience, reduced latency and simplified operations.

REFERENCES

1. Introduction to Content Delivery Networks. White paper. September 2011, Velocix, Alcatel-Lucent.
2. Velocix CDN, A Vision For The Future of Content Delivery. Strategic White Paper. September 2012. Velocix, Alcatel-Lucent.
3. Velocix unified caching. Strategic White Paper. Dec 2012. Velocix, Alcatel-Lucent.

ABBREVIATIONS

BNG	Border Network Gateway
CDN	Content Delivery Network
HLS	HTTP Live Streaming
HTTP	HyperText Transfer Protocol
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPTV	IP Television
MD5	Message Digest
MSO	Multiple-System Operator
NAR	North America Region
OTT	Over The Top
PBR	Policy Based Routing
PoP	Point of Presence
QoE	Quality of Experience
RFC	Request For Comments
SHA	Secure Hash Algorithm
TC	Transparent Caching
TCP	Transmission Control Protocol

Applying Web Principles to the Network

Brian Field, PhD
Comcast Cable Communications

Abstract

The web has been a tremendously successful application within the Internet ecosystem. There are many reasons for this success, but one key reason is likely a result of the use of open source technology. Specifically, the web is built on an open protocol (HTTP) running on an open source application (Apache) on an open source operating system (Linux) running on common, -off-the- shelf hardware. This ecosystem is ideal for innovation—any application developer can test and introduce changes into any of these software domains as appropriate for their specific application. The better changes get rolled into the ecosystem and become available for all downstream developers to use. This pay-it-forward ecosystem is likely a primary reason for the success of web applications.

The web and the Internet run over a network (routing) infrastructure that has clearly provided the foundation for making the web and Internet successful. However, the network ecosystem is not as open as the web ecosystem and this could be stifling innovation in the network space. This paper explores how we can make the network ecosystem more open and provides insights into the value this openness provides to both the network operator and the application developer.

INTRODUCTION

The web has been a tremendously successful application within the Internet ecosystem and there are many reasons for why. The roots of this success have their origins in the open source nature of its ecosystem. Specifically, the web is built on an open protocol (HTTP) running on an open source application (Apache) on an open source operating system (Linux) running on common, off-the-shelf hardware.

This ecosystem is ideal for innovation, as an application developer can test and introduce changes into any of these software domains as appropriate for their specific application. The better, more useful changes get rolled into the ecosystem and become available for all developers to use. This pay-it-forward ecosystem is likely a primary reason for the success of web applications.

Conway's Law

In late 1968, Melvin Conway proposed that "organizations which design systems ... are constrained to produce designs which are copies of the communication structures of these organizations" (1968).

One way to think about what Conway suggested is illustrated in the following analogy. Consider a company that consists of two divisions—one division that makes delicious peanut butter and the second division which makes delicious chocolate. If these two divisions do not have meaningful discussions or interactions, it is unlikely that they will collaborate and make a great confection known as "peanut butter cups".

Current service provider organizations are often aligned where the network engineering teams work in one portion of the company and the application engineers work in a different organization. This separation not only exists on paper, but in where these organizations are located: on different floors, buildings, even cities. Applying Conway's thinking, this logical and physical partitioning is a barrier to collaboration, cross pollination of ideas, and innovation.

This paper suggests that there is great value to be had in the network space if these separate teams were to collaborate, leveraging the technologies that have made the web so successful and applying them to the network space. Specifically, in this paper we identify and apply three web component technologies to the network space and explore the resulting value in such an application.

OPEN SOURCE AND THE WEB'S SUCCESS

As discussed earlier, one of the primary reasons for the web's success is its open ecosystem. This open environment allows the developer to make changes to accommodate specific application requirements at any of the three software layers of the web stack, shown in Figure 1. Good and great ideas often get rolled into the open source package and become available to the entire community. This is an incredibly powerful development environment as it allows innovation and agility at all layers within the ecosystem.

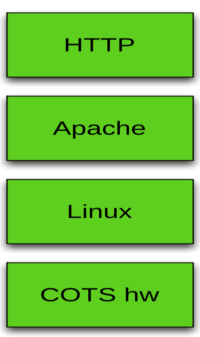


Figure 1: Web stack

Web Model Versus Router Model

Now, consider the router stack. The router stack consists of open protocols (BGP, ISIS, PIM, etc.), but these protocols are managed within the confines of a proprietary command line interface, running within a proprietary operating system, inside a proprietary chassis, running on proprietary hardware ASICs, as shown in Figure 2.

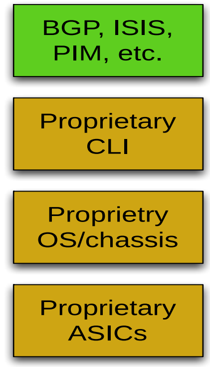


Figure 2: Today's router stack

While the IETF has provided a forum for discussing the details of routing protocols, network operators are limited in their ability to individually experiment and innovate with new concepts due to this existing proprietary router environment. In fact it has been this restrictive ecosystem that has been a motivator for OpenFlow(see Openflow reference), where a set of APIs have been developed to manage this proprietary ecosystem. These APIs are considered by some to be “SDN” (software defined networking), as shown in Figure 3.

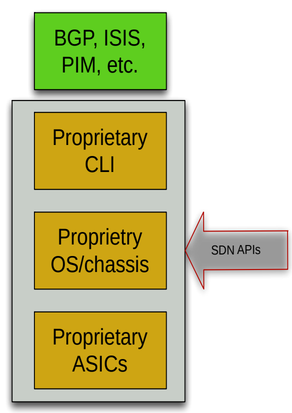


Figure 3: Is this SDN?

But is this really SDN? Or posed differently, is this all we should expect from SDN or can it leveraged for more? The service provider and MSO community may see greater benefits in the network space if we can get the router model to really align with the open source model that has made the web environment so incredibly successful.

To set the stage for this, a discussion is provided about the current service provider ecosystem and how the application technology is being deployed; this really becomes the foundation for an open environment and an innovative ecosystem in the routing domain.

Virtualization in the MSO space

Many MSOs and Service Providers are in the process of augmenting their network infrastructure with data centers full of servers that provide a virtualization environment for applications. Over the last several years it has been common for the application developers to deploy their services on virtualization platforms rather than application-specific hardware platforms. Service providers, including MSOs like Comcast, have been deploying these data center based virtualization platforms to provide an Agile environment for the deployment of their own internally created and managed applications. These virtualization platforms have gone from a centralized deployment, where there are a small number of data centers nationwide, to a more regional deployment, where data centers are located in the regions they serve. One could envision that over time these virtualization platforms will reach yet further into an MSO's footprint, including into many hub sites.

Network Topology for Applications

In simple terms, a service provider or MSO network consists of three pieces: 1) the core network, made up of core routers, with a primary purpose of forwarding packets over very large bandwidth links; 2) the network edge, consisting of routers that provide application/edge features and physical aggregation; and 3) the applications themselves.

The edge routing platform tends to have application-specific features and configurration to support constructs such as ACLs, DSCP policies, QoS settings, etc. As applications change their behavior and what they require from the network, they place new feature demands on the edge router software. As such, the code running on these edge platforms tends to need more frequent code updates. Since these edge platforms are proprietary systems built by vendors, vendors must be convinced to reengineer the software to support any new network feature that applications require. This often means getting the feature on the vendor's roadmap and waiting months or years for delivery.

APPLYING WEB PRINCIPLES TO THE ROUTER ECOSYSTEM

First Application: Router Platform Edge Virtualization

The edge router environment is one where virtualization and open source have a useful applicability. As discussed previously, edge router functionality is required to change rapidly based on application needs. Because of this, edge routers would be well served to run inside a virtual instance. Specifically, there would be great benefit in taking the existing edge router platform and partitioning it into two components: the software component that needs to evolve and change quickly, and the hardware component that performs aggregation and packet forwarding. The edge router software component could be moved into a VM in the virtualization server. Furthermore, router software could come from a router vendor or could be open source based. The open source router paradigm is particularly compelling in that it enables the service provider to innovate, be Agile, and create application-specific features in the routing code but on the service provider's timeframe. Figure 4 details the evolution of the edge router into this new VM-based router paradigm.

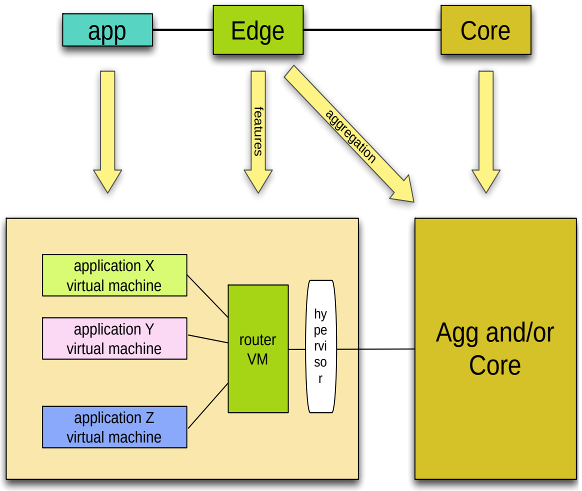


Figure 4: Virtualizing the edge router

In this model depicted in Figure 4, applications exist as VMs, but instead of having connections to the physical edge router, they have virtual Ethernet connections to this new edge router VM. This edge router VM has the set of features found in the non-virtual edge router, but these features are executed within a VM as opposed to operating within a proprietary platform and system.

<u>Result of Virtualizing the Edge Router</u>

Consider this new router VM paradigm, where we have moved the edge routing features into a router VM running in our virtualization platform. The resulting virtualized edge router stack looks nearly identical to the web stack depicted in Figure 5. The virtualized router stack runs as an open source VM on an open source operating system (Linux). This implementation should fundamentally improve the edge routing environment for the same reasons that the web environment benefited: this open source router VM ecosystem will provide platform innovation and agility in the router edge feature space.

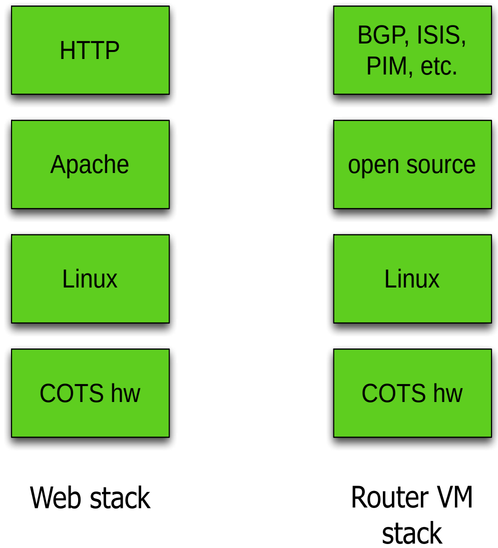


Figure 5: Web and open source router VM stack

<u>The State of Network Control Protocols</u>

A second area of the network ecosystem that might benefit from web ideas and concepts is in the area of network control protocols, specifically BGP. This paper focuses on BGP Link State (Gredler, 2012). BGP Link State is a new address family defined in BGP that carries details about the underlying network topology. BGP Link State encodes the state of the Interior Gateway Protocol (IGP) and makes this information available to applications. It has been proposed that BGP Link State carry not just the network topology, but other topology information, including how VMs are connected into the routing infrastructure and the power topology (what PDUs are being used to power each power supply in each server and router). The VM and power topologies are real-time exposed via LLDP, meaning that BGP Link State will likely carry both IGP information and the contents of LLDP messages (Field, 2013).

Current thinking in how to encode the IGP information into BGP is to follow the existing BGP paradigm: encode the IGP information into BGP using binary protocol constructs, namely, a number of TLVs based on bit and byte structures.

The drivers for BGP's binary encoding system are in part to make the protocol network and processing efficient, which was a valid engineering trade-off for the network ecosystem of the late '80s and early '90s when BGP was first being developed. However, that environment of T1 links and very limited router CPU processing is long behind us. Because BGP would be required to carry very different types of information for three different topologies, more flexible encoding approaches should be considered, approaches that enable agility and innovation.

Second Application: Network Control Protocols

When considering encoding in the web environment, much control and data content is encoded as JSON. While not as compact as BGP binary encoding, JSON encoding is well known and commonly used in the application space, while writing binary-based BGP encoding is relegated to handfuls of network engineers with BGP expertise. To take advantage of the benefits that come from wide adoption and expertise, JSON encoding should be seriously explored for new address families for BGP, such as Link State.

In addition, if encoding the BGP Link State content as JSON has benefits, then consider the possible benefits of refining the BGP data passing primitives (OPEN, UPDATE, NOTIFICATION, etc.) to leverage those that are commonly used in web-based technologies. Specifically, make the BGP protocol RESTful.

In summary the second application of web principes to the network is that new BGP address families use JSON encoding and RESTful primitves. This basic step could be a key enabler to innovation and agility in the network control plane space and be key to empowering applications to make the most of the network.

What Is the Right Network Topology?

In the past, much of the job of a network engineer was to understand where an application would physically plug into the network, how much bandwidth it might consume, and with what end-points this application would communicate. The network engineer would then consider different failure scenarios and, based on this, determine how much capacity was needed in each portion of the network. This task was relatively deterministic in the sense that a network engineer would know with a high level of certainty where an application would exist in the network.

This paradigm does not necessarily apply in an environment where applications can easily be spun up in any corner of the virtualized cloud network. The days of engineering a "custom" network design are behind us. Instead, in the network space engineers and developers need to move to an "instrumentation over engineering" mentality. The general premise in this approach is that no matter how much engineering analysis goes into understanding application traffic patterns as it relates to the physical netwotk topology, the network topology will be inefficient as applications change their requirements, operation, and move around the cloud infrastructure. Therefore, a prudent path to take is to enable much more detailed instrumentation of the network—specifically, provide a much more granular and measured understanding of what each virtualized application is talking to. In a virtualized application environment, one might suggest that a network topology is not inefficient when it runs out of capacity in a portion of the network, but rather that the placement of the application VMs might be less than optimal on the existing network topology.

The shift in thinking proposed here is very much is inline with what the web development community has exceled at—namely analytics.

Analytics

In the web space, analytics are a key part of understanding how well an application is operating and understanding who is using the application. This is done by performing a detailed analysis on the application's logs. In general, each transaction is logged and processed, and relevant information extracted. From this information, the application developers and operation teams can refine the application accordingly.

In the network space, these analytics are performed by mechanisms such as NetFlow. NetFlow provides detailed information about high-level traffic patterns. However, nearly all NetFlow information is sampled, meaning only a very small subset of all packets are actually recorded and processed. Given the performance concerns with attempting to gather too much NetFlow data (and hence overwhelming the routers), sampling one packet per several thousand forwarded is common. In the web space, it would be nearly unheard of to log only one transaction out of several thousand performed.

Given this, per packet analysis is the third application of web principles to the network proposed here.

Third Application: Network Analytics

Recall that the first application of web principles to the network was to drive an open source router instance that can run in a VM. This VM then becomes the interconnect point for virtualization applications. This paper proposes that one of the first features embedded into this router VM software is per packet network analytics.

Specifically, the per packet analytics should track the distance each packet is traversing in the network. This might be grouped into distances such as:

- Traffic that is local to this server
- Traffic that is local to this site
- Traffic that is local to the metro area
- Traffic that is local to the service provider
- Internet traffic

This proposed grouping is depicted in Figure 6.

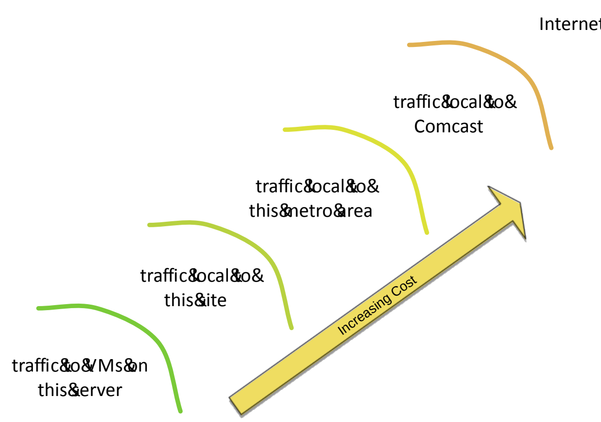


Figure 6: Per packet network analytics

At a minimum, the number of packets and bytes being sent to and from each of these logical domains over narrow time horizons would be tracked. In addition the address family (IPv4 or IPv6), the size of the packets, and possibly other aspects that will help service providers understand the bandwidth-distance usage of this application could be tracked. Why is knowing this information useful? Because rather than attempting to re-architect the network when there is insufficient network capacity—an activity that likely takes weeks and months to complete—service providers may be able to better leverage the existing network resources by re-shuffling the location of the application VMs. By having this per packet level bandwidth-distance usage information over narrow time windows for each application VM, the service provider can derive a very detailed understanding of the impact that moving an application VM will have on the existing network location and to the VM's new network location.

PROGRAM THE NETWORK OR THE PACKET?

The focus on this paper has been applying web principles to the network and correspondingly making the network eco-system more open, agile. and a platform for innovation. Much of the work in the SDN space has been related to "programming the network"; namely, installing network state data from entities outside of the classic router control plane. This clearly enables innovation between the application and network domain, where the application is able to install specific state data into the network. However, when considering router and forwarding technology and evolution, putting additional state data in the router forwarding plane consumes ASIC memory and depending on the amount of state data injected, could place limiting factors on cost effective ways for the router ecosystem to evolve.

An alternative mechanism, called segment routing has been proposed that provides a more flexible way to steer traffic through the network and application space (see segment routing reference). Rather than put state into the network to define a forwarding path for an application, the state instead is inserted into the packet's header (via an IPv6 extension header) by the application or network router for the desired set of application packets. In the segment routing model changing the path that an application's packets takes through the network is done simply by changing the information in the packet extension header, rather than reprogramming state across a number of router devices. Having an open source router edge has the potential to accelerate the deployment of segment routing technology within a service provider or MSO network.

VIRTUALIZATION IS THE ENABLER

The technology that is enabling the approach outlined in this paper is virtualization. This approach takes the virtualization paradigm developed and used extensively by applications and leverage it in the networking space.

Specifically, a proposal to take the edge routing functionality out of the legacy edge router and put it into a VM has been discussed. As part of this migration, the open source ecosystem will be driven to develop production-ready instances of routing code. This open source paradigm then enables innovation and agility in the routing space much like is done in the application space today. Service providers are no longer tied to the development cycles of the existing router vendors and a service provider can innovate on their own technology and timeframes.

The second evolutionary step proposed moving network protocols from their binary format to a format that is more readily extended and where existing tools and software paradigms exist to easily process and develop against. Specifically, we suggest that network protocols, such as BGP Link State, encode information in JSON format and use RESTful primitives.

The third step we propose is to leverage the analytics and Big Data paradigms successfully used in the web space. We specifically propose moving from the "sampled" Netflow paradigm to one where we track per packet bandwidth-distance information. This information then becomes the data that is used as to determine how "inefficient" the application VM's current placement is, and then using this information, VMs can be reshuffled to make better use of the physical network resources.

Is this SDN?

Is the proposed evolution of the network space, as driven by applying web concepts to the network that outlined here SDN? Yes, it is SDN—or minimally it addresses a portion of the SDN problem space. Specifically, the thinking outlined in this paper enables both innovation and agility in the network space by both network operators and application developers, and these are key features in many SDN paradigms.

SUMMARY

In this paper, a number of characteristics that have made the web a success are considered, including an open source paradigm that enables innovation, recasting network protocols to operate using the encoding and primitive mechanisms that are widely used for web applications, and to embrace a paradigm of instrumentation over detailed network engineering via network analytics.

References

Conway, Melvin E., "How Do Committees Invent", *Datamation Magazine*, April, 1968.

Field, Brian, "Exposing network, VM edge and power topology via LLDP and BGP-LinkState – and possible implications", NANOG, June 2013, http://www.nanog.org/sites/default/files/tues.general.field_.topology.32.pdf

Gredler, H., et al, "North-Bound Distribution of Link-State and TE Information using BGP, July 15, 2002, http://tools.ietf.org/html/draft-gredler-idr-ls-distribution-02

Openflow reference: https://www.opennetworking.org/

Segment-routing-reference

https://datatracker.ietf.org/wg/spring/

BEAT IT! HANDLING OBI IN RFOG SYSTEMS

Ayham Al-Banna, Larry Spaete, Amarildo Vieira, Sebnem ZorluOzer, Marcel Schemmann, Tom Cloonan, Dean Stoneback, Venk Mutalik
ARRIS Group, Inc.

Abstract

This article investigates the OBI phenomenon in RFoG systems. Laser behavior leading to OBI events is theoretically described and laboratory experiments to investigate its effect are performed. The results show that RFoG OBI can occur prevailingly and involve more than 90% of the RFoG ONU population. Results also reveal that the preamble of DOCSIS signals is robust to the RFoG OBI problem. The concept of OBI pairs and groups is introduced to offer error-free operation. Finally, potential solutions to the OBI problem are described, which clearly demonstrate that any single-wavelength OBI-free solution for thermally-uncontrolled RFoG system must involve the CMTS support.

1. INTRODUCTION

The RF over Glass (RFoG) technology offers Hybrid Fiber Coaxial (HFC) networks with high capacity potential and extends their life deep into the 2020 decade. MSOs are rapidly rediscovering that RFoG provides a cost-effective means of extending fiber all the way to the customers' premises since it employs the same infrastructure presently used in HFC systems. This technology supports fiber access as a passive optical extension of the same amplitude-modulated (AM) based optical transmission techniques that have been used in conventional HFC architectures.

The evolution of DOCSIS can help RFoG systems thorough the increased capacity offered by newly-created standards such as DOCSIS 3.1. Some of the DOCSIS 3.1 features that help increase the offered capacity include expanded spectrum support, higher order modulations, OFDM, and LDPC.

While RFoG has high potential of offering increased capacities, there are some challenges introduced by the RFoG technology that can limit the overall system performance. One of key challenges that are commonly found in RFoG systems is Optical Beat Interference (OBI). OBI is a signal degradation that occurs when two or more optical transmitters with closely spaced optical wavelengths transmit simultaneously. When two optical sources at wavelengths corresponding to optical frequencies ω_{o1} and ω_{o2} are combined at a detector, a detector output signal is produced around a center frequency (ω_{o1}- ω_{o2}). OBI occurs when the frequency of this noise signal overlaps with other desired RF signals' frequency range, which in turn degrades the signal-to-noise ratio (SNR) of these signals.

This paper is focused on analyzing the RFoG OBI phenomenon and its effect on network performance. This article is organized as follows: section 2 provides a general overview for the RFoG OBI signal. Section 3 introduces the calculation of the Signal-to-Interference Ratio (SIR) for RFoG systems. Wavelength drifting and broadening topics are discussed in sections 4-6. Preamble robustness to RFoG OBI is investigated in section 7. Section 8 analyzes the DOCSIS networks performance in the presence of OBI. Section 9 provides recommendations about the minimum wavelength separation to avoid OBI. OBI avoidance and mitigation is discussed in section 10. Section 11 discusses the interactions between the some of the DOCSIS 3.1 features and RFoG OBI. The article is concluded in section 12.

2. WHAT IS RFOG OBI?

When the fields of two optical transmitters are combined at a non-linear device, such as an optical photodetector, the total optical intensity can be expressed as [1] [2]

$$I(t) = I_1(t) + I_2(t) + I_B(t), \tag{1}$$

$$I_B(t) = 2\sqrt{I_1(t)\cdot I_2(t)}\cdot\cos\left[(\omega_{o1}-\omega_{o2})t + (\phi_1(t)-\phi_2(t))\right]\cdot\cos(\theta_{12}), \tag{2}$$

Where:

- I_1 and I_2 are the intensities of the two optical fields,
- ω_{o1} and ω_{o2} are the frequencies of the two optical signals,
- ϕ_1 and ϕ_2 are the phases of the two optical signals,
- θ_{12} is the polarization angle difference between the two optical signals,
- I_B: beat intensity term due to the square-law nature of the photodetection process.

Observe from equation (2) that the OBI intensity depends on the relative polarization angle of the two beating signals. Parallel polarization is assumed for the analysis of this paper to yield a worst case scenario, where maximum OBI intensity is introduced due to polarization although actual real-world OBI intensities due to polarization are statistically lower. Equation (2) also indicates that the intensity of the OBI signal, which can potentially degrade the quality of US channels, depends on the wavelength (or optical frequency) separation of the beating signals. The optical spectrum of the OBI signal is the convolution of the individual optical field spectra.

Figure 1 shows an illustration of the OBI impact on the upstream RF spectrum assuming an optical field linewidth with Gaussian distribution. The optical beat power spectral density can, depending on the laser linewidth and wavelength spacing, produce low Signal-to-Interference Ratio (SIR) values meaning degraded upstream performance.

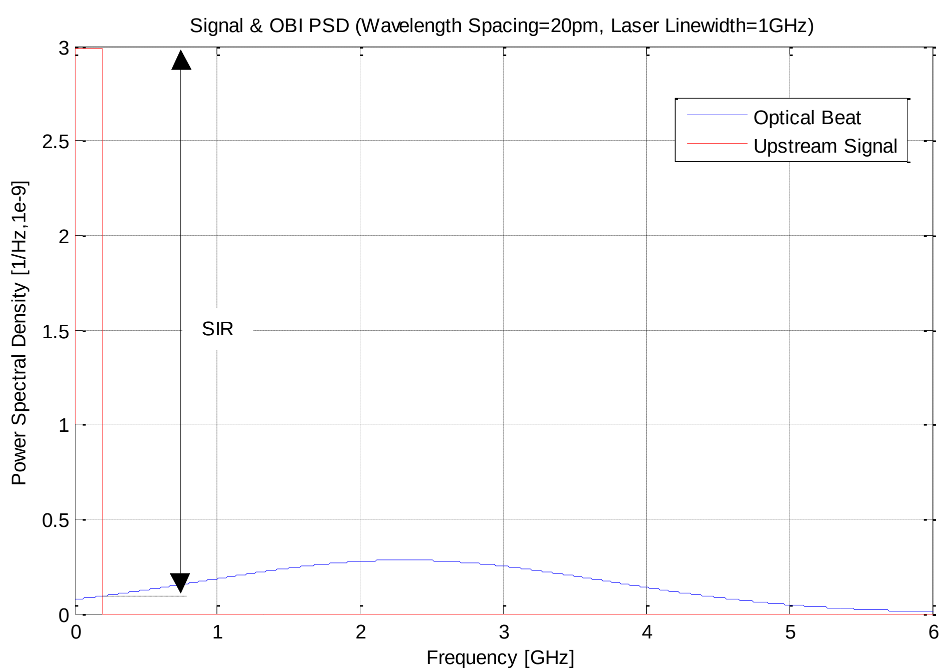


Figure 1. Upstream signal and OBI power spectrum density for 1GHz laser linewidth and 20 pm wavelength spacing

3. SIR CALCULATION IN RFOG SYSTEMS

The calculation of SIR for a desired optical signal in presence of an adjacent optical carrier is introduced in this section. The beat between two optical carriers at the photodetector produces a power spectrum that could be treated as interference to the desired optical signals. The power spectral density of this interference signal depends on the spectrum of the optical carriers and their frequency separation. The resultant SIR is determined by the following equation

$$SIR = 10 \cdot Log\left[\frac{\frac{1}{2}\left\langle m_i(t)^2\right\rangle}{F(\delta f)\cdot B}\right], \qquad (3)$$

Where:

$<m_i(t)>$ represents the detected average signal power (i.e., optical modulation index per channel),

δf is the optical frequency difference of the two sources,

B is the signal bandwidth (after roll-off).

$F(\delta f)$ represents the optical power spectrum of the interference, resulting from the convolution of the two optical fields, as a function of the optical frequency difference (δf). In order to estimate the SIR for the system, the shape of interference optical spectrum, $F(\delta f)$, needs to be specified as well. Therefore, a Gaussian linewidth assumption is made here to represent the modulated laser including the laser chirp during turn on. In this case, $F(\delta f)$ can be expressed, using Gaussian approximation, as

$$F(\delta f) = \frac{13.33}{2\cdot\Delta f\cdot\sqrt{\pi}}\exp\left[-\left(\frac{2\cdot\delta f}{1.2\cdot\Delta f}\right)^2\right], \qquad (4)$$

where Δf is the spectral width of the optical beat spectrum, which is equal to the convolution of the spectra of the two optical sources.

Calculating the SIR value can be further clarified via an example. For instance, consider 17.5% Optical Modulation Index (OMI) per channel, 5.12 MHz signal bandwidth, and 8 GHz laser linewidth operating at 1610 nm wavelength range. Figure 2 shows the SIR calculation as a function of the wavelength spacing. Observe that a minimum wavelength separation of 165 pm would be necessary in order to keep SIR above 44 dB. Note that an optical SIR of 44 dB is large enough to deliver high system SNR values that take the system RF noise into consideration. The delivered SNR values can be as large as 43 dB, which is just enough to support 4K QAM modulation supported by DOCSIS 3.1 [5]. Observe that any wavelength separation less than 165 pm will result in reduced SIR and therefore reduced system SNR values which will not be enough to support 4K QAM modulation. To show the effect of OBI on RF signals, where reduced SNR values are obtained, Fig. 3 is provided below, In particular, Fig. 3A shows high SNR values for two bonded US channels with 60% utilization where two wavelengths, separated nominally by 200 pm, are transmitting simultaneously. On the other hand, Fig. 3B shows lower SNR values due to the OBI resulting from reducing the nominal wavelength separation to 100 pm.

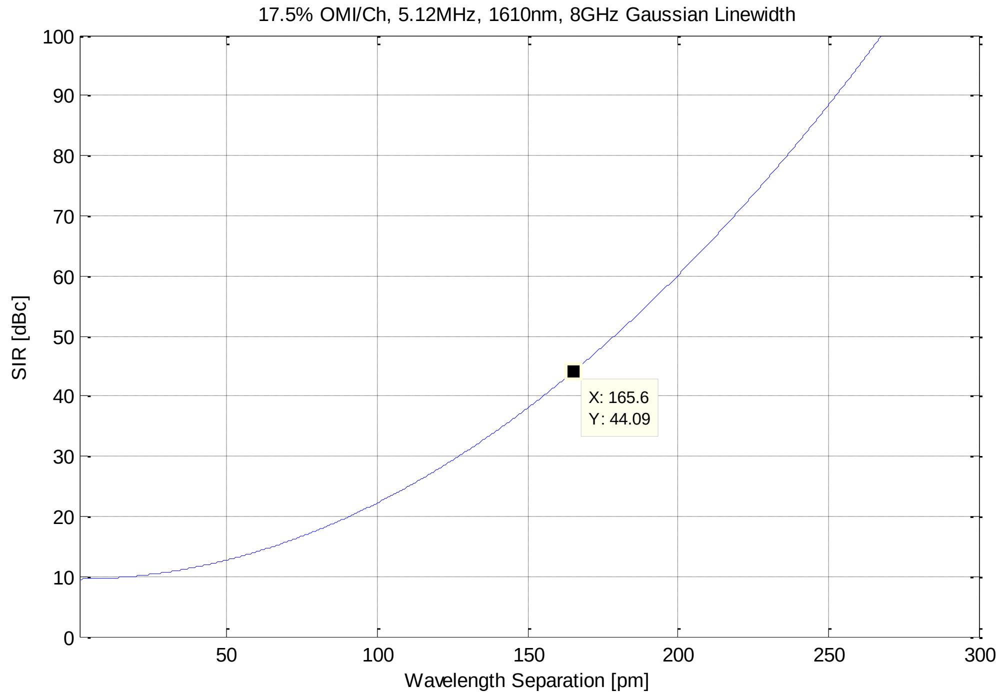


Figure 2. SIR as a function of the wavelength separation for 8GHz Gaussian linewidth

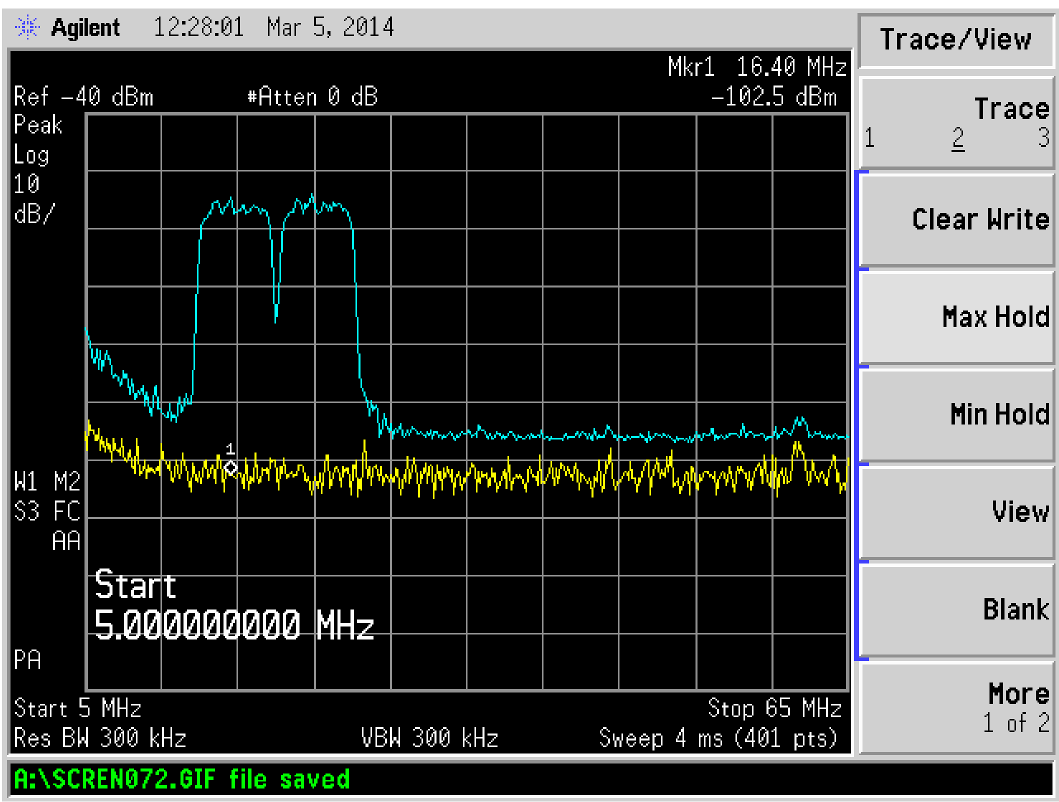


Figure 3A. High SNR values for 2 bonded US channels with wavelength separation of 200 pm

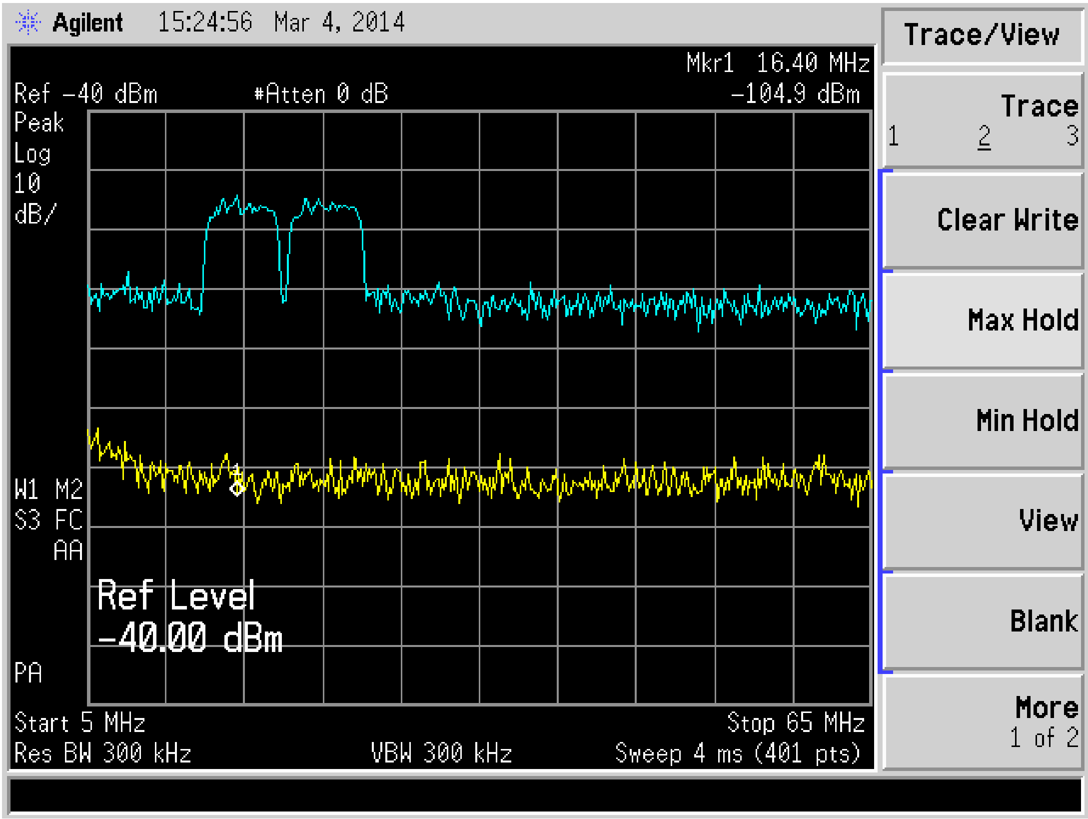


Figure 3B. Low SNR values for 2 bonded US channels with wavelength separation of 100 pm

Note that preliminary RFoG system analyses underestimated the minimum wavelength separation necessary to reduce the SIR impact on the RFoG upstream performance. Numbers between 20pm and 100pm were commonly referred to not too long ago [4], which are less than the actual required wavelength separation of at least 165 pm as shown in Fig. 2. The main issue with previous analyses was not taking in consideration two chirp factors, namely the fast wavelength broadening during the laser turn on and the thermal wavelength drift and modulation-induced wavelength broadening that occurs after the laser turns on. The next section provides in-depth analysis of wavelength drift, broadening, and separation as they are key points for understanding the OBI impact on RFoG systems, which will show that wavelength separation as large as 165 pm may not be sufficient to avoid OBI.

4. FAST WAVELENGTH BROADENING DUE TO LASER TURN-ON

Laser 'turn-on' causing spectral broadening, which is manifested in optical frequency oscillations around the nominal optical frequency, is well known from binary communication systems. The amount of laser chirp or drift in terms of optical frequency variations during the turn-on process can readily exceed the amount of laser chirp that occurs during modulation. It is well recognized that oscillations during the modulation cycle can be reduced if steps are taken to prevent complete turn-off of the lasers during a modulation cycle. This turn-on related laser chirp normally occurs on the sub-nanosecond timescale as photon and electron concentrations in the laser need to balance as laser turns on. However, the situation is different in RFoG systems from such binary systems! Firstly, the relevant time scales of laser turn on/off are much longer (microseconds to milliseconds), secondly, the lasers must be turned off completely in RFoG systems such that the output power in the off state cannot cause OBI with other sources in the system. Whereas the linewidth broadening (i.e., optical frequency variations) due to chirp induced by modulation and laser turn-on is relatively small in RFoG systems, the thermally-induced chirp is large, which can be more than an order of magnitude larger as will be shown later.

5. WAVELENGTH DRIFT DUE TO THERMAL PROCESSES & BROADENING DUE TO MODULATION

Thermally induced chirp is important for processes with slow time constants [3], which are typically longer than 100 nsec. Thermal processes cause wavelength drifting from the nominal wavelength value due to laser heating and cooling during and between transmission bursts. It will be shown later in this section than the laser wavelength drift is fast for the first few microseconds and then slows down after that.

The laser response time to modulation is instantaneous and therefore the laser wavelength oscillates around the drifted wavelength value causing spectral broadening due to modulation. This section discusses the thermally induced wavelength drift and its consequences for data errors caused by OBI.

The ONU in RFoG systems is operated in a burst mode and therefore it is off most of the time on average, which allows the laser to cool during that time. When a burst needs to be transmitted the laser is turned on and heat is dissipated. The semiconductor laser generally has a narrow stripe (width on the order of a few micrometers) that is used to confine light in a single mode so that it can be coupled efficiently to the fiber. Most of the dissipation (series resistance and optical absorption losses) take place in the stripe area. Due to the relevant dimensions of this area and the semiconductor material properties, the thermal response time constant of the laser stripe is in the range of few 100 nanoseconds to microseconds, which is on the order of the length of the preamble of DOCSIS bursts.

The heat generated in the stripe area then needs to be transferred from the laser crystal through solder joints to the heat-sink, which takes place on significantly slower time scales on the order of the duration of the burst transmission and even longer. A quick example is provided below to appreciate the scale of contributions of modulation induced chirp and thermally induced chirp. In this example, assume a typical laser with forward voltage (V_{bias}) of 1.2V that is operated at a current (I_{bias}) of 40 mA, which is mostly dissipated as P_{diss} (48mW) apart from a small amount of produced optical power. Additionally, assume a thermal resistance of R_{th} 90 K/W. This yields a heat-up (dT) of just over 4 Kelvin resulting in a thermal drift dλ of up to 0.35 nm (or 346 pm) as calculated below. Note that the thermal drift calculated here (346 pm) is larger than the SIR-based estimation for wavelength separation performed earlier (165 pm) because the analysis in this section takes the temperature variations into consideration.

$$dT := R_{th} \cdot \frac{P_{diss}}{1000} \qquad dT = 4.32 \quad \text{K}$$

$$d\lambda_{dT} := 0.08 \quad \text{nm per K}$$

$$d\lambda := d\lambda_{dT} \cdot dT \qquad d\lambda = 0.346 \quad \text{nm} \quad . \quad (5)$$

Given an example effective modulation index (μ) of 0.22, a threshold current (I_{th}) of 6 mA, and a laser Frequency Modulation FM response to electrical modulation (FM) of 120

MHz/mA, the modulation induced chirp (or spectral broadening) can be calculated as

$\mu := 0.22$ Effective modulation index

$FM := 120$ MHz per mA modulation response

$$df := (I_{bias} - I_{th})\cdot\mu\cdot FM \quad \text{MHz laser chirp due to modulation} \qquad df\cdot 10^{-3} = 0.898 \quad \text{GHz} \quad . \tag{6}$$

Even for higher chirp lasers, the laser chirp (or modulation-induced spectral broadening) translates to less than 20 pm of width for the modulated laser spectrum. This value is relevant for OBI likelihood of occurrence in the sense that when two lasers are transmitting at wavelengths separated by 20 pm or less, then their modulated spectra partially overlap and severe OBI can occur. Note that the above analysis only applies if no wavelength drifting is present.

However, in the real world, spectral broadening is superimposed on the thermal drift to form the instantaneous laser wavelength value. Since it was shown earlier that the potential thermal drift can be up to 346 pm and the spectral widening (or chirp) is about 20 pm, then the total is about 400 pm. Therefore, it will be safe to assume that wavelength separation of at least 500 pm is necessary to guarantee that OBI does not occur.

6. CHARACTERIZING THE THERMAL LASER WAVELENGTH DRIFT

To investigate the thermal laser wavelength drift, a laboratory experiment was performed, where a 1310 nm Distributed Feedback (DFB) laser was operated at 10% duty cycle to simulate an ONU in a high usage case as shown in Fig. 4. In this experiment, the laser is not producing light during the burst time except for very short amount of time when the wavelength measurements are performed. In particular, the laser was biased just under threshold to induce 6 mW of dissipation during bursts (i.e., not emitting light), while it was off between bursts (0mW). For a small fraction of time within the burst (i.e., <0.5% of burst time, 0.05% overall duty cycle), the laser was putting out light by being biased to just above threshold to induce 13.5 mW of dissipation. Observe that this is insignificant compared to the overall dissipation during the burst.

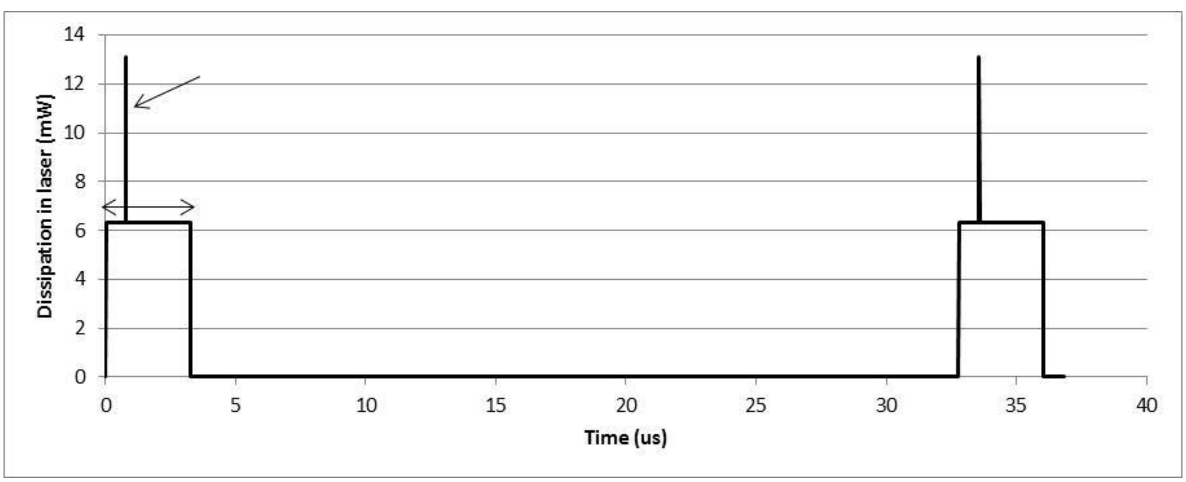


Figure 4. Laser operated with duty cycle of 10% to simulate highly-utilized ONU

Observe that the laser is off most of the time (0 mW), while during the bursts, the laser is turned on just below threshold such that power is dissipated in the device. In this case, the laser is technically not emitting light and therefore optical output power is negligible during this time except for a brief pulse (narrow spike in Fig. 4) that drives the laser just above threshold and therefore emitting light. Wavelength measurements are then performed using a spectrum analyzer.

The wavelength drift due to thermal processes was investigated by changing the location of the pulse throughout the burst duration (horizontal arrow on Fig. 4). In particular, since the wavelength of the DFB laser is linearly related to the cavity temperature, the procedure of shifting the pulse provides a "sample" of the cavity temperature at various times throughout the burst. Figure 5 shows the observed wavelength drift during the burst due to thermal processes, where wavelength measurements occurred over duration of about 200 usec.

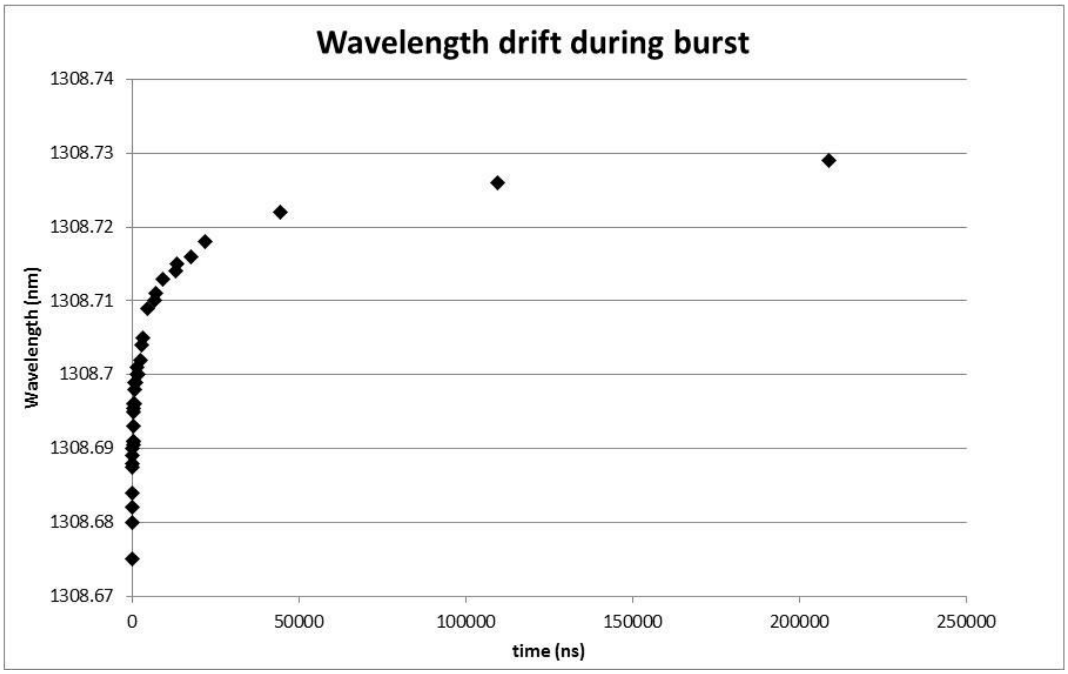


Figure 5. Wavelength drift during bursts

Observe that Fig. 5 captures wavelength drift in the 10 ns-200 usec range, where most of the wavelength drift occurs early in the burst. The curve in Fig. 5 is scaled to reflect operation at 40 mA and plotted on a logarithmic x-axis as shown in Fig. 6.

Again, it can be observed from Fig. 6 that most of the wavelength drift occurs in the first microseconds of the burst time as the laser is heating up. However, slower time constants take over as time progresses and the wavelength continues to drift slowly throughout the burst.

This implies that the probability of an OBI event during the preamble of the transmitting laser is high. However, the duration of such an OBI event is generally short as can be observed from the curve shown in Fig. 7. This curve which shows the duration of an OBI event to cross $\pm$20pm wavelength region at different instant of times during the burst (based on wavelength drift rate at those time instants, which can be calculated from Fig. 6).

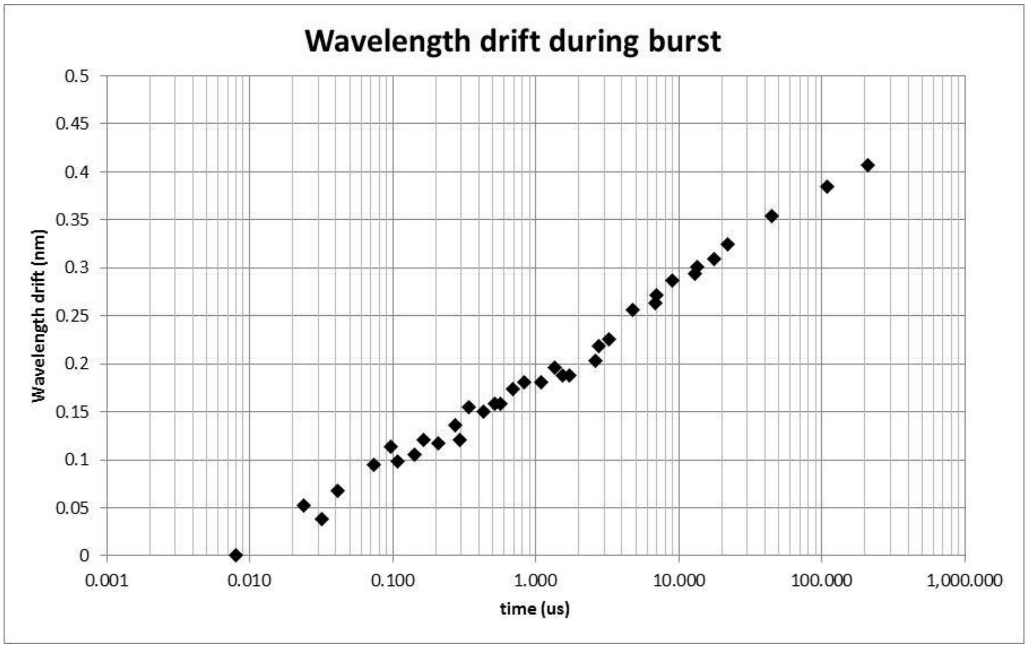


Figure 6. Scaled wavelength drift during bursts (at 40mA bias current) and plotted on logarithmic x-axis

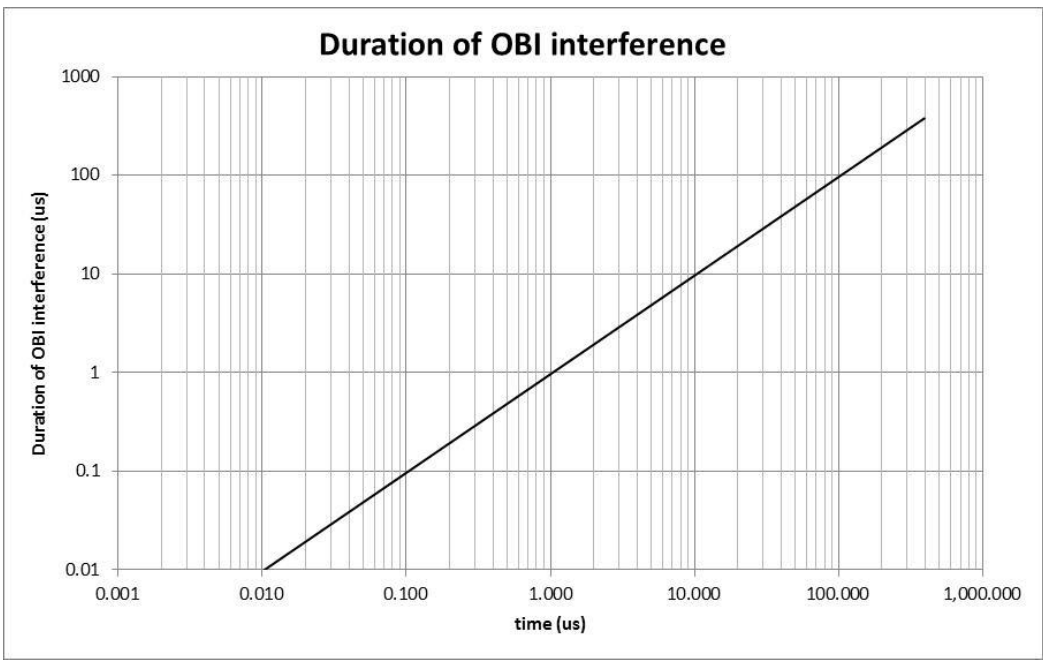


Figure 7. Estimating the duration of OBI events (US interference) throughout the burst time

The next section will describe extensive experiments that were performed to investigate the robustness of the preamble signal to OBI. Those experiments will show that the preamble can survive even in the presence of severe OBI. Observe that for longer time scales into the burst (e.g., 100 usec into a burst), the change rate becomes so slow that OBI events can last more than 100 usec and cause uncorrectable error bursts.

7. ROBUSTNESS OF DOCSIS PREAMBLE TO RFOG OBI

It is important to realize that US packets can be lost due to data and/or preamble corruption. In particular, if US noise or collisions (due to OBI or other reasons) cause significant number of errors in the data portion of the US transmission, then FEC may not be able to recover all of the damaged bits which in turn yields to one or more corrupted codewords (CWs). Lost CWs lead to dropped packets! In the case when excessive noise or collisions corrupt the preamble of an US burst, the receiver may not be able to identify the burst and therefore all data bits carried in that burst will be lost.

The previous sections described the laser wavelength drift that occurs as the laser turns on for transmission. The fast drift in the wavelength occurs over the first few microseconds of the transmission. In general, as the laser heats up for transmission, the wavelength changes by about 200pm over the few microseconds, which will likely cause multiple collisions between the current transmission and other ongoing transmissions. While these collisions can affect ongoing transmissions anywhere in the transmission (i.e., beginning, middle, or end of transmissions), they will affect the current transmission solely during the preamble.

FEC is normally used to protect the data portion of US transmissions from noise. Therefore some errors caused by these short duration collisions can be recovered. On the other hand, the effect of multiple collisions on the preamble is unknown. As a result, an emulation laboratory experiment was designed and performed to investigate the robustness of the preamble of DOCSIS signals to collisions caused by the RFoG OBI phenomenon. This emulation experiment was performed with a cable-only network to control the test environment. The obtained results were validated with an RFoG lab experiment, where real OBI event occurred over fiber links.

In this experiment, the RFoG OBI events were emulated as short-duration wideband impulse noise. The impulse noise was generated with an external signal generator and added to US bursts. Three different impulse noise patterns were generated:

1. Two short-duration impulses (1 usec) with 10 usec in between as shown in Fig. 8.
2. Two short-duration impulses (1 usec) with 30 usec in between as shown in Fig. 9.
3. One medium-duration impulse (9 usec) as shown in Fig. 10.

All of the above patterns have a nominal period of 4 msec. The amplitude(s) of the generated impulse(s) in the above patterns got adjusted such that a specific SNR value (e.g., 0 or -7.5 dB) is obtained during the short impulses. For instance, in the 0 dB SNR case, the impulse noise level was set such that it is equal to the data portion signal level.

The US traffic was transmitted over an ATDMA channel via UGS grants with a nominal period of 4 msec. Very insignificant jitter was allowed for those transmissions. The size of the UGS grant was set to 150B such that two FEC codewords (78B max size) were used to transmit a single data packet (100B each). 10,000 packets were transmitted for each experiment run. The US channel width was configured to be 6.4 MHz which yields 5.12 Msps. The preamble length was 104 bits (or 52 QPSK symbols), which in turn occupies about 10 usec. Different preamble power values (i.e., QPSK0, QPSK1) were experimented to fully characterize the preamble robustness. Observe that a QPSK0 preamble will have the same level as the data portion of the signal. On the other hand, the power of the QPSK1 preamble is 3 dB higher than the signal power.

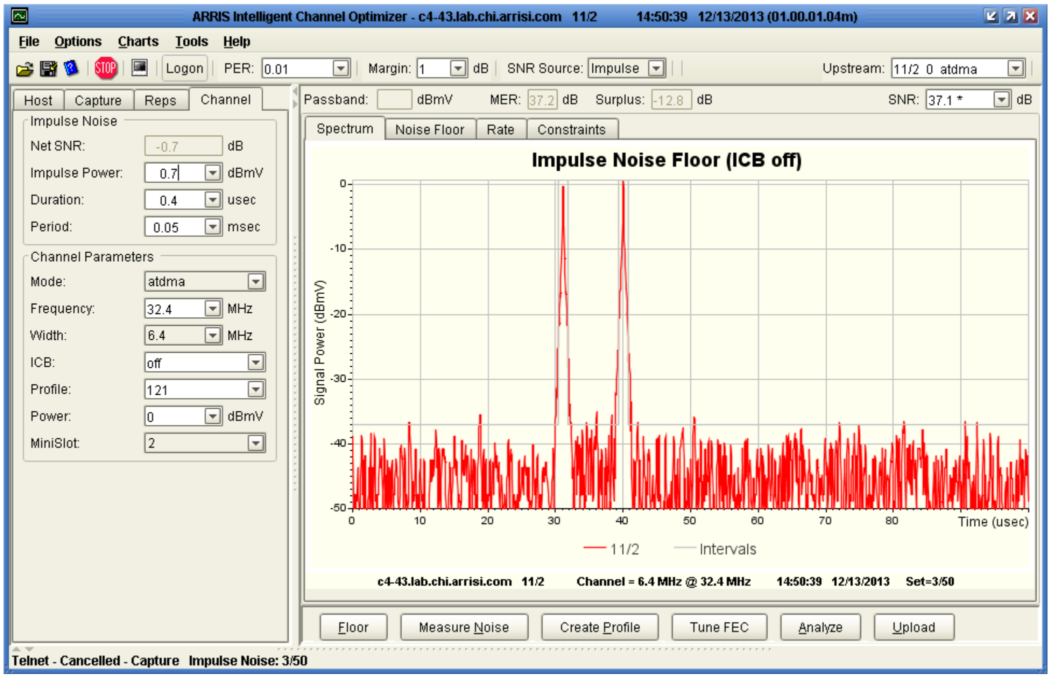


Figure 8. RFoG OBI noise pattern 1: Two short-duration impulses (1 usec) with 10 usec in between. Impulse noise level is 7.5 dBmV higher than the signal level

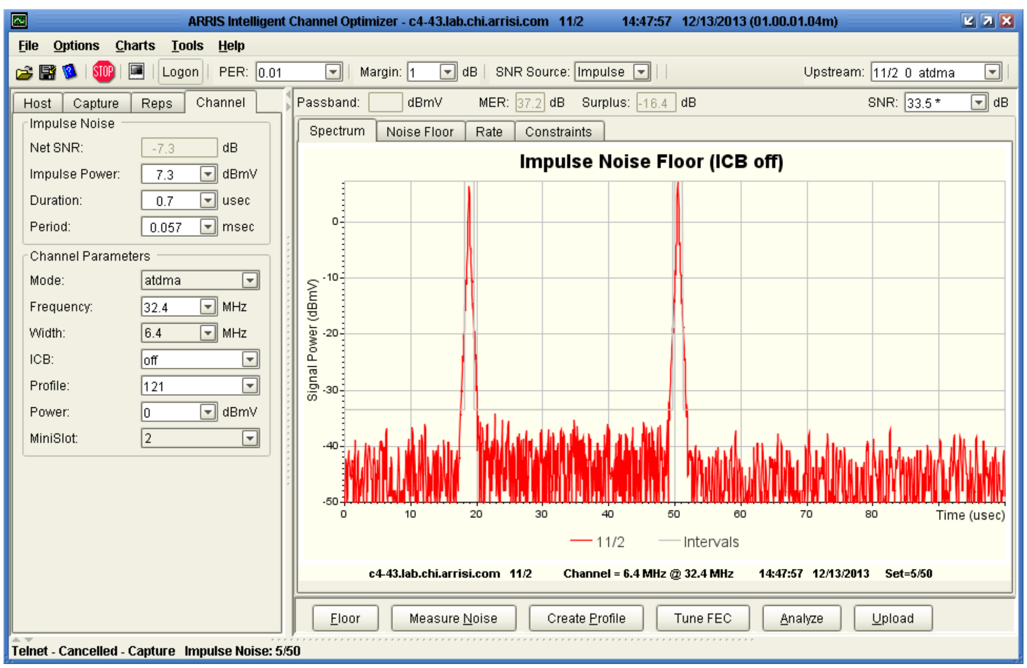


Figure 9. RFoG OBI noise pattern 2: Two short-duration impulses (1 usec) with 30 usec in between. Impulse noise level is 7.5 dBmV higher than the signal level

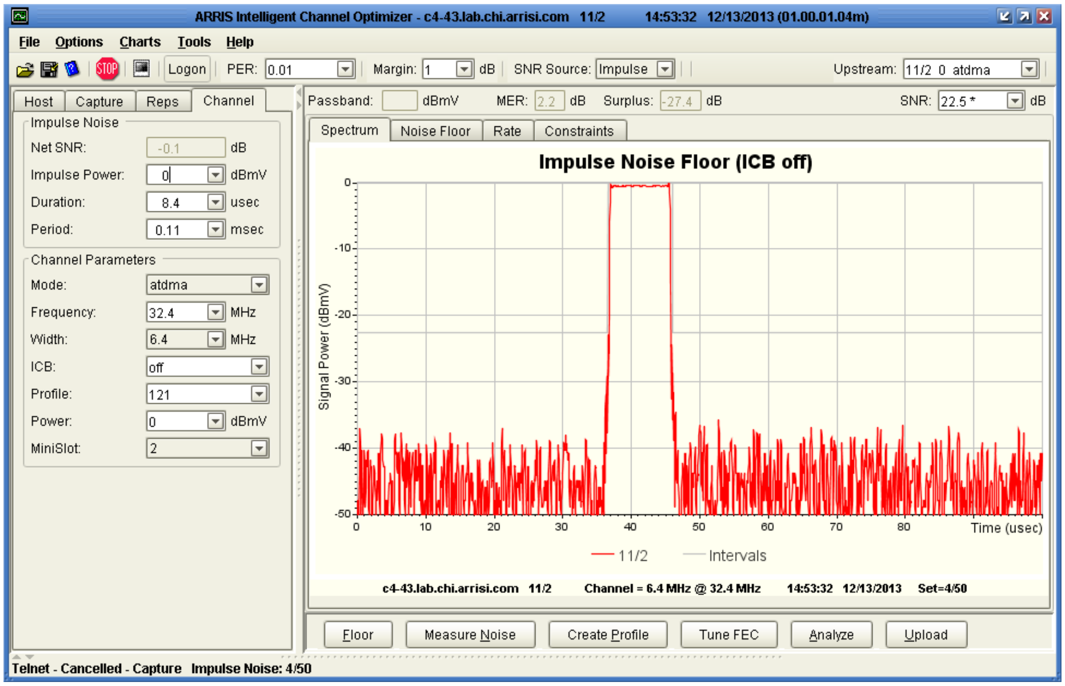


Figure 10. RFoG OBI noise pattern 3: One medium-duration impulse (9 usec). Impulse noise level is 0 dBmV (equal to the signal level)

Since the preamble in the above UGS grant is about 10usec long, the third noise pattern above (impulse width of 9usec) was considered to be the worst case scenario where the majority of the preamble is damaged by an RFoG OBI event (if synchronized). This is not likely to happen in the real world, where OBI events tend to be short in duration although more than one OBI event can affect the preamble. Therefore, noise patterns 1 and 2 are more likely to emulate real-world OBI events.

While impulse noise and UGS data bursts have the same nominal period of 4 msec, the period of the impulse noise was set to a value that is slightly different than the period of the UGS bursts. The slight asynchronization between impulse noise and UGS grants resulted in a situation where the impulse appear to slide across the UGS grants on an Oscilloscope screen as shown in Figs. 11A and 11B. The sliding event on the oscilloscope screen is actually an illusion caused by the fast capture of multiple impulses damaging different (and sequential) parts of different bursts. At any instant of time, the impulse would either corrupt the preamble or the data or does not collide with the UGS signal at all. Running the traffic and noise for a sufficient period of time will yield to a situation, where impulse is corrupting different portions of the preamble of the UGS grants.

For each experiment run, the number of lost packets due to preamble damage (lost bursts) or data damage (unrecovered FEC CWs) was recorded. A summary of the results is shown in Table 1. Observe that the packet loss measurements in the table show that the preamble of DOCSIS signals is very robust to RFoG OBI events. In particular, no packets were lost due to preamble damage even with the worst case OBI noise (pattern 3)! On the other hand, some packets were lost due to OBI events corrupting data FEC codewords. The preamble robustness is also demonstrated by an RFoG lab experiment were real-world OBI events occurred but could not cause packet loss due to preamble damage.

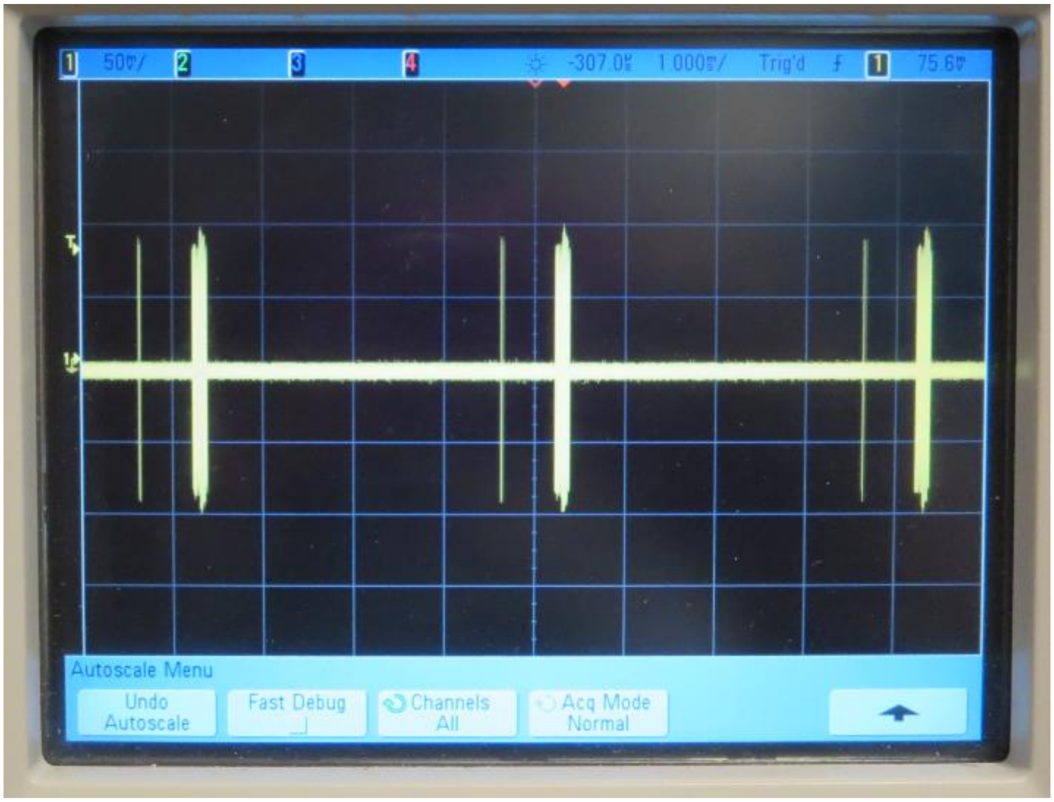


Figure 11A. Multiple impulses getting close to corrupt multiple UGS grants


Figure 11B. Multiple impulses are corrupting multiple UGS grants. Note the signal & noise addition.

Table 1. DOCSIS UGS packet loss due to preamble or data damage caused by RFoG OBI

Test ID	UGS Preamble Type	UGS Modulation Order	RFoG OBI Pattern	SNR (dB)	Corrected FEC CWs	Packets Loss due to FEC	Packets Loss due to Preamble
1	QPSK0	QPSK	1	0	418	0	0
2	QPSK0	QPSK	1	-7.5	634	0	0
3	QPSK0	QPSK	2	0	411	0	0
4	QPSK0	QPSK	2	-7.5	675	0	0
5	QPSK0	QAM-64	1	0	283	0	0
6	QPSK0	QAM-64	1	-7.5	248	28	0
7	QPSK0	QAM-64	2	0	430	2	0
8	QPSK0	QAM-64	2	-7.5	377	18	0
9	QPSK1	QPSK	1	0	409	0	0
10	QPSK1	QPSK	1	-7.5	638	0	0
11	QPSK1	QPSK	2	0	434	0	0
12	QPSK1	QPSK	2	-7.5	657	0	0
13	QPSK1	QAM-64	1	0	302	3	0
14	QPSK1	QAM-64	1	-7.5	246	30	0
15	QPSK1	QAM-64	2	0	396	0	0
16	QPSK1	QAM-64	2	-7.5	377	16	0
17	QPSK0	QPSK	3	0	633	40	0
18	QPSK0	QAM-64	3	0	19	135	0
19	QPSK1	QPSK	3	0	613	50	0
20	QPSK1	QAM-64	3	0	15	136	0

8. PEFORMANCE OF DOCSIS NETWORKS IN RFOG ARCHTECTURE

Earlier sections in this paper provided an overview of the wavelength drifts and broadening and also described the OBI problem. An empirical analysis to show the preamble robustness in the presence of simulated OBI event in a cable-only network setup was also provided. This section introduces a set of laboratory experiments over a real-world setup of fiber network to characterize the OBI phenomenon and its occurrence likelihood in a population of RFoG ONUs. These experiments are also designed to validate the theoretical analyses that were provided in previous sections to estimate the wavelength separation needed between two optical transmitters such that OBI does to not occur. The experiments described below will also verify the conclusions that were made regarding robustness of the DOCSIS preamble based on the cable-only emulation experiment described in the previous section. The performance of DOCSIS networks in the RFoG architecture is determined via these experiments. Finally, some of the OBI mitigation schemes are verified to operate and offer virtually error-free performance in RFoG networks.

This experiment setup was composed of a CMTS, 32 CMs, 32 RFoG ONUs, optical headend transmitter/receiver, and 20 km of fiber links connected as shown in Fig. 12. The nominal DS wavelength was 1550 nm and the nominal US wavelength was 1610 nm.

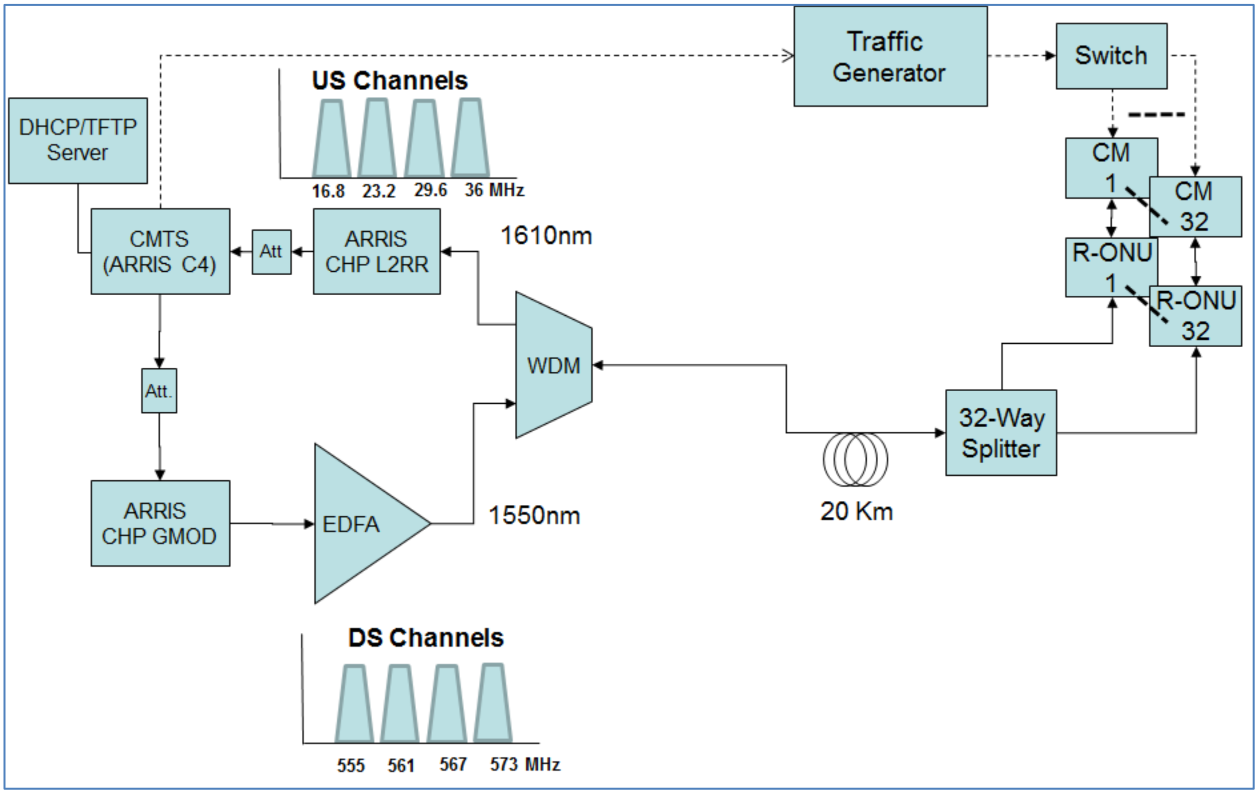


Figure 12. RFoG architecture experiment setup

The CMTS was setup with 4 6.4 MHz US channels operated at QAM64 modulation (T= 12B, CW_size = 81B for short grants, T= 16B, CW_size = 223B for long grants). The first step that was performed is to 'manually' measure the wavelengths of different ONUs using an optical wavelength meter, where the wavelength of the ONU is measured after transmitting a CW tone through the ONU. Given that the wavelength varies over time, the absolute wavelength value will represent a snapshot at a time instant and therefore may not be very definite or descriptive. However, taking wavelength measurements for all ONUs in the same fashion can provide relative wavelength relationships between the different ONUs. Two sample measurements were taken for each ONU: 1) right after the laser is turned on, 2) after 30 seconds. The collected ONUs wavelength measurements are then sorted in ascending order based on the 30 sec reading as shown in Table 2. It is commonly known that less than 20pm inter-wavelength separation is needed for OBI to occur. Therefore, this rule was used to identify the OBI pairs in the ONU population. Later in this section, our analysis will show that it will take much more than 20pm of wavelength separation for OBI not to occur.

Table 2. Wavelength measurements for ONU population

ONU ID	λ 1st value (nm)	λ 30 Sec value (nm)	Wavelength difference based on 1st reading	Potential OBI pairs based on 1st reading	Wavelength difference based on 30 sec reading	Potential OBI pairs based on 30 sec reading
15	1611.041	1611.075				
31	1611.108	1611.142	0.067		0.067	
6	1611.118	1611.157	0.01	6,31	0.015	6,31
18	1611.174	1611.204	0.056		0.047	
12	1611.225	1611.269	0.051		0.065	
13	1611.372	1611.395	0.147		0.126	
4	1611.382	1611.426	0.01	4,13	0.031	4,13
23	1611.44	1611.476	0.058		0.05	
19	1611.532	1611.563	0.092		0.087	
30	1611.573	1611.606	0.041		0.043	
27	1611.655	1611.694	0.082		0.088	
3	1611.675	1611.709	0.02	3,27	0.015	3,27
14	1611.881	1611.925	0.206		0.216	
28	1611.914	1611.949	0.033		0.024	
32	1611.957	1611.99	0.043		0.041	
10	1612.022	1612.052	0.065		0.062	
2	1612.037	1612.089	0.015	2,10	0.037	
24	1612.189	1612.225	0.152		0.136	
29	1612.199	1612.251	0.01	29,24	0.026	
21	1612.305	1612.351	0.106		0.1	
5	1612.354	1612.386	0.049		0.035	
11	1612.408	1612.432	0.054		0.046	
25	1612.536	1612.57	0.128		0.138	
9	1612.732	1612.671	0.196		0.101	
26	1612.732	1612.763	0	9,26	0.092	
7	1613.073	1613.107	0.341		0.344	
8	1613.107	1613.149	0.034		0.042	
17	1613.171	1613.199	0.064		0.05	
20	1613.293	1613.33	0.122		0.131	
1	1613.403	1613.43	0.11		0.1	
16	1613.857	1613.888	0.454		0.458	
22	1614.175	1614.203	0.318		0.315	

The next step was running some baseline experiments. The first experiment had all CMs placed on a single US channel and ran US traffic through all CMs and obtained an error free operation. Afterwards, a second US channel was enabled and all modems were configured to be able to bond across both US channels. US traffic source was configured to achieve about 70% US channel utilization with 218 B packets. To achieve 70%

utilization, each CM was bursting at 1.14Mbps (assuming 26Mbps US channel capacity after DOCSIS overhead is removed). US traffic was run for 5 minutes and the traffic source counts as well as the CMTS FEC counts were monitored after the transmission of the US traffic had completed. The error counts showed that all CMs experienced FEC errors and most CMs had dropped more than 1% of the packets. The results showed packet loss values in the range of 0.01% - 4.99% with uncorrectable FEC codeword rates in the range of 0.01% - 4.19%. Another baseline experiment was performed with 4 bonded US channels with 70% US channel utilization. More OBI occurred between CMs and the results showed that packet loss rates increased to the 8.38% - 25.11% range. In this experiment, the uncorrected FEC codeword rates also increased and fell in the 7.18% - 22.59% range. It is obvious that increasing the number of bonded US channels increases the likelihood of OBI occurrence.

Further experiments were performed, where US traffic is run through CMs and CMTS FEC counts were monitored to identify the exact OBI pairs in the ONU population. The summary of the pairs is provided in Table 3. Note that the data for OBI partners in Table 3 is based on strong OBI relationship between partners. That is, simultaneous transmissions from partners yield significant OBI that causes large amount of FEC errors. Minor (less damaging) OBI can result from simultaneous transmissions from some non-partners. Minor OBI partners can also be included but these minor OBI partners will lead to relationships that tie almost all CMs to each other and it will be difficult to cleanly separate the population into solid partners.

Observe that any attempt to relate the above relationships (partners) to the wavelength measured earlier (in Table 2) using the 20pm rule, results in missing many of the pairs. Therefore, the 20pm rule does not offer high potential for identification of the OBI candidates. Some of the partners in Table 3 are common for different ONUs and therefore the next step was to sort the partners into groups such that simultaneous transmissions of members that belong to the same group will result in damaging OBI. The groups are shown in Table 4 and these groups are independent (ignoring minor OBI relationships), which means that simultaneous transmissions from members of different groups will not produce damaging OBI. Note that Group # 5 represents the group of members that have no OBI partners.

Based on the above grouping, an analysis was performed on the wavelength measurements in Table 2 to relate groups to wavelengths as shown in Table 5.

Note that, in theory, the members of each of the groups above can OBI with each other while members of different groups do not OBI with each other. Therefore, groups could be intelligently assigned to US channels. In this experiment were only 2 US channels are enabled, error-free operation was achieved when groups 0, 1, 3 were assigned to US0 (total of 16 CMs) and groups 2, 4, 5 were assigned to US1 (total of 16 CMs). Observe that members of group 5 (non-OBI candidates) could have been assigned to bond across both channels. Two different 5-minute US traffic runs (37% and 88% channel utilization) were performed and both resulted in zero FEC errors for all CMs.

Table 3. OBI partners for to each ONU

ONU ID	OBI Partner(s)
1	20
2	32, 10
3	27, 30
4	12, 13, 6, 18, 23
5	21, 11
6	15, 18, 31
7	8
8	7, 17
9	26
10	2, 32
11	5, 25
12	4, 6, 18
13	4, 12
14	28
15	6, 31
16	No partner
17	8, 20
18	6, 31, 12, 4
19	27, 30, 23
20	1, 17
21	5, 29
22	No partner
23	4, 30, 19
24	29
25	11
26	9
27	3, 19, 30
28	14, 32
29	21, 24
30	3, 27, 19, 23
31	6, 15, 18
32	2, 10

Table 4. OBI partners in Table 3 are sorted into groups

Group ID	Group Members	Number of Members
0	1, 7, 8, 17, 20	5
1	2, 10, 14, 28, 32	5
2	3, 4, 6, 12, 13, 15, 18, 19, 23, 27, 30, 31	12
3	5, 11, 21, 24, 25, 29	6
4	9, 26	2
5	16, 22	2

Table 5. Relationship between wavelength measurements & groups (based on Table 2)

ONU ID	λ 30 Sec value (nm)	Wavelength difference based on 30 sec reading	Potential OBI pairs based on 30 sec reading	Group ID
15	1611.075			2
31	1611.142	0.067		
6	1611.157	0.015	6,31	
18	1611.204	0.047		
12	1611.269	0.065		
13	1611.395	0.126		
4	1611.426	0.031	4,13	
23	1611.476	0.05		
19	1611.563	0.087		
30	1611.606	0.043		
27	1611.694	0.088		
3	1611.709	0.015	3,27	
14	1611.925	0.216		1
28	1611.949	0.024		
32	1611.99	0.041		
10	1612.052	0.062		
2	1612.089	0.037		
24	1612.225	0.136		3
29	1612.251	0.026		
21	1612.351	0.1		
5	1612.386	0.035		
11	1612.432	0.046		
25	1612.57	0.138		
9	1612.671	0.101		4
26	1612.763	0.092		
7	1613.107	0.344		0
8	1613.149	0.042		
17	1613.199	0.05		
20	1613.33	0.131		
1	1613.43	0.1		
16	1613.888	0.458		5
22	1614.203	0.315		

The above approach was verified with 4 US channels, where groups 0, 4 were assigned to US0, groups 1 &5 were assigned to US1, group 2 was assigned to US2, group 3 was assigned to US3 and ran with 70% channel utilization. The results after 5 minutes run were virtually error free (max of $1.27x10^{-5}$ or 0.001% of uncorrectable FEC).

Another experiment was performed where UGS flows were setup to test the preamble robustness to verify the results obtained by the cable-only emulation experiment. UGS flows were transmitted via all CMs which were bonded on 2 US channels. OBI incurred caused some loss in DATA packets but no bursts were lost due to preamble damage.

Observe that the above experiments were performed with a single packet size of 218B based on packet size statistics that were collected recently. To obtain a controlled environment, the experiments were run with fixed packet rate across all CMs to obtain certain US channel utilization. In reality, various services (e.g., voice, data) are offered over an RFoG architecture, where these services have with different on-off times and packet distributions. These characteristics can affect the way the US utilization and will also impact the likelihood of OBI occurrence. Additionally, various applications have different packet loss, delay, and jitter requirements.

In conclusion for this experiment, it can be noted that from the above experiments that OBI is a prevalent problem that can affect most of the populations. In particular, 30 out of 32 ONUs had OBI partners. While the preamble of DOCSIS signals is robust to RFoG OBI, OBI can cause data damage and corresponding packet loss. Once the OBI pairs or partners were identified, the OBI partner groups were intelligently assigned to US channels to offer virtually error-free operations with multiple channels, where members of each OBI partner group are confined to a single US channels. The CMTS scheduler could be augmented to not limit a group of partners to a single group. In particular, the CMTS can offer an error-free operation in an RFoG environment as long as it scheduled simultaneous transmissions from members of different groups. In a nut shell, CMTS configuration and CMTS scheduler augmentation must be involved in any single nominal wavelength solution to offer an error-free operation in a thermally-uncontrolled RFoG environment.

9. WAVELENGTH SEPARATION NEEDED FOR OBI AVOIDANCE

Earlier in section 5 of this paper, theoretical analyses showed that a minimum requirement of 500 pm of wavelength separation is necessary for OBI not to occur at all. Additionally, analysis is performed in this section to show that wavelength separation values less than 200 pm will result in significant OBI occurrence leading to large amounts of packet loss. In particular, more experiments based on the setup shown in Fig. 12 were performed and a Monte Carlo MATLAB simulation was implemented based on the experimental 30-sec wavelength measurements provided earlier in Table 2. Those wavelength measurements were assigned to the simulated ONUs to estimate the amount of OBI that occurs in the system. In practice, wavelengths continuously change due to drift and broadening as described earlier, therefore, the simulation can be viewed as a snap shot of time with fixed wavelength values.

The MATLAB model is a Monte Carlo tool used to estimate the impact of OBI on data loss based on the wavelength spacing among the ONUs, when two or more ONUs are transmitting at the same time. No FEC correction is assumed in the model. When an OBI event occurs between two ONUs, it is assumed that all channels are corrupted and therefore data from all transmitting ONUs is lost. For this particular simulation, the measured wavelengths from Table 2 were uploaded to the program and the 4 upstream bonded channels were assumed with utilization of 60% to match an experiment that was performed using the setup in Fig. 12. The 60% channel utilization is equivalent to 15.6Mbps, assuming 26Mbps channel capacity. After running the simulation for a long number of data grants (100,000), the percentage of data grants lost is then calculated for each upstream channel. The simulations were run for multiple values of the minimum wavelength spacing needed to avoid OBI (e.g., 145, 165, 185 pm). In particular, the simulation uses the minimum wavelength spacing criteria in such a way that it assumes that OBI affects data grants if two wavelengths are separated by less than minimum wavelength specified in the simulation run (e.g., 145, 165, or 185pm). On the other hand, if two wavelengths are separated by more than the specified minimum wavelength separation, then no OBI is assumed. The results of the simulation are provided in Fig. 13, which show close agreement with the experimental results.

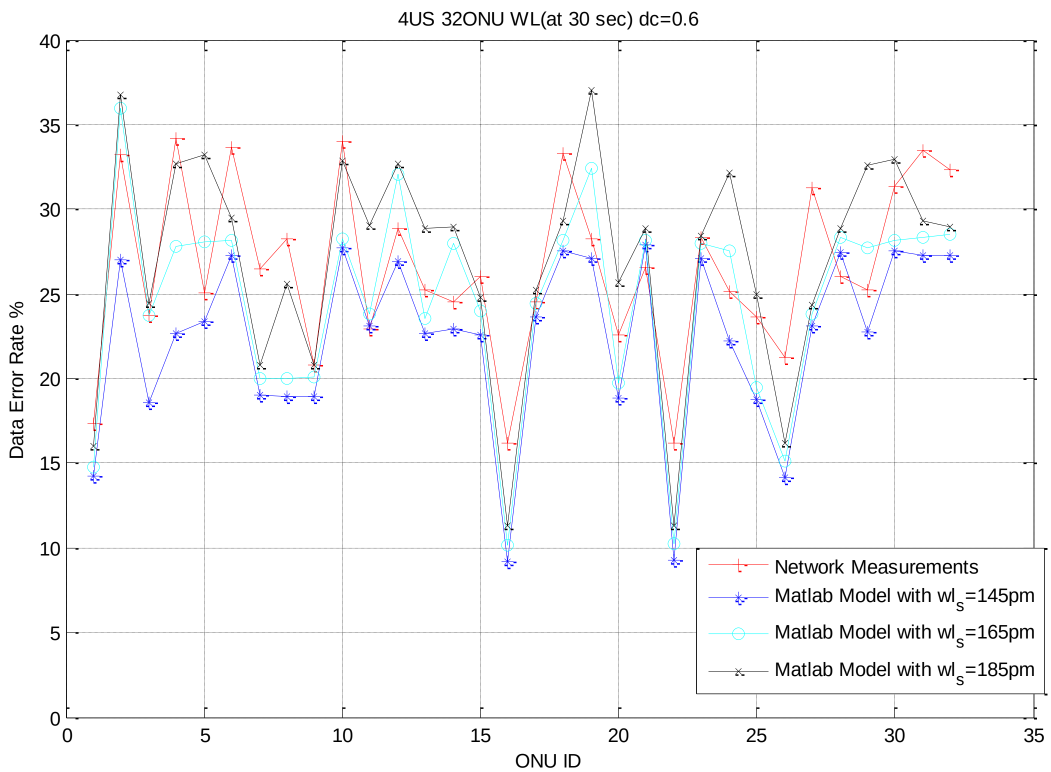


Figure 13. Data error rate measurements & simulations for a 32 RFoG ONU system

Observe that one of the specified wavelength separation values was 165 pm, which is based on a results introduced earlier in Fig. 2. Experimental results in the previous section and simulation results shown in Fig. 13 showed significant packet loss events due to the fact that ideal theoretical 165 pm value is less than the recommended 200 pm minimum separation provided earlier in this section. Moreover, the simulation environment assumed fixed wavelength values although they drift in real world.

Given that OBI can be avoided with wavelength separation values greater than 500 pm and the fact that wavelength separation values less than 200 pm will result in significant packet loss, 200 pm-500 pm wavelength separation values will likely perform well as was shown in section 8, where some level of rare OBI events may occur due to different factors such as temperature and equipment aging.

10. RFOG OBI AVOIDANCE & MITIGATION

There are multiple approaches that can be taken to minimize the effect of OBI in RFoG systems. One approach is to use different wavelengths for different ONUs. While this avoids or eliminates the OBI problem overall, it is not commonly deployed and has significant cost and operations complexities. Additionally, in order to ensure that there is no OBI wavelengths must be separated from each other by at least 500 pm at all times. Since ONUs operate over a wide temperature range and the wavelength varies over temperature, it is difficult to control the wavelengths precisely enough to maintain the required separation while still fitting multiple wavelength channels into the 1610 nm window. Other methods to eliminate OBI is to allow single US transmitter at a time by either placing all ONUs on a single ATDMA US channel or scheduling a single transmitter at a time across all US channels.

Scheduler-based techniques can be developed to mitigate the effect of OBI. In particular, an OBI group/pair-aware scheduler can be designed to schedule US transmitters simultaneously such that OBI is not likely to occur based on the knowledge of the OBI groups/pairs. This will offer the most efficient and error-free network performance.

11. INTERACTIONS BETWEEN DOCSIS 3.1 & RFOG OBI

The introduction of DOCSIS 3.1 can extend the life of RFoG networks. This is enabled by the stretched DS & US spectra that DOCSIS3.1 devices can support. Additionally, the noise-robust OFDM PHY along with the efficient LDPC FEC technologies can support higher modulation orders to offer significant capacity increase compared to earlier DOCSIS technologies.

DOCSIS 3.1 also offers larger US channel bandwidths. In particular, US channels can be as large as 96 MHz. This can be viewed as an advantage or disadvantage depending on the OBI avoidance or mitigation scheme supported by the CMTS. For instance, if the CMTS is avoiding OBI by scheduling a single transmitter at a time across all US channels, then wide channel width will result in significant wasted capacity. When the CMTS avoids OBI by placing all CMs of an OBI group (or all CM population) on a single US channel, larger DOCSIS 3.1 channel width can be viewed as removing the capacity limitations that are introduced, when the transmitter is allowed to send traffic on a single US channel. On the other hand, if the CMTS mitigates the effect of OBI via an OBI pair/group-aware scheduler-based approach, wider channels may introduce more variables to the CMTS' scheduler as the number of CMs that can transmit at any single moment of time rises as the channel width is increased. Note that DOCSIS 3.1 can also support smaller channel widths, which can help reduce the number of potential transmitters at any instant of time.

12. CONCLUSIONS

This article described wavelength drift and broadening and their relation to the RFoG OBI phenomenon. Experiments and analyses showed that complete OBI-free operation requires greater than 500 pm of wavelength separation but an acceptable network performance can be obtained with at least 200 pm of wavelength separation, which is different from the commonly accepted 20 pm recommendation for OBI-free operation. Results showed that the preamble of DOCSIS signals is robust against RFoG OBI such that the system can tolerate short transient OBI events due to fast wavelength drift. It was also shown that the OBI likelihood increases with the number of US channels. Nevertheless, solutions to mitigate and avoid RFoG OBI can be designed to offer error-free performance, where some of these solutions were demonstrated with laboratory experiments. The effect of some of the DOCSIS 3.1 features on RFoG OBI was discussed. The analysis in the paper showed that the CMTS support is critical, when a single nominal wavelength is used for US transmissions, in providing OBI-free performance in thermally-uncontrolled RFoG systems.

REFERENCES

[1] C. Desem, "Optical Interference in Subcarrier Multiplexed Systems with Multiple Optical carriers", IEEE J on Selected Areas in Communications, Vol. 8, n. 7, Pp. 1290 – 1295, September 1990.

[2] T. H. Wolf and N. K. Shankaranarayanan, "Operation of a Passive Optical Network with Subcarrier Multiplexing in the Presence of Optical Beat Interference", IEEE JLT, Vol. 11, N. 10, Pp. 1632-1640.

[3] H. Shalom et. al., "On the Various Time Constants of Wavelength Changes of a DFB Laser Under Direct Modulation", IEEE Journal of Quantum Electronics, Vol. 34, N. 10, October 1998, Pp. 1816-1822.

[4] Calix Access Innovation, "RFoG – The advantages of an 'Intelligent' Micronode," 2009 Fiber to the Home Council Conference and Expo.

[5] Cablelabs, "Data-Over-Cable Service Interface Specifications - DOCSIS® 3.1, Physical Layer Specification, CM-SP-PHYv3.1-I02-140320," http://www.cablelabs.com/wp-content/uploads/specdocs/CM-SP-PHYv3.1-I02-140322.pdf, March 20, 2014.

Big Data – Web Originated Technology meets Television

Bhavan Gandhi and Sanjeev Mishra
ARRIS

Abstract

The development of Big Data technologies was a result of the need to harness the ever-growing zettabytes of data created and consumed on the Internet. Web content, social networking, and user generated content in the form of blogs, images, and videos are just a few examples of the nature of unstructured data that don't fit into the traditional relational database models (RDBMS). In a similar manner, the television eco-system is evolving beyond just the delivery of linear video content to the television. Video experiences are evolving into more complex systems that support delivery of linear and on-demand content on multiple portable devices, capable of local and network DVR, and even supporting targeted content recommendation and advertisement placement. Harnessing Big Data technologies from the Internet world and bringing it into the TV space allows for deeper understanding of user behavior and system performance. This paper provides an overview of Big Data Technologies, and it gives examples an architectural overview of how these technologies can be used in an operator's eco-system.

INTRODUCTION – WHAT IS BIG DATA?

Web Originated

Terms typically used to characterize Big Data are volume, velocity, and variety [1, 2, 3]. Volume refers to the sheer amount of data that currently exists and is being created continuously. Velocity is the speed with which this data can be processed and analyzed for a meaningful use. Variety is simply the different types of data that can be collected. From the perspective of the Internet, zettabytes (2^{70} bytes) of data are created and stored in the form of images, documents, tweets, videos, maps, social interactions, etc. This data has to be analyzed and classified with speed so that end users can find and retrieve relevant information to accomplish their tasks. Much of this data is either unstructured, or in cases where there is inherent structure, the variety of the data leads to the need to support multiple structures.

The advent of the Internet and the volume of data that needs to be managed, stored, and analyzed led to the emergence of modern day Big Data technologies. Yahoo and Google have developed breakthrough technology on processing Internet scale disparate data and storing the processed results on distributed commodity servers to enable high availability and scalability. The Google File System (GFS) [4] and Big Table [5] showcase early pioneering work that was developed for storing data at Google. This is the pre-cursor to some of the mainstream, open source big data related technologies that are being developed and evolved by the open source community like Apache. Apache HBase [6] and Apache Hadoop Distributed File System (HDFS) [7] are open source equivalents of Big Table and GFS respectively.

Given that companies like Google, Yahoo, and Facebook are storing petabytes (2^{50} bytes) of data, there is a need to access this vast repository of data to process the information and derive meaning quickly. MapReduce is a framework that works on top of HDFS and HBase; it allows for parallel access, incremental, and distributed processing of huge amounts of data to derive useful meaning that can be used by subsequent services or applications. Hive and Pig are higher-level languages that allow for programming or scripting of MapReduce jobs. Researchers at Facebook developed Hive to allow SQL-like queries against their Big Data sets [8, 9]. Pig was developed by Yahoo staff to reduce the complexity of programming MapReduce jobs.

Big Data for Television

The television eco-system is migrating from dedicated hardware / software solutions to more flexible sets of services that allow customized and interactive experiences for end-consumers. The power of Big Data technologies from the web world can be harnessed to create measurable and customized experiences in the television space. This is especially true of experiences that are migrating from the single primary screen, in-home, linear viewing experiences to multi-screen, on-the-go, on-demand content consumption experiences. There are essentially three stakeholders that benefit from Big Data and corresponding data science: the operator, the content provider, and the end consumer. The operator benefits by having meaningful information for monitoring and optimizing system performance for capacity planning and measuring impact. The content provider can use this information to optimize and understand how their programming is being used by the end-consumer. Finally, the end consumers benefit from better interactive and targeted experiences that are more meaningful for them or their household.

A well-instrumented TV eco-system mirrors the web world in terms of volume, velocity, and variety of data that is collected, albeit not at the same scale. Volume of data is collected from loosely coupled systems and modules that span application, control, and delivery services. With disparate services comes data variety. All collected data has to be analyzed with velocity to provide meaningful and useful information to subsequent applications impacting user interaction, or systems impacting operator systems. The analyzed results can be provided through application programming interfaces (APIs), or through visualizations and operational dashboards.

ANALYTICS SYSTEM ARCHITECTURE

Analytics systems are comprised of three major components: data collection, storage, and analysis. A high level architecture of a general analytics system is given in Figure 1.

Data Collection

The key to any analytic system is collecting data, which is federated across a number of services within an eco-system. This requires instrumenting network, client services and applications, or equipment to generate and report events, which are then sent to an event intake mechanism. From an applications and a services perspective, event reporting should be a lightweight process where the complexities of ingesting events are abstracted away from the event collection probes.

At ARRIS, we have implemented an event-reporting library (in Java) that abstracts away the event collection logistics from the instrumented service. The instrumented service needs to

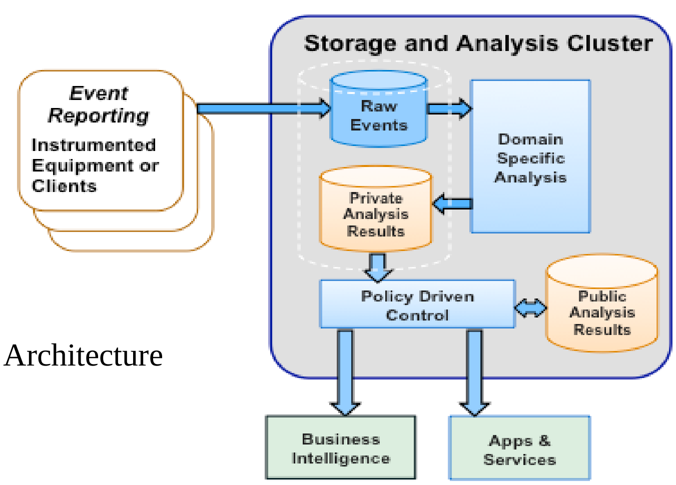


Figure 1 - High Level Analytics Architecture

define the events that are meaningful within its context and "report" them either using an abstracted library or directly to an event intake mechanism. Keeping this lightweight for the instrumented services, it is best to collect raw or minimally processed events, hence placing much of the compute burden on the analytics system.

We are currently using Spring XD [10] as our real-time event intake mechanism. It is an open source project being developed under the Spring framework umbrella. Spring XD, a distributed event intake mechanism supporting HTTP, TCP, and Syslog among others, allows for processing of these events prior to storing and persisting. We define one Spring XD stream per event type or event source.

<u>Storage & Persistence</u>

Once events are ingested, the question becomes one of the database technologies that the system should employ. The answer is quite simply a function of system requirements and the types of data that are being housed. The choices are either relational database management systems (RDBMS - SQL), or non-relational (NoSQL).

For storing and subsequently processing events from a variety of sources, where there is no obvious cohesive relational structure, a non-relational database such as HBase or files stored directly on HDFS indexed with Hive are obvious choices. As previously mentioned, this stems from Big Data technologies from the Internet. This allows for storing raw or minimally processed events for subsequent domain specific analysis. The result of any analysis or processing that is done on the data can also be stored in the same DBMS

The flexibility of Hadoop makes it a great choice for systems collecting information from loosely coupled, disparate sources, where processing typically needs to be done across the events to derive meaningful domain specific insight.

<u>Data Analysis & Processing</u>

Domain specific analysis needs to be done to accomplish targeted goals of the system. This is application dependent and determined by an articulated set of key performance indicators. For example, in the web domain, search and retrieval end up being important tasks. More complex tasks may include either content or social recommendations. Similarly, in the TV space, the data science applied determines the efficacy of the solution.

A Hadoop / HDFS system is designed to run MapReduce jobs for pulling and processing information. Higher-level languages that allow MapReduce scripting are Pig, Hive, and Scalding. Processing may be a simple statistical processing, correlating, and tabulating information across various individual services, or it may require machine learning techniques to do predictive analysis to profile human or system behavior; predicting user behavior facilitates end-user interaction, or predict system behavior / trends to circumvent system limitations.

A typical data analysis system houses sensitive user or subscriber information, which has to be protected and secured. A policy-based access control mechanism is imperative for ensuring that subscriber-specific information is exposed only to authorized external applications or business intelligence tools

Analysis and processing information have three primary goals:

- Create visuals and dashboards to provide business intelligence,
- Provide application-programing interfaces for downstream usage to drive system behavior or user enhancements, and
- Derive atomic and aggregate data sets on which additional analysis and exploration can be performed.

This is the data science or intelligence that rides on top of the Big Data technologies to derive business value.

ANALYTICS FOR DIGITAL AD INSERTION

Dynamic ad insertion (DAI) is a prime example of an application where Analytics plays a differentiating role increasing operator efficiency, measurement, and relevant ad targeting. Following the principles and architecture outlined previously, a data analytics system for DAI is comprised of instrumenting data sources for event collection, event intake, persistence / storage, and analysis. The results of the analysis are exposed via an application interface or they are exposed to a business intelligence tool for visualizing system information through appropriate dashboards.

Figure 2 is an architecture that is specifically designed for a DAI system. In this example, we are following an SCTE-130 based approach for DAI targeting multi-screen devices based on subscriber / user profiles [11].

The sources of event information are:

- ADS – Ad Decision Service for capturing ad placement request, response and notification events
- ADM – Ad Decision Manager for content delivery events. In the case of HLS, this may include manifest related information that is delivered to a client / application.
- CIS – Content Information Service for processed content and ad metadata
- POIS – Placement Opportunity Information Service for ad placement location events.
- Clients – Where possible actual client / application events are collected to get actual content and ad consumption information directly from the consumer facing application.

In addition to these operational events, bulk data imports are needed to correlate collected events with actual subscribers and programming. For this, we collect Subscriber Management System (SMS) information and guide information (EPG).

The events and bulk data are ingested via an

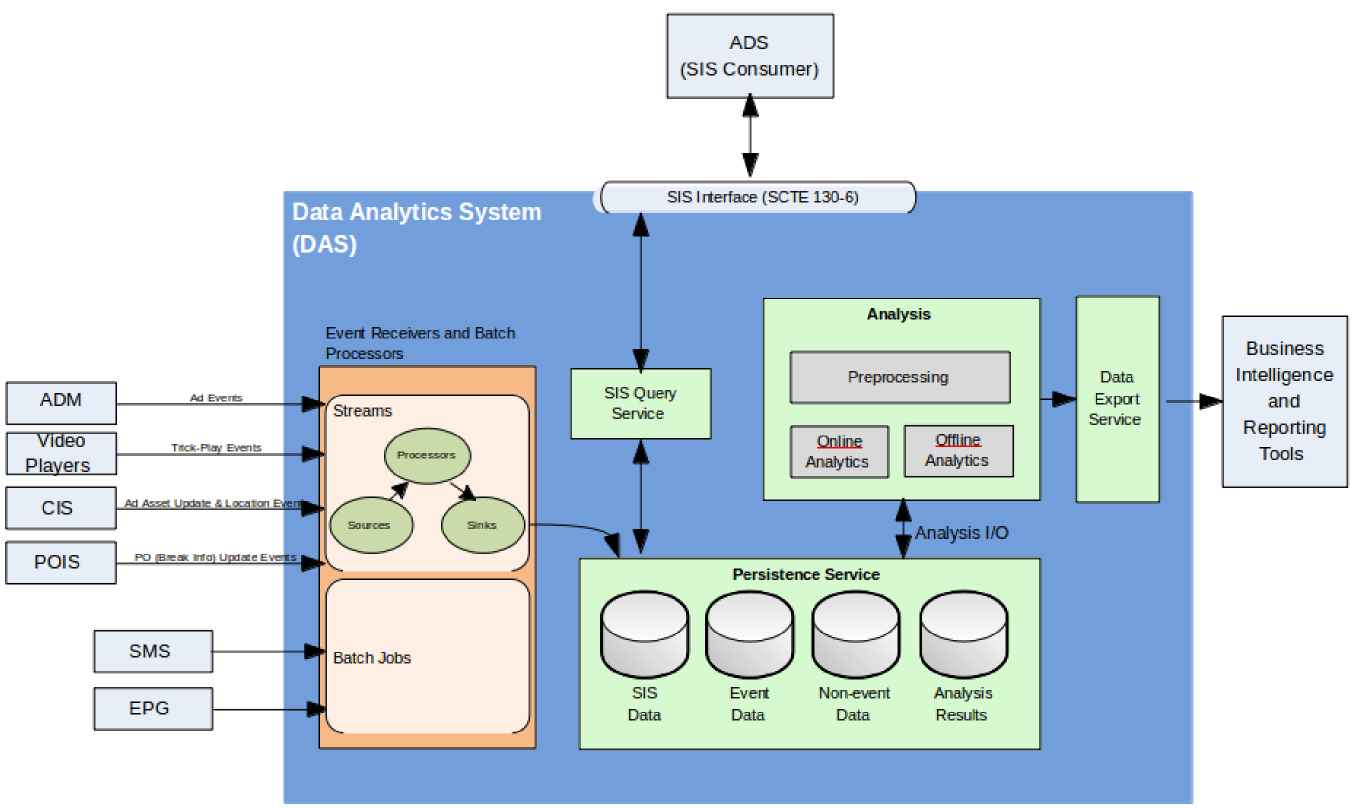


Figure 2 - Analytics Architecture for DAI

event intake module, which has the ability to do some filtering and processing of incoming data before storing it into our Persistence / Storage service. The collected information is processed using a variety of statistical techniques to derive user / usage profiles, and to collect historical information about DAI system usage. The user profiles are exposed via an SIS compliant interface so ADS systems can do appropriate ad targeting based on this user information. Historical performance and measurement information for content and ad delivery is visualized through appropriate dashboards for Business Intelligence.

In a typical DAI analytics system, the value is in:

- Classifying system behavior and usage. This has operational benefits for understanding how the system is being used. This is useful for operators, content providers, and advertisers.
- Characterizing subscribers so that appropriate insight is provided when targeting specific ads for better effectiveness.
- Understanding network and system limitations by analyzing and getting insight into trends of system usage and comparing them to current capacity has system is able to support. This has implications on scalability, storage, and bandwidth management.

DEPLOYMENT MODELS

The architecture supports various deployment models, either dedicated hardware-based or cloud (public or private). To enable cloud deployment, the services are developed to be hardware agnostic and can be scaled up or down as needed. Except in certain cases such as database interactions, the communication between different services is restricted to REST or SOAP (with preferably JSON, or XML, encoding) for simplicity. As new nodes can be added or removed, services are addressed by a proxy service.

The architecture is analogous to any Internet-scale web application: a distributed n-tier system. Unlike a standard web application that handles millions of requests/response from users, the DAI application handles millions of events from different components of the TV eco-system, which we denote as event sources. So as not to add additional burden on the event sources, we ensure that events are consumed asynchronously, and any additional filtering or

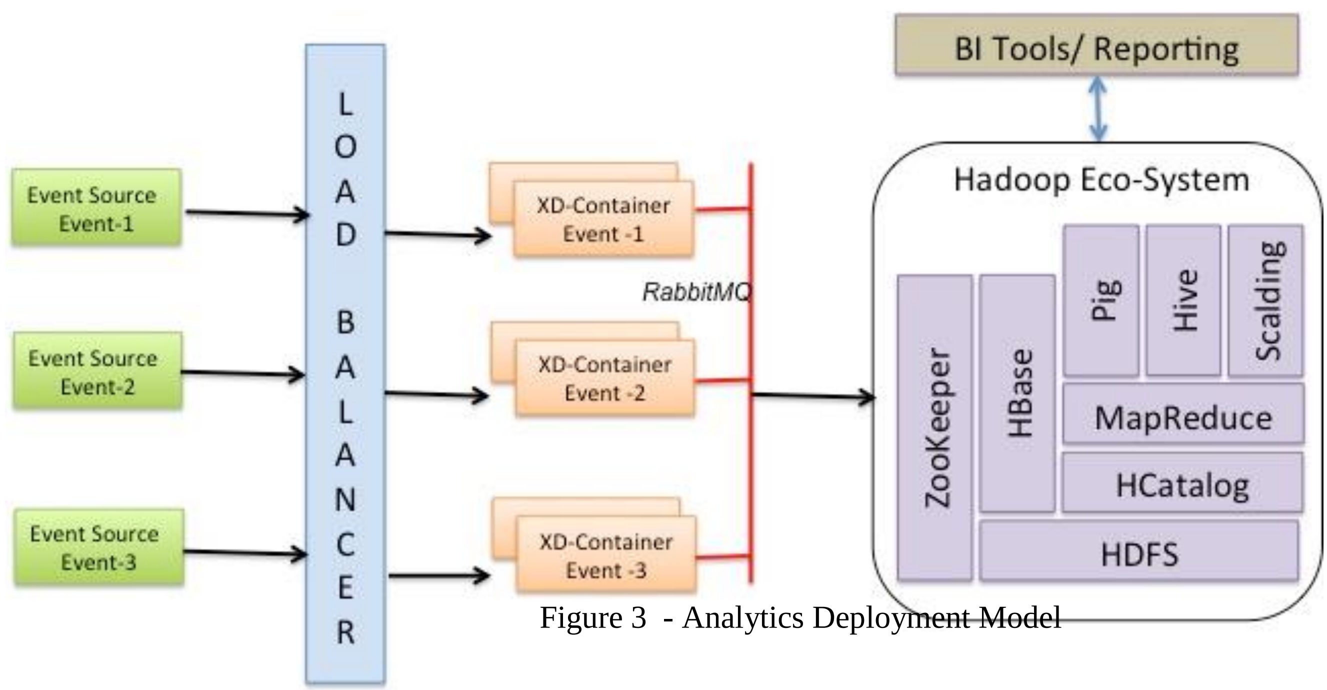


Figure 3 - Analytics Deployment Model

processing burden lies on the event consumers.

Special attention is given to high availability and fault tolerance throughout the system and especially at event consumer and storage/persistence layers. The Spring XD layer is deployed behind a software load balancer that allows us to deploy replica of each Spring XD container – one container per event type. Similar approach is taken at storage layer. Moreover, the Hadoop infrastructure (HDFS/HBase etc.) is already equipped with high availability and fault tolerance. Figure 3 depicts a typical deployment model.

ACKNOWLEDGEMENTS

This paper represents work that is being done by the Applied Research Center in ARRIS' Network & Cloud business.

CONCLUSION

Big Data technologies were developed primarily to target specific needs of the Internet. However, as our TV delivery systems grow more complex and as the technologies evolve to be more web-like, there is direct applicability and flexibility in using these technologies for television systems. A dynamic digital ad insertion system is just one example that can employ Big Data with appropriate analysis to provide more meaningful user interaction and targeting, operator measurement, and drive network efficiencies.

REFERENCES

1. Arthur, L. (2013, August 15). What is big data?. Retrieved from http://www.forbes.com/sites/lisaarthur/2013/08/15/what-is-big-data/

2. Couts, A. (2012, September 25). Breaking down 'big data' and Internet in the age of variety, volume, and velocity. Retrieved from http://www.digitaltrends.com/web/state-of-the-web-big-data/

3. Laney, D. (2001). 3D data management: Controlling data volume, velocity, and variety. In Application Delivery Strategies. Stamford, CT: META Group, Inc.

4. Ghemawat, S., Gobioff, H., & Leung, S. (2003, October). The Google file system. ACM - SOSP '03, Bolton Landing, NY.

5. Chang, F., Ghemawat, S., Hsieh, W., Wallach, D. A., Burrows, M., Chandra, T., Fikes, A., & Gruber, R. E. (2006, November). Bigtable: A distributed storage system for structured data. OSDI'06: Seventh symposium on operating system design and implementation, Seattle, WA.

6. Apace hbase. (2014, March 17). Retrieved from http://hbase.apache.org

7. Apache hadoop. (2014, March 14). Retrieved from http://hadoop.apache.org

8. Thusoo, A., Sarma, S. S., Jain, N., Shao, Z., Chakka, P., Anthony, S., Wyckoff, P., & Murthy, R. (2009, August). Hive - a warehousing solution over a map-reduce framework. ACM - VLDB '09, Lyon, France.

9. Apache Hadoop, Hive, and Pig on Google compute engine. (n.d.). Retrieved from https://cloud.google.com/developers/articles/apache-hadoop-hive-and-pig-on-google-compute-engine

10. Spring xd. (2014). Retrieved from http://projects.spring.io/spring-xd/

11. American National Standard, Society of Cable & Telecommunications Engineers (2011). Digital program insertion – advertising systems interfaces: Part 1 – advertising systems overview (ANSI/SCTE 130-1 2011)

BIG DATA DIFFERENTIATORS: A QoE PREDICTIVE MODEL FOR VIDEO SDN

Hui Zhang
Conviva

Abstract

Providing a positive viewing experience is critical for content owners relying on advertising and subscription-based revenue models to capitalize on the opportunity present in online video. However, with more viewers turning to online sources for consumption, understanding the quality of experience (QoE) and providing an optimal QoE becomes more critical. In this paper, we will discuss a possible online video predictive model for improving QoE using global intelligence extracted from analyzing billions of streams in real time utilizing a big data processing platform. While predictive modeling has been used in conjunction with big data to analyze historical and current trends for countless other disciplines, its application in digital media delivery has not yet been explored. Our research shows that on-line predictive modeling can provide tremendous value to those looking to monetize and enhance the viewer experience.

INTRODUCTION

Online video streaming is one of the most important applications on the Internet. Today, more than 57% of Internet traffic is video and the percentage is predicted to reach 69% in 2017 [1]. At the same time, users are demanding better and higher quality video (e.g. HD and Ultra-HD or 4K video) [2][3][4]. Ensuring video Quality of Experience (QoE) is becoming, and will continue to be, a challenge for both video content publishers and service providers.

Existing protocols that enable Internet video streaming assume two fixed end points (e.g., a video server and a client streaming the video) and varying resource availability between, or at, the two end points. The key mechanism to achieve better quality is to rapidly *adapt* or react to congestion and/or apply changes in resource availability along the path or at the end points. One example of this approach is the set of adaptive bitrate algorithms that have been implemented in many video players, often highly optimized for specific streaming protocols. While these adaptive solutions have served us well in the past, they become sub-optimal as video streaming services become more sophisticated and new opportunities emerge.

Towards a Video Software-Defined Network

Most scalable and reliable Internet video streaming services have many control knobs for adjusting video playback quality across different layers in the network stack that call for cross-layer optimizations. For example, many video services are implemented using multiple servers, typically distributed geographically, e.g., Content Delivery Networks (CDNs). For each video session, a video player can select one of many possible servers from which to start streaming. If the duration of the video session is long, it is possible to switch servers during the lifetime of the session. The control plane is further complicated by recent trends in content providers utilizing multiple CDNs [5] and/or CDN federations [6]. In addition to streaming server selection, for an adaptive bitrate video streaming protocol, the initial bitrate needs to be selected and the player needs to continuously adapt the bitrate during the video playback. We argue that by leveraging and extending the recently proposed Software-Defined Networking (SDN) approach [7][20], it may be feasible to select the best Internet route on a per-video-session basis. Furthermore, we will not be surprised if other knobs, such as video encoding profiles [8], client or server initial TCP window size [9], or even the transport protocol itself, can be opened up for control.

The Case for a Data-Driven Predictive Model

The new opportunities call for cross-layer optimizations where the decisions involving different control knobs need to be considered together, instead of separately. This causes the decision space to grow exponentially, making it extremely difficult for a reactive protocol to do the best possible job. Indeed, it will take a long time for such a protocol to converge to a good decision. Next, we briefly discuss the challenges faced by existing protocols.

Typically, these protocols use static initial configuration parameters that are often sub-optimal. For example, adaptive streaming protocols usually start with a statically configured bitrate. If this bitrate is too low, the protocol might not even be able to reach the optimal rate by the time the video has ended (e.g., for a 30s or 60s news clip). Additionally, such pre-configured bitrate may be sub-optimal during periods of congestion or compromised bandwidth.

Even after an initial decision is made, when these protocols react, they don't always make the optimal decisions, which may further impact user experience. For example, in case of congestion, an adaptive bitrate protocol may switch up to a bitrate that cannot be sustained, and, as a result, the user may experience re-buffering.

In this paper, we make three arguments. First, given the fundamental limitations of reactive approaches, we argue for an alternative predictive approach, which aims to accurately predict the outcome of making a particular choice, e.g., will a stream be able to sustain a particular bitrate? In theory, a perfect prediction would allow protocols to use "optimal" configuration parameters and make "optimal" decisions. For example, it would be possible to exactly pick the largest sustainable bitrate for a video stream at start time.

Second, in order to accurately predict the outcome of a given choice, one may be tempted to use an analytical approach to model the environment, the streamer, the network, or some combination thereof. However, we believe this is infeasible due to the huge complexity of the delivery ecosystem, and this is also unnecessary. We instead argue for a data-driven empirical approach to leverage the information available from other players, streams or connections, i.e., use the performance experienced by other "similar" sessions to predict the performance for a given session. For example, if sessions located at some organization can sustain 2Mbps on average when streaming from a CDN, then it's likely that a new session from the same organization will be also able to stream at 2Mbps from the same CDN.

A Cloud-Based Big Data Solution: V-SDN

To this end, we propose a global control plane architecture – or Video SDN (V-SDN) – that continuously collects data from various sources, e.g., the quality of current and historical video sessions, and uses this information to maximize quality of other sessions. There are several challenges to implement such a system. First, we need an Internet-scale solution. To provide some perspective, we collect data from over 4 billion streams per month, and at each point in time, we are collecting data from up to 2 million concurrent streams. When applying to all Internet videos, that number becomes at least one order of magnitude bigger. Second, we need to process this information and make decisions in real time, i.e., milliseconds or sub-millisecond response time. Third, to make the best decisions, we need on-line prediction models to accurately model the session performance as new data is continuously collected.

We present the design and early experience with deploying such a control plane architecture, called video Global Optimization (GO). Conviva has built a cloud-based data platform using a data hub architecture [10], where data ingested from different sources are stored in a distributed file system, and different tools built on top of the data provide batch processing, stream processing, search, graph computation, ad-hoc analytical querying, time-series analytics, and statistical analytics computation. The data platform is designed to be distributed, horizontally scalable, and highly available. With this data platform, applications such as GO do not need to worry about the distributed nature of the system and are able to explore the data more freely. GO uses the quality-related information (e.g., current bitrate, re-buffering rates, start time, etc.) which is continuously collected from each streaming video client. Next, GO processes the quality information in real time, and, based on this information, provides hints to clients about the best bitrate or server to start with or switch to. For ease of deployment in today's Internet ecosystem, we make two simplifying assumptions. First, for a given session, GO selects the server and Internet path at the CDN granularity, instead of server granularity. Second, GO currently makes decisions at the start of a video session, and not in the middle of the session. These simplifying assumptions reduce both the frequency and the number of decisions GO needs to make. Despite these restrictions, our experience with deploying GO to optimize video streaming across several video sites shows that it is a highly advantageous step towards a fully featured global control plane.

In particular, due to the flexibility of our data platform architecture, it would be easy to extend GO to handle additional data sources, to make new types of decisions (e.g., select an encoding format), or to implement different algorithms.

GO System Overview

As shown in Figure 1, GO consists of two main components: (i) a backend that collects and processes the information about the video quality across all clients, and (ii) a client library that collects quality information at each client and sends it back to the backend. In particular, the client library monitors the states of player and network condition, summarizes them in the form of quality samples, and sends these samples back to the GO backend.

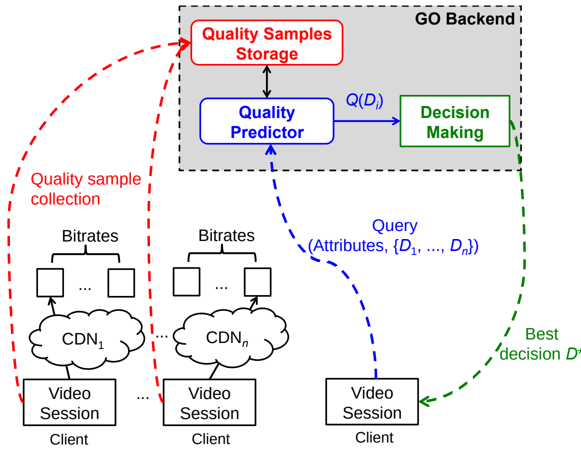


Figure 1: Go System overview

The GO backend uses the streaming capability of the data platform to process the quality samples received from clients in real time, and predict the quality outcomes of future sessions for each possible decision. The GO system makes two decisions for each session: the initial or starting bitrate (most videos today are encoded at multiple resolutions), and the initial CDN to stream the content from (many content providers today are on or moving towards multi-CDN setup). The initial CDN and bitrate are chosen from a pre-determined set of options and they can be changed at any point during a session. Note that to play a video that is encoded at n different bitrates and stored on m CDNs, there are m × n options to chose from.

To pick one of these options, GO predicts the quality metric for each option, and then picks the one that corresponds to the highest quality. In particular, GO uses real-time, quality-related information from every client currently streaming video to predict the quality of another user, and uses this prediction to select the initial bitrate and CDN for a new user.

Quality Samples

As explained above, the client library collects various pieces of quality information, summarizes this information into quality samples, and sends these samples to the GO backend. More precisely, a quality sample contains a set of quality measurements and attributes. Based on the quality measurements, GO computes a set of quality metrics. In this paper, we focus on four industry-standard video quality metrics that have been shown to impact user's engagement [11][5]:

1. Buffering ratio: The percentage of time a session spends in buffering state, i.e., waiting for the player's buffer to replenish with enough data to continue the playback.

2. Join time: The time it takes to start playing the video from the time the user clicks the "play" button.

3. Average bitrate: Many of today's video players support adaptive bitrate switching to effectively react to changes in the bandwidth availability. The average bitrate of a session is simply the time-weighted average of the bitrates used in a given session.

4. Video start failures: Some sessions fail to start playing due to various reasons, including content unavailability, or CDN server overload.

In addition, a quality sample contains a

Name	*Description*	*Number of Unique Values*
ASN	The Autonomous System (AS) Number that the client IP belongs to	15K from multiple countries
Video asset	The video asset being streamed	80K
Content provider (site)	The video site providing the content	279
Initial CDN	The CDN that the video session starts with	19 including ISP, commercial, in-house
Initial bitrate	The bitrate that the video session starts with	15K video assets with multiple starting bitrates
Connection type	Type of last-mile connectivity	7 including WWAN, DSL, fiber-to-home

Table 1: Number of unique values for various session attributes

large number of client and video session attributes. Table 1 summarizes some information about these attributes. Results are derived from quality samples from over 800 million sessions or views (both successful and failed) over a one-month period.

CHALLENGES

Before we delve into solutions, we first discuss some of the challenges to building such a video quality prediction system. First, we describe the scalability challenges and then we discuss the algorithmic challenges in managing prediction errors.

Internet-Level Scalability

A global intelligence system needs to be scalable in order to make a large number of real-time decisions. This is especially true given the combinatorial nature of the session features, i.e., the number of feature combinations increases exponentially with number of base features. At the same time, the decisions need to be made in an extremely low (sub millisecond) constant time so that the decision making process doesn't adversely affect the video session.

Managing Prediction Error

Any prediction has error. The natural question is: where do errors come from and how do we manage them? We roughly follow the decomposition of prediction error developed in [12], specializing it to our setting. In quality prediction, there are four sources of prediction error:

1) Estimation error caused by limited data: All things being equal, more data will give a more accurate prediction because predictions are less impacted by random fluctuations in the available data. Given a large number of attributes and the combinatorial nature of the attributes (the number of partitions grow exponentially with more attributes), estimation error can be a serious problem in practice.

2) Bias due to missing or unused information: The bias occurs when one does not observe (or use) an attribute that is important for prediction. Bias is not alleviated by gathering data from more sessions, but by gathering more attributes from each session. There is a fundamental tradeoff between estimation error and bias that we need to address: more features reduce bias but also reduce the number of samples within each feature set and potentially increase estimation error if not designed carefully. On the other hand, the common approach to handle estimation error is to increase number of samples by aggregation, thus increasing bias.

3) Unavailability of recent data: In a practical system, there are delays in measuring, sending and processing quality samples. If conditions change rapidly, there may be no quality samples sufficiently close to the session under prediction. This is an extreme example of estimation error. In this case it may be necessary to model the evolution of the video ecosystem over time in order to extrapolate to the current time. Figure 4 shows per-minute quality variability. It shows that even with sufficient data, the mean value of quality samples belonging to sessions in the same partition could vary significantly. This clearly indicates that any practical algorithm has to be running in real time (on the order of one minute or less).

4) Noise: Outcomes may be affected by non-deterministic inputs that could not reasonably be observed or predicted by any system. For example, performance may be affected by exponential back-off at the data link layer, by the congestion generated by cross traffic at the network layer, or by a randomized algorithm somewhere in the networking stack. Though prediction error induced by such factors technically falls under the

category of bias, it is more useful to think of it as noise. Noise implies that some degree of prediction error is inevitable.

ALGORITHM

In this section, we present a practical algorithm that addresses the challenges presented in previous section.

The tradeoff between estimation error and bias naturally leads us to consider a class of algorithms that compute average quality outcomes for partitions of sessions under different feature sets, and then dynamically chooses the feature set that seems to work well for a given session. We propose the following basic structure for the algorithm. The algorithm chooses a set of features offline based on analysis. Then the algorithm starts collecting quality metrics for all the feature sets. Based on the quality metrics collected, each feature set will be given a weight so that the weights reflect statistical properties of the feature set and are thus updated when new quality samples are collected. Finally, when receiving a new session, we estimate the performance for each available decision using a weighted sum. Intuitively, the algorithm works as follows: for a given session, we look at the quality metrics of all the sessions that match this session exactly. If we have enough data and the metrics are statistically stable, the decision will be mostly based on that. Otherwise, we will have to remove a feature and look for quality metrics with similar sessions, etc.

The next question then becomes how to design a weighting scheme to balance between estimation error and bias to minimize overall prediction error.

George et al [13] considered this problem and proposed an algorithm called WIMSE (short for Weighted Inverse Mean Squared Error). As the name suggests, the weights w_i are chosen to be the inverse of an estimate of the mean squared error (MSE). Inverse-MSE weighting has the following desirable property: If the mean quality pi of each group is statistically independent and the mean squared error is estimated exactly, then the resulting prediction is an optimal estimator in the sense that it has the minimal overall prediction error among all samples drawn from the same distribution [13]. While in practice neither condition is met, George et al have found that WIMSE can nevertheless work well in most cases.

The basic algorithm still faces other practical challenges. However, due to space limitation in this paper, we concentrate on only the major challenges the algorithm faces.

DEPLOYMENT

We leverage recent advances in data systems to build our data platform, around Hadoop [14] and the Berkeley Data Analytics Stack (BDAS) ecosystems [16]. From a high-level design, we use HDFS as our primary storage layer and Spark as our primary computation engine. We also use other tools to provide both HA (high availability) and data processing abstraction such as batching (Spark [15]), streaming (Streaming Spark [15]), search (Solr [17]), analytical querying (Shark [18]), statistical analytics (R [19]), etc.

We implemented the GO system on top of our data platform using the streaming capability of the platform to train the model and produce decision tables every minute. The decision tables are then sent to a number of decision makers distributed across multiple locations. Finally, when a query comes into a decision maker, it runs the algorithm using the decision table and responds within milliseconds.

The GO system is currently deployed with multiple premium content publishers. However, there are several practical difficulties of evaluating the performance of GO in production deployment environments. First, most publishers do not want to perform A/B testing when they understand that the performance of some of the streams may not be optimal when they are grouped by the non-optimized version of the algorithm. Second, with all publishers, there are usually additional business considerations beyond the goal of optimizing the QoE of the video streams. For example, when using multiple CDNs, a publisher could get a lower price from a particular CDN if it would allocate more than a certain percentage of its total traffic to the CDN. This CDN allocation policy would put additional constraints on the GO optimization algorithm. Consider a scenario where a publisher has three CDNs X, Y, and Z, and has a minimum committed usage percentage on X and Y. This would mean that even if CDN Z was the best performing CDN based on the prediction algorithm, beyond certain percentage of traffic, no additional streams would be allocated to it.

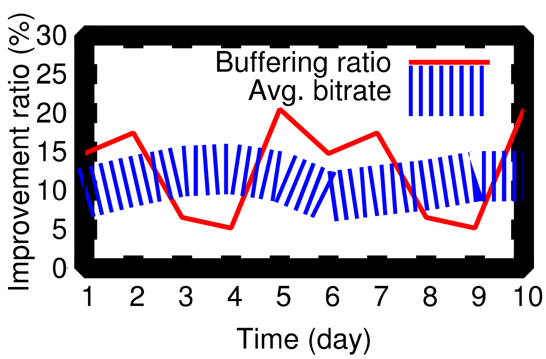


Figure 2: Performance improvement over time compared to baseline algorithm

Buffering Ratio and Average Bitrate: In Figure 2, we show the percentage of improvement of GO over the baseline (randomized decisions) with two performance metrics: buffering ratio and average bitrate. In particular, for buffering ratio, we show the percentage of reduction of buffering ratio for all sessions served by the GO algorithm in each day as compared to all sessions served by the baseline algorithm on the same day; for average bitrate, we show the percentage of increase of average bitrate over the entire session duration for all sessions served by the GO algorithm in each day as compared to all sessions served by the baseline on the same day. We present the comparison over a continuous time period of 10 days. There are several points worth noting. First, both metrics are improved simultaneously with GO as compared to the baseline. In contrast, with normal adaptive bitrate protocols without special optimization, the improvement of one metric usually results in the deterioration of the other. For example, the reduction of the buffering ratio usually comes together with the reduction of the average bitrate also. With GO, both metrics are improved simultaneously. Second, the performance improvement varies daily. The most likely explanation is that it is due to the CDN performance variation. To understand this

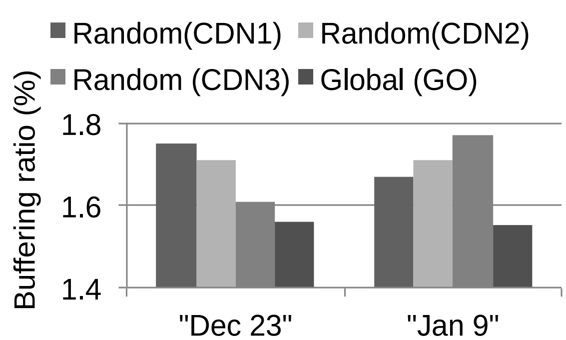
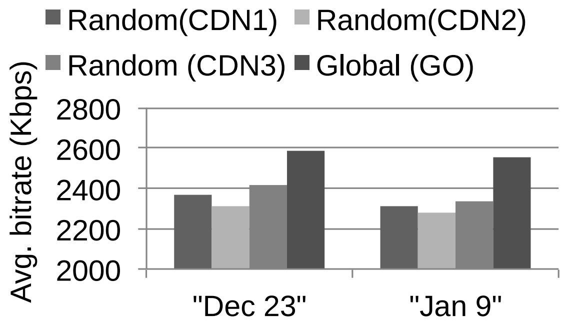


Figure 3: Buffering ratio for different CDNs at different points in time, illustrates CDN performance variability

Figure 4: Average bitrate for different CDNs at different points in time, demonstrates very much the same phenomenon as Figure 2

better, we compare the performance of all sessions under GO and the performance of all sessions on each of the three CDNs under the baseline algorithm. This is shown in Figures 3 and 4 respectively with buffering ratio and



average bitrate as performance metrics respectively. For each figure, we show the comparison in two separate days. Some additional points to note: first, the performances of sessions for different CDNs under the baseline do vary, with respect to both buffering ratio and average bitrate. Second, the sessions under GO perform better than the sessions for even the best of the three CDNs. This suggests that GO is not only looking for the best CDN on average, but also differentiates CDN performance in finer granular partition across time and space. In addition, if one compares the relative performance for each individual CDN, the ranking varies between the two days. In particular, with respect to the buffering ratio, CDN3 is the best on Dec 23, but CDN1 is the best on Jan 9; with respect to the average bitrate, CDN3 is the best on both days.

Interaction with the adaptive bitrate protocol: Since GO in this deployment only selects the bitrate and CDN at the beginning of each session and the HLS protocol controls the bitrate adaptation for the duration of the session, we would like to understand how the initial selection decisions by GO impacts the future adaptation decisions made by HLS. The number of bitrate switches per session made is a good indicator of how closely GO is able to select the ideal bitrate for a session. A good initial selection would result in a lower number of bitrate switches in the future. Figure 5 shows the comparison between GO, a static selection policy (traditional initial bitrate selection algorithm) and an algorithm that selects the starting bitrate at random. As shown in the figure, GO outperforms either case. Also note that static selection is as bad as picking a bitrate at random!

DISCUSSION

While our results look promising, the magnitude of the improvements may appear underwhelming. However, we believe that this is not due to the lack of potential of the approach, but it merely comes down to the

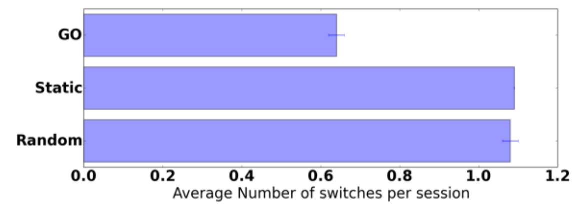


Figure 5: Average number of switches per session for GO initial bitrate selection, static initial selection (1200Kbps) and random selection

inherent limitations that are typical to any first instantiation of a radical approach. In the remainder of this section, we consider some of these limitations which, when removed, will result in much higher improvements.

Mid-stream selection: Currently, GO makes decisions at the sessions' start times. For a long session this may not be optimal, as, for example, the quality of the selected CDN may degrade during the session's lifetime. We are adding the ability to make the decisions during the mid-stream, as well. Note that in this case, the prediction is equally important, as switching to a new CDN is not guaranteed to increase the quality, especially if the quality degradation is due to the last mile or due to the client inability to render at a high bitrate.

Leveraging network and CDN information: GO makes decisions based on client side information only. While clients provide the most accurate information regarding the quality experienced by users and at the final viewing stage, this information may not always be optimal when making decisions. The decision process can be considerably improved if GO were to leverage information from other entities in the distribution ecosystem, including CDN servers, caches, switches and routers. Using such information, GO could significantly improve the prediction accuracy. For example, GO could learn much faster that a CDN server is overloaded by getting load information directly from that server than inferring this information from clients that experience quality issues when connected to that server.

Finer grain selection: Currently, GO selects the resource at the CDN granularity. This means that GO does not do much if the CDN redirects the client based on its location (e.g., IP address) to a set of congested servers. However, if the client were able to specify the servers to stream from, GO could avoid the overloaded servers and dramatically increase the quality. We believe that we will soon have the ability to perform such fine grain selection, as CDNs are incentivized to expose such information to clients. Indeed, a CDN operator will prefer that a quality-impacted client move to other servers in the same CDN rather than migrate to a different CDN. Furthermore, an ISP CDN that also runs its own software on set-top boxes or other user devices would be in the perfect position to run a GO-like algorithm that makes decisions at server granularity.

CONCLUSION

As the Internet infrastructure becomes more complex, the potential number of congestion and failure points will only increase. In this paper, we have shown that despite this increasing complexity, a predictive model for QoE, leveraging an Internet scale control plane architecture (GO), gives online video providers the ability to deliver an optimal viewing experience for their content. With new SDN technologies being developed and deployed to expand the capabilities of devices within the video ecosystem, we believe that a GO-like solution will be able to make even more granular optimizations and better control the delivery environment. Content publishers and service providers are in an excellent position to utilize a V-SDN architecture, such as GO, to ensure the integrity of their business as the industry moves to an Internet TV model.

REFERENCE

[1] Cisco Visual Networking Index: Forecast and Methodology, 2012–2017

[2] http://gigaom.com/2014/01/06/netflix-4k-ultra-hd-3d/

[3] http://mashable.com/2014/01/03/youtube-4k-ces/

[4] http://vimeo.com/tag:4k

[5] A Case for a Coordinated Internet Video Control Plane. Xi Liu, Florin Dobrian, Henry Milner, Junchen Jiang, Vyas Sekar, Ion Stoica, Hui Zhang, Proceedings of ACM Sigcomm, 2012

[6] Content Delivery Network (CDN) Federations, How SPs Can Win the Battle for Content-Hungry Consumers, Cisco Whitepaper

[7] The Road to SDN, Nick Feamster, Jennifer Rexford, Ellen Zegura, ACM Queue, December 30, 2013

[8] Dynamic Video Transcoding in Mobile Environments, Bo Shen, Wai-Tian Tan, and Frederic Huve, IEEE Multimedia, 2008

[9] Increasing TCP's Initial Window, IETF RFC 6928

[10] Big Data Requires a Big, New Architecture, Dan Woods, Forbes

[11] Understanding the Impact of Video Quality on User Engagement, Florin Dobrian, Asad Awan, Dilip Joseph, Aditya Ganjam, Jibin Zhan, Vyas Sekar, Ion Stoica, Hui Zhang, Proceedings of ACM Sigcomm, 2011

[12] A unified bias-variance decomposition, Pedro Domingos, Proceedings of ICML 2000

[13] Value Function Approximation using Multiple Aggregation for Multiattribute Resource Management, Abraham George, Warren B. Powell, Sanjeev R. Kulkarni, JMLR 2008

[14] http://hadoop.apache.org/

[15] http://spark.apache.org/

[16] https://amplab.cs.berkeley.edu/software/

[17] https://lucene.apache.org/solr/

[18] http://shark.cs.berkeley.edu/

[19] http://www.r-project.org/

[20] http://www.opendaylight.org/

CABLE NETWORK MANAGEMENT INFRASTRUCTURE EVOLUTION

L. Alberto Campos, Jennifer Andreoli-Fang and Vivek Ganti
Cable Television Laboratories

Abstract

An approach to enable advanced troubleshooting, granular analysis and service quality of experience assessment is presented. The use of topology information in the identification of each cable network element along with granular information of the element configuration and health is proposed. This technique covers multiple layers including the service layer. At the physical layers street addresses, taps and amplifiers are used to identify impairment location. All layers are leveraged to measure network and service reliability, service degradation and to quantify quality of experience. A cable infrastructure implementation is described as an example.

BACKGROUND

Historically cable networks consisted of a multitude of services that developed relatively independent from each other. There have been always considerations for coexistence since the services are sharing the same physical environment but their designs had limited relation. Analog video services initiated as a basic transport of broadcast signals over the cable infrastructure. The introduction of digital video had only to be cognizant of the legacy analog video channelization in addition to fidelity requirements for suitable coexistence. DOCSIS® went a step further as in addition to adopting the channelization of legacy services, DOCSIS® also adopted the same downstream physical layer as the digital video downstream. This provided synergies that led to common downstream transmission systems called Edge-QAMs and the modular systems that are described in the M-CMTS specifications. The HFC telephony systems have been for the most part an adaptation of circuit switched telephony systems to the cable environment. The limited relation of these service platforms has also resulted in separated isolated management systems with different performance indicators and management parameters. The industry has yet not gone through the trouble of examining the relation of all these indicators and parameters to realize potential simplifications in network management systems. Recently with the advent of DOCSIS® 3.1 and the implied migration to a common IP platform for all services, there is the potential of introducing a unified set of performance transport metrics and performance thresholds with significant enhancements regarding how the networks are managed and provisioned.

Moreover, the capabilities of the tools to handle big data systems and perform complex analysis on them have significantly improved in recent years. In the cable environment we have been introducing more elaborate network maintenance techniques such as Proactive Network Maintenance, which leverage granular plant health probing and data correlation and analysis.

This paper proposes a strategy for advanced network monitoring and management to support end-to-end service analysis, on an individual subscriber service basis. This leverages the collection of data in a granular fashion from all cable monitoring and management systems. A key goal of this approach is to develop a management system capable of assessing end-to-end service performance, quantifying quality of experience, determining network health, network reliability and determining the resources to provide services. The latter is also useful to assess the resources remaining available to provide other services. Provisioning of business services in general and backhaul services in particular are use case scenarios of particular interest.

LAYERED APPROACH EVALUATION

In order to assess end-to-end service performance, the assessment will have to encompass/traverse multiple layers. The network management systems are traditionally specific to a section of the network and to a layer of the network. These management systems are typically isolated from each other. There is a need to extract information from all these layers and correlate such information to assess end-to-end service performance. Figure 1 shows a schematic representation of the typical components in a cable network.

Figure 1 shows a collection of components that are used to provide the services that are typically delivered through a cable TV network. By no means one unified management system is used to manage the components represented in Figure 1. Yet, a single service traverses different sections of this network that is the under control of different management systems.

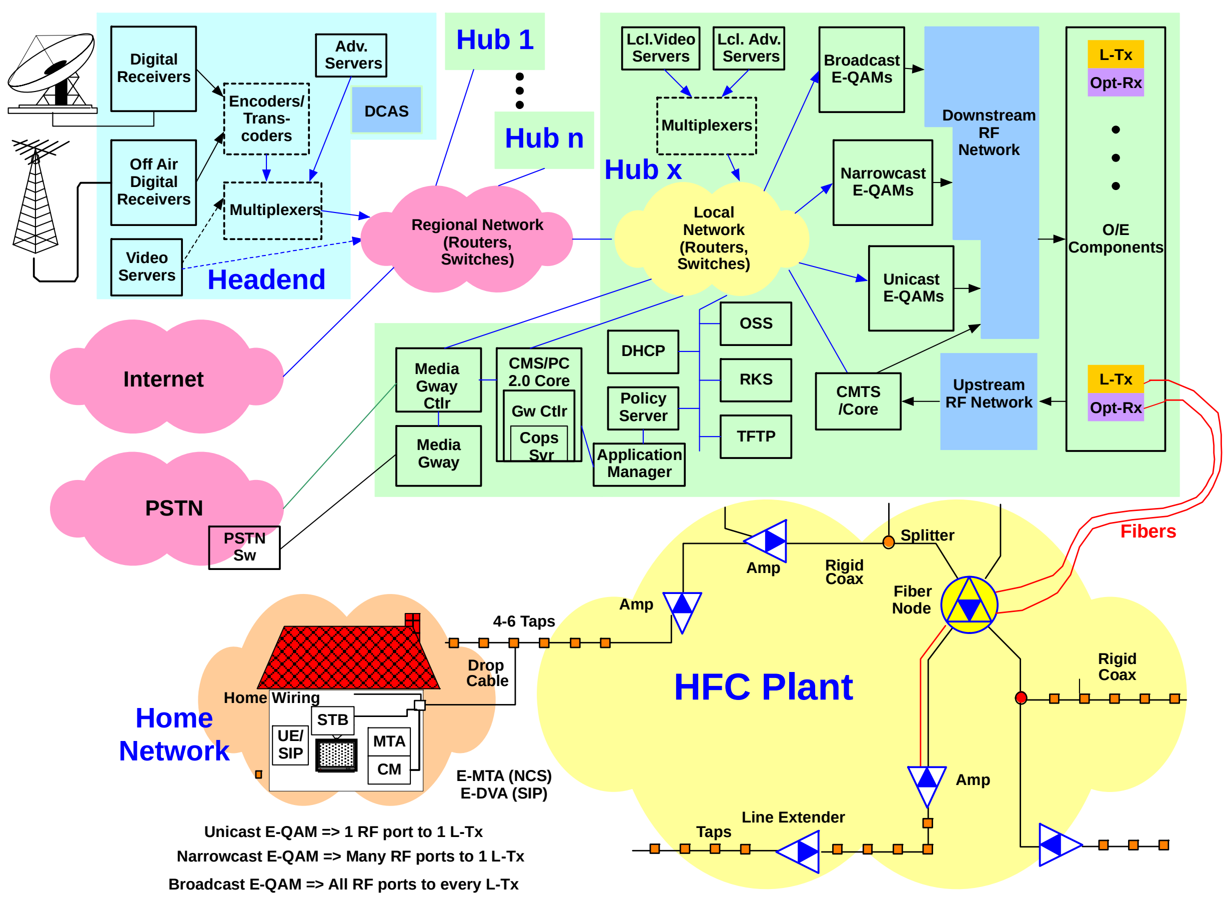


Figure 1 – Cable Network Elements for Service Delivery

In order to analyze an end-to-end service, one has to select the subset of the network components that participate in the delivery of that service. This subset of the network has a corresponding set of management systems that for end-to-end service delivery would have to be coordinated in a cohesive fashion. To facilitate analysis, one can further divide this service into sub-services. For example, a further segmentation for voice services would be call-signaling and call-payload transport. A segmentation for data sub-services could be data-service provisioning, data-authentication and data-payload transport.

Figure 2 shows a subset of the network containing the end-to-end paths of a data-service payload transport example and a video service payload transport example. These end-to-end service paths are also divided based on the layers they traverse. For practical analysis purposes we divide these service paths into four layers. The physical (PHY) layer, the link or medium access control (MAC) layer, the TCP/IP layer and the services and applications layer. As shown in Figure 2 there are different network elements that participate depending on the layer.

In the data-payload transport service example the PC and the data server (in light blue) are the elements that participate in the services and applications layer while for the video-payload transport example a video-terminal and a video server (in red) are the elements that participate in the services and applications layer.

Examining the MAC layer participant elements in the data-payload transport service example, the elements involved are:
PC, cable modem, Edge-QAMs, CMTS core, local network, regional network, public internet and data server.

In the case of the PHY layer participants of the data payload transport service example, the elements involved are:
PC, cable modem, taps, amplifiers, splitter, optical node, laser transmitter and optoelectronic receiver, downstream and upstream RF networks, Edge-QAMs and CMTS core, local network, regional network, public internet and data server.

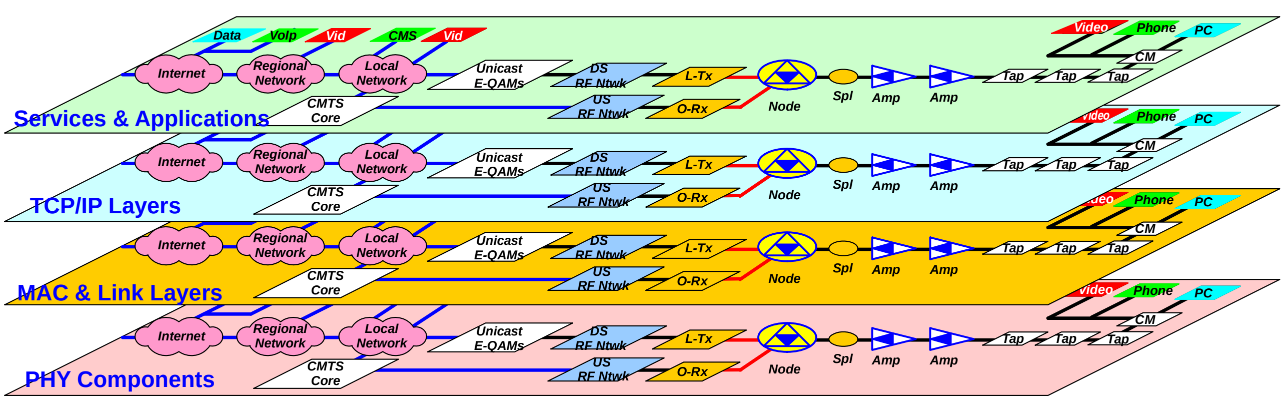


Figure 2 – Cable Network Subset Dissected in Layers

Based on the available tools that can troubleshoot issues at different layers, different elements get involved in the analysis process. There are elements that participate in more than one layer. These are elements that can be leveraged as anchor-elements which enable the correlation of data across multiple layers.

There are many systems that currently participate in the analysis of service performance. Each system typically has its own performance metrics. Many of these metrics are service layer metrics, symptoms gathered in these upper layer metrics are also used for issues that may be happening in the lower layers. Accessing directly the appropriate lower layer would be more efficient but many times the management system of the upper layer service doesn't have access to the lower layer. Breaking the silos of management systems and coordinating the management information enable the implementation of end-to-end service management.

Network Connectivity Relationship Among CATV Elements

The way management systems are used, highlights the importance of establishing element connectivity relations and understanding in which layer this connectivity relation takes place.

In the HFC portion of the network, the traffic follows the general paths from the Headend or Hub to the end device. In this portion of the network, which follows a tree and branch topology, there is a simple way to describe the connectivity relations using the "Long Name convention" which was proposed in [1].

This long name convention for PHY Layer uses the device naming technique that describes the path between the device or component starting from the element where the service originates to where it ends. In the case where there is a common point for all elements using that layer, a common northbound device that is uniquely described can be used. In the case of an upstream RF transmission to the CMTS, the optical node can be used as the common, uniquely identified northbound device where the tree and branch topology converges like the optical node (Figure 3). In this notation, subscripts are used to indicate branching out of multi-port devices.

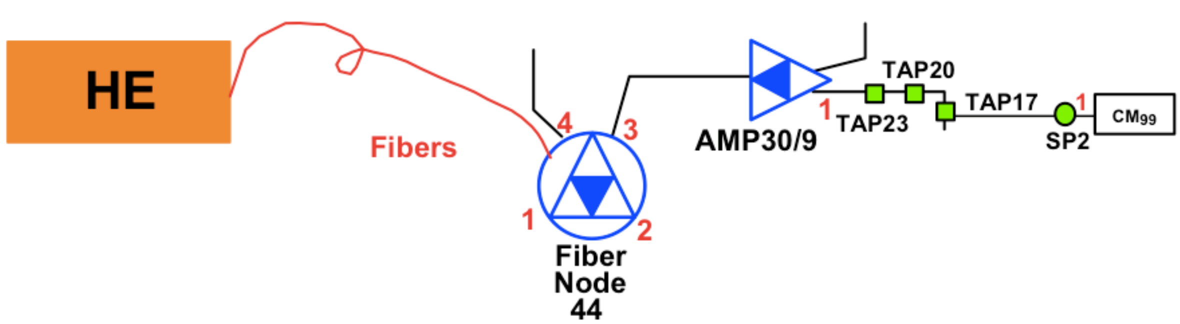


Figure 3 – PHY Layer Long Name Convention

In this case the cable modem with unique name CM99 would also be known in its long name convention as:

$FN44_3$-AMP30/91-TAP23-TAP20-$TAP17_1$-$SP2_1$-CM99

and the 23dB Tap can be identified by it's long name convention;

$FN44_3$-AMP30/91-TAP23

This naming convention representation is suitable for the PHY layer that follows a tree and branch topology. The naming convention is also suitable for a sequential path of components being traverse and for point-to-point topologies. It is however, not suitable to capture all possible connection and routes for mesh topologies enabled in the MAC and TCP/IP layers, but it is suitable for tracing the communications path for a particular service or application through the TCP/IP, MAC and PHY layers. This occurs as a direct consequence of the IP routing principle that avoids routing loops.

Therefore for describing mesh topologies in the MAC and IP layers, the entire connectivity relationships are required. The level of complexity increases since the connectivity relationships must be kept for all layers. The advancements in computational power and analysis capabilities that come along with big data, have made feasible tackling problems with this level of complexity. In the case of service layer analysis for a particular service rendered to a device, a long name convention can also be used. A device subscribed to a service or sub-service may have a long name in the IP layer another one in the MAC layer and a third one in the PHY layer. For example the naming convention in the MAC layer for a Video Download Service to a video terminal with an integrated STB is;

LocalVideoSvr33- $EthSwitch17_{15}$-$EQAM2_{12}$-VideoTerm99

Practical Approach to Granular Service Performance Measurements

It is many times easy to leverage the ever increasing computational power. Nevertheless this carries also complexity of implementation and in operations. It is desirable to strike a balance between granularity and simplicity.

In the multilayer management environment we have in our systems and services, there are significant dependencies between the different performance metrics and thresholds. As the management systems become unified through either interactivity among them or through a master management system that is able to access them, significant simplification on the way we determine health and performance of our networks becomes feasible. As a result of this unified management approach, the silos between the performance metrics databases originated from separate management systems are broken down.

If a metric used in a system is dependent on a root metric, then only the root metric should be used. A review of all performance metrics across the different management platforms is required. The abundant metrics currently available on the different systems lead to operational expenses that could be avoided through a network management optimization process.

There is a multitude of performance metrics that could be considered when evaluating broadband services and applications [2]. Figure 4 shows an example of broadband services and applications and their most relevant performance metrics.

Services & Applications	General Metrics						Service Specific Metrics						
	Instant Availability	Packet Loss	Latency	Jitter	Peak Rate	Sustainable Rate	Controllability	Avail. Content	Switching Speed	Capacity	Feature Richness	Drop Call %	Competence
Channel Surfing			H						H				
Data BE					H								
Video Conferencing			H	H		H							
VoD		H				H		H					
Music Services		H				H		H					
Gaming Client		H	H										
Gaming Server		H	H										
SuperDownloading					H	H							
SuperUploading					H	H							
Storage										H			
Tele-Medicine	H	H	H		H	H				H			
Tele-Education		H	H			H							
Tele-Commuting						H							
Tele-Library					H			H		H			
Customer Support													H
Home Security	H										H		
Home Remote Ctrl	H						H				H		
Telephony			H	H							H	H	
Video SD		H				H		H					
Video HD		H				H		H					
Video Low Res		H				H		H					
Business Services	H	H	H	H	H	H							

Figure 4 - Multiple Performance Metrics used for Different Services

In Figure 4, the red cell with an H indicates high relevance of the performance metric for that particular service or application.

Collapsing Multiple Performance Metrics to Common Performance Index

As mentioned above there is an advantage in reducing the number of performance metrics. If a relationship between the performance metrics can be determined, this relationship can be leverage to reduce the number of performance metrics. All possible performance metric relationships ought to be explored. There may be relationships between percentage voice call-drops and packet loss or relationship between video macro-blocking and packet loss. If a strong relation is determine only one parameter can be used.

In the DOCSIS® 2.0 and 3.0 HFC environment for example, there is an intuitive correlation between packet loss, latency and jitter performance. These are fundamental objective network performance metrics through which other upper layer performance metrics may be derived. We proceed to explore the relationship of these metrics.

A DOCSIS® 3.0 network consisting of a CMTS and CMs is setup. In this network loading of the channel is performed and the latency, jitter and packet are measured. Figure 5 shows the behavior of these metrics with channel loading.

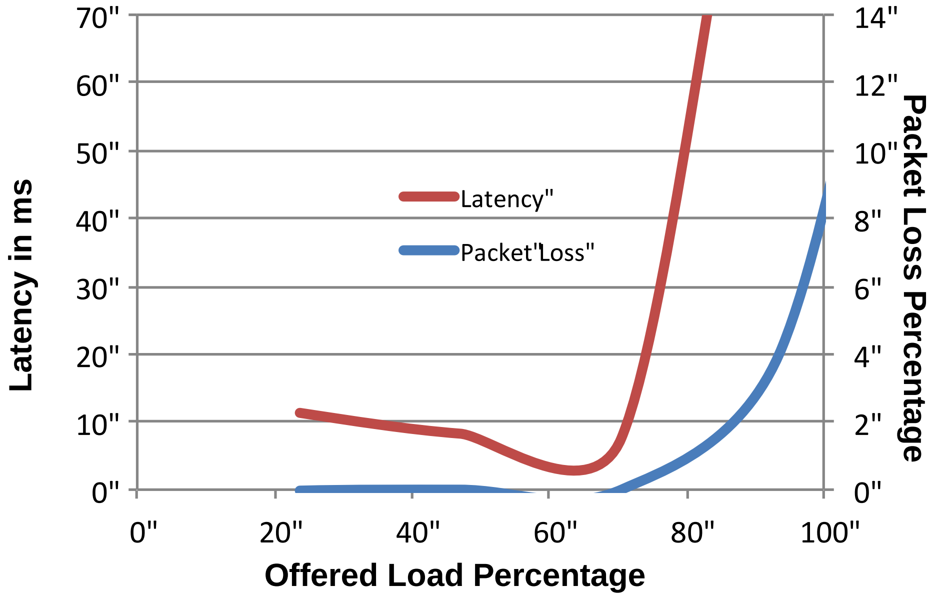


Figure 5. Correlation of DOCSIS Performance Metrics

Each curve correspond to a particular packet size, an additional curve based on an average packet size distribution can be generated. The resulting graphs indicate that within the DOCSIS® environment there is a strong correlation between packet loss, jitter and calibrated latency. Such a relationship enables the use of a unified performance index by collapsing multiple performance metrics into a common performance index.

For example, the main performance metric for gaming is latency. Let's assume that the threshold for unacceptable latency in gaming is 20 msec. Let's also assume that the main performance metric for video streaming is packet loss and the threshold for packet loss is 10^{-6} PER.

The sample performance table below shows the relation between packet loss, jitter and latency for a specific DOCSIS® network.

Performance Index	Packet Lloss % Lost/Sent Ratio	Latency ms
1	0.01	9
2	0.2	10
3	0.5	20
4	1	40
5	2	80

Table 1 - Performance Index Equivalency Chart

Having a smaller set of performance metrics defined the performance of all broadband services is very attractive. This could limit complexity in assessing performance on a per service basis, yet enable the granularity of performance on a per user per service basis.

GRANULAR DATA MODEL DEVELOPMENT AND AUTOMATION OF ANALYSIS

Some of attributes foreseen for future network management systems are those of higher granularity and capable of end-to-end assessments. In addition to these or to enable these, it is proposed here that the future network management systems also have to be aware of the network connectivity relations and the network layers the different management systems influence.

The higher granularity has to be assumed in two dimensions. One dimension relates to the increase granularity in element coverage (a larger set of components enter the management systems). A second dimension relates to the increased level of detail in which all the components are managed. As mentioned above it is beneficial to introduce network layer awareness and connectivity awareness. Most of the problems in the network even though they may be symptomatic across multiple layers, their cause is layer specific. Network layer awareness allows better assessment of network health and better troubleshooting of problems. In the network connectivity is a concept that is layer specific. What is view as connected in the physical layer is not necessarily view as connected in the TCP/IP layer or in the services and applications layer (Figure 2). Connectivity awareness in the databases containing the elements allows among many other things for automation of service reliability estimation, impairment localization, network performance assessment, capacity planning etc.

A couple of examples of data model elements that fit the desired goals of granularity, connectivity awareness and network layer awareness are included. The goal is to have this data model description of all of our network elements to fulfill the desired capability for end-to-end analysis.

Figure 6 shows the physical layer portion of the CMTS element model that has the characteristics proposed in this approach. It is intended to have a similar description for all network elements that participate in the physical layer. Figure 1 and Figure 2 illustrate the diversity of components that are encompassed in the granular database of cable network elements.

An important component of the data model is the downtime counter which is used to assess reliability. Another key component of this proposed model is the connectivity relation for each interface. This, in the case of the physical layer, would allow the manipulation of elements to automatically determine location of impairment or reliability. A third fundamental ingredient of the data model is the use of latitude and longitude or other means to determine location. This model uses location based on the type of element. There are three element types identified in the data model. One is a single location type, such as node or termination. A second is a linear type to represent as a linear device like a cable that has two vertices an location information for two points. The third is polygon type to represent a device covering an area such as a network or a linear multipoint element such as a cable conduit that requires a path to be specified.

To facilitate management there is some grouping proposed in this model. These are done through the association to a node, to a hub and to a system. The above is a sample approach not intended to be all inclusive but to demonstrate the required structure and to be expandable following the same structure.

Element	CMTS	
ElementID	CMTS4	Unique Within Fiber Node
P. Fiber Node Assoc.	NA	Unique within Hub
P. Hub Assoc.	Hb10	Unique within System
P. System Assoc.	S5	Unique
Name	ArrisC4-2	
Element Type	Node	Polygon/Cloud Node Link Termination
Location Vrt 1	Lat-Lon	
Layer 1	Physical	
Physical Element Type	Node	
Number of Interfaces	9	(ports)
InterfaceID	10	Unique within Element
Vrt Assoc	1	1 Vrt/Node, 2 Vrt/Link, Multiple Vrt/Polygon or Cloud
Connection		
Interface Rank	Primary	Primary/Secondary
Interface Category	Fiber	Coax/Fiber/CatN/Wireless
Interface Type	FC/PC	FC/PC, APC, JS
Wavelength(s)	1550	nm
Power Level(s)	10	dBm
Tx/Rx Mode	Full Duplex	Tx Only, Rx Only, Full Duplex, Half Duplex
Downtime Counter		sec
InterfaceID	43	Unique within Element
Vrt Assoc	1	1 Vrt/Node, 2 Vrt/Link, Multiple Vrt/Polygon or Cloud
Connection		
Interface Rank	Secondary	Primary/Secondary
Interface Category	Coax	Coax/Fiber/CatN/Wireless
Interface Type	F	F/KS
Level wrt Primary Fwd	NA	dB
Level wrt Primary Rev	NA	dB
Tx/Rx Mode	Rx Only	Tx Only, Rx Only, Full Duplex, Half Duplex
Downtime Counter		sec
Intended Paths	10-11, 10-20, 11-20, 11-10, 20-10, 20-11, 40-10, 40-11, 40-20, 41-10, 41-11, 41-20, 42-10, 42-11, 42-20, 43-10, 43-11, 43-20	

Figure 6 - Layer 1 Granular Data Representation for CMTS Data Element Model

Figure 7 shows the MAC layer and the TCP/IP layers portion of the CMTS element model that has the characteristics for this proposed data model. For the purposes of illustration only a few of the many interfaces in the CMTS data element model are shown. It is worth noting that each layer has connectivity information that is likely different the other layers of the same element. This is highlighted in Figure 2 as the participation of elements in the different layers is different and the resulting connectivity within a layer varies.

Elements such as a splitter, a tap, a fiber node, a coaxial cable segment participate are only described in the PHY layer, while elements like a cable modem or an Ethernet switch are described in the PHY layer and the MAC layer. An element like a CMTS or a home WiFi router are described in the TCP/IP layer in addition to the lower layers and elements like video servers, an EMTA or a PC are described using all layers.

Layer 2	MAC		
MAC Element Type	Node		
Number of Interfaces		9	(ports)
	MAC InterfaceID	1	Unique within Element
	Phy IF Assoc	10	
	Connection		
	Interface Rank	Primary	Primary/Secondary
	Interface Category	Eth	EPON, GPON, Eth, DOCSIS, 802.11, LTE
	Interface Type	10G-Eth	10BT, 100BT, 1G-Eth,10G-Eth,100G-Eth,
	MAC Address		
	Downtime Counter		sec
	MAC InterfaceID	6	Unique within Element
	Phy IF Assoc	43	
	Connection		
	Interface Rank	Secondary	Primary/Secondary
	Interface Category	DOCSIS	EPON, GPON, Eth, DOCSIS, 802.11, LTE
	Interface Type	D3.0	D1.0, D1.1, D2.0, D3.0, D3.1
	MAC Address		
	Downtime Counter		sec
Intended Paths		Any-to-Any	
Layer 3	IP		
L3 Element Type	Node		
Number of L3 Interfaces		12	
	L3 InterfaceID	1	Unique within Element
	MAC IF Assoc	1	
	Connection ???	???	
	Interface Rank	Secondary	Primary/Secondary
	Interface Category	IP	IP, other
	Interface Type	IPv4	IPv4, IPv6
	IP Address		
	Subnet		
	Class		
	Gateway		
	Downtime Counter		sec
	L3 InterfaceID	12	Unique within Element
	MAC IF Assoc	1	
	Connection ???	???	
	Interface Rank	Secondary	Primary/Secondary
	Interface Category	IP	IP, other
	Interface Type	IPv4	IPv4, IPv6
	IP Address		
	Subnet		
	Class		
	Gateway		
	Downtime Counter		sec

Figure 7 - Layer 2 and 3 Granular Data Representation for CMTS Data Element Model

The data model just described has the characteristics of being; granular, layered, connection oriented and end-to-end analysis friendly. Next we exploit the use of the database and the relationships among its elements to implement in advanced network management and monitoring tasks.

USE CASE SCENARIOS LEVERAGING NETWORK MANAGEMENT APPROACH

Several use case scenarios of advanced network maintenance and management exercises that leverage the proposed data model are described.

First Use Case Scenario – Data Network Performance Analysis

In the sample scenario where gaming and video streaming bundled services are simultaneously supported, the minimum performance index supporting both services is performance index 4. In a group of devices that are evaluated the performance index varies around a mean. The mean is calculated from the group of devices that share certain topological characteristics (a node, MAC domain, a branch within the coaxial network etc.)

Figure 8 shows a sample data collection of different serving groups corresponding to multiple nodes and the multiple DOCSIS MAC layer domains under evaluation. The measured performance of end devices from different serving groups are distributed across a performance index range. The color coding is such that if two sigmas (standard deviations) of the population are below the performance threshold it is red. If two sigmas are above the performance threshold is green and yellow in other cases.

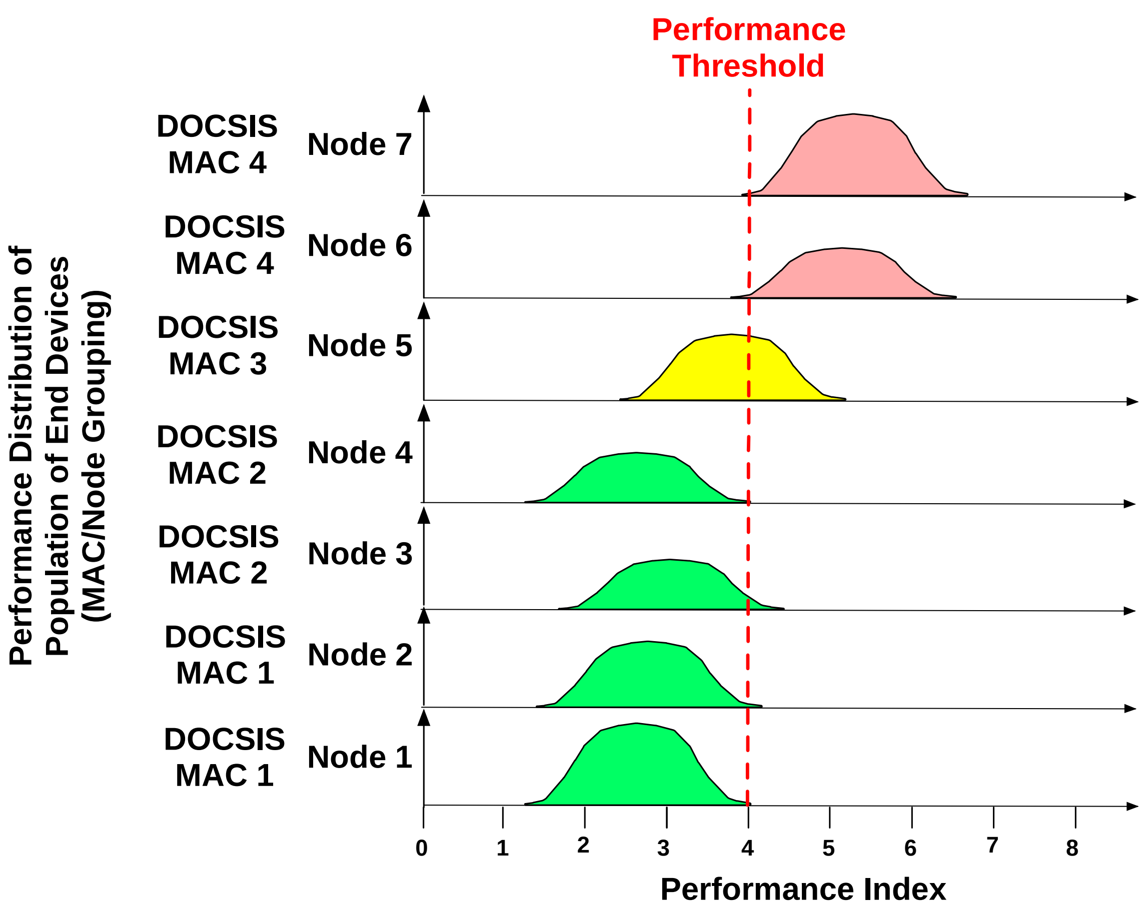


Figure 8 - Gaming and Video Streaming Performance Assessment

Figure 8 indicates that the latency problem extends across two fiber nodes, two isolated physical layers which share a common MAC layer domain. This symptom provides a high likelihood that this problem is a MAC or IP Layer problem. The analysis to troubleshoot and isolate the problem can be automated through relational database manipulation when performance metrics and topology information are stored in a granular database.

Figure 9 shows an example of how a hub network configuration. This configuration schematic highlights how the performance information from Figure 8 can be used to determine the source of the problem. A database with connectivity information across different layers with the associated performance data can be used to determine the source of the problem.

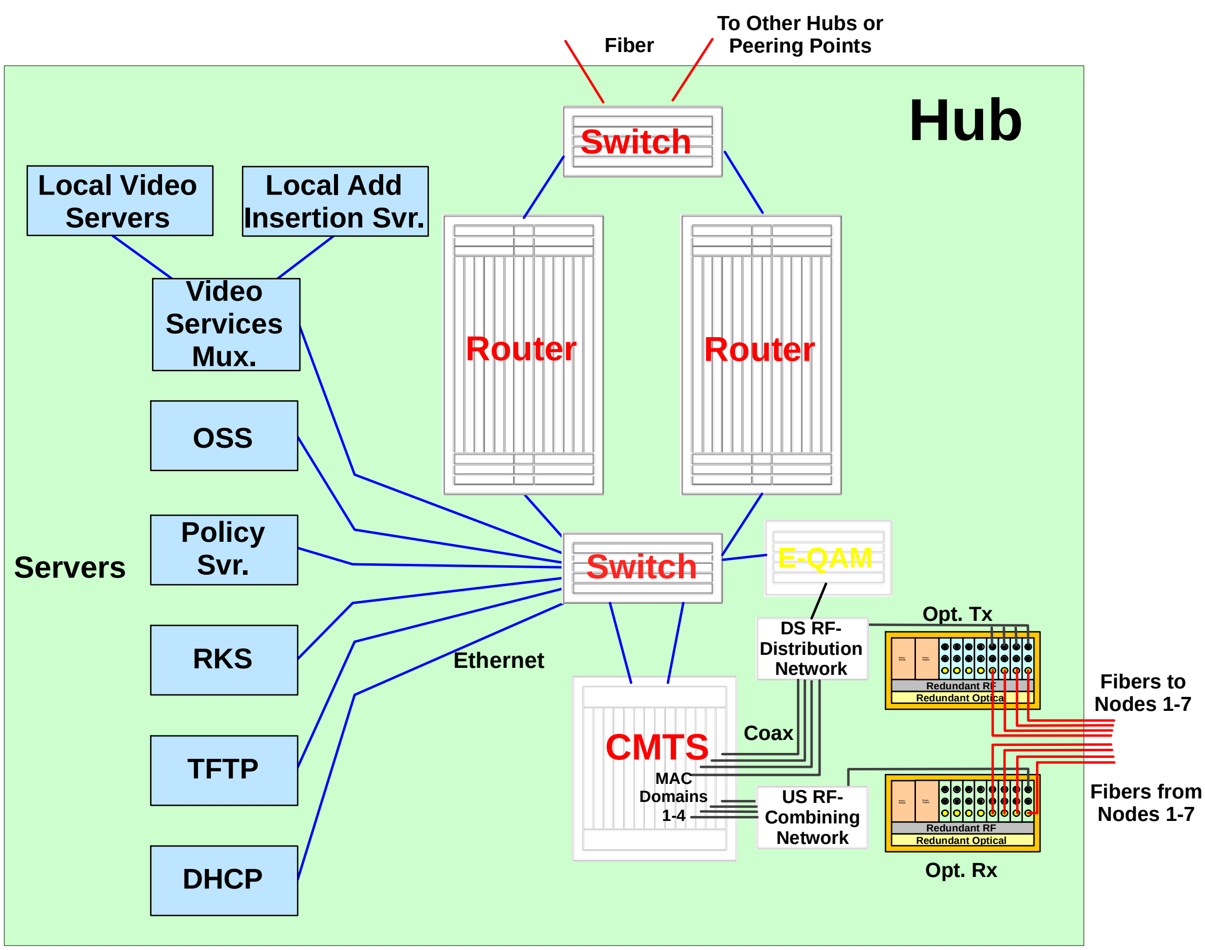


Figure 9 - Hub Video and Data Network Components

Subscriber service performance results of the network portion playing a role in delivering data services across different nodes within a Hub are shown in Figure 8. These performance results correspond to the specific Hub shown in Figure 9. The symptoms of nodes 6 and 7 are not observable in other Hubs of the regional aggregation network shown in Figure 10. In addition, traffic records indicate significant usage in nodes 6 and 7. This leads to believe that the cause of the problem is a layer 2 congestion issue within the MAC Layer domain shared by nodes 6 and 7.

The connectivity information of the regional network shown in Figure 10 could have been used to troubleshoot problems in the regional network. Many time when analyzing end-to-end services, the performance of the service has to be traced back from the local cable networks to the aggregation networks and potentially to the egress points of the cable network into other providers network and the public internet. It is important to discriminate the source and location of the problem of the problem so that action can be taken. In the case of problem responsibility lying outside the network this could range from communication to the third party network provider to reassessing peering arrangement that could address performance issues detected.

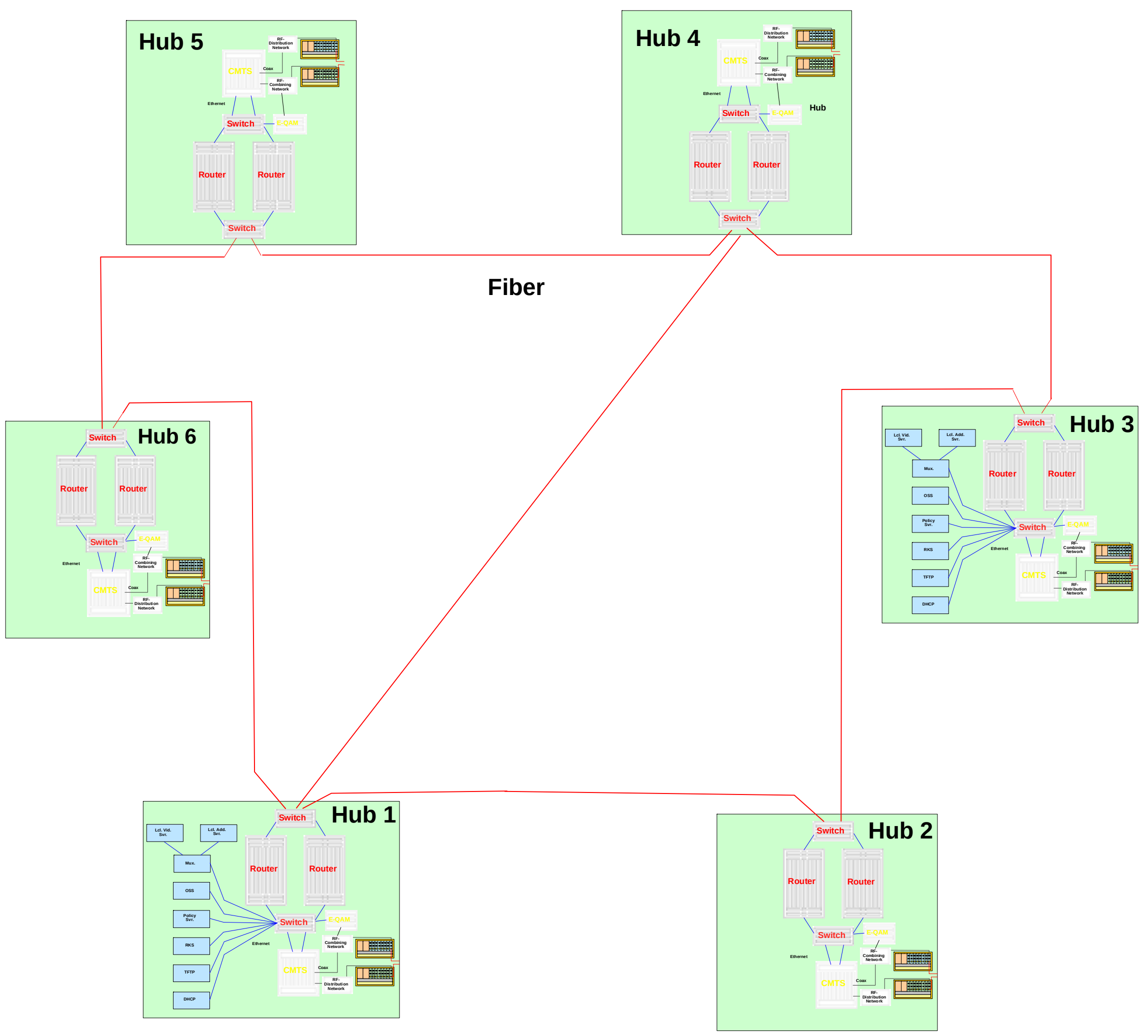


Figure 10 - Regional Cable Network

Second Use Case Scenario – Detection and Localization of Linear Distortions

Pre-equalization analysis has proven to be very effective at detecting and localization US impairments [3],[4]. This process involves analysis and correlation of data from CMs that share the same RF environment such as the same RF channel and optical node. The distortion signatures obtained from the CM pre-equalization coefficients are unique for each distortion within the node. Grouping the CMs that share the same impairment combined with topology information allows the determination of where the problem originates. Topology, hence device connectivity information as an integral part of the management databases becomes key at automating the fault location processes and determining impairment impact.

Figure 11 shows the linear distortion grouping results obtained from a node. This amplitude distortion versus frequency view of the channel by group of impacted CMs represent only a subset of the CMs with distortion.

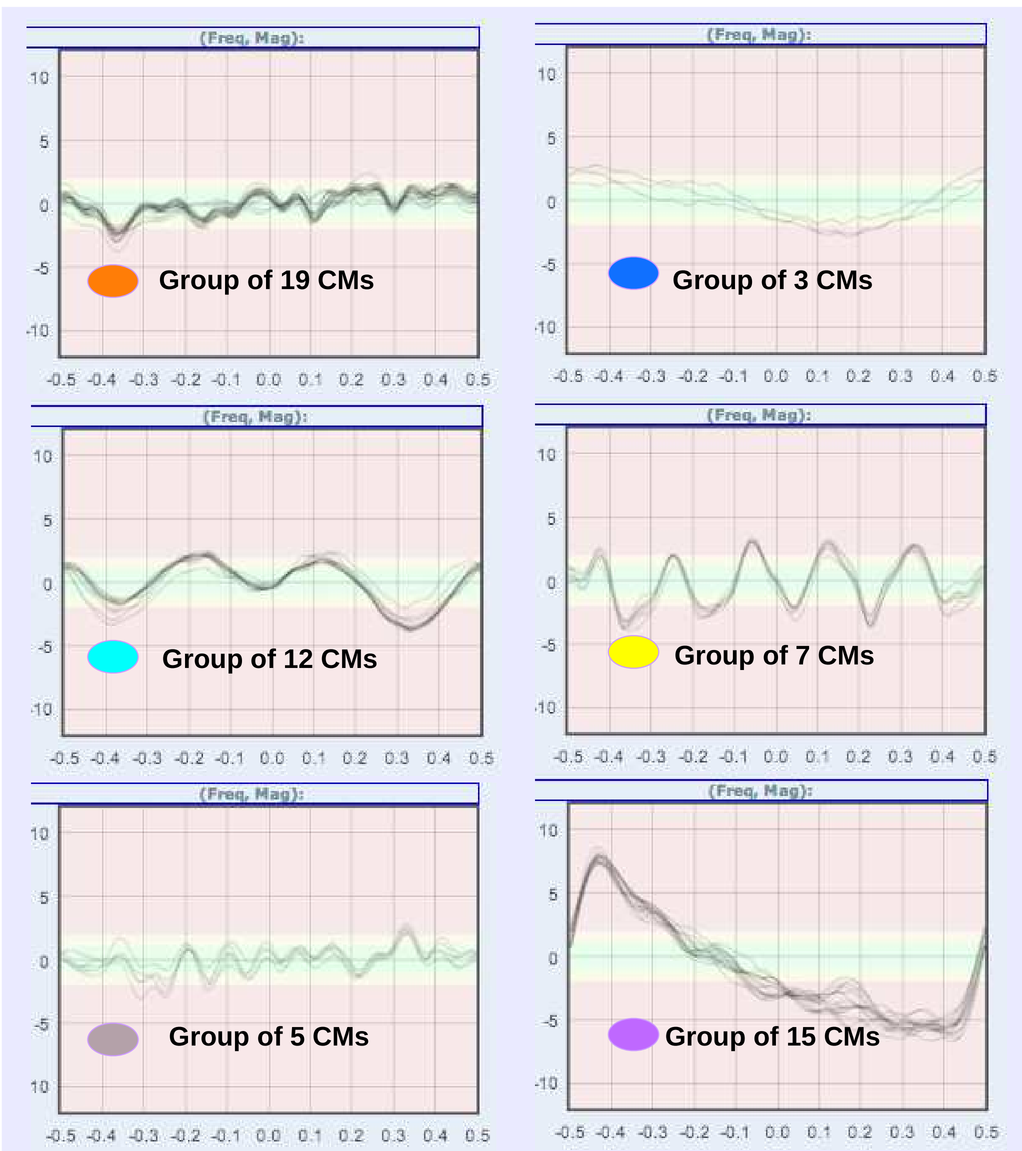


Figure 11 – Correlated Equalization Data from CMs in Node

Individual CMs with distortion are not shown in Figure 11 for brevity, but this determination is also important because that indicates an impairment in the drop/home portion of the network which should be addressed by a installer rather than a line technician.

Figure 12 shows a logical representation of an optical node with CMs that are color-coded based on the linear distortion grouping conducted for Figure 11. It is clearly seen that when the impairment correlation is combined with topology information, the determination of the location of the impairment is intuitive. It is the boundary (indicated by red Xs in Figure 12) between the topology regions that share the same impairment and the ones that don't.

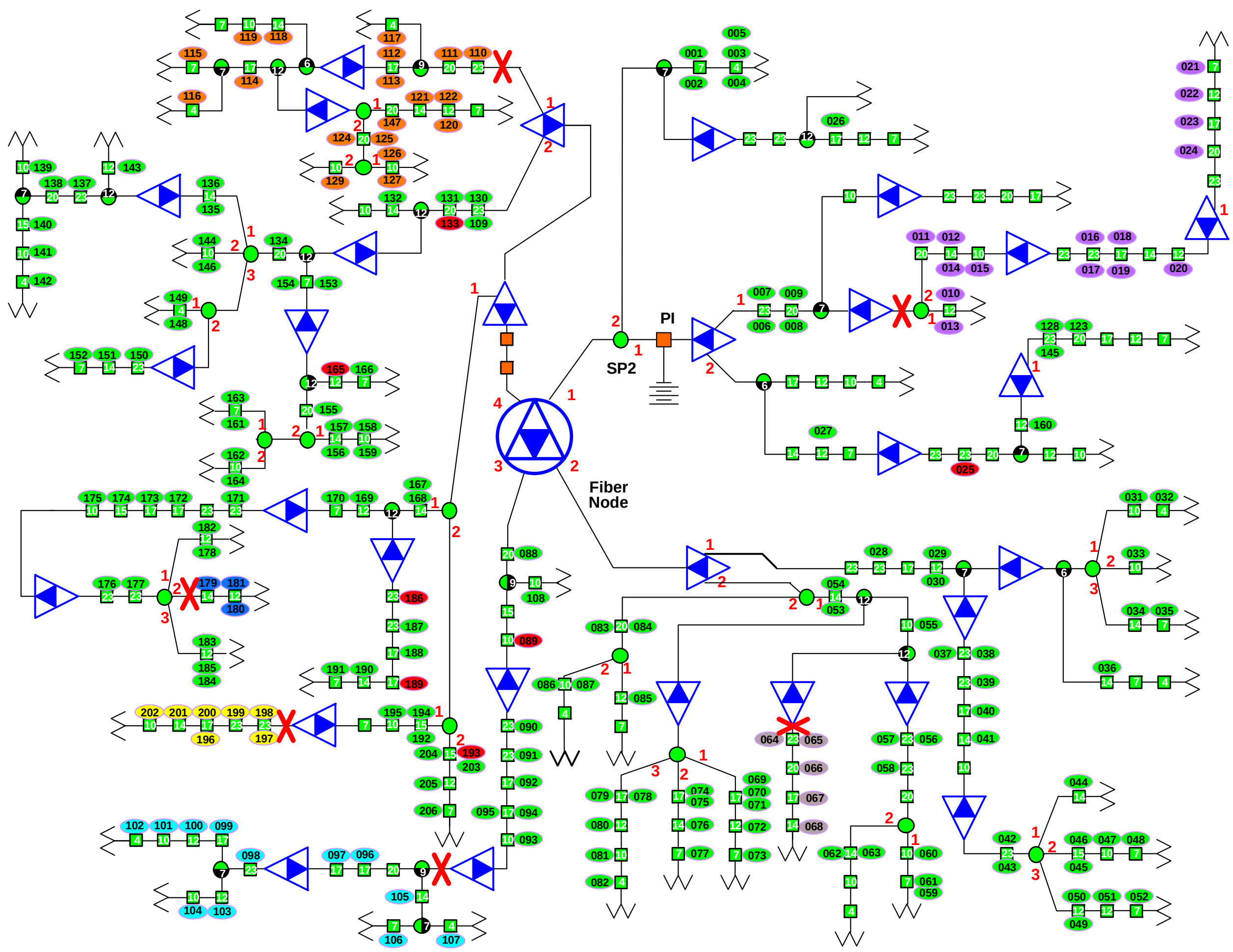


Figure 12 - Fiber Node Topology Representation Grouping Subscribers with Common Distortion

The color code used in Figure 12 is as follows; Green CMs have negligible distortion, Red CMs have measurable distortion but their distortion does not correlate with distortion of any other CM. Other colors indicate measurable distortion that correlates with the distortion of the CMs with the same color code.

Since this grouping technique also indicates how many CMs share the same impairment. Assessing severity by the number of CMs affected or type of customers affected can be implemented. Typically CM pre-equalization coefficients fully compensate for linear distortion impairments. When there is full compensation there is no other means for detecting that something is wrong with the network. This also facilitates a proactive network maintenance strategy since the pre-equalizers buy time for the operator to decide when to fix the problem. Incorporating distortion grouping information in a topology aware database can be used to automate the impairment discovery process.

<u>Third Use Case Scenario – Assessment of Resources to Optimize Node Splitting</u>

Capacity in cable networks is typically measured with coarse granularity in time and by grouping individual user consumption on a node basis. An hour or 15 minute time granularity may hide through averaging short term capacity events. Likewise aggregate node consumption average out the traffic requirements few specific users may have. When traffic trends in a node project capacity starvation in a node, a node split is planned. This node split is typically done based on number of end devices covered.

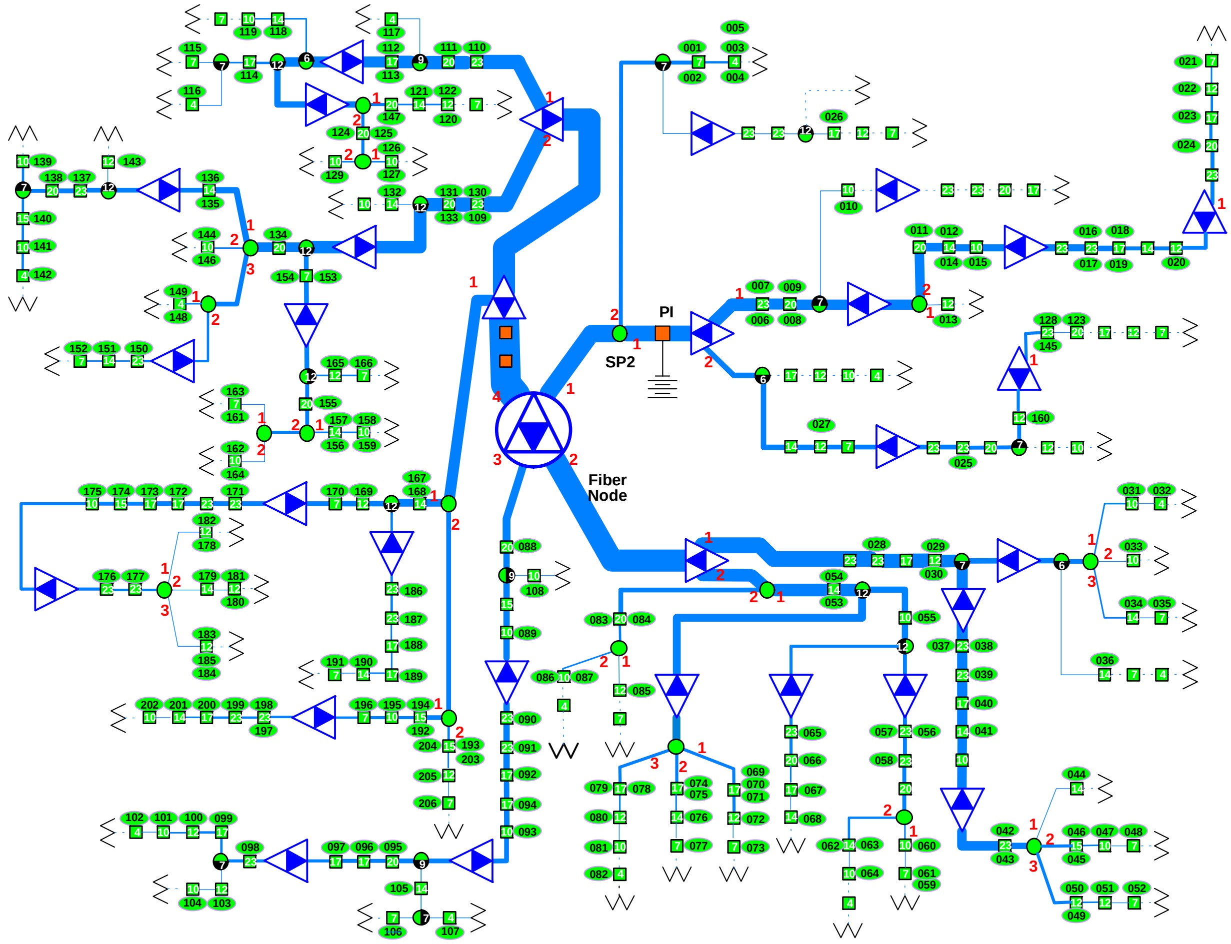


Figure 13 - Fiber Node Topology Representation Showing Subscriber Data Traffic Consumption

A different approach would be to have a node divided based on equal consumption. This can be implemented by monitoring end device consumption and correlating this information to physical network topology [4]. End device consumption information included as a data model element of the end device allows the automatic determination of traffic consumption for the nodes and any physical subset of the node.

Figure 13 highlights such an approach. In this case the individual consumption of users is aggregated as flows superimposed to topology to determine the actual consumption through each physical path. The time domain granularity increase the accuracy in determining that a channel is congested.

Service Availability and Reliability

A variety of use cases which benefit from granular network management and leverages connectivity and network layer awareness. In addition to impairment detection and localization, capacity planning, and node splitting applications, it can also be used for service availability, reliability and quality of experience assessment. In addition the insight in the network topology and the resources knowledge allows the determination of what type of business services can be provided and what SLAs can be met.

Statistical Process Control

Another potential application for granular data analysis can be derived leveraging the concepts of statistical process control promoted by Western Electric [5]. In this use-case the end-devices that are “out of control” or that not follow the expected behavior within 1, 2 or 3 standard deviations (sigmas) are identified within the topology and the same correlation exercises can be used to determine where within the network the problem originated.

Feedback Mechanisms

The use of subjective feedback from customers is very important. The increasing use of social networks as feedback tools can also be leveraged to determine customer satisfaction. Is that feedback concentrated in certain areas or is it generic to a system or a region? The same techniques to handle objective metrics and determine problem location can be also be used with subjective feedback.

The CM, STB and MTAs are quite intelligent and great probes to leverage for feedback on network health and to measure performance on services and applications. The most important characteristic of these devices is their ubiquity across the network. They are already in the customers home and can provide a wealth of information. This information can be included in the granular databases and used to manage the network. In many instances the cable industry has not taken advantage of many useful parameters available on these devices.

Operational Practices

The increasing amount of data gathered in the field requires also a change in the cable industry operational practices. We have to make sure that our databases reflect accurate information of our network. In particular for maintaining the HFC network, some manual processes are required. As part of the maintenance practices, cable technicians are changing and replacing components in the field. They are adjusting amplifiers, measuring plant performance and are exchanging end devices from customers. All these events have to be recorded to calculate reliability, estimate impact under impairments, to know the actual components in the network, to estimate resources such as available taps ports and available fiber, etc.

The technicians would have to enter greater amount of information to keep track of the changes in the level of detail that is envisioned. Technicians would also have to be empowered update the databases or through an automated verification process a technicians proposed changed would quickly validated by a supervisor.

Data Model Structure Standardization

Cable system operators have similar architectures, components, services they provide, operational practices, etc. Based on the similar environment and needs, it is advantageous to use similar databases using the same data model structure. The elements in the data model may vary from operator to operator but if it has the same structure, common management techniques can be implemented. In this paper several characteristics of a data model have been advocated. The main ones have been; granularity both in number of elements covered and in detail of description of each elements described, connectivity awareness, network layer awareness and end-to-end analysis friendliness. Standardization of a data model with these properties that is also extensible, is very useful for the cable industry. Standardization enables use of similar tools and processes and reduces the cost of implementation of management systems.

CONCLUSION

This paper promoted a significant increase in the level of detail used to manage the cable network. It proposes to achieve a balance between reduction of data, by collapsing the different performance metrics into a few metrics and an increase in the number of elements and element parameters being managed. The analysis that is shown in DOCSIS® 3.0 indicates that two important performance metrics such as latency and packet loss are strongly correlated. Other performance metrics, in particular the ones used by the upper layers can be correlated with the fundamental lower layer performance metrics.

An in-depth knowledge of the network is also advocated. It is important to understand that the network is a fundamental part of the user experience. The amount of detail information proposed for collection along with the ubiquity of all cable end devices, provide significant insight into the network, the services and the customer. Control and management of the network enable management of the user experience. An example of a comprehensive granular data model has been shown. This data model is layered and is connection aware. Connectivity information is leveraged in use case scenarios for localizing linear-distortion impairments, estimating resources for node splitting applications and determining congestion points in the network. All of these use cases through the granular data model lend themselves to automation.

The unified management system approach presented here, which promotes breaking down the management system silos, facilitates end-to-end analysis. Leveraging a standard data model structure is fundamental for using similar tools across the industry and to leverage economies of scale.

The cable network has to be friendly to evolution and innovation, and for that it has to be open to leverage its resources for services and applications. A management system that is highly automated, comprehensive and granular enables that open environment.

ACKNOWLEDGEMENT

The authors would like to thank Robert Cruickshank III for his encouragement and insightful feedback in writing this paper.

REFERENCES

[1] "Pre-equalization based Pro-active Network Maintenance Methodology", SCTE Cable Tec Expo 2008 by Alberto Campos, Eduardo Cardona and Lakshmi Raman (June 2008, Philadelphia, Pennsylvania)

[2] "Holistic Approach to Evaluating Quality of Experience" SCTE Cable-Tec Expo 2011, L. Alberto Campos, Jennifer Andreoli-Fang, (November 15-17, 2011, Atlanta, Georgia)

[3] "Progress on Pre-equalization Based Proactive Network Maintenance" SCTE Cable-Tec Expo 2011, L. Alberto Campos, Eduardo Cardona, Tom Williams, Larry Wolcott, Phillip Chang, Todd Szuter, (November 15-17, 2011, Atlanta, Georgia)

[4] "Proactive Network Visibility: Utilizing Service and Device Mapping for Outage Response and Network Planning" SCTE Cable-Tec EXPO 2008, Robert F. Cruickshank III (June 2008, Philadelphia, Pennsylvania)

[5] "Statistical Quality Control Handbook" by Western Electric, AT&T Technologies, 1984 2nd Edition.

Cloud-Based DVR and Multiscreen Support Strategies – Optimizing Storage and Transcoding

Ludovic Milin, Carol Ansley
ARRIS

Abstract

This paper presents extended and recent usage statistics from in-home multi-room/multi-tuner DVR deployments with an emphasis on characterizing subscriber behavior.

These statistics are then extrapolated to formulate sizing assumptions for a typical cloud-based DVR deployment, including ingest, storage and streaming capacities.

INTRODUCTION

The advantages of Cloud, or Network, based DVR have been well documented over the past five years. From a cost and operational perspectives, it is perceived as overall cheaper to deploy and operate than in-home DVR, requiring no truck roll to deploy or fix in-home DVR. An additional benefit is the flexibility resulting from a virtually unlimited number of tuners and completely scalable storage space. An nDVR service offer can be changed overnight to add virtual tuners or allow the customer to easily increase the amount of storage they are paying for without any physical changes in the home. Once the storage and playback infrastructure is in place, it also provides easy extension to multi-room, multiscreen services as well as offering potential OTT integration. An operator can seamlessly augment, then replace legacy in-home DVR with new services utilizing the new nDVR resources that can offer a number of ways for the operator to increase its revenue sources, from easy roll out of advanced features like TSTV, Catch-up TV, Pause live TV, to controlling and monetizing advertising on playback, for example preventing ad skipping or replacing existing ads with better targeted alternatives, providing detailed and complete visibility into subscriber playback viewership behavior, and increasing ARPU at high margins for additional recording tuners, disk space, etc.

To date there have been few full scale network DVR deployments and it is difficult for both vendors and operators to accurately predict the amount of storage space or the video ingestion and playback capacities required. While this can be worked around through progressive and controlled rollouts, accurate planning for facilities and operations as well as budgeting for the entire deployment can be difficult. In this paper, we present some study results from a current whole-home DVR deployment with Buckeye that relate well to nDVR considerations.

MULTI-ROOM/MULTI-TUNER DVR USAGE STATISTICS

The results below come from an analysis of the DVR activity for 1,014 households across 8 representative days in 2013. Each home has 6 tuners available for recordings, and had between 1 and 6 IP set top boxes used to playback the recorded content. Each recording and each playback event was recorded in the utility logs of the system allowing later study and research into user behavior, such as this study.

Recording Behavior

Two areas can be studied on recording behavior: what did people record, and when did they record it. The first topic is important because of the current state of content licensing regulations. For optimum storage efficiency, a single common copy would be kept of every unique piece of content, but under current regulations and court decisions, a single copy model can only be used if the content provider gives the operator authorization. The choices that people make about what content they are interested in recording affect the relative value of those licensing arrangements to an MSO who is contemplating deployment of nDVR.

The second topic, when do the recordings tend to happen, affects the content ingestion scaling necessary for an nDVR deployment. If the operator must keep separate copies of a piece of content for each subscriber that records it, that affects the overall storage capacity clearly, but it also affects the scaling needed to drive a potentially very large number of recording simultaneously into multiple individual subscriber virtual disks.

Looking first at what content people chose to record, the chart below categorizes the content by unique title. In this view, the dominance of popular content can clearly be seen.

Table 1

	Unique Programs	Percentage of Programs Recorded	Copies	Percentage of Recordings
Recorded once	5585	40%	5585	9%
Recorded twice	2745	20%	5490	8%
Recorded 5 or fewer times	11889	85%	21804	34%
Recorded more than 5 times	2128	15%	42896	66%

Slightly more than 14,000 unique programs were recorded across the 1,014 households, but the distribution of those recordings was heavily weighted toward the most popular content. Only 15% of the unique programs accounted for more than 66% of the actual recordings. In chart form, the information can be seen in Figure 1.

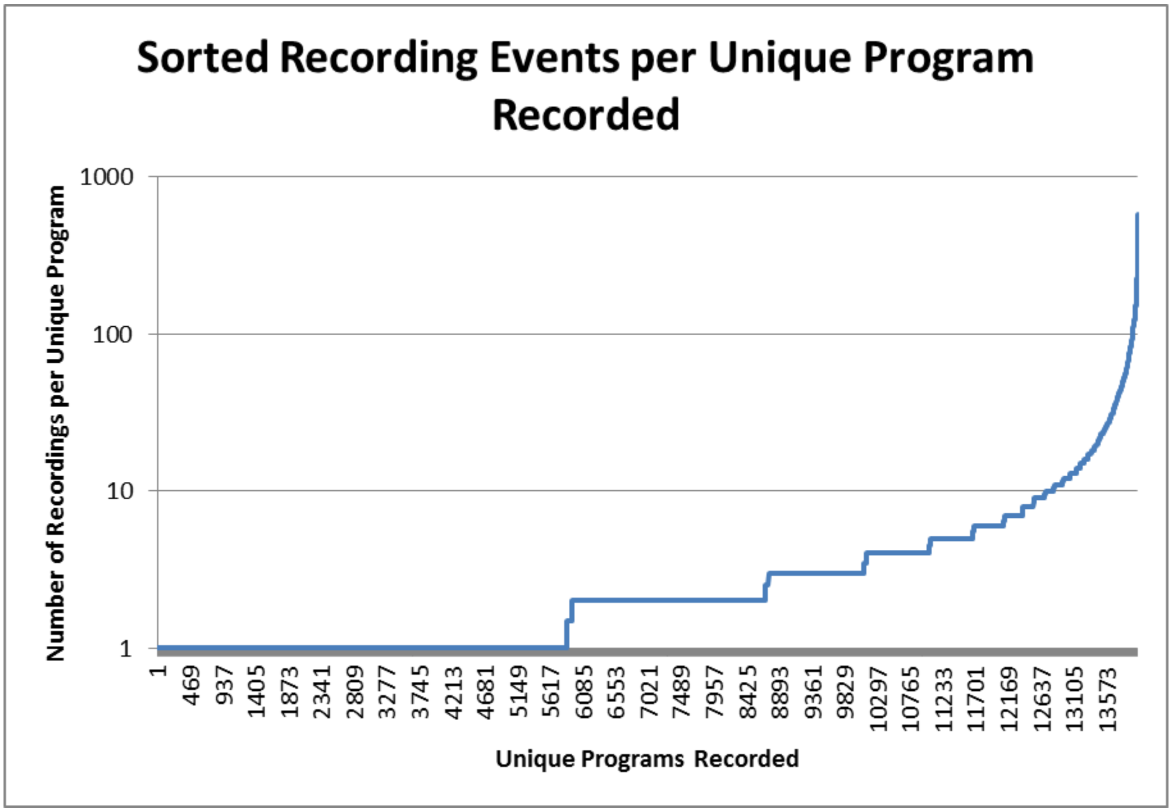


Figure 1

A logarithmic scale is used to make the data more accessible since a linear plot just hugs the axes.

Turning to a consideration of when recording activity happened, the observation that the most popular material dominated the recordings is also reinforced by an analysis of the times when recordings occurred in Figure 2. Recording activity peaks hugely during primetime. Smaller peaks also appear in the afternoon. Since subscribers have the ability to record up to 6 simultaneous programs, Table 2 below will provide details on how many concurrent sessions to expect per subscriber.

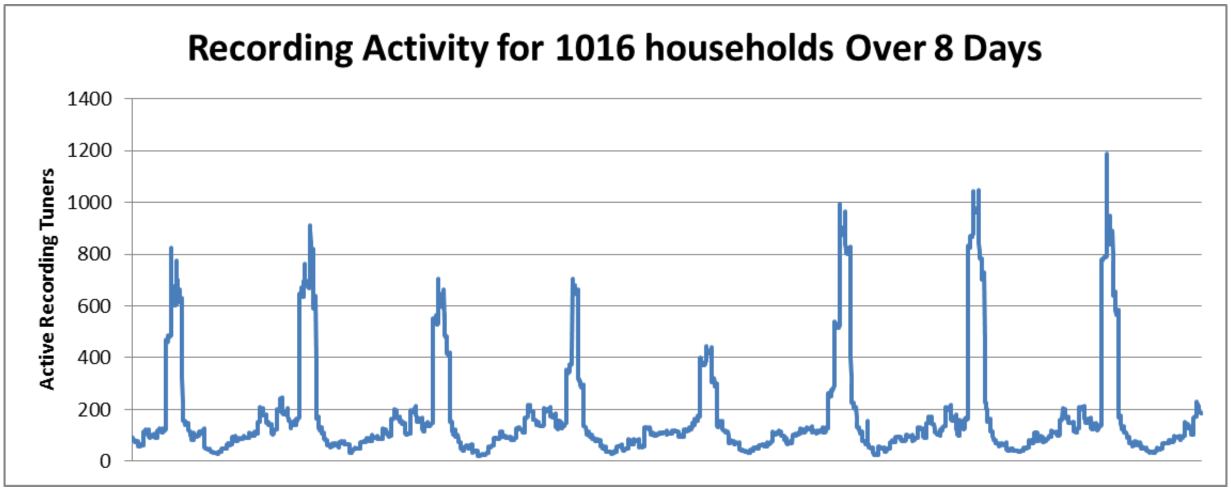


Figure 2

If we consider individual household's behavior, we see similar patterns with peaks of usage during primetime, but when individual users are looked at, there is much more seeming randomness to their behavior.

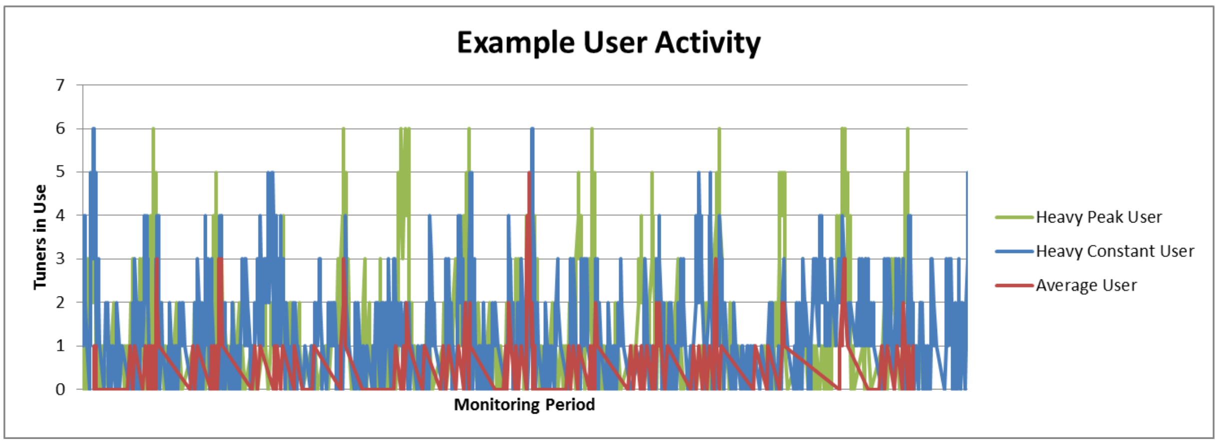


Figure 3

These subscribers were chosen as examples of some typical user behavior. The average subscriber tends to record a few programs every day and rarely uses more than 3 tuners. It is interesting to note also that the distribution of subscriber use is skewed with the arithmetic average or mean some distance away from the median value, which is the value in the center of the data. For this data set, we found the following:

Table 2

Number of Sessions	One Tuner	Two Tuners	Three Tuners	Four Tuners	Five Tuners	Six Tuners
Average	49.7	24.7	11.2	4.9	1.9	0.4
Median	35	14	3	0	0	0

The average subscriber made over 90 recordings over the 8 day period. The spread between average and median implies that the relationship between the heaviest users of the DVR feature and the lightest users is non-linear. When we looked at the amount recorded, the relationship was also seen.

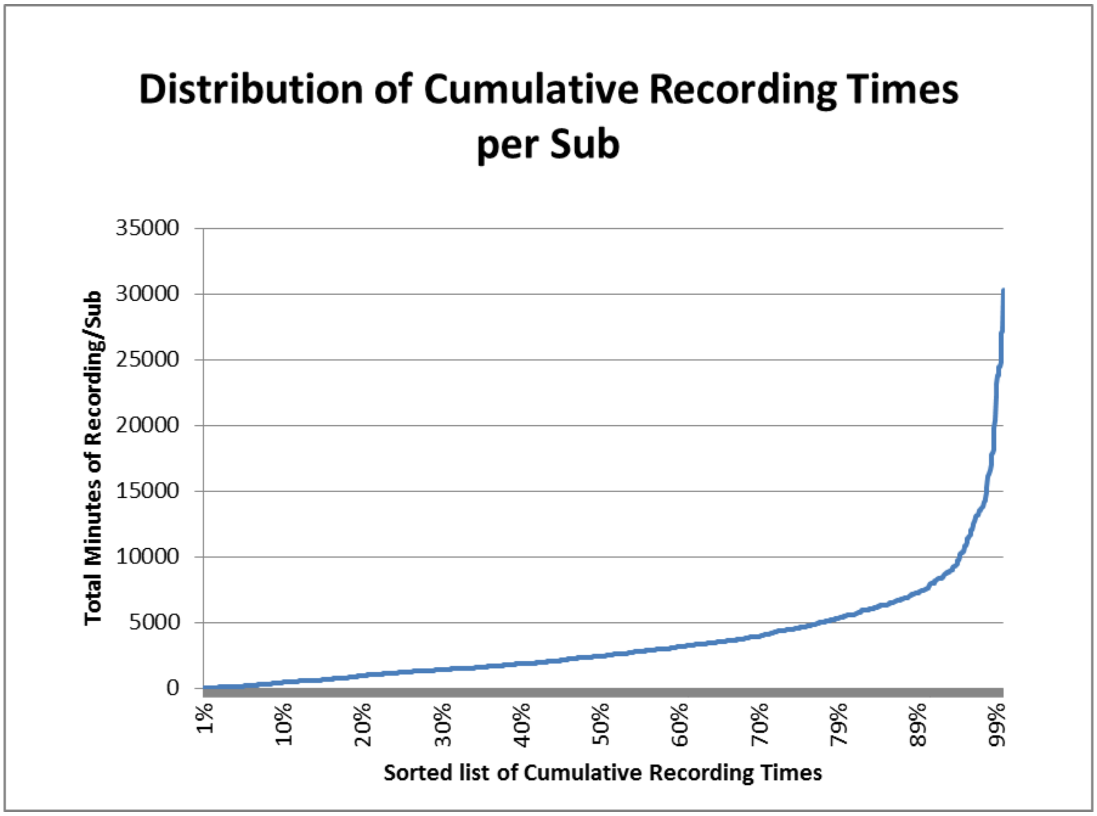


Figure 4

The practical impact of these results is that an nDVR system should be scaled assuming that at least 5 to 10% of the subscribers will use the system to its published limits for active tuners and storage capacity, even though the majority of users will probably use it relatively lightly. On the positive side, about 80% of subscribers recorded fewer than 5000 minutes of content in the 8 day period, which, translated to the worst case of full resolution AVC HD content, only represents approximately a maximum of 300GB of storage. Not that the STB in their possession allows significantly more than this, so there is no bias introduced by hardware limitations. Also note that for those subscribers recording more content over the time period than their STB hard disk capacity allows, they either recorded a significant number of shows in SD resolution, or they deleted content on a regular basis to make room for new recordings.

Playback Behavior

Playback behavior also can have substantial real world implications. The number of streams active simultaneously scales both the nDVR playback infrastructure required to generate and manage the streams as well as the network bandwidth required to carry that video to the end subscribers.

Playback activity also peaked during primetime, with lesser amounts of activity during the rest of the day. Below is a chart showing DVR viewing activity spread across the week.

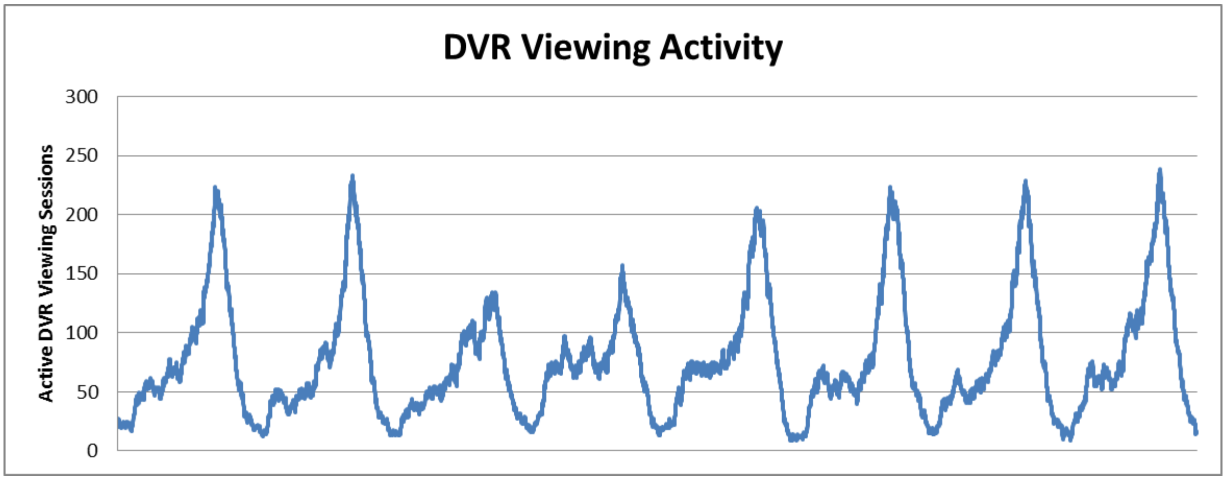


Figure 5

Overall the number of playback sessions varied greatly across the subscribers studied. The majority of the subscribers recorded more content than they played back; overall there were 64,701 recordings made with 53,950 playback sessions for a usage ratio of 77.5%. But, again the individual subscribers' behavior across this week varied widely as shown in Figure 6.

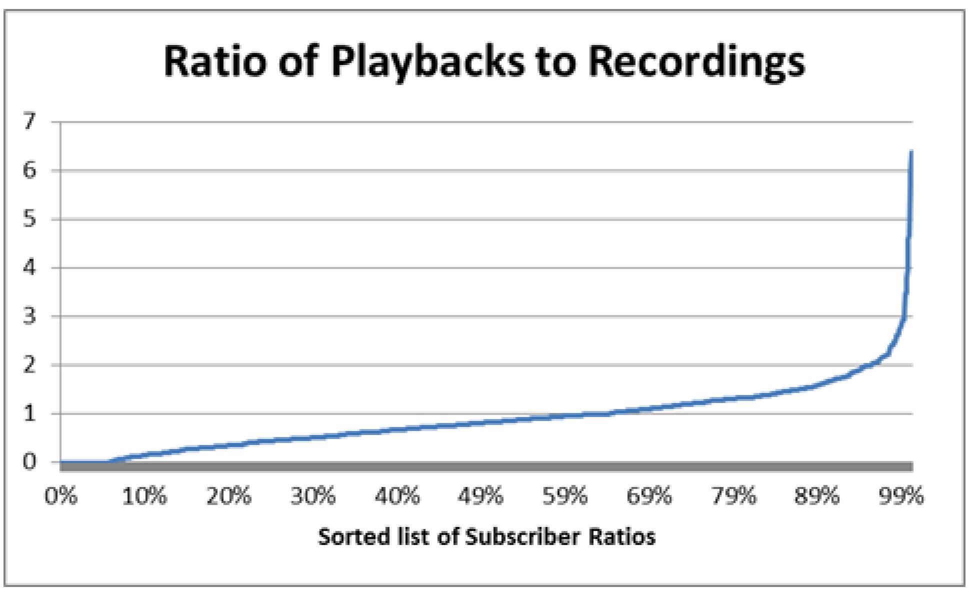


Figure 6

Figure 6 compares the ratios of playback to recordings for the 1,014 households in the study. One issue for this study is that it can take only a snapshot of the activity during the week. Some households showed only recording sessions, others showed only playback sessions – these households were generally light users. Overall about 70% of subscribers record more content than they watch. If the busiest households for DVR activity are considered the ratio of playbacks to recordings is 96.8%.

Age of the Assets

While it is not currently possible to track exactly the time between a recording and its playback or the time between a recording and its deletion, it is possible to report on the number of assets recorded and viewed during the same week. For the period under analysis the number of events in each of the recording and playback categories is shown in Table 3.

Table 3

Events	Total	Unique	Viewed in week	Not Viewed in week	Viewed in week, recorded earlier
Recorded	64,701	14,017	41,737	22,964	
Playback	53,949	11,283			12,212 4,190 Unique

Fewer than 22.6% of the playback events are for assets recorded over a week prior, and out of those less than a third are for unique content. This would appear to make 7 days a good initial candidate for migrating undeleted recorded assets to archived storage.

Recording and Playback Sessions Duration

The distribution of the recording session duration is shown in Figure 7. As one would expect, two major peaks at 30 and 60 minutes (plus 3-4 minutes of automatic buffer at both ends of the asset) account for almost 70% of all recordings. A minor spike can also be seen at 2 hours.

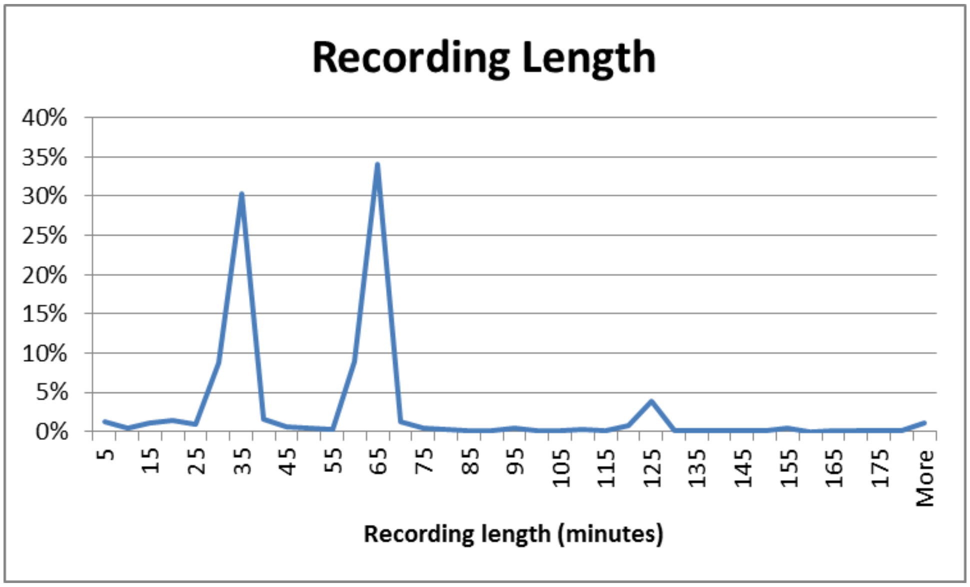


Figure 7

The distribution of playback session duration in Figure 8 on the other end is not so well delineated. It appears that a large number of sessions are abandoned within the first 5 minutes. Then there is a spread to the left of each of the 30 and 60 minutes duration peaks which can likely be attributed to subscribers skipping ads during playback and terminating the session prior to screen credits or similar closing content.

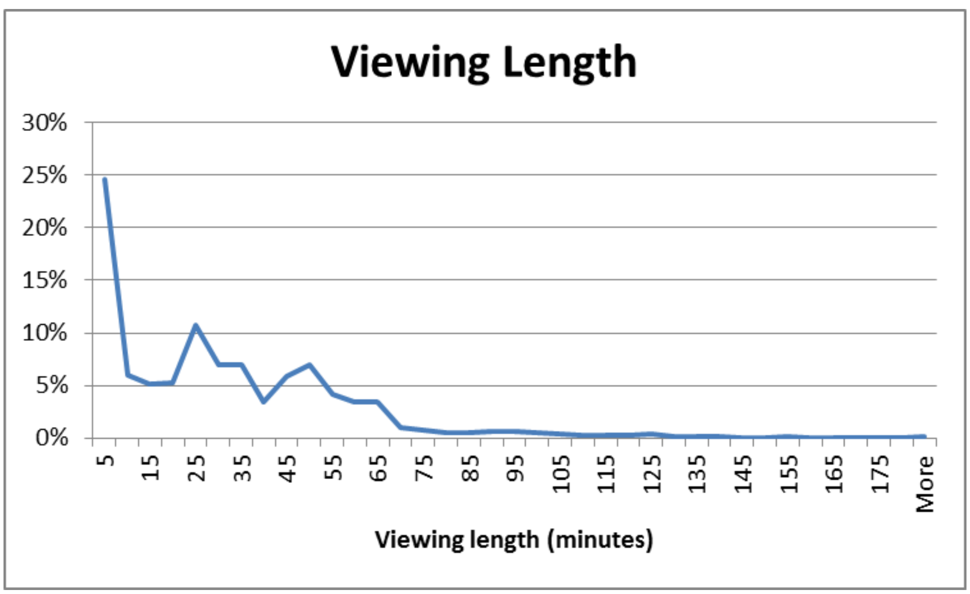


Figure 8

SIZING A TYPICAL CLOUD BASED DVR DEPLOYMENT

Based on the home DVR data in the previous section, we can propose sizing models for ingest, storage and playback capacity. The initial assumption is that all content is stored as unique/personal copies.

Storage Requirements

One approach is to size the storage based on the individual quota allotted to each subscriber. However, there are some economies of scale to be had when considering the statistical behavior of a large population of subscribers. As shown in Figure 4, 90% of subscribers typically use less than 400GB of their local 500GB storage. When looking at a large system with a well-tuned archival mechanism, the savings will add up very rapidly and translate either into less upfront CAPEX spending from purchasing 20 to 25% less storage capacity, or additional revenue opportunities by being able to offer increased storage capacity options for a fee to subscribers.

I/O Capacity

At peak time, Figure 2 shows that there are at least as many active recording sessions as there are subscribers, possibly up to 20% more. From a recording perspective, an ingest system should be designed for being able to record at least 8-9Mbps per subscriber, assuming HD AVC content, and two to four times more when planning multiscreen support with multiple profiles being recorded simultaneously. However, Table 1 shows that most of the content is actually duplicate sessions for popular programs, so the network I/O requirements decrease significantly, by approximately 70% to about 3Mbps, but the burden on content replication and disk writes does not go down.

On the playback side, Figure 5 shows that the number of active sessions only reaches about 25% the number of subscribers; on average this nDVR system should be designed to support an average of about 2Mbps per subscriber disk read and network I/O for HD AVC.

Content De-Duplication

When moving the recorded content to archive after a few days, seven days appearing to be a good initial setting based on the data in the previous section, one could possibly de-duplicate the recorded sessions and move to a single/shared copy store. There are, however, a few things to remember when planning this: since every subscriber can typically customize the offset for start and stop recording times, every recording of the same show could actually be a slightly different asset, and as a consequence the de-duplication process may not be as efficient as initially hoped.

Based on the numbers in Table 3, with a total of only 14,017 unique assets out of 64,701 recordings, one would expect to save approximately 78% of the required storage space from de-duplication. Even if this is only performed when archiving, the savings should be substantial.

Adding Multiscreen Support

Multiscreen support typically requires storing and streaming multiple bitrate versions of each asset. More profiles typically provide higher resolution and better video quality where possible, and on the other end also support more degraded network conditions. Additional to the increased transcoding and transport costs, increasing the number of profiles also impacts nDVR storage, ingest I/O and replication capacities, so subscriber experience must be weighed.

We first consider a particular operator five 16:9 AVC profiles use case as described in Table 4.

Table 4

	Resolution	Bitrate (kbps)
Profile 1	1280 x 720, L4, 29.97FPS	6,250
Profile 2	1280 x 720, L3.1, 29.97FPS	3,480
Profile 3	768 x 432, L3.1, 29.97FPS	1,660
Profile 4	640 x 360, L3.1, 29.97FPS	1,175
Profile 5	512 x 288, L3.1, 29.97FPS	940

In this scenario, multiscreen support essentially doubles the amount of I/O required from 6.25 Mbps to 13.5 Mbps, and for a 30min asset increases storage from 1.4GB to 3.0GB. A typical scenario with a larger number of AVC profiles is shown in Table 5.

Table 5

	Resolution	Bitrate (kbps)
Profile 1	1920x1080, High, 29.97FPS	8,300
Profile 2	1280 x 720, High, 60FPS	8,300
Profile 3	1280x720, Main, 29.97FPS	4,600
Profile 4	1280x720, Main , 29.97FPS	3,000
Profile 5	864x486, High, 29.97FPS	2,500
Profile 6	864x486, Main, 29.97FPS	2,000
Profile 7	640x360, Main, 29.97FPS	1,200
Profile 8	640x360, Main, 29.97FPS	900

In this case, the overhead of multiscreen support is significantly higher, requiring over 3.5 times more storage per asset, 6.9GB versus 1.9GB.

In order to provide the best video quality, but avoid a significant increase in storage requirements, one could decrease the number of profiles offered as the assets get older and moved to archive. Another approach could be to archive only the highest resolution and bitrate version of the asset and use just in time transcoding should there be a request from a multiscreen client. The cost of transcoding is decreasing rapidly with the introduction of Intel I7 GPU based

COTS servers which will soon make just in time transcoding an economical and scalable alternative.

CONCLUSIONS

Using detailed home based DVR usage data, we were able to provide insight into typical DVR utilization and provide recommendations for initially sizing a cloud based DVR deployment. Most of the content recorded is popular content, limited to 10-20% of the channels, with multiple individual copies, so it is important to consider either negotiating content rights for those channels allowing shared copies, or implementing a de-duplication mechanism when archiving for considerable storage savings. 10% of the subscribers will use 100% of the capabilities given to them but on the other end 80% will use less, some far less, so there is an opportunity for savings by over-subscribing resources. Content older than a week appears to be rarely watched so it can be archived with little impact on subscriber experience.

Continued analysis and staying abreast of existing and upcoming network DVR deployments will help refine these recommendations and allow for local variations of content, multiscreen component, and subscriber behavior. This in turn will enable building stronger and more accurate business cases for nDVR deployments.

ABBREVIATIONS AND ACRONYMS

CDN Content Delivery Network
COTS Commercial Off-The-Shelf
CPE Customer Premises Equipment
DVR Digital Video Recorder
HD High Definition television
HDD Hard Disk Drive
HTTP HyperText Transfer Protocol
MSO Multiple System Operator
RAID Redundant Array of Independent Disks
SD Standard Definition television
STB Set-top Box
TCO Total Cost of Ownership
VOD Video On Demand

COMPREHENSIVE TESTING FOR CLOSED CAPTIONING IN VIDEO CPE

Joe Frank, Doug Howell, Howard Marron, Asha Nayak, Benny Pruden, Phil Reeves, Jin Zhou

Abstract

Closed Captioning is a key element of the technologies which increase accessibility for MVPD customers. Over the years it has been a struggle to validate that the elements of the MVPD video pipeline correctly handle the Closed Captions included in the video stream. In the last 3 years, Comcast has made an extremely focused effort to construct a comprehensive framework for the validation of CC handling in the Video Delivery Network with initial emphasis on its video CPE (DTAs and STBs).

PROGRAM OVERVIEW

Comcast is taking a number of steps to improve the accessibility of its products. On the operational side, this effort includes:

- Implementation of a national network-wide effort to monitor all video streams for Closed Captioning (CC) impairments
- Upgrading of all guide programs to improve the ease of access to and use of accessibility features
- Rigorous testing of CC performance on its Video Customer Premise Equipment (VCPE)

This paper will focus on the engineering behind the VCPE test effort.

CC Testing on Comcast VCPE

In mid 2011 Comcast began an effort to improve the video and operational performance of its Video Customer Premise Equipment (VCPE). The first functional area selected was Closed Captioning (CC) performance.

The goal was to eliminate CC issues in video CPE in our video production network. The scope includes all associated programs: Cavalry (DTAs), legacy STBs and the X1 Platform. The target timeframe was the releases planned for deployment in 2013.

This effort would require participation across all software/firmware providers for each product:

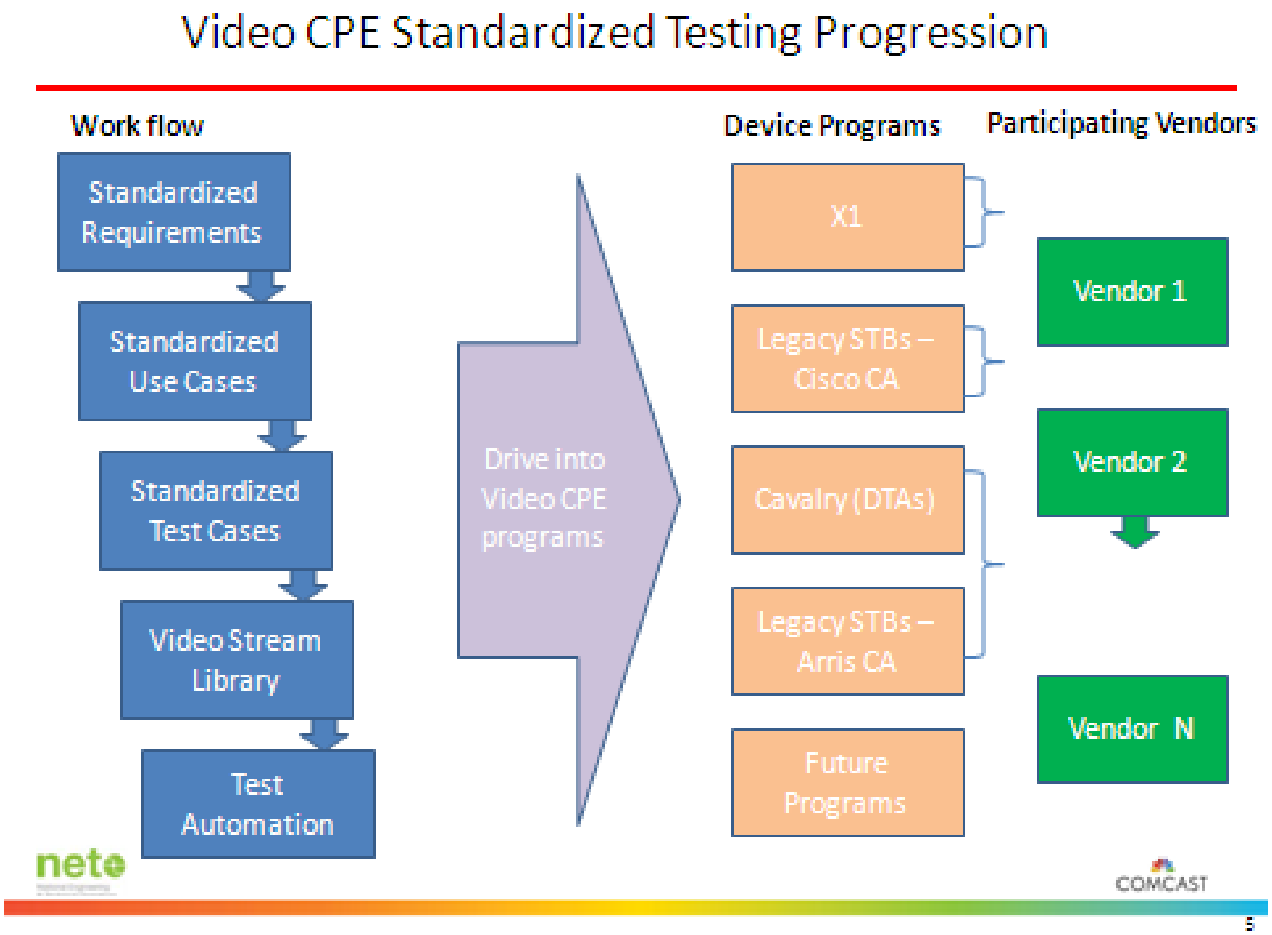
- External vendors for DTA and STB firmware
- Partner guide development/test teams
- Comcast guide development/test teams
- Comcast program teams (devices and guides)

TEST DEVELOPMENT PROGRESSION

The intent was to engineer our way out of the uneven performance of the past. From requirements to test cases. The key elements of this effort:

- Requirements analysis – what are the products supposed to do?
- Use cases – at a high level what functions must be validated?
- Test cases – exactly what steps are needed to execute on the use cases?
- Video Stream Library – testing by hopping around the live channel map was inadequate. We needed a set of streams to comprehensively exercise CC functionality in the Comcast video streaming environment – these predefined streams permit reproducible manual and/or automated testing
- Automated testing – it quickly became obvious that manual testing was an inefficient use of time and test resources.

The test development process flow is depicted on the following page.



REQUIREMENTS: CLOSED CAPTION

The foundation of the CC requirement for VCPE consists of the CEA-608 and CEA-708 specifications. In addition to passing CC data through to the SD video outputs in the VBI, the CC data must be rendered on the graphics overlay plane on all HD outputs.

Early HD STBs initially based their requirements on the profile defined by the FCC Report and Order FCC 00-259 for DTV receivers. Subsequent to that order, FCC Report and Order, FCC 12-9 governing, among other matters, enhanced closed captioning went into effect on 01/01/2014, requiring additional settings, presentation modes, and caption preview capability.

As part of its renewed focus on accessibility, Comcast made the decision to extend these requirements to older devices not formally covered by the new order. Each of our VCPE guide programs and device families was evaluated against these more stringent requirements and then any required feature development and/or defect repair was added to the plan for the target releases.

TEST STANDARDIZATION

In the past, VCPE testing was designed and implemented in guide oriented silos. Testing focused primarily on the guides and since the guides were unique per VCPE family, the testing per family was developed in isolation, often independently of any pre-existing test plans/cases and with varying degrees of comprehensiveness when measured against the requirements.

Prior to this effort, test cases were developed directly from a myriad of requirements. The test cases were then executed on the independent guide/VCPE families. This resulted in test duplication and inconsistencies among the various families. To address this, development of use cases was prioritized and formalized to streamline the process. Test cases were derived from the use cases which were *then* applied to the guide/VCPE families

After collecting the requirements, they were coalesced into use cases derived from each specific requirement. These use cases, generic and device-independent, became the master requirements for the test engineers.

For example, CC use cases are very similar across the various VCPE. They pertain to 708, 608, and SCTE20 pass through. The CC functionality is then normalized across pertinent orthogonal test planes such as EAS, video formats, DVR, parental control, PPV, etc. The intent is to ensure that CC functions correctly across each plane.

The use cases shielded the test engineer from the burden of organizing the scattered requirements. Test engineers could now focus on what they did best -- write test cases. The use cases describe the test environment, provide a high level description of the functional sequence of events and conclude with a description of the expected outcomes.

The next phase in the standardization process was to develop a set of generic test cases from each uses case. Usually multiple test cases can be traced back to each use case. Like the use cases, the generic test cases would also span the device families. They would be the test case superset from which device/guide specific subsets could be drawn. The generic test cases break the use cases down into individual test steps. The individual test steps are described at a high enough level to be independent of guides and guide key sequences and yet generic enough to be readily applied to the actual guide/device pairings.

The final step in the test standardization process was to create the test cases specific to each device family, taking into considerations the guide-specific displays, menus, and menu navigation.

The test standardization process permits consistent in-depth comprehensive test coverage across all VCPE families. Starting with the actual requirements and generating cases avoids the trap of testing to a guide or a device and not to the real requirement. The use cases can also be readily prioritized for varying levels of test coverage depth per development program phase. And lastly, of course, this structure provides an excellent basis for test automation.

VIDEO STREAM LIBRARY

A necessary complement to the development of rigorous use/test cases is an environment in which those cases can be executed repeatedly and consistently. Thus a comprehensive stream library is required.

In the past, live feeds were used for testing audio, video, and closed captions. Although live feeds have their merits particularly with regards to replicating plant conditions and a production environment, the feeds do not lend themselves to repeatability and consistency. This is not to say that live feeds should not be used, but rather they should supplement the library.

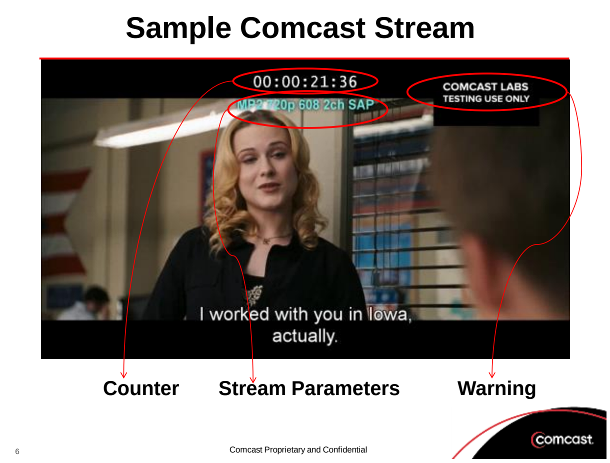
A video stream library significantly improves efficiency. It can be time consuming and challenging, if not impossible, to search through channels and locate content with specific features such as close captions for CEA-708 only, CEA-708 with CEA-608, CEA-608 only, and SCTE20. Just as importantly, having content with precise and known captions is imperative in ensuring that no captions are dropped. In addition, having known content with prescribed features such as video format (1080i30, 720p60, 480i30) and audio format (AC3 stereo, AC3 5.1) permits verification of captions in the varied stream environment of the device. A video stream library is fundamental to a deterministic test environment.

Just as important as having comprehensive feature-laden content, the encoders used to create the streams are vital to the test system. By creating streams which use production equipment and production encoding profiles, the streams can quickly confirm that VCPE has no issues with production configured encoders. The production streams should be created for both linear and VOD applications using the encoders and profiles specific to the delivery vehicle. Once again, bit rates, GOP structure, video formats, and audio formats are all designated and known.

Stream generation and capture are important components of our custom Comcast video stream library. However, purchased streams provide a valuable supplemental capability. Comcast has purchased streams from Sarnoff Labs which offer the ability to drill down deep into CC functionality. Streams are available which thoroughly exercise MPEG syntax, video formats, audio formats, and CEA-708 CC functionality. It is worth noting that these purchased streams only focus on one function at a time where the Comcast-generated streams incorporate the varied functionality already present in the real cable distribution plant.

Our Comcast video stream library is the fundamental enabler for our test standardization strategy. The video stream library provides the consistency and repeatability required for a rigorous standardized test process. Our feature rich Comcast streams offer tremendous efficiency as they are used to simultaneously exercise multiple functions permitting simultaneous test execution.

Probably the most important aspect of our stream library is that it enables test automation. Without them, an automated test process would be impossible or at best impractical.



TEST AUTOMATION

Closed Caption test automation development could now begin with the three necessary building blocks: use cases and test cases from the standardization effort; and video streams from the video stream library. Assessing CC compliance in all the applicable video environments within the Comcast footprint and across the variety of VCPE devices in an automated method became the primary automation goal. Manual testing is not a feasible solution, due to the numerous permutations associated with 25+ device models, multiple guide releases, and the large number of streams.

The CC automation breaks down into two primary areas defined by the streams: testing against all required video formats and configurations using the 10-minute Comcast streams and testing implementation compliance against CEA-608 and CEA-708 CC specifications with the Sarnoff stream library. The Comcast stream testing suite addresses topics including audio splicing, video splicing, CC transitions between types, resolution changes, video encryption, delivery type (VOD, PPV, trick-play, live), and encoding type (MPEG2, MPEG4). Sarnoff stream based testing validates the CEA-608 and CEA-708 protocol specific functionality including fonts, colors, location, timing, scrolling and data rates.

Automating the testing posed several hurdles. The largest issue was finding an automation tool that had an HD capture solution able to handle the high frame rate updates of the streams and also able to analyze the results in a reasonable amount of time. The second obstacle was creating and routing the multitude of required video configurations for the Comcast streams to the applicable boxes within the Comcast test environment. The content had to be provided in encrypted and unencrypted forms on both controller types, DAC and DNCS and across different device types (DOCSIS STB, legacy STB and DTA). Additionally all the streams needed to be ingested into the VOD systems and guide data generated to allow for VOD and DVR functionality.

The automation solution also is required to automatically select a video to stream and to receive the video with all the permutations of settings needed for the test. To accomplish this, a pool of video streamers was established in which each is addressable and available for reservation by the automation tool. Each streamer can play any of the appropriate video files from the stream library. The selected IP video stream is routed to the edge devices for the appropriate test systems. The edge devices (both Cisco and Arris systems) then replicate the video, and insert unencrypted and encrypted copies into the channel lineup. The automation tool then tunes to the appropriate channel to get the proper test configuration. In addition, the system is configured with mock guide data and all videos are ingested into the VOD system giving the automation a way to replicate VOD, DVR and trick-play functionality.

Test tool selection

Several test tools were investigated; each being tested against a representative set of streams. The selected tool, Witbe QOE Robots, demonstrated the required HD capture capabilities, a modular approach to test development, satisfactory OCR performance and a reporting solution that could be integrated with Comcast’s existing test framework. The tool selection also aligned with automation efforts in other parts of the company so that we could share resources and best practices as necessary.

Test design

Reusability and portability were driven into the test design from the beginning. With over 25 different device models to be tested and the multitude of different resolutions, guides and other test parameters, the test block architecture was required to be very configurable and reusable. Using a reusable automation test block system combined with test configuration files, the test architecture is able to reuse a majority of test blocks between test permutations. This reduces the amount of work necessary to change any test. Small configuration changes allow for scripting to cover large populations of tests.

We also had to account for the minor differences in CC implementation among the various implementations of the CEA-608 and CEA-708 specification. The specification has some ambiguity about locations and performance that the test results must account for. Allowances had to be made in the verification of timing, position and fonts. This was accomplished by creating device specific configurations or creating search parameters in the applicable algorithms able to accommodate variation -- i.e. fuzzy location searches and timing variances. This flexibility allowed tests to adjust dynamically without having to be modified for each device.

The general flow of the automated closed captioning test is:

1) Validate device health
2) Set relevant guide settings (CC, fonts, menus, resolution)
3) Start video play-out from streamer
4) Tune to appropriate channel
5) Validate stream video matches expected
6) Validate CC
7) At end of test, undo guide settings as appropriate

Each Comcast stream has an associated test data file that the automation utilizes for validation. Inside the file is a listing of caption text, the number of times each caption should appear and where on the screen it should be rendered. The automation script will synchronize the video to the test data and then commence validating all of the captions within the ten minute sequence. Validation of each closed caption is done via OCR. Using the defined appearance time from the configuration file the script will search within a configurable time window for up to three simultaneous captions. The best OCR result within the window is used for the result. After completion of the sequence, the automation will compute a pass or fail for the test and create an output file that has all the expected and detected results. A nominal ten minute test stream sequence contains about 300 closed captions.

The Sarnoff stream based test cases are designed to validate the functionality of the CEA-608 and CEA-708 CC specifications. Therefore a different rigor is applied to the testing. The Sarnoff streams are designed for efficient specification compliance by a knowledgeable person observing the video output of the device. The video is highly dynamic and covers many features simultaneously.

The test blocks use image comparison to compare known good images against set top output to verify that fonts, colors, location, timing and presentation are all correct. Depending on the test requirement, the automation will compare one or multiple images simultaneously on the screen. Simultaneous image compare allows us to verify that frames are timed correctly to the video time source. Due to the high change rate of video and closed captioning, testing of the Sarnoff streams can take several iterations through a video loop to complete a single program descriptor.

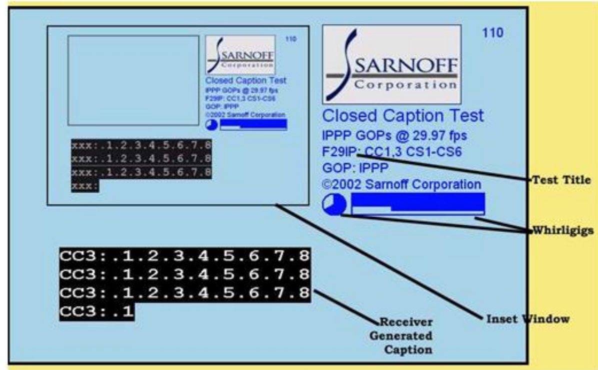
Test Development

Test development is run just like a typical Comcast agile software development project. Teams are on a two week scrum schedule. Tests scripts are kept in source control and are run through integration/ingest cycles as part of the release process.

After completion of the initial phase of automation we have several lessons learned from the difficulties encountered.

- OCR has inherent repeatability issues - to reduce the number of false positives & negatives a large population of tests must be used.
- Rendering across device families may appear the same to the human eye, but automation is subject to small variations in timing and location.
- Compliance to a specification that was not rigorously enforced initially can lead to many defects when testing first occurs

Automated Testing - Comcast Stream



Automated Testing - Sarnoff stream

VENDOR PARTICIPATION

The final step in implementing this program was to drive the automated standardized testing upstream into our CPE vendor community (both internal and external). The benefits are significant:

- Earlier detection leads to less costly defect repair (the engineers are still "on the job" and have immediate access to reproducible defects)
- Earlier detection leads to a shortened elapsed time (reduced repetition of the development cycle phases)
- Drive cost out of the formal Comcast test process

As did we, our vendors quickly realized the need to automate this testing. Some vendors chose to "port" our tests and scripts to their automation systems. Others chose to adopt the same tools and hence take advantage of the work we had done and gladly shared. Either way, this automated testing has been migrated back up the pipeline.

CONCLUSION AND FORWARD LOOK

Comcast has instituted and organized a thorough Closed Captioning test suite comprised of use cases, test cases and a comprehensive video stream library. By automating the testing of Closed Captioning using carefully selected tools and an integrated framework, we hope to eliminate VCPE induced flaws in the presentation of CC services to our customers.

This framework and test catalog is in a continuous improvement mode, expanding and fine tuning the test methods, test stream library and testing scope to incorporate new and improved technologies as they become available. Future areas of expansion include new video services such as:

- IP streams
- MPEG4 HD
- HEVC encoding
- Transition/spliced streams
- Descriptive Video Services (for enhanced accessibility) automation

Beyond addressing the expansion of video stream variety, areas that Comcast may apply the benefits of this rigorously engineered framework and technology include:

- Future Comcast Video CPE
- Customer owned, operated, and maintained devices
- Centralized Video Delivery Network elements (encoders, ad-slicers, etc.)
- End-to-end testing of all elements in the video delivery network

AUTHORS

Device Engineering
Joseph_Frank@cable.comcast.com
Douglas_Howell2@cable.comcast.com
Benny_Pruden@cable.comcast.com
Philip_Reeves@cable.comcast.com

Device Test
Howard_Marron@cable.comcast.com
Asha_Nayak@cable.comcast.com
Jin_Zhou@cable.comcast.com

Create Data-Driven Solutions to Optimize IP Content Delivery and Identify Revenue Opportunities

Brennen Lynch and Anukool Lakhina
Guavus, Inc.

Abstract

Consumers are increasingly turning to IP-delivered content as their primary form of media entertainment. This shift from traditional media consumption is presenting cable operators with new challenges in providing consistent quality of service (QoS) to a multitude of subscriber devices well beyond the traditional set top box. These quality challenges are paired with revenue opportunities as cable operators now have more platforms with which to interact with their subscribers. In both cases, QoS and revenue, the value can often be found within the data.

This paper explores leveraging existing data sources to enable Cable Operators to improve customer quality and drive business value from IP-delivered content services. In the transition from traditional QAM delivery to CDN architecture, IP-based video servers and clients generate copious performance and usage data that can be fused with subscriber reference data, as well as legacy usage trends, in order to derive key QoS and revenue insights. Data assets are being generated from IP networks continuously and at huge volume, and from a wide variety of platforms. Correlating and analyzing this network-generated data for key metrics related to subscribers, CDNs, or QoS issues, is a challenge to achieve in real-time. Operators need to draw out correlations that have genuine actionable insights and enable line of business users to make the best possible decision at any given moment, while simultaneously maintaining day-to-day operational requirements of delivering consumers' primary content over IP.

Such a solution provides an ongoing method to take action upon insights that enable cost reduction through CDN optimization, increased QoS on multiple devices and IP STBs, and new revenue opportunities through targeted marketing opportunities.

The Widening Discrepancy between Network Delivery & Consumer Demand

Recent IPTV offerings such as Netflix, Hulu, & HBO Go, have blurred the consumer (& network) distinction between channels and programs. Traditional content rights limited to cable are being re-tooled, as with the Dish/Disney deal that enables Dish to live stream Disney, ESPN & ABC via their Dish Digital IPTV service. Consumer expectations are increasingly demanding a revolution from plain linear programming to an *a la carte* video on demand structure that offers highly interactive viewing experiences. Consumers still demand a smooth and high-quality service, of which over-the-top (OTT) providers are attempting to satisfy with offerings such as Netflix's Super HD. Despite the fact that the industry is phasing in paid peering agreements with these content providers, such bandwidth-hungry services are severely taxing current network infrastructure and delivery mechanisms.

Service operators are adapting in a number of ways such as rolling out competitive services; HFC plants are beginning to support 10Gbps capacity and home gateway devices are supporting new media handling and translation capabilities. IP delivery is a much more efficient use of the HFC infrastructure (due to the statistical multiplexing capabilities (ie inherent SDV) and advanced codecs such as HEVC). However unless the service operator runs fiber to the home, there is still a key fundamental challenge of providing a smooth, high-quality IP video experience that could be comparable to current QAM delivery of linear TV. From a business perspective, executives are focused on retaining existing customers and attracting new ones. Therefore implementing stateful business processes/monitors and identifying-prioritizing high-value network operations are key differentiators for the cable operator aiming to offer competitive IP video services.

Architecture Considerations

This paper is not a focus on network architecture that supports IP delivery of content, nor a focus on the big data platform required to enable the use cases that are explored. However it is relevant to bring the architecture and management techniques into a high level perspective, to allow for association of big data use cases with network conditions.

Network Architecture

Cable operators generally deliver content via IP to the headend/hub. It is at the last mile where data and video services differentiate in their delivery. There are a number of inhibitors to an immediate switch to all-IP, of which the foremost is the capacity requirement. Transition from QAM video channels to DOCSIS is a costly venture (as DOCSIS channels generally cost 8x more than QAM). Although advanced codecs with IP video can provide 4-5x efficiency gain (H.265 HEVC) that reduces the number of channels required to carry IP video, this still does not offset the 8x cost of the DOCSIS channel over QAM [11].

The converged cable access platform (CCAP) is a solution that is being developed to enable the operator to move gracefully to an all-IP network. CCAP combines the eQAM for digital video with the CMTS for data into one device (or logical entity, if virtualized), allowing for a greater density per port, leading to a huge reduction in operating cost. The CCAP performs a number of functions such as providing a control/management plane (for IPDR, DOCSIS, RF, QAM MIBS PCMM, & edge resource management), as well as processing for RF, DOCSIS, & QAM. A virtualized CCAP environment would allow these compute processes to occur in the most efficient location as determined by the controller (currently all these functions reside in the headend) [11]. Ultimately, this virtualized CCAP will create spatial & power-related efficiencies as well as supporting scalability of engineering & service delivery – such as IP video. By centralizing an IP/MPLS control plane, the operator has ultimate control over resource allocation and enabling end user-driven policies (such as optimizing for QoE).

The distributed caching architecture of a CDN network allows an operator to optimize content flow. The distributed servers can request content on behalf of the subscriber, and then retain and distribute to many geographically local subscribers. Typically these are purpose-built appliances in the datacenter, engineered for peak demand (meaning they sit idle for much of the time). Furthermore, they request content based on static metrics such as hop-count, and do not take into account optimizing for congested links. Those routing-based use cases can be targeted now, by integrating a real-time analytics platform. By integrating with CCAP on a per-flow basis, and an end-to-end optimization/ QoE scheme can be implemented. The platform components required to drive this scheme can not only support other marketing/media – based use cases, but will be positioned to drive highly-optimized, virtual content caching on standard hardware as the industry looks to transition away from fixed appliance content caching.

Platform Architecture

Carrier-grade availability of massive streaming & batched data fusion must be supported by a unique set of requirements. Even just collecting network data requires the ability to pull data from the edge through the backbone network in real time, without taxing mission-critical infrastructure. Furthermore, in order to be useful, this petabyte-scale, high-velocity network data needs to be dynamically fused and correlated with static & reference data sets to produce key causal relations. Thus, a departure from the traditional 'store-then-analyze' approach must be explored.

A highly available, 'compute-first' architecture (including edge processing) is the cornerstone of this *new data fabric* that supports this paper's use case explorations (see figure 1 for sample architecture). By using lightweight compute nodes that are optimized for low-latency, high-throughput transactions, data can be normalized and fused at the point of collection. This ultimately allows the data to be presented to the stakeholder in real-time, enabling an environment of triggers & actions (also providing an underlying data framework for fueling SDN & SON capabilities).

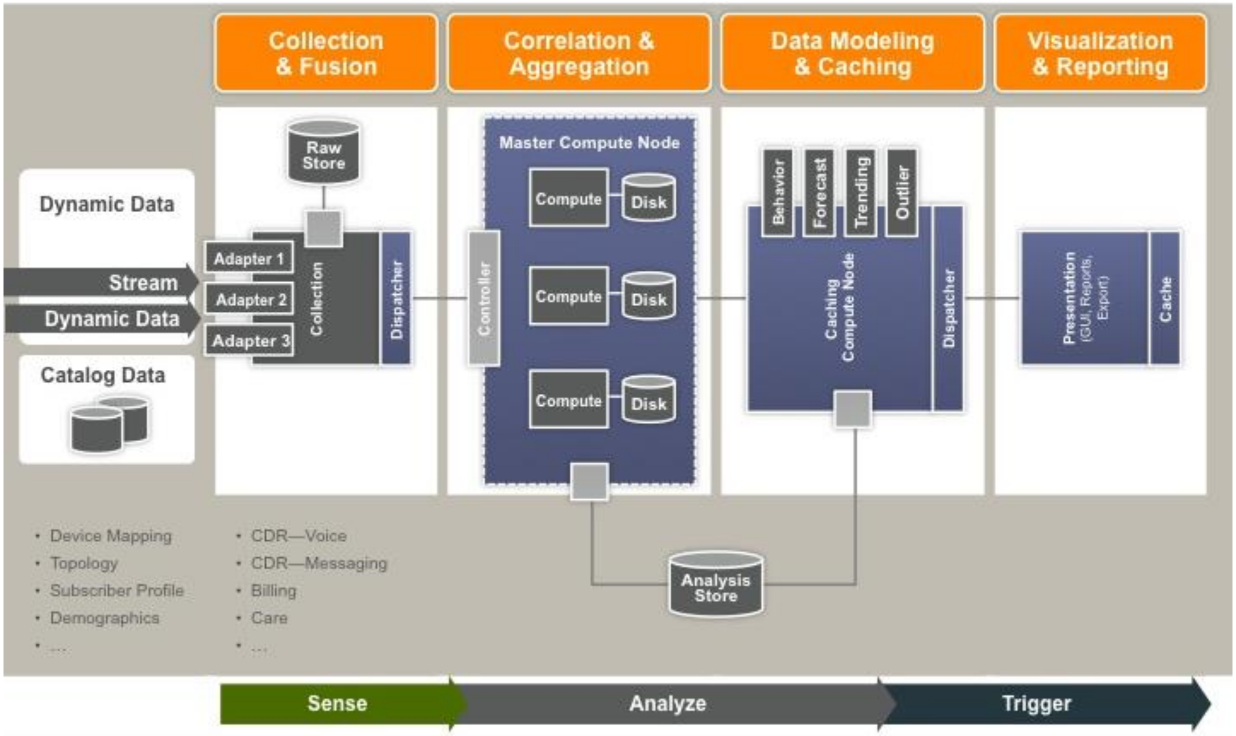


Figure 1: sample architecture of a big data fabric that will support real-time decisioning. Processing & normalization occurs at the edge, followed by centralized compute & store. Data is then modeled, cached, & presented via UI or trigger to support the stakeholder's use case.

Data Sources

Player & Application Logs

As the purveyors of the "last mile", service providers are in the unique position of being able to offer efficient IP services deployed at the edge. Consumers increasingly are streaming content to multiple devices beyond the traditional set top box (such as computers, mobile phones, tablets..). This player data (generated by the end subscriber) are the raw customer & device interaction metrics- around how subscribers use the network and how devices are performing. Fusing this raw data with relevant datasets enables the operator to calculate and monetize high value insights that will be expounded upon further (such as churn risk, profitability calculations, or targeted marketing). Service providers require a dynamic ability to deliver high QoS to these many devices, in order to satisfy the consumer's expectation of consistent experience regardless of medium or delivery mechanism.

Service providers with managed CDNs have the opportunity to use the application data (which may be procured from a third party) generated by this video serving equipment to support the delivery of high QoS to these many devices. Fusing the application data with player data allows the operator to perform dynamic network optimization with status & trend information by content, location, node health, link congestion, and other CDN-related metrics. In a business context product engineers are provided transparency towards granular demand and campaign effectiveness. Further, marketing teams can dynamically target advertisements to increase media sales revenues. The data generated from serving IP content is becoming a rich and highly valuable source of truth as the industry transitions towards CDN architecture from the traditional head-end service delivery.

Creating a service provider solution around analyzing player and application data, in abstraction or in fusion, supports a number of use cases, which can be categorized as follows:

Use Case Category	*Data Focus*
1. Marketing & Product Engineering	player data
2. Network Optimization	application data
3. Targeted Advertising	fusion of player & application data
4. External Systems Feed (to support greater intelligence in existing systems)	fusion of player & application data

Figure 2: use cases as they pertain to data focus

The use cases pertaining to each category will be detailed in the following sections.

Player data and application data each provide a unique measure of analysis that leads to actionable discoveries. Both prime data sources must be fused with relevant catalogues, location databases, as well as categorization databases in order to resolve all data to necessary metrics such as subscriber and topology. These two unique measures are complementary and logically related as two parts of one whole IP Video application (with the necessary reference/catalog databases).

IPTV Sources- Player Data

Generated by the end subscriber, player data is mediated and aggregated from a number of raw data sources including guide-based and data generated from the delivery software located at the client player. To gain full perspective fusion must also occur with any geolocation service (enabling ISP identification) and information around subscriber DVRs, entitlements, and purchase activity.

Data Specification: the following reports in figure 3 provide an indication of the data that is used (system, data fields, format ie json or binary, & richness will vary per-operator)

Player Data Report Examples	*Description*
ContentStarted & ContentEnded	Content played including device, UUID,..
ContentUpdate	Update to content stream including place-shift
CrashReport & ErrorReport	Details of error, device hardware/OS, content,..
IdentityRecord	Device, subscriber, & session identification
HardwareRecord & SoftwareRecord	Hardware & device, software/platform & OS
ContentViewed	Source/type/position for block of viewed content
UserExperience	Traffic, QoS metrics, playback quality
ContentAuthorization	Privileges associated with a piece of content
ProfileRecord	Profile usage including ABR bitrate, codec,..
EventRecord	User action while playing content

Figure 3: example player data reports & attributes

Grouping>	|--------SUBSCRIBER DATA--------|		|------------IP TRAFFIC / PACKET DATA------------ |			|---------NETWORK---------	
Dataset ->	**BSS/Billing**	**Demographics**	**DPI (Full or Partial)**	**Netflow**	**CDN Logs**	**NMS / SNMP**	**OSS/Service Assurance**
System ->	*BSS*	*BSS*	*OSS*	*OSS*	*OSS*	*OSS*	*OSS*
Media Buying	Required	Required					
CDN/App	Good				Required	Good	
IPTV/Ops	Required	Good	Good	Good	Required		
Data Type ->	*Semi-static*	*Semi-static*	*Dynamic*	*Dynamic*	*Dynamic*	*Dynamic*	*Dynamic*
Refresh ->	*Weekly*	*Weekly*	*5 mins*	*5 mins*	*15 mins*	*15 mins*	*15 mins*
| | | | | | Key fields | | | | | |	Account # Video/HSD/DV Plan Type Address (geo link) Inventory ID Home score	Age Gender Dwelling Household type Income Level Ethnicity Other 3rd party attributes	Headers Packet data User Agent	Ingress interface Source IP Dest IP Protocol Source Port Dest Port Type of Service	Record ID (UUID) Timestamp (start & end) Identity/Software Record QoS Metrics Error Reports bytesSent/Rec CDN Hostname	Topology	Protocol Fault Record SLA QoS

Grouping>	|------------NETWORK-------------|		|----------------------------VIDEO SERVICES (LINEAR/IP)--------------------------------				
Dataset ->	**CPE/Device**	**Network Inventory**	**Channel Mapping**	**Tuning Events**	**VOD**	**Player Logs**	**Advertiseme Guide**
System ->	*OSS*	*OSS*	*OSS*	*OSS*	*OSS*	*OSS*	*OSS*
Media Buying	Good		Required	Required	Good	Good	Required
CDN/App	Good	Good					
IPTV/Ops	Good		Required	Required	Good	Good	
Data Type ->	*Semi-static*	*Semi-static*	*Semi-static*	*Dynamic*	*Dynamic*	*Dynamic*	*Semi-static*
Refresh ->	*weekly*	*weekly*	*weekly*	*15 mins*	*15 mins*	*15 mins*	*daily*
| | | | | | Key fields | | | | | |	Account # Serial # Device type Device model Device manufacturer LOB	Account # Node/Inventory ID System ID Hub name Headend name CMTS name Controller name Video controller TSAP Address Voice Switch Node Telemetry Non-HFC (Access, wifi, FTTH)	Channel lineup Timestamp Region Identifier *(ie Tribune/Rovi)*	Program Program schedule Device definition Session start Session end Media Device Media Viewer Type Media Viewer Presentation State Event Segment Event Play Control Event App State Event	Content requested Price SD HD *Clickstream*	Session records Player Interaction logs	Channel mapping Local Broadcaster Media request Advertisement ID Duration Start time End time Play success

Figure 4: data matrix detailing the data sources and data fields as they pertain to each use case channel: Media Buying (targeted advertising), CDN/Application, IPTV/Operations.

Other data sources may either be required for subscriber identification, or may provide a rich set of business rules and states that can be used to enrich any solution. These data sources are presented in figure 4.

CDN Sources- Application Data

Video-serving equipment or CDN delivery nodes generate the application data relevant to content distribution. Application data can be aggregated, processed, and correlated from a number of raw logs or data sources.

The primary source of data is from Delivery Node logs (specific to vendor such as ALU (Velocix), Cisco or internal/proprietary logs). The data enables resolution of traffic to subscriber or further granulated into delivery (unicast, ABR, etc). In addition, by generating content resolution charts for both constant bit rate (CBR) and adaptive bit rate (ABR) encoded media (requiring associated catalogue data and access logs), the operator can effectively analyze the most efficient and highest-quality method to deliver content. In this manner network engineers can monetize data to implement the most efficient processes and expansions.

Additional data sources to provide a full perspective can include proxy logs from upstream nodes to provide visibility of network utilization from the delivery nodes to the upstream devices; catalog data such as Asset Management System databases; and lastly, Session Manager (created when users select an asset to download) and License Manager logs used to correlate content accessed to subscribers. Fusion of the above-referenced application data with subscriber/topology-resolution data such as billing/CRM enables near real-time analytics to deliver timely insights on the performance of CDNs. A summarization of the data sources relevant to the use cases is detailed in figure 4.

Use Cases

Marketing & Product Engineering

The analytics solution built on player data provides a number of unique insights around traffic trends, subscriber engagement and content metrics. Although an operator's specific needs may vary and the use cases can be adapted as such, there are a number of consistently high-value solutions that can be supported:

- Identify & analyze subscriber cohorts
  - Of join duration, joining months, promotional offers, partners, & price-points with anomalous (high/low) Customer Lifetime Value (CLV)
  - Of promotional offers & packages with anomalous (high/low) Trial-to-Subscription rates
  - Of transitions/ movements across lifecycle stages to evaluate business health
- Monitoring Package performance across trial-to-subscription rates, subscriber cohorts, & price-points
- Identify & analyze churn rates for different subscriber cohorts

of partners, promotional offers, price points & CLV
  - Monitoring service availability, issues & customer support for different subscriber cohorts
  - Monitoring financial projections and actuals for different subscriber cohorts of partners, promotional offers, & languages
  - Creating demographic & behavior profiles per-program for advertisement monetization

To reveal subscriber QoE (ie traffic anomalies through consistency & continuity) traffic is resolved to subscriber, then topology, to reveal potential drivers. Marketing solutions around churn enable near real-time action on high-risk subscribers. As subscribers begin to exhibit behavioral variations, those deviations via machine learning can be processed into a churn risk (based on extrapolation of previous subscribers' behavioral trends who have already churned). Traffic metrics on broader terms enable dynamic product/network engineering efforts such as variant analysis (how changing one traffic metric - ie delivery - will affect others).

Content discovery to the subscriber or to any part of the network allows behavioral analysis around specific product engineering, as well as targeted marketing and campaign activity. For example, subscribers can be categorized into certain states based on activity, trials (in-trial/ex-trial & activity within trial period), transaction history, subscription package & tenure, etc. The operator can then cluster the subscribers based on attribute & state variations - from the context of key business measures. This would allow the operator to analyze profitability, demand, success, and identify product-engineering decisions driven by data produced by the operator's customers.

Analysis of player data generated by interactions with a Service Provider's IP Video service enables a rich set of actionable insights. An appropriate software application enables executives and other operators to closely monitor, analyze, and identify future trends in the business. It is an open solution that benefits from additional data sources for budget overlay calculations and specific costs. On a real-time basis, it can support campaign visibility and optimization tools. Churn analysis modules (based on predictive machine learning algorithms) can take advantage of the streaming subscriber-based data to enable continuous targeting of high-risk subscribers with save techniques. The solution supports Marketing, Product Engineering, Customer Care, Sales, and

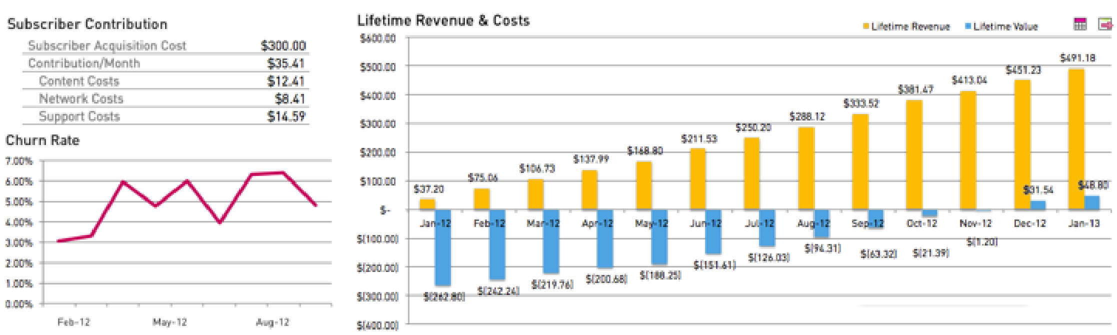


Figure 5: transient view of a subscriber's profitability exposes break-even point. Summarizes cost-contribution, churn rate, & CLV for a selected subscriber cohort.

Financial executives in regular monitoring and course-correction strategies.

Network Optimization

Content analysis may compare performance of different assets over selectable time periods and allows a marketing or engineering user to: analyze usage patterns, anticipate demand, and provision the CDN efficiently to meet customer requirements at minimal cost. In addition to identifying popular assets, access information, etc., a solution could drill further into the content measures. By analyzing the subscriber access to ABR/HLS- or CBR- delivered content, an operator can increase efficiency and QoE through intelligent data-driven engineering decisions. For example, by running a relative commonalities clustering algorithm on the highest subset of Mean Opinion Score, or network elements that are experiencing QoS issues, the underlying contributors can be exposed for targeting.

Approaching CDN-generated data from a geographical distribution of client access enables the marketing user to deliver titles on a per-demand (down to geographical market) basis. Network engineers can analyze best locations for network expansions to capitalize on demand and increase the efficiency of expansion.

Device-level analytics is becoming increasingly important as subscribers access more content on devices (tablets, smartphones, etc.) other than their primary television/STB. Product engineers can prioritize support of operating systems based on near real-time demand of content. Classifying popularity amongst devices and content-per-device allows for data-driven marketing to users of popular devices & contents.

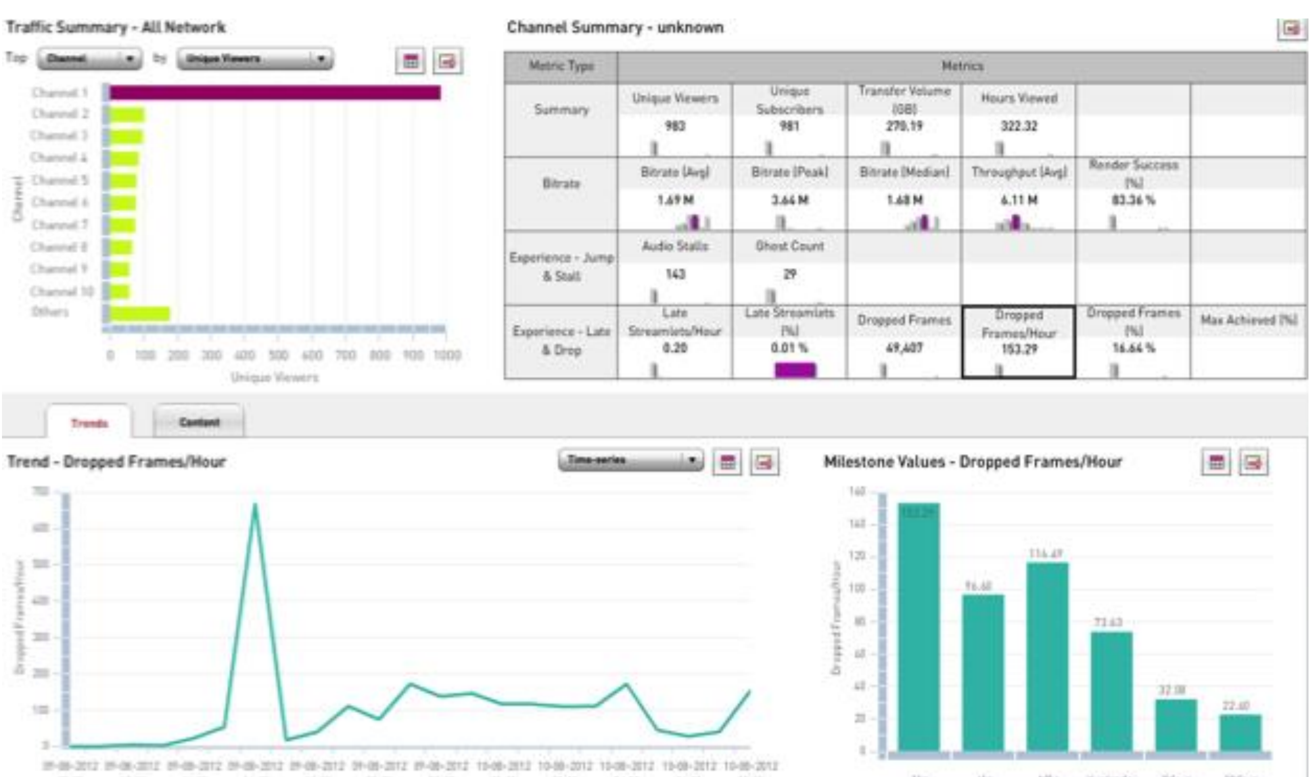


Figure 6: centralizing QoS metrics on a real-time basis allows anomalies to become visible. Note the exposed anomaly around dropped frames; in a fully virtualized CCAP environment, this type of data could be exposed to management systems to re-allocate CDN resources based on the underlying reason for dropped frames (exposed through a relative clustering algorithm)

Network resource utilization is an important perspective enabled with application data. By quantifying metrics such as specific CPU utilization of nodes/caches, HTTP transactions, cache hit ratios etc, the user is provided transparency towards demand on network resources. This perspective allows network engineers to decide when to expand the capacity of the network, view how the CDN is routing requests, and detect possible problems with delivery nodes and caches (see figure 6).

As a summarization, these metrics may be correlated & measured:

- Summary metrics (viewers, subscribers, tonnage, hours)
- Bitrate-related metrics
- Experience- jump & stall – consistency related metrics
- Experience – late & drop – continuity related metrics
- Content (per ISP, CDN, device, channel, service,..)

To reveal the following:

- What impacted the user's experience the most, & why? (see figure 7)
- Traffic & volume distribution by location
- Any corresponding effect of modifying any one metric
- What was the trend & frequency histogram to identify anomalous spikes/troughs?

Selective grouping and intelligent fusion of application data with relevant peripheral datasets enables an extremely powerful solution for the CDN operator or Service Provider who can access the relevant CDN data third party. Delivery of these insights in a drill-down approach allows the operator to capitalize on market opportunities, identify important trends, and identify/prevent problems that could have impacted network performance, customer satisfaction and loyalty.

Targeted Advertising

The dynamic ability to target advertisements and characterize eyeballs is an extremely valuable function enabled by the IP delivery of content. Ad slots no longer need to be sold via number of viewers or via the relatively small sample size of the Nielsen audience measurement analysis. By centralizing subscriber interactions with demographic and other reference data associated with the industry ecosystem (figure 8), a full scope involving 100% of the subscribers' anonymized viewing behavior can be analyzed and provided

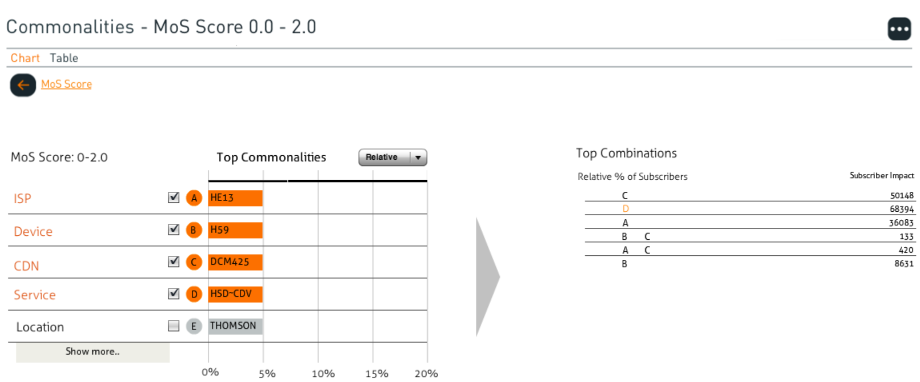


Figure 7: relative clustering algorithm exposes top contributors to Mean Opinion Score influencers.

as justification for pricing, as well as selling ad slots that might generally be undersold (i.e. monetization of long-tail programming).

IP video providers can support a granular Audience Measurement scheme that enables a number of actions such as

Audience Measurement:

Monetize programmatic buying: programs operate in an environment where they are targeting different audiences, or may be competing during the same time (on different channels). The provider can easily partition

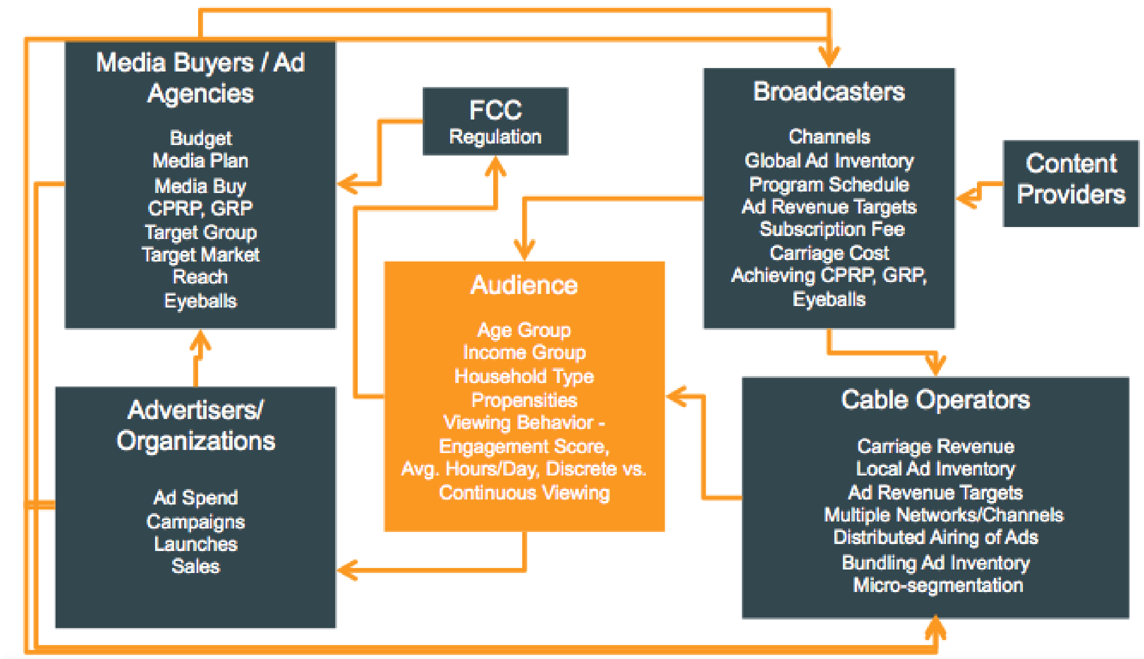


Figure 8: Data sources infrastructure available for Targeted Advertisement use cases. [12]

monetization of programmatic buying, or via- microsegmentation to power niche genre monetization. This in turn supports two related modules – Media Planning, which enables equivalent plan offerings (in case of inventory unavailability) and programmatic analytics to maximize the effective rate, and Yield Management, which enables the provider to construct proposals that ensure relative profitability as well as ensuring uniform consumption of deals by saving top-shelf inventory.

premium inventory via the best performing programs ($ revenue / ratings). Following this, the performance can be monitored and trended to impact any pricing changes. This also identifies the best adjacent programs to mitigate opportunity loss.

Applying viewership analytics for micro-segmentation: audiences belong to different age groups, household types, income groups etc., and they depict different viewing habits such as watching ads on mute, or continuous vs. discrete viewing. Through a big data fabric that is able to provide real-time data fusions and clustering with such a high dimension of cardinality, the provider can achieve micro-segmentation. Viewing habits of the top segments (with higher purchasing power) can be tracked and identify highly engaged programs. To mitigate opportunity loss, the provider can check for best adjacent programs. Further, to monetize the niche genre, the provider can identify hidden values (through eyeball analysis) and therefore charge a premium.

Media Planning:

Monetize bundling opportunities across networks while offering a programmatic bouquet: cable operators carry channels from multiple networks (increased breadth vs. a broadcaster). Thus operators can leverage all of the property/program combinations from cross-channel networks to maximize the value of a programmatic buying bundle. The best combination can be extracted from a suitable budget, while bundling premium & slow-moving inventories.

Offer an equivalent plan to the media buyer in the case of inventory unavailability: in peak season or for popular programs, there is usually a crunch for inventory. Media buyers have KPIs that must be satisfied; therefore, options are required in order to prevent a loss of revenue opportunities. By targeting eyeball makeup instead of the program, the provider can identify a mix of programs and time-bands that serves as an equivalent of the bouquet offered by the media buyer.

Constructing proposals that ensure profitability: the provider can build a proposal while optimizing profitability and saving premium inventory. This will achieve an increase in profitability as well as an avenue to satisfy the KPI benchmarks such as CPRP or eyeballs. The provider can then offer the media buyer a good mix of programmatic and RODP buying options.

Allocating ad spots to maximize revenue and satisfy the business constraints: proper allocation of advertisement slots is required to clock the revenue and to honor the deal. Business constraints also must be taken into perspective – for example, no two spots of competitive brands may be aired back to back. Therefore a programmatic vis a vis RODP mix is to be served. This allows the provider to achieve an increase in revenue with optimized allocation of spots, as well as achieve more efficiency and accuracy with automation than manual allocation. Furthermore, the provider can generate a list of all the best 'make good' options for the dropped ad slots.

Ensuring uniform consumption of deals by saving premium inventory: inventory planning is required to have a visibility in future inventory booking levels. Executing this exercise on multiple deals simultaneously will help guide future sales efforts. This enables the provider to achieve visibility into future inventory booking levels, and to achieve uniform consumption of deals while saving premium inventory.

Measure viewership across TV, web & mobile to optimize programmatic media allocations: as viewers consume content on different devices, different program preferences can be attached to the devices based on viewer behavior. This identifies most popular genre/programs across device category. The profile of the device can then be examined vis-a-vis the audience, to extract high-value micro-segments to be targeted for programmatic buying.

Utilizing engagement score to enable programmatic buying for niche, slow-moving genres/programs: certain audience segments might not contribute to the viewership of the popular programs/niches, but they demonstrate loyalty in viewing a niche program/genre (ie travel, infotainment and news). This opens up opportunities to look beyond CPRP and GRP, and give rise to loyalty and engagement score based pricing. The engagement score can be capitalized on to create value for niche genre/program by contextually positioning brands / ads targeting such audience groups.

More specifically, it allows the service provider to target programs by:

- Effective Rate *(rate per 10secs of advertisements)*
- Reach *(# of eyeballs viewing the program)*
- Engagement Score *(avg % of the program viewed by target audience)*
- Subscriber behavior & demographics

Pivoting on a specific metric will each target a specific use case set:

Effective Rate: in short pivoting on this metric can allow the provider to uncover hidden assets – or those assets that may have an undervalued effective rate. Drilling down on programs with a high Engagement and a high Reach produces a set of programs that is highly valuable to viewers. By pivoting on the effective rate, programs quickly surface and those with a low ER can be targeted for revenue maximization, with the underlying data being presented as justification for further monetizing these undervalued programs. Proving a program has 80% eyeballs that watch it week after week can be highly valuable justification. Over time, the provider can track this target market engagement and correlate with behavioral/demographic data.

Reach: simply, this allows a provider to capitalize on large target audiences. By drilling down on programs that have a low engagement score and a low effective rate, then pivoting on high reach, provides opportunities to capitalize on eyeball-based contracts. Furthermore it allows for the potential to take action on increasing the engagement through exploration of eyeball-based behavior & demographic (and therefore, increasing the effective rate).

Engagement Score: programs that have a highly engaged target audience present

Combination #	Effective Rate	Reach	Engagement Score	Output	Business Case
1	High	High	High	Top Programs	Normal behavior
2	High	High	Low	Widespread Programs	Already capitalizing; can be sold on eyeball-based contracts (quick mileage)
3	High	Low	High	Highly engaged target audience	Good: already capitalizing on high engagement score
4	High	Low	Low	Risky	May impact demand; may be replaced by another program/go off air
5	Low	High	High	Hidden Assets	Identify potential programs to monetize by increasing ER
6	Low	High	Low	Widespread Programs	Opportunity to capitalize on eyeball-based contracts (quick mileage)
7	Low	Low	High	Highly engaged target audience	Opportunity to increase the ER for the programs with highly engaged target audience
8	Low	Low	Low	Ignore	At extinction stage; may be replaced/ go off air soon; normal behavior

Figure 9: Differing combinations of Effective Rate, Reach, & Engagement score have unique business case outputs. [12]

opportunities to increase the effective rate. By providing data justification around the target market, specific ad plans can be sold to monetize this highly engaged long-tail segment.

As a summarization of the use cases described above, figure 9 describes eight combinations of metrics that will each power a specific business case and will each provide a unique output.

Yield Management:

Implementing stateful inventory/yield management allows providers to create targeted media plan bouquets through historical analysis and relevant references. This maximizes ad revenues, by minimizing deadweight loss (ie from overbooking at a single price point – see figure **x**) and by monetizing programs by per-eyeball behavior and makeup (ie long-tail programming). This module supports two distinct use case channels – (1) Generation of Program Bouquet; (2) Deal Evaluation.

Generating a program bouquet based on inputs/constraints will consume the entire budget from the media buyer (reduce deadweight loss), while saving premium inventory and satisfying the constraints from the media buyer. Inputs from the media buyer can range from the Budget, the GRP (gross rating point), CPRP (cost per rating point), Reach/Eyeballs, Engagement, Campaign Period, and the underlying demographic makeup. By providing demographic analysis of all the eyeballs, the media buyers can target their intended audience through a bouquet of programs (including long-tail) rather than opting

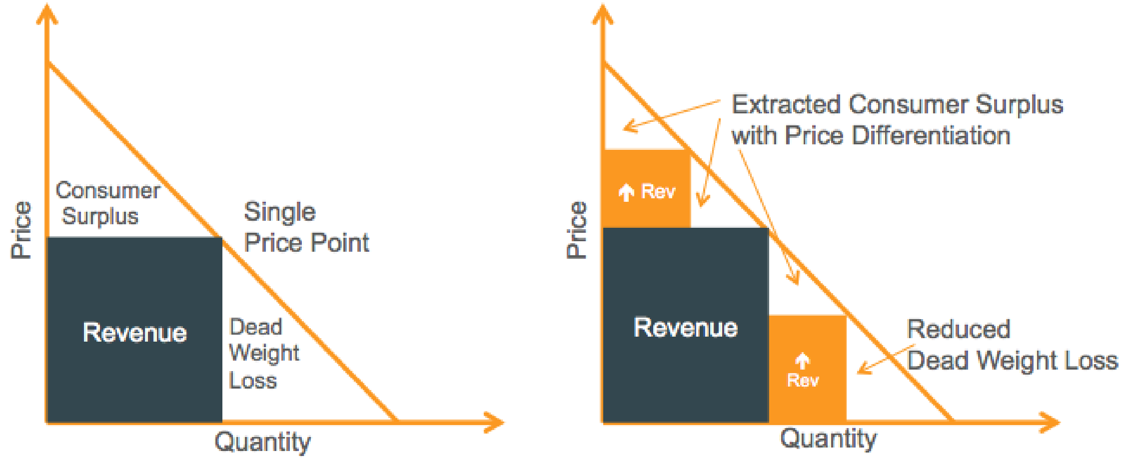


Figure 10: single price-point results in deadweight loss. By offering multiple price points (with underlying data as justification), the operator can reduce deadweight loss. [12]

for the most popular programs (which are generally in higher demand than supply).

(2) Subsequently deal evaluation will evaluate the profitability of a deal, based on user-set business constraints. Less profitable records can subsequently be removed, to free up inventory. In the case of unavailability, the provider can suggest an alternative bouquet.

CONCLUSION

With the eventual transition to an all-IP architecture, the increasing depth of data available will provide a significant competitive advantage to the operator who chooses to leverage it. Application data and player data provide a strong set of complementary analytical insights. Fusion of both player data and application data into a structured software application can enable the service provider to identify a number of actionable insights from a per-subscriber level for IP Video services and a content/network level for CDN deployments. Customer experience metrics allows CDN decisioning; viewer measurement enables marketing to perform ad decisioning; and network engineers can leverage asset analytics for network optimization.

As service providers transition towards all-IP, they are also moving towards a virtualized environment. CCAP is beginning the transition towards a centralized control plane; one that allocates functions in a stateful manner. Structuring the data appropriately – with business context use cases in mind such as optimizing for QoE –, and within a big data fabric that supports low latency, will support these stateful control functions. Furthermore, even before the control functions become completely virtualized, an application can be structured in closed-loop format to perform necessary network optimizations based on anomalous activity (changing serving CDN, adjusting bitrate, etc) and marketing operations (interfacing with ad decisioning engine).

REFERENCES

[1] Altman, Ross, and Bruce Robertson. "Learn a Few Lessons on Application Architecture From Netflix." *Gartner* (n.d.): n. pag. 7 June 2013. Web.

[2] Bhalla, Vivek, and Jonah Kowall. "Use These Best Practices to Manage IP Voice, Video and Unified Communications Deployments." *Gartner* (n.d.): n. pag. 21 Mar. 2013. Web.

[3] Brovedani, Ric. "Making Any Screen Your Main Screen Reality: Simplifying the Complexities of TV Everywhere Services." *SCTE Cable-Tec* (2013): n. pag. 21 Oct. 2013. Web.

[4] Cerqueira, E., M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin. "An Overview of Quality of Experience Measurement Challenges for Video Applications in IP Networks." *Wired/wireless Internet Communications: 8th International Conference, WWIC 2010, Luleaa, Sweden, June 1-3, 2010: Proceedings*. By R. Serral-Gracia. Berlin: Springer, 2010. 252-63. Print.

[5] Horrobvin, John, and Gitesh Shah. "Pioneering IPTV in Cable Networks." *SCTE Cable-Tec* (2013): n. pag. 21 Oct. 2013. Web.

[6] Inácio, António Pedro, Rui Santos Cruz, and Mário Serafim Nunes. "Quality User Experience in Advanced IP Video Services." *Annals of Telecommunications - Annales Des Télécommunications* 68.3-4 (2013): 119-31. Print.

[7] Jammeh, E., I. Mkwawa, A. Khan, M. Goudarzi, L. Sun, and E. Ifeachor. "Quality of Experience (QoE) Driven Adaptation Scheme For voice/video over IP." *Telecommunication Systems* 49.1 (2012): 99-111. Print.

[8] Kowell, Jonah, and Joe Skorupa. "ADCs and APM Solutions Will Combine to Optimize Delivery Intelligence." *Gartner* (n.d.): n. pag. 18 Oct. 2012. Web.

[9] Ladid, Lafif, and Ioannis Chochilouros. "The Impact of IPv6 on Video-to-Video and Mobile Video Communications." *Artificial Intelligence Applications and Innovations: AIAI 2012 International Workshops: AIAB, AIeIA, CISE, COPA, IIVC, ISQL, MHDW, and WADTMB, Halkidiki, Greece, September 27-30, 2012, Proceedings*. N.p.: n.p., n.d. N. pag. Print.

[10] O'Donovan, Paul. "Market Trends: Digital STBs, Video-over-IP Content Drives High-Volume Production Worldwide, 2013." *Gartner* (n.d.): n. pag. 30 Aug. 2013. Web.

[11] Parikh, Samir. "The Video Conundrum: Migrating to IP Video." *SCTE Cable-Tec* (2013): n. pag. 21 Oct. 2013. Web.

[12] Patel, Ashutosh. Guavus, Sr. Business Analyst.

[13] Patrick, Charlotte. "Market Insight: CSPs Are Pushing Forward with New Strategies to Fully Monetize Customer Data." *Gartner* (n.d.): n. pag. 1 Aug. 2013. Web.

[14] Petri, Gregor. "Tech Go-To-Market: Four Strategic Options for CSPs to Explore Cloud Computing Opportunities." *Gartner* (n.d.): n. pag. Web. 2 Oct. 2013.

[15] Rickard, Neil, and Andrew Lerner. "Four Steps to Optimize Your Network for IaaS." *Gartner* (n.d.): n. pag. 20 Dec. 2013. Web.

[16] Sharma, Akshay K. "Emerging Technology Analysis: Cloud-Based Mobile Content Delivery Networks Can Help CSPs Avoid Becoming 'Bit Pipes'" *Gartner* (n.d.): n. pag. 19 Sept. 2011. Web.

KEY ACRONYMS

- ABR: adaptive bitrate encoding
- CBR: constant bitrate encoding
- CCAP: coverged cable access platform
- CDN: content delivery network
- CLV: customer lifetime value
- CPRP: cost per rating point

- CSP: communication service provider
- DOCSIS: data over cable service interface specification
- GRP: gross rating point:
  HFC: hybrid fiber coaxial
- HEVC: high efficiency video encoding
- IP: internet protocol
- IPDR: internet protocol detail record
- KPI: key performance indicator
- MIBS: management information base
- MoS: mean opinion score
- OSS/BSS: operational / business support systems
- OTT: over the top
- PCMM: PacketCable MultiMedia
- QAM: quadrature amplitude modulation
- QoS/QoE: quality of service / experience
- RF: radio frequency
- RODP: run of day-part
- SDV: switched digital video
- SDN: software-defined network
- SON: self-organizing network
- STB: set-top box

DELIVERING THE NEXT GENERATION OF CONTENT: TECHNOLOGIES TO SUPPORT ULTRA HIGH DEFINITION TELEVISION

Wade Wan
Broadcom Corporation

Abstract

Ultra High Definition Television (UHDTV) technology, also known as UltraHD or 4KTV, is already being embraced by content producers in both film and television. As broadcasters begin to investigate UHDTV as a way to significantly improve the viewer experience, operators will need a way to deliver premium content at higher resolutions than possible with current services.

This paper will discuss recent technical advances that will allow cable providers to efficiently and cost-effectively deploy UHDTV offerings in the near term. It describes how new technologies and industry standards, such as HEVC/H.265 video compression, BT.2020 color gamut, HDMI 2.0 digital interface, and HDCP 2.2 content protection must be supported in next-generation set-top boxes to enable cable providers to deliver next generation content.

INTRODUCTION

Generally, UHDTV refers to doubling the horizontal and vertical resolutions, as compared to the 1080p and 1080i High Definition Television (HDTV) formats. In addition to increased image resolution, increased pixel density enlarges the end user's field of vision and promises a more realistic and immersive viewer experience.

Within the industry, there is general agreement that UHDTV services should increase both the bit depth and frame rate of video to further differentiate it from current HDTV services, both technically and commercially [1]. Increasing the bit depth enables the transmission and display of more brightness levels and colors, and increasing the frame rate reduces motion judder artifacts that appear when high motion video is transmitted at insufficient frame rate.

Encoder and decoder implementers have started supporting an UHDTV format of progressive scanned video with 3840x2160 pixels, 60 frames/second, and 10-bit pixel depth, which is a significant quality improvement over current HD formats such as interlaced scanned video format with 1920x1080 pixels, 60 fields/second, and 8-bit pixel depth.

As shown in Table 1, the uncompressed pixel rate of this new UHDTV format is a 10x increase over existing HDTV formats. It would be inefficient, if not impractical, to deploy UHDTV services using the existing methods used for HDTV services.

Table 1: Comparison of HDTV and UHDTV Formats and Uncompressed Pixel Rates

Parameter	HDTV Format	UHDTV Format
Spatial Resolution (pixels/frame)	1920x1080	3840x2160
Pixel Depth (bits/pixel)	8	10
Frame Rate (frames/sec)	30	60
Uncompressed Pixel Rate (Mbits/sec)	497.664	4,976.640

VIDEO COMPRESSION

The significant increase in the video data rate poses a key challenge to new UHDTV service offerings, since transmission bandwidth is a valued and costly resource. The resulting increase in the compressed bit rate does not scale by the same 10x factor as the uncompressed pixel rate, with the exact ratio being content and operating-point dependent.

To address this challenge, the HEVC/H.265 video compression standard [2-4], ratified in January 2013, provides a 2x improvement in coding efficiency compared to the AVC/H.264 video compression standard and a 4x improvement compared to the MPEG-2 video compression standard.

The example bit rates in Table 2 show how compressed bit rates scale for UHDTV formats. Note these examples are estimates with actual bit rates in practice depending on the content and encoding complexity. As the table indicates, there is an approximate 150% to 250% increase in compressed bit rate for each UHDTV bitstream as compared to an HDTV bitstream, even after incorporating the 2x increased coding efficiency provided by the HEVC/H.265 video compression standard for the UHDTV formats. While there have been UHDTV demonstrations using AVC/H.264, it is generally accepted that any realistic UHDTV service cannot use AVC/H.264 and must utilize HEVC/H.265 for the increased coding efficiency it provides.

Table 2: Bit Rate Examples

Resolution	Codec	Bit Rate
HDTV 8 bit 1920x1080p30	AVC	6 Mbps
UHDTV 10 bit 3840x2160p24	HEVC	10 Mbps
UHDTV 10 bit 3840x2160p60	HEVC	15 Mbps

COLOR GAMUT

Modern display capabilities already exceed current color gamut standards such as ITU-R Rec. BT.709 chromaticity standard used for HDTV [5]. To effectively support the color reproduction capabilities of future display technologies like OLED, an expanded color gamut is needed. ITU-R Rec. BT.2020 was announced in August 2012 to define such a color gamut [6].

A comparison of the BT.2020 and BT.709 color gamuts is shown in Figure 1. Both color gamuts have the same reference white but BT.2020 has a wider gamut to represent more colors, and thus requires a larger bit depth to properly sample/represent the range of colors within the gamut, which is one reason why BT.2020 is defined for higher bit depths (10-bit and 12-bit) than BT.709 (8-bit and 10-bit). The transition to 10-bit coding enables the use of a wider color gamut. Until now, this would not have been possible without most likely generating artifacts from a sparse sampling of a larger color gamut with 8-bit coding.

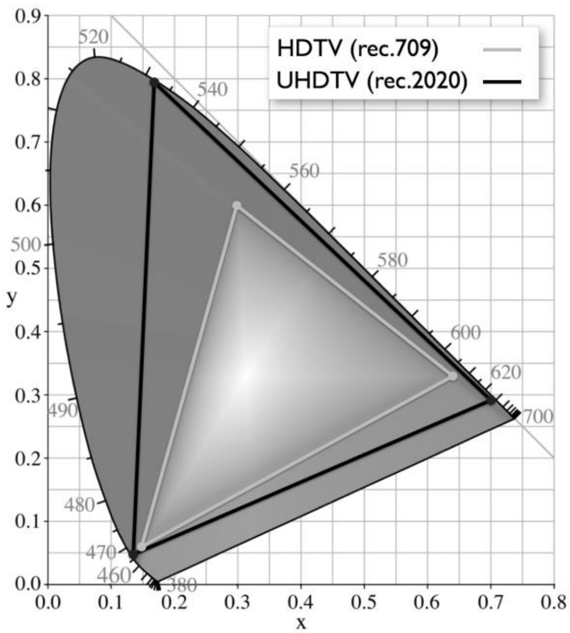


Figure 1: HDTV and UHDTV Color Gamuts

DIGITAL INTERFACE

The increase in the uncompressed or decompressed data rate of UHDTV video is quite significant, especially when using a digital interface to connect two devices, such as a set-top box and a display. While other alternatives are available, such as DisplayPort developed by the Video Electronics Standards Association (VESA), High Definition Multimedia Interface (HDMI) is still the most natural and prevalent interface for connecting high definition components.

Until recently, HDMI interfaces had limited UHDTV format support. More specifically, because the maximum throughput across the interface was limited to 10.2 Gbps [7], a single HDMI 1.4b interface was only capable of supporting UHDTV formats when the pixel data was 8-bit and the frame rate was less than 30 frames per second.. Conceptually, this limitation could be overcome by using multiple HDMI interfaces. However, this soulution would be technically problematic since it could possibly lead to synchronization issues between the interfaces and problems due to the increased cost of duplicating cables and connectors.

To address this limitation, the HDMI 2.0 specification announced in September 2013, increased the bandwidth supported across the interface to 18 Gbps [7]. A comparison of the 4K formats supported by the previous (1.4b) and new (2.0) versions of the HDMI specification is shown in Table 3. Note that the new specification adds support for UHDTV resolutions at 50/60 Hz over a single interface, thereby potentially avoiding implementers and end users struggling to use multiple cables and connectors to handle the increased throughput of UHDTV services.

Table 3: 4K Formats Supported By HDMI 2.0

4K formats	Previous HDMI 1.4b	New HDMI 2.0
8 bit 24/25/30 Hz RGB/4:4:4	YES	YES
10 bit 24/25/30 Hz RGB/4:4:4	NO	YES
8 bit 50/60 Hz RGB/4:4:4/4:2:0	NO	YES
10 bit 50/60 Hz 4:2:0	NO	YES

CONTENT PROTECTION

High-bandwidth Digital Content Protection (HDCP) is a digital copy protection and digital rights management specification for securing audiovisual content between devices. HDCP version2.2, which can be mapped to HDMI interfaces [8], was released in February 2013. It is important to note that this version was made "clean" by not being backward compatible to previous HDCP versions 2.0 and 2.1, which were identified as possibly having a security breach in August 2012.

Additional content protection measures, such as watermarking, might be implemented to protect high-value UHDTV content. However, securing the HDMI digital interface with HDCP 2.2 is expected be the minimum requirement for securing the transmission of UHDTV content and satisfying the distribution requirements of content providers. As an example, MovieLabs has published a content protection specification listing HDCP 2.2 (or better) as a link control/protection requirement [9].

CONCLUSION

The promise of delivering UHDTV services offers an exciting opportunity to provide viewers with a new, immersive experience. The original concept of doubling horizontal and vertical resolution from current HDTV services has expanded to also include increasing the bit depth and frame rate to differentiate UHDTV services from HDTV services.

A key challenge has been the increased bit rates that result from the UHDTV pixel rates. The HEVC/H.265 video compression standard represents a technological milestone for video coding with its 2x improvement in coding efficiency as compared to the previous state of the art AVC/H.264 video codec. With transmission bit rate being a key consideration, the practical solution is for UHDTV deployments to use HEVC/H.265 to minimize bit rates.

The challenge with delivering UHDTV services is not limited solely to the video codec. It also requires advances in other parts of the system pipeline. This paper has highlighted three new technologies and industry standards that were developed for use in UHDTV systems. The BT.2020 color gamut expands the color gamut to better match the capabilities of current as well as future capture and display equipment. The HDMI 2.0 digital interface can better support 4K formats. HDCP 2.2 content protection can secure the HDMI interface in order to protect UHDTV content.

Now that these technologies have matured into industry standards, cost-effective, single System-on-Chip (SoC) solutions have been designed and are being used to power next-generation set-top boxes [10,11]. This will allow cable providers to efficiently and cost-effectively deploy UHDTV equipment and services in the near future.

REFERENCES

[1] ITU-R, "The present state of ultra-high definition television", Recommendation BT.2246-2, Nov. 2012.
[2] ITU-T, "High Efficiency Video Coding", Recommendation H.265, Apr. 2013.
[3] G.J. Sullivan, J.-R. Ohm, W.-J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. Circuits Syst. Video Technol., Vol. 22, No. 12, Dec. 2012.
[4] J.-R. Ohm, G.J. Sullivan, H. Schwarz, T.K. Tan and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)", IEEE Trans.

Circuits Syst. Video Technol., Vol. 22, No. 12, Dec. 2012.
[5] ITU-R, Recommendation ITU-R BT.709-5 : Parameter values for the HDTV standards for production and international programme exchange, Apr. 2002.
[6] ITU-R, Recommendation ITU-R BT.2020 : Parameter values for ultra-high definition television systems for production and international programme exchange, Aug. 2012.
[7] FAQ for HDMI 2.0, http://www.hdmi.org/manufacturer/hdmi_2_0/hdmi_2_0_faq.aspx
[8] Digital Content Protection LLC, "High-bandwidth Digital Content Protection System - Mapping HDCP to HDMI", Revision 2.2, Feb. 2013.
[9] MovieLabs Specification for Enhanced Content Protection, v1.0, 2013.
[10] Broadcom Unveils World's First UltraHD TV Home Gateway Chip, Press Release, http://www.broadcom.com/press/release.php?id=s732069, Jan. 2013.
[11] Broadcom Proliferates UltraHD Set-top Box Availability with New HEVC Chipsets, Press Release, http://www.broadcom.com/press/release.php?id=s790302, Sep. 2013.

Discussion of DOCSIS 3.1 Deployment Migration

Carol Ansley   Jeff Howe
ARRIS Group, Inc.

Abstract

As DOCSIS 3.1 prepares to begin its rollout over the next several years, many questions have to be analyzed to determine the most efficient transition methodologies. This paper considers the problem of network transition. It will look at the services currently deployed and consider how to overlay new D3.1 services. The factors considered in the model will include the HSD usage of current customers over D3.0 and D2.0 modems, and several different models of new D3.1 customers and their potential usage patterns. The D3.1 usage patterns could include commercial customers like small businesses or larger customers that might use D3.1 tiers for data backups off-hours. Another group of potential D3.1 users could be IPTV subscribers. The different D3.1 user groups would potentially have different performance expectations and different usage patterns that might allow a more efficient deployment alongside the current customers. Finally, the long-term transition of existing customers should be considered.

The modeling of various scenarios would also consider the different modes of deploying the new D3.3 tier, overlay or new spectrum, as a part of the modeling process.

Introduction

After months of standards committee work, DOCSIS3.1 specifications have been released, and the cable industry's vendor community is hard at work developing chips, software and systems for the coming deployments. But, what form will those deployments take? This paper explores some options available to operators as they consider how to roll out this new technology.

DOCSIS 3.1 Considerations

DOCSIS 3.1 incorporates some fundamentally new technologies to enable superior throughput over the cable plant. Orthogonal Frequency Division Multiplexing, or OFDM, is used instead of traditional single carrier Quadrature Amplitude Modulation, SC-QAM.

Downstream Details

The Media Access Control layer has also changed from a general broadcast downstream to a more complex and efficient directed transmission. Broadcast mode is still possible, but now subsets of cable modems sharing transmit path characteristics can be addressed as well. This allows modems which are experiencing less noise impairment to use higher modulation orders for higher throughput. The channel width for a D3.1 Downstream (DS) channel has changed as well. A D3.1 channel can range in size from 24MHz up to a maximum size for a single channel of 192 MHz. The Downstream spectrum range for D3.1 is changed from an earlier version. The required lower end is moved to 258 MHz, and the high end can now reach up to 1.218 GHz, with an option to reach to 1.794 GHz.

An additional feature is that data flows destined for D3.1 modems can be bonded across D3.0 and D3.1 channels. This capability can allow a new D3.1 section to be opened up for use while still allowing D3.1 modems to potentially utilize some or all of the D3.0 band to pass additional content.

Since D.31 modems are also required to be able to access D3.0 channels, multicast content present on D3.0 channels does not have to be repeated in the D3.1 band.

Upstream Details

In the upstream direction, the MAC changes have been less dramatic than in the DS, but the PHY changes are still significant. OFDM has also been adopted in the US, and the minimum and maximum channel widths have been changed to 6.4 MHz and 96 MHz. The required frequency range has also shifted. A D3.1 CM and CMTS are required to support an US that can reach at least 85MHz, with an option to go up to 204 MHz.

Current Traffic Behavior and Predictions

The possibilities for user migration onto D3.1 are potentially shaped by usage patterns. In a recent study by Sandvine®, they described upstream and downstream usage patterns on wired links.[i] This section examines some of the recent studies and reports on bandwidth usage to provide an informed basis for speculation about future trends.

Downstream Traffic

An important finding was that downstream usage is more homogeneous that it had been in the past. In some previous years, a small percentage of users accounted for a disproportionate proportion of bandwidth use, for example in 2011, 1% of the users accounted for about 25% of the downstream traffic.[ii] The latest report shows that now 1% of the users only account for about 10% of overall downstream traffic.

This shift in the concentration of users' traffic can be partially explained by the increase in streaming video. The same Sandvine report also showed that Netflix® users accounted for over 30% of the bandwidth usage during primetime peak hours. YouTube users accounted for almost 19% of the peak bandwidth. In fact the overall estimate for streaming video during primetime was over 67% of the downstream bandwidth. This percentage has been steadily increasing over the years (graph here). Currently Netflix subscribers are between 30% and 40% of the general population depending on whether you rely on PriceWaterhouseCoopers' July 2013 study or USA Today in September of 2013, so their use of 30% of primetime bandwidth seems roughly proportional.

Based on the trends seen in the marketplace, the consumer preference for streaming media in the home is likely to keep growing. The only disputed area of streaming media is whether bandwidth streaming to applications loaded onto multi-purpose devices will surpass that streaming to consoles and smart TVs. The bandwidth for the mobile devices is usually smaller due to the smaller form factor, but there are many more of those devices in the consumers' hands.

Upstream Traffic

Turning to upstream bandwidth usage, the peer to peer traffic surge that dominated upstream traffic engineering has begun to decline. BitTorrent and other file sharing applications are now about 42% of the total upstream traffic at peak in the latest Sandvine report references earlier. This level represents a substantial decline over years past, for example in the Sandvine 2011 report BitTorrent alone represented 52% of the upstream traffic at during primetime.

The slow decline of file sharing applications has also decreased the level of user asymmetry in the upstream, though it is still greater than it is in the downstream. The top 1% of upstream users still account for over 40% of traffic.

Looking toward the future, the upstream traffic mix may continue to fragment. Some real-time applications have been growing, such as FaceTime. Cloud storage applications, such as Dropbox, have also been growing. As the trend toward cloud everything continues, the consumer will expect to be able to use the cloud for storage and processing of personal information, as they become used to working with that sort of technology in their businesses. That trend will undoubtedly increase the variety of application consuming upstream bandwidth.

D3.1 Migration Strategies

The new features incorporated into D3.1 allow much flexibility in designing a network migration strategy and also present many options for the end configuration of the target network after the migration is complete. As a first principle, this paper assumes that as much as possible a new D3.1 deployment should not disrupt existing customers, while still enabling the operator to take advantage of the higher bandwidth services that can be provided by deploying D3.1.

Contention between services is already intense for all available upstream and downstream spectrum. To squeeze in new D3.1 spectrum into the current bandwidth in most systems is challenging. We will discuss some options for spectrum selection in the Upstream and Downstream sections, since the available choices for each band differ significantly and have profound implications in the amount of upfront work needed to enable the expanded spectrum selections of D3.1.

Aside from the spectrum allocation and availability questions, another question will be explored, what services and customers can be most profitably be used to drive the deployment of D3.1? The changing dynamics of customer traffic profiles discussed earlier play into consideration of what customers can be most advantageously targeted for D3.1 modem deployment.

D3.1 Upstream Migration Strategies

Migration strategies possible for upstream D3.1 present some interesting challenges. The D3.1 specification implicitly assumed that substantial changes can be made to the current HFC network. To accomplish the changes, downstream subscriber services will also be affected.

In order for a D3.1 upgrade to be most effective, the outside plant needs to be upgraded to support at least 85 MHz of upstream spectrum, a substantial increase from the current 42MHz in North American and 65MHz in Europe. The specification supports an expansion of upstream spectrum up to 200MHz, though at this point few operators are convinced that upstream spectrum needs that much expansion. The upstream upgrade will involve moving any existing downstream channels out of the new expanded upstream band, and may also drive the need for DTA deployment if they are not already in place. When the downstream channels, usually video, are removed, a consumer education campaign may be required if DTAs are not in use, since the lowest video channels have for many years been preferentially used for basic tier service. Any outside plant elements with a diplexer or any form of two way active operation may need modification or replacement. This change will include fiber nodes, amplifiers, line extenders and even the humble house amplifier.

Since D3.1 allows exclusion regions in the middle of a channel, the existing legacy DOCSIS channels can be left in place in the upstream to allow legacy modems to continue to operate effectively. A single D3.1 channel can be defined encompassing the spectrum occupied by the legacy channels without disturbing them. In this manner, the D3.1 channel can make use of both the lower portion of the spectrum which is typically unusable by SC-QAM channels and the higher portion of the spectrum made available by the expansion to 85 MHz.

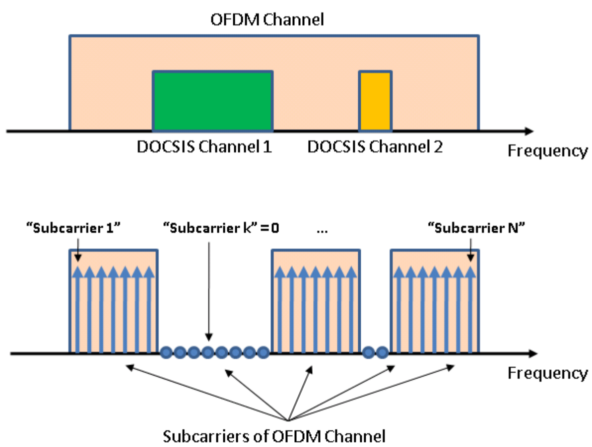


Figure 1 Illustration of OFDM Overlap with Legacy Channels[iii]

As part of a migration strategy to D3.1, customers on the highest speed tier could be moved to the new D3.1 channel once the plant modifications are completed. A D3.1 channel can provide more data throughput than legacy DOCSIS channels, and moving the highest bandwidth customers to the new channel will provide them with more available bandwidth, and will free up bandwidth for the other customers left on the legacy channels. It is expected that initially only a small number of customers will need to be moved to the D3.1 channel. Alternatively, higher usage customers could be identified and invited move to the new channel. As an example, if a substantial fraction of the BitTorrent users in a service group could be moved, it would have an out-sized impact in the bandwidth available to the customers remaining on the older channels since 30% or more of the primetime traffic would move with them.

As more customers are migrated over to the D3.1 channel, available bandwidth on that channel may become an issue. In order to make more bandwidth available to those customers, a next step would be to channel bond those customer modems across the D3.1 and legacy channels. This change will make more bandwidth available to the D3.1 modems. Since the number of legacy customers is also decreasing as they are slowly moved to D3.1 devices, the subscribers remaining on older modems should not be adversely impacted by this change.

It is likely that the D3.1 channel will support a higher modulation order in the spectrum used by the legacy channels. D3.1 offers an alternative to using channel bonding to make use of the spectrum occupied by the legacy channels. D3.1 allows that spectrum to be time multiplexed between the SC-QAM channels and the OFDM channel. However, due to the D3.1 OFDM framing structure, this mechanism will likely not be an efficient use of bandwidth until a sufficient number of D3.1 modems is present to assure that the majority of OFDM frames are filled with data. Before the number of D3.1 modems are available to make this method efficient, the channel bonding of OFDM and SC-QAM channels can continue to be used.

Longer term, as bandwidth demands continue to increase, the plant can be upgraded to support a 204 MHz upstream split and a second OFDM upstream channel added. The two OFDM channels can be channel bonded together, and provide a customer with a 1Gbps upstream service. The second upgrade will again require migration of affected downstream channels, and an upgrade of active elements in the network. Hopefully, network elements developed for the 85MHz upgrade can incorporate simple cost-effective methods to support a later 200MHz upgrade. Alternatively, an operator may elect to make only a single upgrade but push up to the 200MHz boundary. This decision will probably be motivated by the demographics and demand predictions seen by the individual operator.

D3.1 Downstream Migration Strategies

The Downstream features within D3.1 provide the same channel bonding and channel squelching that were discussed in the US migration section. The potential use of the squelch feature may not be as widespread in the DS as it promises to be in the US since there are few, if any, areas of vacant channel in the current DS bands. The channel bonding feature may be used to allow traffic spreading between D3.0 channels and D3.1 channels if new D3.1 spectrum is used to augment the current DS bandwidth available, but without sufficient bandwidth within just the D3.1 spectrum to support new service tiers that may be targeted to receive D3.1 CPE (Customer Premises Equipment) devices.

A more expedient course may be to place a new D3.1 band above the current services, above 750MHz. Using this spectrum may be precluded if there are network elements such as fiber systems or amplifiers that will block these frequencies, but the standard choice for most HFC outside plant deployments for some time has been 1GHz elements, even if no STBs or CMs deployed are capable of using this spectrum. If the new D3.1 band is deployed without having to work around existing services by using previously unused spectrum, it may be possible to deploy a D3.1 channel with sufficient bandwidth to avoid the complication of effectively requiring some portion of the D3.0 bandwidth to be reserved for new D3.1 users.

Turning to potential subscriber migration issues, many MSOs have been pursuing business services with increasing success in the past few years, but to provide competitive service, MSOs have often had to run fiber drops, which may not be cost effective for small to medium business (SMB) opportunities. The higher modulation rates of D3.1 offer new service capabilities that can dovetail well into the SMB market segment. New CPE units for D3.1 will probably be more expensive than standard D3.0 modems at least until volumes become large because a D3.1 modem is defined in the specifications as a D3.0 modem with an additional D3.1 Media Access Control layer (MAC) and Physical layer (PHY). Though many MSOs would like to aggressively deploy the new units, similar to the way the D2.0 and D3.0 CPE units were deployed in advance of the widespread availability of headend equipment and frequency allocation, the additional cost may pose an obstacle in early deployments. SMB customers can offer an excellent target for those early deployments. Their bandwidth requirements will fit well within the expected D3.1 performance and can be used to justify a higher price for the higher performing service.

Another group that may be advantageously deployed with new D3.1 modems could be comprised of the heaviest users within the residential ranks. With the shifting of traffic patterns for DS users, the candidate users are not as simple to target as they might have been in years past. Users who make frequent use of high resolution streaming media may be a useful group to consider for the first migration, but their ranks are fairly large. Additionally, streaming media protocols typically shift their resolutions in response to the presence or absence of bandwidth. Moving some streaming users to the new D3.1 band may not provide much bandwidth relief to the legacy band if the remaining users’ streaming clients soak up much of that newly available bandwidth. If the majority of the heavy streaming users are shifted to the new band, then the legacy users will have substantial bandwidth relief during prime time.

Conclusions

After several years of discussing and debating D3.1 features, the first CPE and headend units will arrive soon. Migration planning needs to begin in earnest informed by the latest trends in user behavior as well as technical and financial feasibility. The greatest challenges to deployment of D3.1 services are found in the new upstream since extensive plant work will be needed – downstream deployments may be simpler.

On the customer migration side, business customers, since they may have both increased upstream and downstream bandwidth requirements, are excellent targets for the first D3.1 customer deployments. Assuming streaming media continues to gain in popularity, its residential users can only increase their domination of the downstream bitstream which will tend to create a broad class of users that may be good candidates for a D3.1 transition group in the downstream.

Overall D3.1 presents challenges and opportunities as the industry tries to deploy new technology while keeping ahead of the consumers whose bandwidth uses are constantly changing.

References

[i] "Global Internet Phenomena Report", **Sandvine**, 2H2013, as found at https://www.sandvine.com/downloads/general/global-internet-phenomena/2013/2h-2013-global-internet-phenomena-report.pdf.

[ii] "Global Internet Phenomena Report Spring 2011", **Sandvine**, as found at http://www.wired.com/images_blogs/business/2011/05/SandvineGlobalInternetSpringReport2011.pdf.

[iii] Ayham Al-Banna, "WiMAX Links and OFDM Overlay for HFC Networks: Mobility and Higher US Capacity", 2010 Spring Technical Forum, NCTA-SCTE, (May, 2010).

DISPOSABLE DOMAINS

Yizheng Chen, Manos Antonakakis, Wenke Lee
Georgia Institute of Technology

Abstract

In recent years DNS has been increasingly leveraged to build and scale highly reliable network infrastructures. In this paper, we will introduce and analyze a new class of domains, which we refer to as disposable domains. Disposable domains appear to be heavily employed by common Internet services (i.e., Search Engines, Social Networks, Online Trackers etc.), and they seem to be automatically generated. They are characterized by a "one-time use" pattern, and appear to be used as a way of "signaling" via DNS. While this is yet another "creative" use of the DNS to enable new Internet applications and efficient scaling of services, little do we know about the size and DNS caching properties of this family of domains.

To shed light on the pervasiveness and growth of disposable domains, we present a study of their characteristics based on live DNS traffic observed at Comcast, in a city that serves millions of end users. We found that disposable domains increased from 23.1% to 27.6% in all queried domain names, and from 27.6% to 37.2 % among all resolved domain names daily, and more than 60% of all distinct resource records observed daily in modern DNS traffic are related to disposable domains. We discuss the possible negative implications that disposable domains may have on the DNS caching infrastructure, resolvers validating DNSSEC transactions, and passive DNS data collection systems.

INTRODUCTION

Domain Name System was originally designed for mapping a human-friendly domain name to a machine-readable IP address. Over the years, people have used DNS in new ways to make their services more agile and scalable. However, they all had unanticipated and sometimes negative impact as the following three examples shows.

The first example is using DNS to select a Content Delivery Network (CDN) server that is closest to client. When a CDN sees a DNS request for content, it will return a CDN server IP address that is closest to the requester IP, and with small load at the time. Since what CDN sees is the IP address of the DNS server that user's machine is configured to, not the user's IP address, the effectiveness of such approach depends on how close users are to their local DNS servers. Researchers [1] have shown that 64% of associations of user's and the local DNS server's IP addresses are in the same Autonomous System. However, only 16% of associations are in the same network-aware clusters, from the perspective of BGP routes. The second example is browser prefetching to speed up webpage loading performance [2]. When a user is entering search queries, the browser will look up unfinished search queries as possible domain names and pre-resolve all the domain names before user finishes typing. The design of prefetching is used for web objects as well, to minimize the delay user perceives while browsing. However, an unanticipated negative impact from that is DNS prefetching could potentially leak user's privacy by exposing the search terms in just the DNS queries. The last example is NXDOMAIN redirection for displaying commercials. Parked domains are often redirected to advertisement pages to monetize existing users for the old domain name. The practice of doing that was called "DNS lie" [3] [4]. It has always been controversial of whether ISPs should do that, since advertisement page is not the page users intend to look for.

```
load-0-p-01.up-1852280.mem-251379712-24440832-0-p-50.swap-236691456-297943040-0-p-44.3302068.1222092134.device.trans.manage.esoft.com
load-0-p-49.up-1066332.mem-118550528-17743872-0-p-49.swap-186757120-347877376-0-p-35.3300639.1643250616.device.trans.manage.esoft.com
load-0-p-90.up-41144.mem-193540096-523649024-0-p-19.swap-56713216-477921280-0-p-11.3303042.3049260335.device.trans.manage.esoft.com
load-0-p-08.up-117864.mem-76529664-15839232-0-p-29.swap-13049856-529776640-0-p-02.8551447.2050639502.device.trans.manage.esoft.com
load-0-p-01.up-122977.mem-76460032-16359424-0-p-29.swap-13180928-529645568-0-p-02.8551447.2050639502.device.trans.manage.esoft.com
load-0-p-01.up-12664453.mem-195096576-117325824-0-p-39.swap-541405184-536096768-0-p-50.5001772.2852986008.device.trans.manage.esoft.com
load-0-p-05.up-2968675.mem-405557248-302886912-0-p-39.swap-91910144-442724352-0-p-17.3300672.2763414838.device.trans.manage.esoft.com
load-0-p-56.up-9190020.mem-112308224-14741504-0-p-43.swap-49680384-493146112-0-p-09.8120531.946954102.device.trans.manage.esoft.com
load-0-p-38.up-1852942.mem-253808640-26693632-0-p-50.swap-236720128-297914368-0-p-44.3302068.1222092134.device.trans.manage.esoft.com
load-0-p-13.up-9160910.mem-108138496-15101952-0-p-41.swap-48463872-494362624-0-p-09.8120531.946954102.device.trans.manage.esoft.com
```

(i)

```
0.0.0.0.1.0.0.4e.135jg5e1pd7s4735ftrqweufm5.avqs.mcafee.com
0.0.0.0.1.0.0.4e.13cfus2drmdq3j8cafidezr8l6.avqs.mcafee.com
0.0.0.0.1.0.0.4e.13kqas3qjj46ttkdhastkrdsv6.avqs.mcafee.com
0.0.0.0.1.0.0.4e.13pq3hfpunqn1d51pmvbdkk5s6.avqs.mcafee.com
0.0.0.0.1.0.0.4e.13qh71bf782qb54uzz9uhdz4mq.avqs.mcafee.com
0.0.0.0.1.0.0.4e.13vw6p3bwdrpilru9g3ffnjdft.avqs.mcafee.com
0.0.0.0.1.0.0.4e.141cl3lsue1evf162v8879z8r6.avqs.mcafee.com
0.0.0.0.1.0.0.4e.149r3hs7fi6p23z2p66t3n5smj.avqs.mcafee.com
0.0.0.0.1.0.0.4e.14akprculurmj5bp9bg16pfj2i.avqs.mcafee.com
0.0.0.0.1.0.0.4e.14fnwz97bjp3n2fzgq4f4nnagj.avqs.mcafee.com
```

(ii)

```
p2.a22a43lt5rwfg.ihg5ki5i6q3cfn3n.191742.i1.ds.ipv6-exp.l.google.com
p2.a22a43lt5rwfg.ihg5ki5i6q3cfn3n.191742.i2.v4.ipv6-exp.l.google.com
p2.a22a43lt5rwfg.ihg5ki5i6q3cfn3n.191742.s1.v4.ipv6-exp.l.google.com
p2.a22antzfkdg5g.nay6cy6qq26fr64b.544760.i1.v4.ipv6-exp.l.google.com
p2.a22antzfkdg5g.nay6cy6qq26fr64b.544760.i2.ds.ipv6-exp.l.google.com
p2.a22bc6fi6edwk.qa2gdjd72sdbycs5.199480.i1.ds.ipv6-exp.l.google.com
p2.a22bc6fi6edwk.qa2gdjd72sdbycs5.199480.i2.v4.ipv6-exp.l.google.com
p2.a22bc6fi6edwk.qa2gdjd72sdbycs5.199480.s1.v4.ipv6-exp.l.google.com
p2.a22cax6c5l5h2.7s2llcerkgtvdu5f.632143.i2.ds.ipv6-exp.l.google.com
p2.a22cax6c5l5h2.7s2llcerkgtvdu5f.632143.s1.v4.ipv6-exp.l.google.com
```

(iii)

Figure 1. Three examples of disposable domain names from eSoft, McAfee, and Google.

As the Internet has evolved over the years, more service providers, such as popular search engines, social networks, and online trackers, began to use a new class of domain names, that we call disposable domains. Disposable domains almost seem to be a natural result of people seeking even more agility and scalability for their Internet services. Using disposable domains, service providers don't need to set up any dedicated infrastructure for their service, but to simply overload DNS with customized protocols. We will discuss the properties of this new class of domain names, specifically focusing on their algorithmically-generated zone structures and their low cache hit rates obtained from a cluster of recursive DNS resolvers operated by Comcast, a large north-American ISP.

This increase in use of disposable domains may have unanticipated negative effects on day-to-day DNS operations for large ISPs. For instance, a large number of DNS requests for disposable domains could fill up the cache of recursive DNS resolvers. Such an event may cause premature cache evictions of non-disposable domains, which would degrade DNS service for the ISP. In turn, these premature evictions may inflate the traffic between the DNS resolvers and authoritative name servers, a phenomenon that could be very costly for ISPs in a DNSSEC-enabled recursive environment. Lastly, disposable domains increase the storage requirement for passive DNS data collection systems, and could potentially degrade database query latency.

In the rest of the paper, we will first show some examples of disposable domains and discuss their properties. Then we will provide supporting evidence on how disposable domains are currently used by large service providers. Lastly, we will discuss possible negative implications that the growth in disposable domains may have on the DNS caching infrastructure, DNSSEC-validating resolvers, and passive DNS data collection systems.

MINING DISPOSABLE DOMAINS

In this section we will define disposable domain names and we will provide some real world examples of their use in case studies. Then, we will discuss the prevalence of disposable domains. We define disposable domain names as successfully resolved domain names that have the following properties:

1). Their name strings are automatically generated.

2). The median cache hit rate for the resource records of child domain names under a zone that facilitates disposable domain names is low or close to zero. In other words, the resource records under that particular zone are only observed once, or a handful of times, when they are in the recursive DNS servers' cache.

Case Studies

Figure 1 shows three examples of what we define as disposable domain names. The eSoft (i) domain names are used as a storage communication channel that reports CPU load, machine up time, memory usage and swap disk usage. The McAfee [5] (ii) domain names are used for file reputation queries on behalf of McAfee's Global Threat Intelligence File reputation Service. This is yet another case of using the DNS as an information storage communication channel. Lastly, Google's IPv6 experiment domains [6] (iii) are queried by browsers of selected users that perform cryptographically signed background requests after receiving their search results. The background requests record IPv4 and IPv6 addresses, image request latency, and User-Agent strings.

Examining the zone structures from Figure 2 shows that 1) disposable domain names tend to have same number of periods ("."), 2) at certain places between two periods, the labels are "random-looking". The structure property reflects how zone operators parse and use different parts of disposable domains for different purposes or transfer different information, by using algorithm-generated strings.

In addition to zone structural properties, disposable domains typically have very low or sometimes zero cache hit rates. Usually, over 90% of cache hit rates from disposable domains are zero. On the other hand, cache hit rates of non-disposable domains follow a closer to linear cumulative distribution, and the median cache hit rate would be around 40%. In general, resource records of disposable domain names are used only once or up to a few times while they are in the recursive cache, which results in the ***overall*** low cache hit rate distribution for domains under disposable zones.

Measurement Results

We built a disposable domain miner system to automatically mine disposable domains. The technical details of our system can be found in [7]. Over the period of a year, we found 14,488 zones that use disposable domains, with a confidence of more than 90%. Disposable domains are used by various industries, including popular websites (e.g., Google, Microsoft), Anti-Virus companies (e.g., McAfee, Sophos, Sonicwall, Mailshell), DNSBLs (e.g., Spamhaus, countries.nerd.dk), social networks (e.g., Facebook, Myspace), streaming services (e.g., Netflix), P2P services (e.g., Skype), cookie tracking services (e.g., Esomniture, 2o7.net), ad networks (e.g., AdSense, Bluelink Marketing), e-commerce business (e.g., Paypal, ClickBank), etc.

Disposable domains are not only widely used currently, but are also increasingly being used. For unique domains being queried by clients, the percentage of disposable domains increased from 23.1% to 27.6%. Also, of the daily resolved unique domains the percentage of disposable domains grew from 27.6% to 37.2% over the year of 2011. From traffic during 11/28/2011 to 12/10/2011, we observe that the number of new disposable domains seen every day is always high, around 5 million to 7 million. However, the number of new non-disposable domains dropped from 13 million to 1.6 million. So after one day, more than 50% of new domains seen daily are disposable, and after 13 days, more than 80%

of new domains seen daily are disposable, since new disposable domains are constantly generated. Moreover, the volume of unique disposable resource records daily increased from 8,111,274 (02/01/2011) to 29,738,493 (12/30/2011), during which 33,704,127 were observed on 11/14/2011. The percentage of daily unique disposable RRs increased from 38.3% to 65.5%.

DISCUSSION

In this section, we will discuss possible negative effects of using disposable domains. We will discuss their impact on DNS caching, DNSSEC-enabled resolvers, and passive DNS databases, so that the operational community can anticipate them and plan ahead in case changes to current DNS operations are needed.

DNS Caching

As disposable domains are increasingly used, the cache of recursive DNS servers may be filled up with entries that are highly unlikely to be reused. Assuming a typical Least Recently Used cache implementation with fixed memory allocation, during periods of heavy load, queries to disposable domains may cause some useful non-disposable domains to be prematurely evicted from the cache. In turn, this may have the effect of unfairly inflating the traffic between the DNS resolvers and the authoritative name servers responsible for the evicted non-disposable domains, thus increasing the query-response latency.

DNSSEC

There will inevitably be more pressure on validating resolvers when DNSSEC becomes more widely deployed. Validating signed responses requires higher CPU usage, and increased memory needs due to DNSSEC specifications [8] [9] [10]. Disposable domains will naturally, and potentially dramatically, increase this pressure on validating resolvers. In fact, each queried disposable domain may require an additional signature validation whose result will never be reused. Also, the cache must not only store the disposable RRs, but also their signatures. This problem may be mitigated in part if the authoritative servers responsible for the disposable zones register disposable domains under a single signed wildcard domain, from which the disposable domains are synthesized.

pDNS-DB

Passive DNS database systems (pDNS-DBs) have recently been adopted by computer security and networking communities as an invaluable tool to analyze security incidents, monitor and troubleshoot DNS operations, and develop dynamic reputation systems [11] [12]. Disposable domains have the effect of increasing pDNS-DB storage requirements, and potentially their query-response latency, depending on the implementation. In fact, we found that after bootstrapping a pDNS-DB with 13 days of resolution traffic, 88% of all unique resource records in the database are disposable, and new RRs related to disposable domains make up more than 94% of all the new distinct RRs observed daily. The problem can be mitigated by filtering disposable domains and storing a single wildcard domain in the pDNS-DB.

REFERENCES

[1] Z. M. Mao, C. D. Cranor, F. Douglis, M. Rabinovich, O. Spatscheck, and J. Wang. A precise and efficient evaluation of the proximity between web clients and their local dns servers. In *Proceedings of the General Track of USENIX ATEC*, 2002.

[2] S. Krishnan and F. Monrose. DNS prefetching and its privacy implications: when good things go bad. In *Proceedings of USENIX Workshop on LEET*, 2010.

[3] P. Vixie. What dns is not. Queue, (10), Nov. 2009.
[4] N. Weaver, C. Kreibich, and V. Paxson. Redirecting DNS for Ads and Profit. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, 2011.
[5] McAfee. Faqs for global threat intelligence file reputation. https://kc.mcafee.com/corporate/index?page= content&id= KB53735, 2013.
[6] S. H. Gunderson. Global IPv6 statistics: Measuring the current state of IPv6 for ordinary users. In *Proceedings of the Seventy-third Internet Engineering Task Force*, 2008.
[7] Y. Chen, M. Antonakakis, R. Perdisci, Y. Nadji, D. Dagon, W. Lee. DNS Noise: Measuring the Pervasiveness of Disposable Domains in Modern DNS Traffic. In *proceedings of the 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, 2014.
[8] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Dns security introduction and requirements. http://www.ietf.org/rfc/rfc4033.txt, March 2005.
[9] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Protocol modifications for the dns security extensions, rfc 4035. http://www.ietf.org/rfc/rfc4035.txt, March 2005.
[10] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Resource records for the dns security extensions. http://www.ietf.org/rfc/rfc4034.txt, March 2005.
[11] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster. Building a Dynamic Reputation System for DNS. In *Proceedings of USENIX Security Symposium*, 2010.
[12] M. Antonakakis, R. Perdisci, W. Lee, D. Dagon, and N. Vasiloglou. Detecting Malware Domains at the Upper DNS Hierarchy. In *Proceedings of USENIX Security Symposium*, 2011.

Distribute or Die: Hybrid Fiber Coax Hanging by a Cable

Samir Parikh
Gainspeed, Inc.

Abstract

The hybrid fiber coax (HFC) networks currently in place were designed to carry broadcast analog video, not the wide ranging high-speed data (HSD) services, IP video products, teleconferencing, cloud computing and other business services they need to deliver today.

There is enough capacity potential in today's hybrid fiber coax (HFC) networks to satisfy bandwidth demand, even with a 40% to 60% annual growth rate. However, a new architectural approach is required in order to unlock this potential.

This paper will explore how cable operators can re-imagine their networks and extend CCAP to its next logical step by virtualizing and distributing functions and services. Doing so will dramatically reduce power and space consumption in the hub and headend, lower costs, speed up service velocity and enable new technologies such as DOCSIS 3.1.

INTRODUCTION

Cable networks are being overwhelmed by a deluge of IP Video traffic. As a result, cable operators face escalating costs to deliver bandwidth and are forced to invest heavily in equipment and infrastructure just to maintain network performance levels. Hybrid fiber coax (HFC) networks were designed to carry broadcast analog video, not the wide ranging high-speed data services, IP Video products, teleconferencing, cloud computing and other business services they need to deliver today. New services and growing bandwidth demands are driving a 40% to 60% average annual compound growth in capacity requirements.

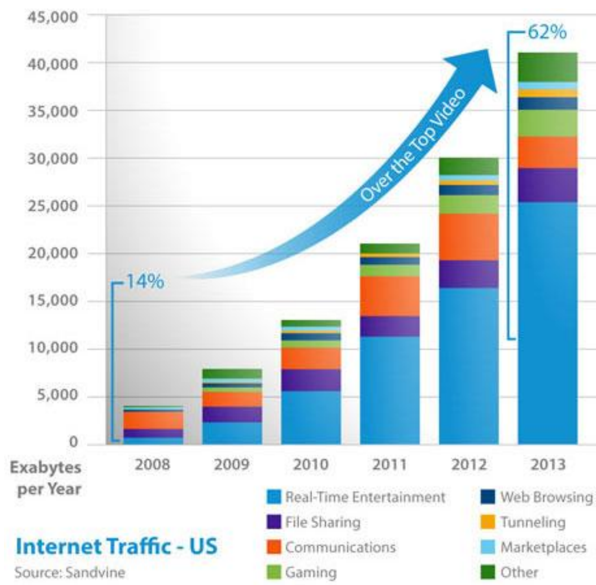


Figure 1: US Internet Traffic Profile

IP Video now accounts for over 62% of all Internet traffic – 67% of downstream traffic – and growing. These trends show no sign of fading and have radically changed everything.

At the time the Converged Cable Access Platform (CCAP) specification was conceived, operators were seeing tremendous demand for both IP services as well as legacy video services (video on demand and switched digital video). CCAP, which combines the physical edge QAM and CMTS devices to support both of IP and legacy video services on a single platform, made a lot of sense. However, just as platforms based on the CCAP specification are beginning to come to market, everything has changed. Growth in legacy video has dropped dramatically and IP Video (both operator provided as well as

Over-the-Top (OTT)) has become the dominant consumer of Internet bandwidth. Pressure from OTT is driving an accelerated shift away from QAM video to IP video, pushing the limits of DOCSIS capacity available in the network today.

The good news is that with DOCSIS 3.0, cable operators have close to 6 Gbps of raw IP capacity on a 1 GHz HFC plant. (DOCSIS 3.1 will further increase this.) The bad news is that there are numerous hurdles in unleashing this capacity:

- Existing CMTS have very limited scale – 8-16 downstream (DS) channels per port. Even the newer CCAPs only scale up to 32 DS channels per port. Scaling a service group (SG) to 128 or 158 DS channels will therefore require a lot of equipment in the headend. Perhaps in a couple of generations, CCAP systems will come to fulfill the promise of a single port per SG supporting full-spectrum services.
- Most networks only support up to 860 MHz of spectrum, and the bulk of this spectrum is allocated for legacy video services (analog video and digital QAM video – broadcast and narrowcast). Upgrading the network to 1 GHz or higher is a costly undertaking.
- Spectrum allocation changes require manual processes. RF combining and splitting networks need to be adjusted. Power levels need to be re-balanced in the headend and in the fiber node. This acts as a natural rate-limiter for capacity increases in the network.

SCALING CHALLENGES

Spectrum conversion to DOCSIS in conjunction with node-splits has been the modus operandi for operators for years now. But continuing down this path is unsustainable without a major change in the architecture. With each node split, the strain on headend resources (power, cooling, space) is being exacerbated.

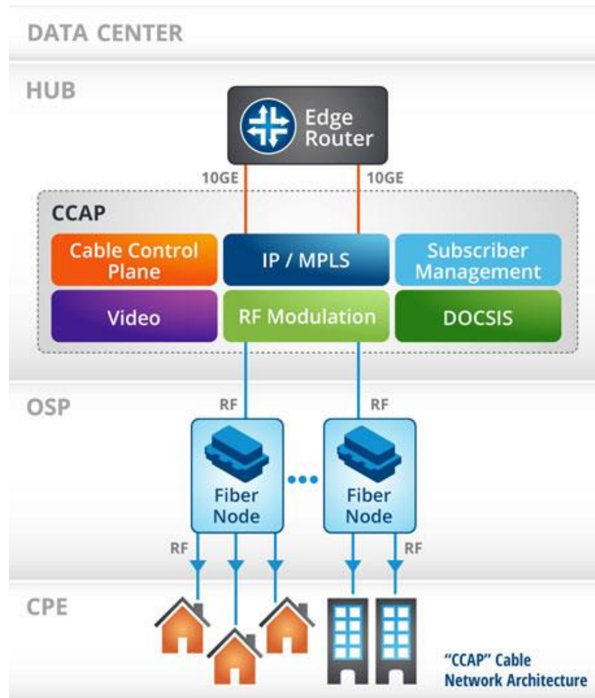


Figure 2: Network Architecture with CCAP

Let's take a very simple example to illustrate the problem operators face in managing facilities. In this example, an operator is going to add 1,288 new SGs in a given headend over a 7-year period. Figure 3 below shows the cumulative space required over this 7-year period.

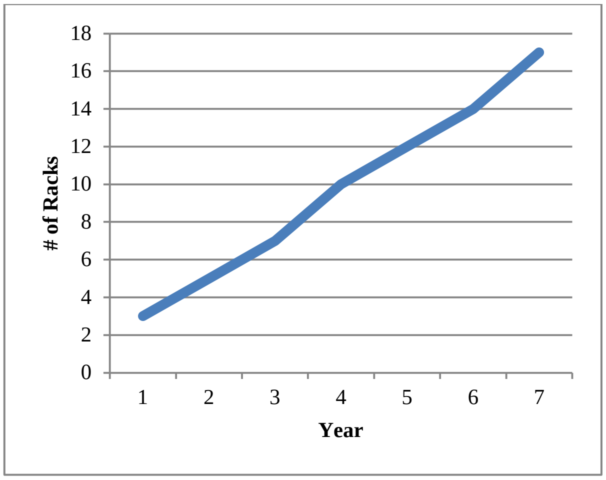


Figure 3: Cumulative Rack Space for Centralized CCAP Deployment

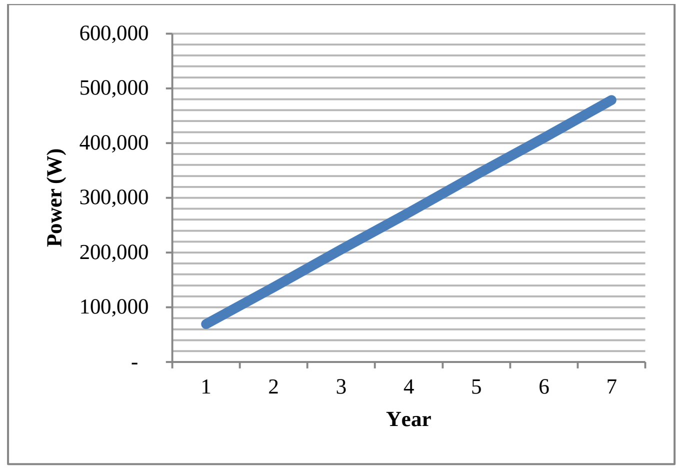


Figure 4: Cumulative Power Consumption for Centralized CCAP Deployment

A total of 17 racks are fully utilized after 7 years to accommodate this growth, and this is a very optimistic scenario as it makes the following assumptions:

1. The CCAP system supports 64 SG per device.
2. The CCAP supports full-spectrum services per DS RF port so that there is no need for RF combining and splitting. This requires the CCAP to support:
   a. 158 DS channels per port
   b. Flexible allocation of services to DS channels
3. The CCAP supports narrowcast and broadcast video encryption – the CCAP ECMG system is in place to enable support for Motorola Digicipher II and Cisco PowerKEY encryption schemes.
4. Operators collapse *all* services onto the CCAP device and retire existing QAM and CMTS platforms.

Achieving all of these requirements really isn't possible today, so the actual space needed will be a lot greater. Figure 4 shows the cumulative power consumption for this example, which again is optimistic given the constraints listed above.

Table 1 below has the raw values for the charts in Figure 3 and 4.

	YEAR						
	1	2	3	4	5	6	7
# of Racks	3	5	7	10	12	14	17
Power (kW)	69	136	205	272	342	409	478

Table 1: Raw Numbers for CCAP Power and Space Requirements

In order to address this challenge, the industry needs to develop and embrace a new architecture which minimizes power, cooling and space requirements in the headend. The architecture must leverage the power of software to provide operators with a flexible mechanism to reallocate spectrum. Furthermore the architecture must not force operators to rip and replace the legacy video EQAM infrastructure and integrate a new solution into the video back-office. It must also eliminate analog transport and use standards-based Ethernet to lower transport costs and improve fiber utilization. Finally, the architecture must lower costs (CAPEX and OPEX) and increase service agility and performance. The industry needs **Virtual CCAP**.

VIRTUAL CCAP

So what is a Virtual CCAP? At its core, it is a CCAP. It delivers all of the features and functions of a CCAP – all existing services are supported without modification, all existing customer premise equipment (CPE) works without modification, and existing back-office systems remain in place and unmodified. However, Virtual CCAP takes CCAP to the next logical step by distributing functionality and virtualizing control and management.

Rather than viewing CCAP as a physical platform, it can be viewed as a collection of base functions that support the services cable operators offer. These base functions include:

1. Cable Control Plane
2. IP/MPLS Control and Forwarding Plane
3. Subscriber Management
4. Video (QAM) Processing
5. DOCSIS Processing
6. RF Modulation

Software Defined Networking (SDN) technologies enable us to move away from the notion that all of these functions must be co-located in a monolithic system that is centralized in the headend.

Taking these base functions and bringing them together in a virtualized environment enables operators to leverage best-of-breed products for each function. The resulting eco-system will provide operators with the most feature-rich and reliable, yet lowest cost, solution.

Placement of these base functions can now be rearranged, moving each to the device and location in the network that optimizes the overall deployment.

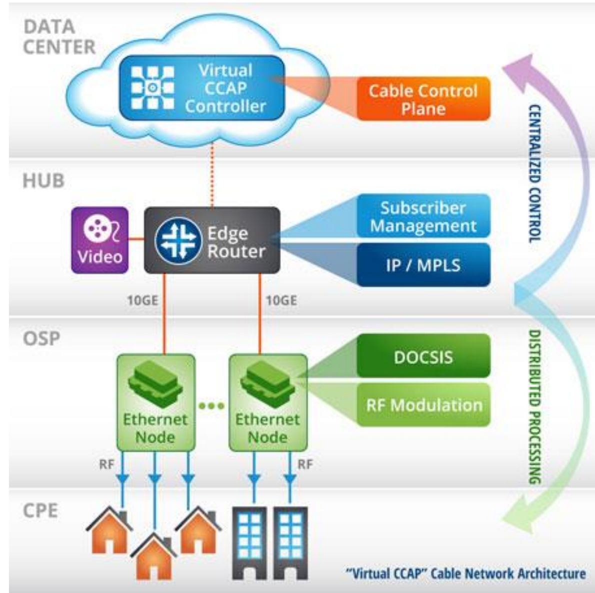


Figure 5: Network Architecture with Virtual CCAP

Cable Control Plane functions are collected into a single package and incorporated into the Virtual CCAP Controller. The Virtual CCAP Controller can be placed in any data center in the network, as it is a control and management entity only.

IP/MPLS and **Subscriber Management and processing** are collapsed onto the Edge Router in the hub. The Edge Routers commonly deployed in operator networks have all the native capabilities to support this; they just lack any cable awareness. This is addressed by the Virtual CCAP Controller.

Video (QAM) Processing stays in the hub, but migrates to Ethernet to take advantage of digital transport to the Ethernet Node. This enables operators to maintain the infrastructure that is already deployed, maximize the usable life of the equipment and avoid the costly integration of a new platform into their video back-office systems.

DOCSIS Processing (MAC and PHY) as well as **RF Modulation** is moved into the node, creating a new class of nodes – Ethernet Nodes. This eliminates analog optical transport and moves everything to digital, Ethernet transport. This improves RF performance, reducing customer calls and truck rolls and setting up the network for deployment of DOCSIS 3.1. The data-carrying capacity of the fiber is increased and operators can now converge their HFC transport systems with their Carrier Ethernet transport systems.

Looking at the same deployment scenario described earlier, it is clear that the Virtual CCAP solution dramatically changes the space and power equation.

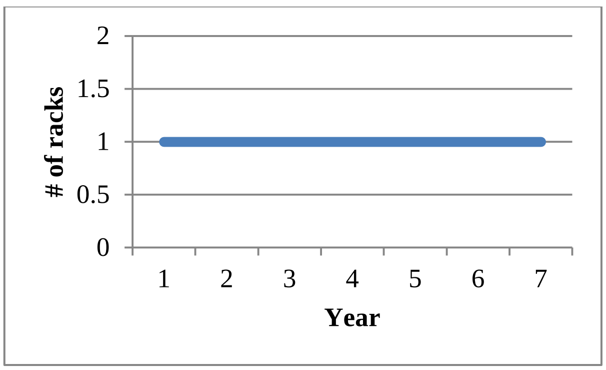


Figure 6: Cumulative Rack Space for Virtual CCAP Deployment

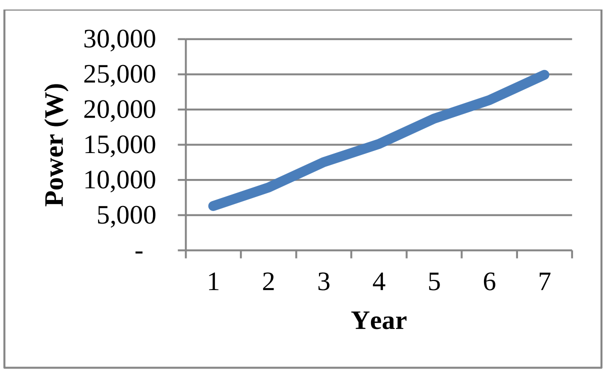


Figure 7: Cumulative Power Consumption for Virtual CCAP Deployment

	YEAR						
	1	2	3	4	5	6	7
# of Racks	1	1	1	1	1	1	1
Power (kW)	6.3	8.9	12.5	15.1	18.7	21.3	24.9

Table 2: Raw Numbers for Virtual CCAP Power and Space Requirements

Comparing the Virtual CCAP deployment against the Centralized CCAP deployment, operators can realize the savings described in Table 3 below.

	YEAR						
	1	2	3	4	5	6	7
Cumulative Power Saving (thousand's $)	82.9	167.6	254.1	338.7	425.2	509.9	596.4
Space Savings (Racks)	2	4	6	9	11	13	16

Table 3: Power and Space Savings of Virtual CCAP

This dramatic reduction in power and space enables operators to consolidate headends / hubs. Operators can eliminate smaller facilities and centralize the equipment in larger headends. This not only saves hundreds of thousands of dollars of costs related to ongoing site maintenance, it eliminates the need to spend millions of dollars to augment smaller facilities to accommodate the growing power and space demands of CMTS/CCAP devices.

DOCSIS 3.1

A Virtual CCAP clearly provides enormous financial and technological benefits compared to a traditional CCAP for DOCSIS 3.0. The difference becomes ever more pronounced with DOCSIS 3.1, especially with respect to the ease of transitioning.

DOCSIS 3.1 is designed to enable 10 Gbps downstream (DS) and 1 Gbps upstream (US) transmissions on the existing coax plant. There are a host of challenges in achieving this:

1) Improving end-to-end signal-to-noise (SNR) in order to achieve 4096 QAM modulation in the downstream and 1024 QAM modulations in the upstream
2) Increasing the top-end spectrum in the DS to 1.2 GHz
3) Changing the upstream split from 42 / 65 MHz to 200 MHz
4) Reallocating spectrum from existing services (analog and digital video, DOCSIS 3.0) to use for D3.1

The single largest impediment to achieving the necessary SNR for ubiquitous support of 4096 QAM in the DS and 1024 QAM in the US is the analog optical transmission system. The specific values vary based on a host of factors (age of system, top-end frequency supported, number of wavelengths carried on a single fiber, length of fiber, etc.), but replacing the analog optical transmission system with Ethernet provides a 3-7 dB improvement in end-to-end SNR. This is crucial to achieving the higher order modulation rates to make the cost of migrating to D3.1 worthwhile.

A traditional monolithic CCAP doesn't address the issues with analog optical transmission. It actually requires the use of the existing analog optical transport systems, whereas a Virtual CCAP transitions the optical transport to Ethernet.

With a traditional monolithic CCAP deployment, there are multiple changes that operators must perform in order to deploy D3.1 with the new frequency splits in the US and DS.

1) Replace existing splitter-combiner networks
2) Replace the analog optical transport systems
3) Replace the fiber nodes
4) Replace amplifiers
5) Replace linecards in CCAP to D3.1 capable versions

By contrast, in a Virtual CCAP deployment, all of the headend work is eliminated. A simple diplexer adjustment in the Ethernet Node changes the US split. Amplifiers will still need to be replaced to support the US and DS frequency range changes. The rest is just a system configuration change to reallocate spectrum to D3.1. The Ethernet transport infrastructure does not need to be touched at all; it already has the capacity required to support D3.1.

Alternatively, operators can also choose to deploy D3.1 within the existing US and DS frequency plan in order to take advantage of improved spectral efficiency while minimizing the outside plant changes.

In this scenario, for a traditional monolithic CCAP deployment, the following changes would need to be done:

1) Replace linecards in CCAP with D3.1 capable versions
2) Manually adjust RF power levels at the analog optical transport system in the headend and the fiber node in the outside plant to accommodate spectrum allocation changes.

For a Virtual CCAP deployment, a simple configuration change re-allocates spectrum to D3.1 and adjusts the RF power balance. No hardware needs to be replaced, no need to visit the node to manually adjust power levels.

CONCLUSION

As operators evolve their network infrastructure to address ever-growing and changing consumer demand, they need a new network paradigm to satisfy bandwidth requirements. Existing products limit the network architecture, increasing costs, limiting flexibility and impeding service deployment.

Technology advances in SDN and NFV have opened up new network architecture options. Virtual CCAP is a prime example of that. A Virtual CCAP deployment enables operators to support all of their existing services while lowering overall CAPEX and OPEX. It addresses the facilities challenges faced by operators, streamlines capacity additions to the network and sets the network up for a simple migration to DOCSIS 3.1. This is done in a way that seamlessly integrates with existing back-office systems and processes, and works with all current and proposed CPE.

It provides the foundation for operators to evolve their business and address the growing threat from FTTH and OTT competitors.

EPoC - THE UPCOMING IEEE P802.3bn STANDARD FOR EPON PROTOCOL OVER COAX OVERVIEW AND IMPACT

Mark Laubach
Broadcom® Corporation

Abstract

Following the interest from both the Chinese and North American cable industries, the IEEE Standards Association approved a project under the IEEE 802.3 Standard For Ethernet Working Group with the scope of extending the EPON protocol over coaxial (EPoC) networks. The P802.3bn EPoC Task Force is in progress, developing a physical layer standard for symmetric and asymmetric operation of up to 10 Gb/s on point-to-multipoint coaxial distribution plants comprising either active or passive media. EPoC objectives, key technical and performance considerations, and resulting architecture are presented. Select comparisons of EPoC with the DOCSIS 3.1 PHY including common component architectures are noted. The resulting PHY standard will enable several deployment models, increasing the number of options available to cable operators as part of their future service growth.

INTRODUCTION

Ethernet Protocol over Passive Networks (EPON) was first standardized in 2004 as part of the IEEE 802.3 Working Group's [1] Ethernet in the First Mile (EFM) initiative. Both 1 Gbit/s and 10 Gbit/s EPON standards are incorporated in the 802.3-2012 - IEEE Standard for Ethernet [2]. The P802.3bn Task Force [3] extends the 10Gbps EPON point-to-multipoint solution over cable operator coax networks via a new physical layer. Initial interest for EPoC developed in China and followed shortly by North America. The motivation in both markets is different. In China, Hybrid Fiber-Coaxial (HFC) solutions are deployed extensively, with a typical architecture of placing the fiber node on the MDU campus. *CableLabs®* DOCSIS® [4] while becoming increasingly more popular, is not as widely deployed as in other parts of the world. Cable operators use government funded optical fiber in the streets and are looking at the opportunity of transparently extending existing EPON services over legacy coaxial cabling present in a large majority of older and new multi-tenant/dwelling units (MxU). In North America, High Speed Data (HSD) residential services are provided using DOCSIS technology. Some cable operators are increasingly deploying EPON to primarily capture business market services and cellular backhaul. Metro Ethernet Forum (MEF) [5]. service performance and contracted Service Level Agreements (SLAs) are pro forma.

Both Chinese nd North American operators appear to have a similar desire for the simplicity of Ethernet at gigabit speeds and for extending the life of existing coax cable networks. Key requirements for EPoC will be maintaining unified management and Quality of Service (QoS). DOCSIS Provisioning of EPON (DPoE™ [6]) is straightforward to extend managed EPoC systems. IEEE P1904.1 Service Interoperability in EPON (SIEPON) is also available for seamless management [7].

Additionally, EPON and EPoC directly support the step-wise migration from back-end legacy MPEG-2 transport to MPEG-4 video distribution via IP over Ethernet. In the future, a large Ethernet-based gigabit pipe to the home and business will be fundamental for cost-effective growth and evolution. When available in the future, a cable operator will then have a choice to deploy IP over

DOCSIS 3.1 or over EPON/EPoC based on their business needs.

Service and Application Architecture

Figure 1 shows an example of the target applications for a mix of EPON and EPoC technologies that leverages fiber-deep access architectures of current networks and the extension to existing coaxial distribution infrastructure.

For EPON, the IEEE Ethernet 802.3-2012 standard [2] defines the MAC and PHY sublayers for a service provider OLT and a subscriber ONU. Two wavelengths are used on the fiber for full-duplex operation, one for continuous downstream channel operation and another for upstream burst mode operation. The OLT MAC controls time-division sharing of the upstream channel for all ONUs via the Multipoint MAC Control Protocol (MPCP).

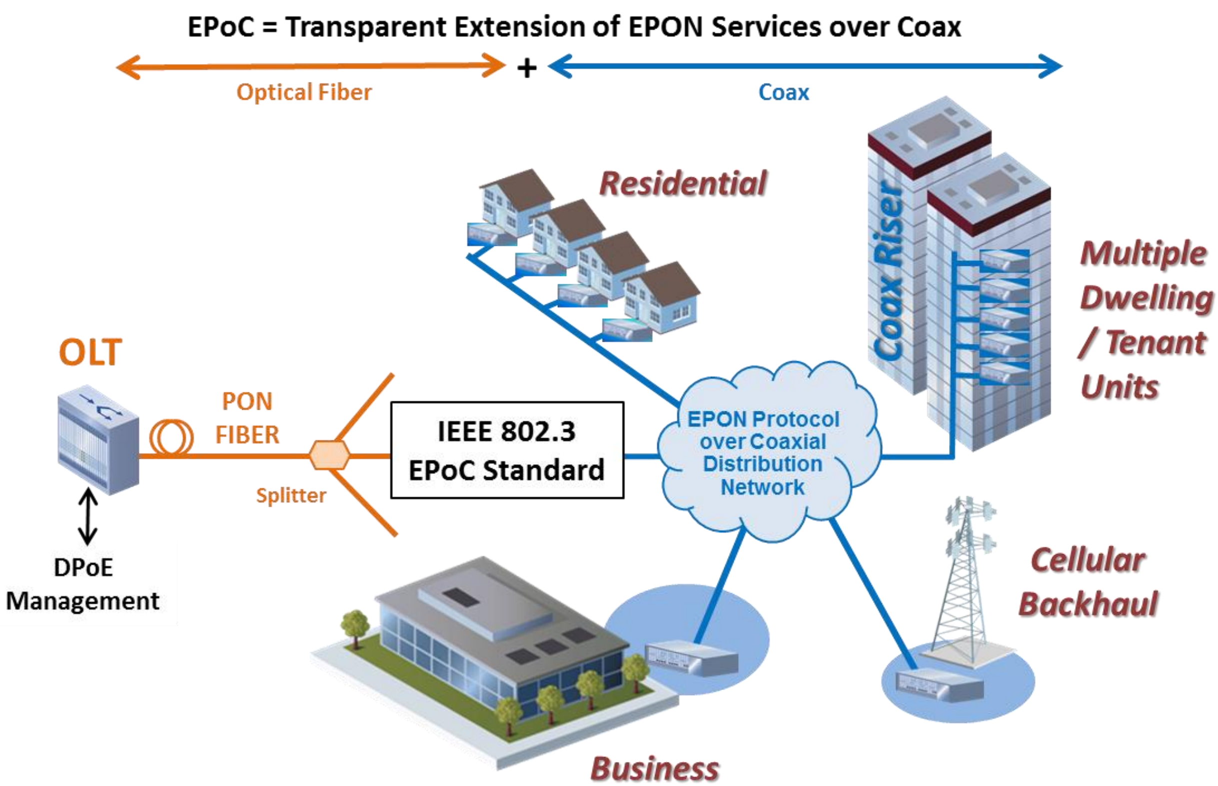


Figure 1. EPoC Applications for Extending EPON Services over Coax

Figure 1 illustrates an existing EPON Optical Line Terminal (OLT) connection through a Passive Optical Network (PON) made up of optical fiber and passive splitters. EPoC services are expected to extend the same range of services as that of EPON: for example, Residential, Business, Multiple Dwelling, and Cellular Backhaul. DPoE Management will be extended to EPoC.

EPoC Architecture

Similarly, the EPoC architecture consists of a service provider Coax Line Terminal (CLT) and a subscriber Coax Network Unit (CNU). The EPoC CLT and CNU MAC sublayers will be substantially similar to respective layers found in the OLT and ONU. Minimal augmentation to MPCP will be necessary to account for new packet burst overheads. Similar to the DOCSIS system, the EPoC PHY employs a full-duplex point-to-multipoint architecture, and the downstream and upstream communication channels will utilize RF spectrum as assigned by a cable operator for their coax network.

EPoC Topology

Figure 2 illustrates the three topology models for EPoC. Each model follows a *Node + N* approach, where N represents the greatest number of amplifiers on a single path between the Coax Line Terminal (CLT) and a Coax Network Unit (CNU). N varies from 0 (e.g., *Node + 0*), and represents a passive network or segment, to a preferred depth of 3. Higher N values are also considered. A traditional HFC architecture is included in the topology.

Similar to DOCSIS 3.1 PHY OFDM, EPoC will require the allocation of dedicated RF spectrum in both downstream and upstream directions. Provisioning of in-band service requires co-existing with other cable operator services in either direction (e.g., video, data, voice) without mutual interference As an operational note on differences in upstream use, EPoC is the sole user of its assigned spectrum, where DOCSIS 3.1 OFDMA may share the same upstream RF channel with legacy DOCSIS via TDMA scheduling of different burst types.

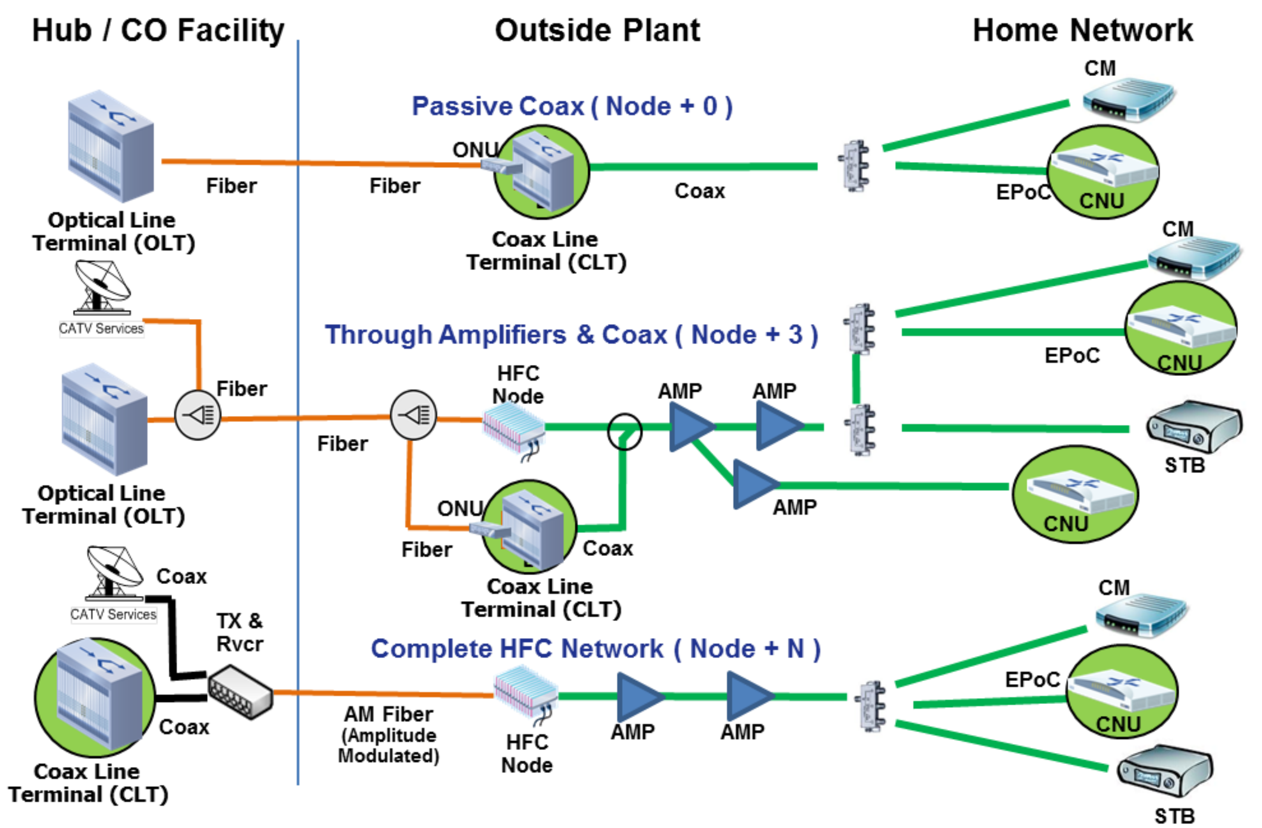


Figure 2. EPoC Topologies

In a practical implementation, the conversion from EPON digital fiber to EPoC coax will be implemented as a single unit device that is placed in the field adjacent to the fiber node. For example, the conversion device in Figure 2 is shown by an ONU coupled with a CLT.

System Models

Two system models have been under discussion in the EPoC Task Force, as shown in Figure 3. The first is a CLT with one more CNUs directly interconnected by a coaxial distribution network. This system model follows the scope of the standard. Product realization would be a CLT that is colocated in a headend as illustrated in the *Complete HFC Network* topology shown in Figure 2.

The second and more preferable system model is enabled by the EPoC standard, but is outside the scope of the IEEE 802.3 Working Group. This model is a traditional EPON with an OLT and multiple ONU devices together with one or more Fiber Coax Unit (FCU) converters that interconnect between PON with a coax distribution network using EPoC. This second model permits CNUs to appear as ONUs to DPoE, and is, therefore, the manageable equivalent to EPON ONUs.

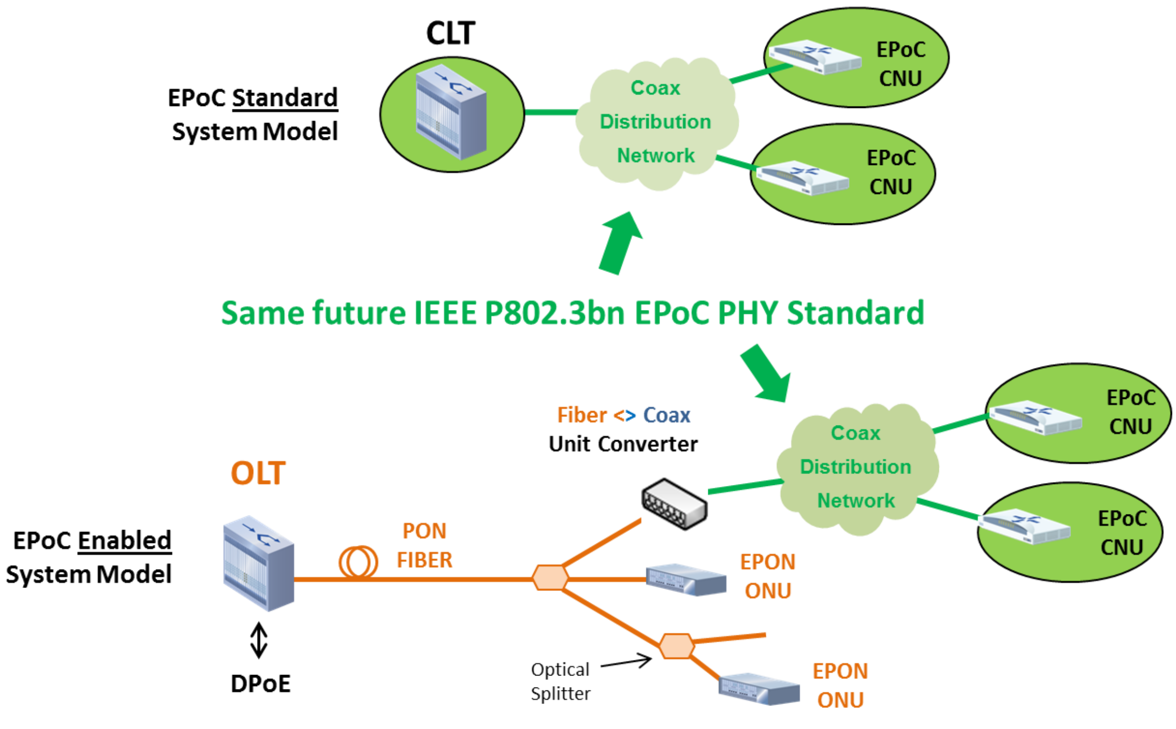


Figure 3. Two EPoC System Models

The FCU has two implementation architectures that can be considered. The first is a repeater FCU, where the OLT MPCP is transparently extended over the coax network. Here, each CNU's MPCP is part of the single MAC domain that includes both CNUs and ONUs. The OLT provides scheduling and MAC control over both the optical media and the coax distribution network. The repeater FCU converts between EPON PHY frames and EPoC PHY frames as the PCS encoding, i.e., the FEC and line encoding, are different. The EPON preamble and LLID are retained, and therefore the same over both media.

The second architecture is a bridge FCU. In this case, the PON port on the FCU and each ONU is a member of the OLT's MPCP domain on the PON with an LLID assigned by the OLT. On the coaxial network port, each CNU is a member of the FCU's MPCP domain with an LLID assigned by the FCU. There would likely be no requirement for LLID values to be retained across the bridge. Rather, the LLID would become part of the internal frame forwarding mechanism. QoS flows may be individually retained; however, it may make sense to aggregate flows in the upstream to support OLT scheduling of traffic from the FCU to the OLT.

The architecture and specifications of both the repeater FCU and bridge FCU are beyond the scope of this paper. It is expected that details would be forthcoming from the cable industry as the IEEE P802.3bn standard work progresses towards finalization.

DETAILS OF IEEE P802.3bn EPoC

The work of the IEEE P802.3bn EPoC PHY Task force is in progress at this time, and the progress of the Task Force can be viewed on the project website located at [3].

The IEEE Standards Association provides an open public process that consists of consensus building and a technical selection process. In IEEE 802 and 802.3, participation is by an individual, and technical consensus is indicated by a 75% or greater vote of Yes versus No on decisions at face-to-face meetings and on ballots. The product of the P802.3bn activity is the Task Force Draft. When complete, this draft is reviewed and subsequently approved through one or more cycles of the 802.3 Working Group Ballot process. Upon approval, this draft is submitted to IEEE 802 as a Sponsor Ballot, and the review ballot cycle is completed again using a larger ballot group. Upon passing the Sponsor Ballet, the result is then reviewed by the IEEE SA New Standards committee. When approved, the work will become the published standard. In terms of prestandard availability, an approved Working Group draft has been shown to be generally stable. However, as with any formal process, a nonapproved draft at any stage is returned to the Task Force for more work, and the process starts again.

P802.3bn maintains a timeline on its website at [3]. The last update from November 2013 anticipates an approved Working Group draft by the end of 2014 with a published standard in Mid 2015

IEEE P802.3bn Project Documents

Each authorized project under the 802.3 Ethernet Working Group has three guiding project documents. The first is the IEEE-SA Project Authorization Request (PAR) which sets the scope, the project area (e.g., amendment to the 802.3 Ethernet standard), and other details. The PAR is the agreement between the IEEE SA, the 802 LMSC, and the 802.3 Working Group. Upon issuance of a PAR, the 802.3 Working Group charters a Task Force. The second document is Criteria for Standards Development (CSD) also known as the *5 Criteria* or *5Cs*. The CSD/5Cs is a requirement the 802 LMSC and the 802.3 Working Group. The third document is the set of Objectives that constitute an agreement between the 802.3 Working Group and the Task Force. Access to past and current project documents for 802.3 is available at [1] and specifically for P802.3bn EPoC at [3].

IEEE P802.3bn Objectives

Developed as an output of the Study Group phase of the EPoC project, the objectives focus the effort within the scope of the PAR. Objectives are developed by the consensus of the participants and then considered and approved by the 802.3 Working Group. Objectives may change over the lifetime of the project as the Task Force develops its understanding and, approach, and after input from the 802.3 Working Group or its members.

The following is the list of current P802.3bn objectives developed by the participants consisting of individuals representing cable operators, equipment and component manufacturers, and others.

1) Specify a PHY to support subscriber access networks capable of supporting burst mode and continuous mode operation using the EPON protocol and operating on point-to-multipoint RF distribution plants made up of either amplified or passive coaxial media.

2) Maintain compatibility with 1G-EPON and 10G-EPON as currently defined in IEEE Std. 802.3 with minimal augmentation to MPCP and/or OAM if needed to support the new PHY.
3) Define required plant configurations and conditions within an overall coaxial network operating model.
4) Provide a physical layer specification that is capable of:
    a. A baseline data rate of 1Gbit/s at the MAC/PLS service interface when transmitting in 120 MHz, or less, of assigned spectrum under defined baseline plant conditions.
    b. A data rate lower than the baseline data rate when transmitting in less than 120 MHz of assigned spectrum or under poorer than defined plant conditions.
    c. A data rate higher than the 1 Gbit/s baseline data rate and up to 10 Gbit/s when transmitting in assigned spectrum and in channel conditions that permit.
5) PHY to support symmetric and asymmetric data rate operation.
6) PHY to support symmetric and asymmetric spectrum assignment for bidirectional transmission.
7) PHY to support independent configuration of upstream and downstream transmission operating parameters.
8) PHY to operate in the cable spectrum assigned for its operation without causing harmful interference to any signals or services carried in the remainder of the cable spectrum.
9) PHY to have:
    a. A downstream frame error ratio better than 10^{-6} at the MAC/PLS service interface.
    b. An upstream frame error ratio better than $5x10^{-5}$ at the MAC/PLS service interface.
10) Define Energy Efficient Ethernet operation for EPON Protocol over Coax PHYs.
11) Mean Time To False Packet Acceptance (MTTFPA) at least equal to 1.4×10^{10} years.

EARLY TASK FORCE DECISIONS

EPoC first met as a Task Force in September 2012 in Geneva, Switzerland. The first technical decisions were: the use of Orthogonal Frequency Division Multiplexing (OFDM) modulation in the downstream channel, OFDM Access (OFDMA[1]) in the upstream channel, the use of Low Density Parity Check Coding (LDPC), up to 4096 QAM in the downstream, up to 1024 QAM in the upstream, an OFDM channel size of 192 MHz with a sample clock of 204.8 MHz, and the ability to combine multiple 192 MHz OFDM/A channels for higher capacity. Later in its progress, the Task Force narrowed for use of a single 8K Fast Fourier Transform (FFT) with a subcarrier spacing of 50 KHz. For the 192 MHz channel, 3800 subcarriers are usable.

Availability of RF Spectrum

The trend in cable will be to convert existing non-IP services running over legacy modulation to IP services running over next-generation modulation: OFDM/A + LDPC. Neglecting overheads, to achieve a 10 Gbit/s data rate using a 12 bit/s/Hz (4096 QAM) modulation rate will require over 830 MHz of spectrum. Transitioning towards this future will be done in steps defined by each cable operator based on their business plans. With protocol overheads, P802.3bn targets a downstream data rate of 1.6 Gbit/s at the MAC/PLS service interface per 192-MHz OFDM channel in baseline channel conditions

[1] OFDM/A will be used hereafter unless there is a need to differentiate downstream OFDM from upstream OFDMA.

(Refer to Technical Decision #40. See [3].). The minimum downstream RF spectrum allocation for P802.3bn is 24 MHz (the same for DOCSIS 3.1 OFDM). Note that the *baseline channel conditions* are based on an approved channel model [3].

The upstream RF spectrum may use less than 24 MHz initially, and the configuration is intended to be flexible with current split configurations used nationally and internationally (e.g., 5-42 MHz, 5-85 MHz, etc.) and adapt as the cable industry evaluates and implements possible shifts in the split (e.g., 5-120 MHz, 5-234 MHz).

P802.3bn internally will be configured to support frequency and configuration values for up to 5 GHz of RF spectrum. The standard will initially support the frequency ranges as shown in Table 1.

Table 1. EPoC Channel Frequencies

	Frequency Range
Downstream	54 – 1212 MHz
Upstream	5 – 234 MHz

The overlap of the downstream and upstream, diplexer requirement, and the lowest frequency used will be determined by cable operator requirements. P802.3bn will not specify diplexer requirements. The downstream top end to 1212 MHz appears at this time to be a reasonable limit for expansion of existing HFC architecture downstream passband. In the future, upstream and downstream and active and passive topologies beyond 1212 MHz can be accommodated by P802.3bn.

Note that in order to support more than 1.6 Gbit/s, P802.3bn will need to support additional channels. See Section *Up to 10 Gbit/s*.

Challenges of RF Spectrum

The selection of OFDM/A for future modulation has the advantage of being more robust as well as adaptable for creating wide channels that may be flexibly provisioned. Subcarriers may be turned off, meaning no energy is transmitted in that subcarrier. Turned off subcarriers are termed *excluded*. Excluded subcarriers are used for sizing the RF spectrum of the channel, creating notches to work around legacy channels, as well as for well-known interference, such as LTE ingress. A group of adjacent excluded subcarriers is called an *exclusion band* or an *edge band*, depending on location. The minimum size exclusion band is 1 MHz. Allocation of used spectrum versus excluded spectrum can be changed at any time by cable operator provisioning as needed.

Another common challenge for many modulations is resilience to micro-reflections and Inter-Symbol Interference (ISI). An extended OFDM/A modulation symbol includes a cyclic prefix, which essentially is duplicate overhead of an end of the OFDM symbol added to form the transmitted signal. CP size is set during provisioning based on local system conditions and has ranges from 5% to 25% time overhead. There is an additional parameter called windowing that is used to shape the sharpening of the subcarrier edges. There is one CP size and one window size for an OFDM/A channel.

The provisioning of a downstream OFDM/A channel includes the setting of: channel frequency, excluded subcarriers, modulation rate per subcarrier, CP, and window size. In addition, a portion of the non-excluded subcarriers will be allocated for the PHY Link with the remaining non-excluded subcarriers assigned for data. Continuous pilots, as required for OFDM receiver synchronization, will likely be allocated as part of an algorithm based on available subcarriers.

The Need for a PHY Link Channel

OFDM/A is well known in wireless networks and other media, but this is the first time Ethernet[2] has been used for EPoC. In contrast to other Ethernet physical layer (PHY) modulations, OFDM/A has a much higher degree of configuration flexibility and complexity. The permutations on a downstream channel hunt procedure would be significantly higher than the way it is done now, for example, with traditional DOCSIS QAM channels that are located on well-known frequencies.

The EPoC Task Force specifies a PHY Link channel made up of a contiguous subset of the subcarriers available in an OFDM/A channel. A 400 KHz data subchannel is configured and made up of eight (8) adjacent subcarriers. The PHY Link may be placed anywhere within the OFDM/A channel. The modulation rate is 16 QAM. Downstream data is framed on a repeating cycle of 128 symbols, where each cycle begins with a well-known preamble value[3]. Other subcarriers adjacent to the eight main subcarriers are used to convey OFDM/A pilot information. The combination of the above permits a CNU to quickly locate, synchronize with, determine the CP size, and then decode downstream PHY Link messages.In addition, there is a requirement to place the center frequency in increments of 1 MHz permitting a cable operator to provision placement aligned with well-known frequency locations, such as a local channel plan.

The PHY Link cycle of 128 symbols also defines the OFDM frame cycle on the downstream data subchannel. Included in each PHY Link message is a Next Codeword Pointer (NCP) that indicates the starting bit of the first codeword in the next cycle of the data subchannel. The NCP is used for helping to quickly establish frame synchronization after a downstream channel loss event (e.g., burst noise) as well as permitting FEC codewords to straddle cycles; i.e., an FEC codeword is not required to be aligned to the start of the cycle.

CNU Initialization Overview

After locating and decoding the downstream PHY Link messages, a CNU learns of the both the downstream OFDM data subchannel configuration and the upstream OFDMA subchannel configuration, including the upstream PHY Link configuration. At this point, the CNU can participate via the PHY Link with the CLT for the link auto-negotiation process before the CNU is allowed to go *linked,* connecting the CLT and CNU MAC entities. Note that the PHY Link subchannel communications are separate from the main data subchannel communications; i.e., the PHY Link is transparent to the EPON / EPoC MAC.

PHY Link management messages include: PHY Discovery, fine ranging, PHY Link messages, and wideband probes. The first message sent by a CNU is a PHY Discovery in response to a broadcast invitation sent by the CLT. This first response message is used to determine initial range timing as well as CNU MAC address identification. The CLT then assigns a unique PHY identifier and transitions to unicast messaging to direct the new CNU into a fine ranging mode, responding to probes as well as responding to PHY Link messages. Fine ranging permits adjustment of CNU power, symbol timing, and OFDMA frame timing. Wideband probes permit adjustment of pre-equalization coefficients to better match the CNUs signal to the channel response characteristics. When complete, the CLT will transition the CNU to a *linked* status used by Ethernet to signal the link and new station

[2] In parallel, OFDM/A were selected earlier in 2012 by the CableLabs DOCSIS 3.1 project.
[3] The preamble value is distinguished from that used by the DOCSIS 3.1 OFDM downstream PHY Link.

availability. At this point, the data channels of downstream and upstream are enabled, permitting EPON MAC operation.

Link details such as downstream and upstream PHY rates and CNU capabilities are communicated via station management.

Ongoing PHY Link Operation

After initial station discovery, the PHY Link continues to monitor the PHY operational details of each CNU. Periodic adjustments for power, timing, pre-equalize coefficients, etc. should be a normal and ongoing part of operation.

A given channel configuration is known as a profile. The PHY Link channel will be used to communicate updated channel profiles as well as coordinate switch-over time. It is expected that cable operators will make use of the flexible nature of OFDM to periodically adjust subcarrier configuration to adjust for changes in plant conditions, impairments, spectrum use, etc.

The Downstream Tx PHY Path

Figure 4 presents a high-level block diagram of the P802.3bn downstream CLT physical layer for the data path. The PHY Link processing path is omitted from this description. Following strict 802.3 layering, the MAC is separated from the PHY by the MAC/PLS service interface. This also serves as a reference point for performance measurements. The downstream channel implements point-to-multipoint access in continuous mode; the CLT transmitter is heard by all CNU receivers.

The Physical Layer is composed of three distance sublayers: the Physical Coding Sublayer (PCS), the Physical Media Attachment Sublayer (PMA), and the Physical Media Dependent sublayer (PMD). The PMD connects directly to the Coax Cable Distribution Plant via a Medium Dependent Interface (MDI) commonly known by the cable industry as the standard F Connector. P802.3bn has allocated its functions as described in Figure 4.

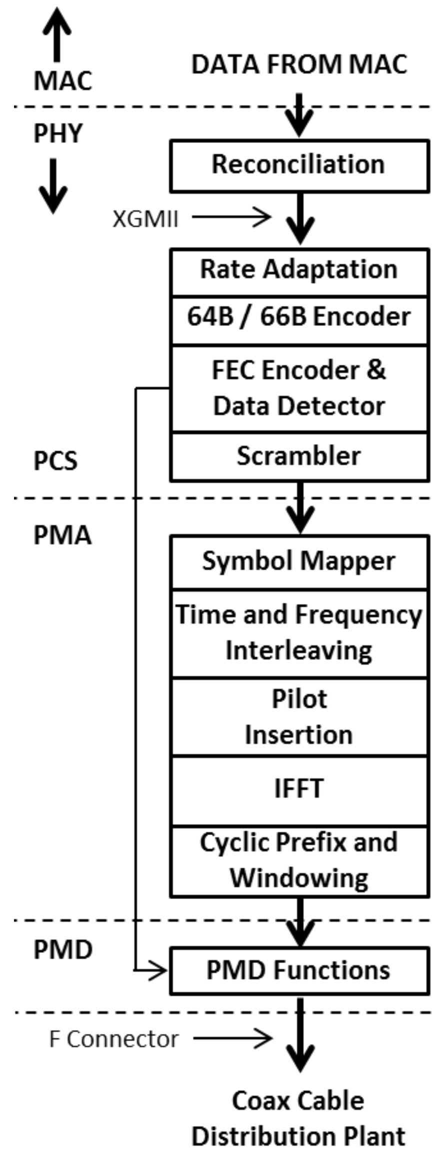


Figure 4. Downstream CLT Tx PHY Block Diagram

User data from the MAC layer is passed as standard Ethernet frames with a traditional length of between 64 and 1500 bytes including the Frame Check Sum (FCS).EPoC will accommodate newer lengths of up to 2000 bytes. Each Ethernet frame includes a 32-bit CRC.

The reconciliation function is the same used by the EPON standard. Its purpose is to prepend an 8 byte EPON preamble containing the assigned Logical Link Identifier (LLID), other information, and a separate CRC. The LLID is required to properly support the virtual MAC entities used by the OLT (CLT) MPCP and each ONU (CNU) MAC entity.

The XGMII is the 10 Gbit/s Media Independent Interface (MII) used by the 10G-EPON standard. It establishes a 10 Gbit/s transfer rate from the MAC to the PHY. In order to accommodate FEC overheads and the required InterFrame Gap (IFG), the EPON MAC layer inserts idles between MAC packets to match the overheads and establish timing. These idles are transferred across the XGMII as part of the constant rate. EPoC will make a minimum augmentation to the EPON MPCP to adjust the idle insertion for both the different EPoC PHY and FEC overheads as well as to effectively lower the overall downstream rate to match the rate of the configured OFDM channel. Cable operators will first deploy EPoC with a minimum of 24 MHz or more, dependent on their cable plant configuration. As the service mix changes and more RF spectrum is allocated to EPoC services, the higher channel rate will be matched by changing the amount of idle insertion by the MAC. The Rate Adaptation removes any idles as required for matching the actual configured line rate.

The 64B/66B encoder is a modified version of the 10G-EPON, where the output is compressed to 65 bits per block rather than 66 bits per block. The line encoder provides a means to multiplex control and data information as well as establish a standard block framing for use by the receiver decoder.

The output of the line encoder is then processed by the downstream FEC encoder. P802.3bn Task Force has selected an LDPC rate 8/9 (14400, 16200) as the single long-sized code for the downstream. The FEC encoder places 220 whole 65-bit line encoded blocks in the information word accompanied by a 40-bit CRC. This CRC is used to meet the 802.3 Working Group Mean Time To False Packet Acceptance (MTTFPA) criteria. There are 36 bits of unused information word that will be removed via shortening.

The output of the FEC encoder is processed by a conventional streaming bit Scrambler. The scrambler is initialized at the beginning of each PHY Link cycle.

The resulting output crosses from the PCA to the PMA and is processed by the Symbol Mapper function. The operation of mapping PCS bits to OFDM symbols and subcarriers relies on configuration information that includes the following: excluded and non-excluded subcarrier status, continuous pilot subcarrier use, bit loading per subcarrier, QAM mapping per subcarrier, scattered pilot repeating placement per cycle, and PHY Link channel subcarriers. Not shown in Figure 4 is the parallel processing of the downstream PHY Link channel via the same (or similar) Symbol Mapper function. The output is a complex I and Q value per subcarrier (IFFT bin).

The Pilot Insertion function places both continuous pilot information and cycle-repeating scattered pilot information into the Symbol Mapper output before being processed by the Inverse Fast Fourier Transform (IFFT) process.

On conversion to digital OFDM symbols, the Cyclic Prefix (CP) and Windowing function adds the necessary CP overhead and symbol shaping. The resulting digital information is then processed by the PMD. Typically, most standards do not specify PMD functionality as specifics of Digital to Analog conversion, and analog processing is left to the vendor. Rather the power and spectral requirement of the electrical output at the F-Connector are very well defined.

The Upstream Tx PHY Data Path

The P802.3bn Task Force has many decisions pending in its upstream PHY path functions and architecture. The material here should be considered as a general overview of topic areas with the details subject to change as the Task Force completes its work.

Figure 5 presents a high-level block diagram of the P802.3bn upstream CNU physical layer for the data path. The PHY Link processing path is omitted from this description. The upstream channel implements multipoint-to-point access based on TDMA scheduling from the CLT MAC. Scheduling is based on time, and each CNU MAC is allocated time on the wire for its next burst. Upstream bursts in EPON frequently contain multiple concatenated MAC packets and EPoC will contain them as well.

To overcome inefficiencies of a single transmitter in an OFDM channel, EPoC uses OFDM/A and introduces a one dimensional (1D) to two dimensional (2D) mapping algorithm together with an upstream framing and timing structure that permits data as well as channel probing with PHY Link Discovery and Ranging. P802.3bn adds a process of employing three upstream FEC codewords and their combination that maximizes burst efficiency.

Overall, upstream PHY processing is very similar to downstream processing. The key differences are:

1) The FEC encoder is replaced with an FEC codeword builder that also includes MAC burst detection,

2) The symbol mapper, interleaver, and OFDM Framer together build a two dimensional framing structure composed of Resource Blocks, and

3) A configuration and timing function controls OFDM/A framing and wideband probes used for per-CNU channel response analysis.

FEC Codeword Builder

The P802.3bn Task Force selected three upstream codewords sizes and rates for the upstream channel: LDPC (14400, 16200), same as the downstream, LDPC (5940, 5040) rate 28/33, and LDPC (1120, 840) rate 3/4.

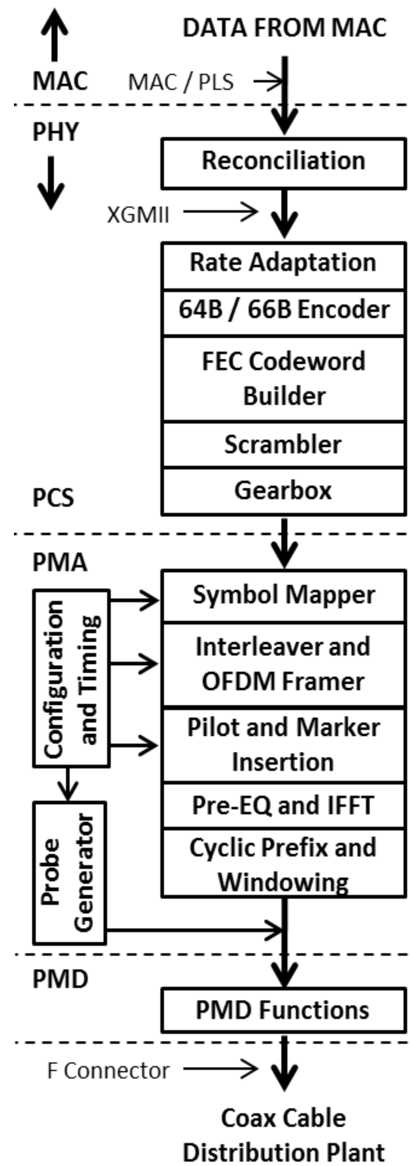


Figure 5. Upstream CLT Tx PHY Block Diagram

These are referred to as long, medium, and short respectively. All codewords support shortening. The Task Force selected a codeword filling approach where the ordering of codewords is concatenated for the upstream burst based on the actual MAC burst length. One of the challenges caused by the 802.3 layering model is that the length of the upstream MAC burst is not known by the PHY. The FEC codeword builder detects the start and stop of a burst and then determines how many and which FEC codewords will be needed for the burst. The basic algorithm is presented in [8] and uses a threshold of 6601 bits to determine the lower limit on long shortened codewords and 1601 bits for medium codewords:

- If there are enough bits to create a full long codeword, do so. Keep doing this until there are not enough bits left.
- If there are now enough bits to create a shortened long codeword (subject to the thresholds above), do so and end the burst.
- Otherwise, if there are enough bits to create a full medium codeword, do so. Keep doing this until there are not enough bits left.
- If there are now enough bits left to create a shortened medium codeword (subject to the thresholds above), do so and end the burst.
- Otherwise, if there are enough bits to create a full short codeword, do so. Keep doing this until there are not enough bits left.
- Use whatever bits remain to create a shortened short codeword and end the burst.

The DOCSIS 3.1 upstream PHY uses a similar upstream codeword building approach. However, EPoC includes a CRC-40 in the information payload and subsequently has different thresholds.

OFDM/A Framing and Resource Blocks

In EPON, each CNU is allocated upstream transmission time via the GATE message. The message specifies both *grantStart* and *stopTime* values[4]. For EPoC, this linear time-on-the-wire sequence of bits is converted to parallel using 1D-to-2D mapping onto Resource Blocks within the OFDM/A frame. A frame is the definition of a matrix of subcarriers by symbols where a repeating sequence is known by both the CLT and each CNU.

A Resource Block (RB) is a frequency and time grouping of Resource Elements (RE) defined by a dedicated set of (N) contiguous subcarriers and a consecutive number of (M) symbols defined by the OFDMA frame length. RBs are nonoverlapping. A CNU may be assigned to transmit in one or more contiguous Resource Blocks in an OFDMA frame. A Resource Element (RE) is a one-subcarrier by one-symbol element that is allocated within a resource block and used to convey a portion of the upstream signal; e.g., data, pilot, or burst marker information.

The P802.3bn Task Force is considering RBs where N may be 1, 4, or 8 subcarriers, and M may be 8, 12, or 16 symbols. M also represents the interleaver depth. For any given OFDM/A frame, the number of RBs and their allocation is static within a cycle and assigned via PHY Link management messages. All RBs in the same frame must have the same value of M. There is no requirement for all RBs in an OFDM/A frame to have the same value of N, simply that the CLT and CNU have the same RB configuration knowledge. As an example, three uses of an N=4, M=8 RB configuration that may be used during a CNU burst are shown in Figure 6. In this figure, the following conventions are used: **B** denotes a

[4] Refer to [2]: Section 77.3.5 Gate Processing.

begin burst marker, **E** an *end* burst maker, **P** a pilot, **C** a complentary pilot, and **D** is data.

As an example, a MAC burst is transferred from the PCS to the PMA. The Symbol Mapper is aware of the RB timing and starts to use the first RB by inserting a beginning of burst marker into the appropriate RE designated B. Pilot REs for P and C are identified for future insertion in all RBs. As the burst continues, D REs are used to convey data. The bit loading of a D RE is determined by configuration. All D REs have the same bit loading within an RB. The Symbol Mapper continues filling RBs until an end-ofburst is detected in the processing of the PCS data. The RB containing the end-of-PCS burst will be indicated by inserting an end-of-burst maker in the appropriate REs designed E. The first bit of data of the PCS burst should be aligned with the first available D RE in the first RB. The RB containing the end-of-burst will likely have padding in the trailing D REs to account for normal end-of-data burst misalignment with the end of the RB.

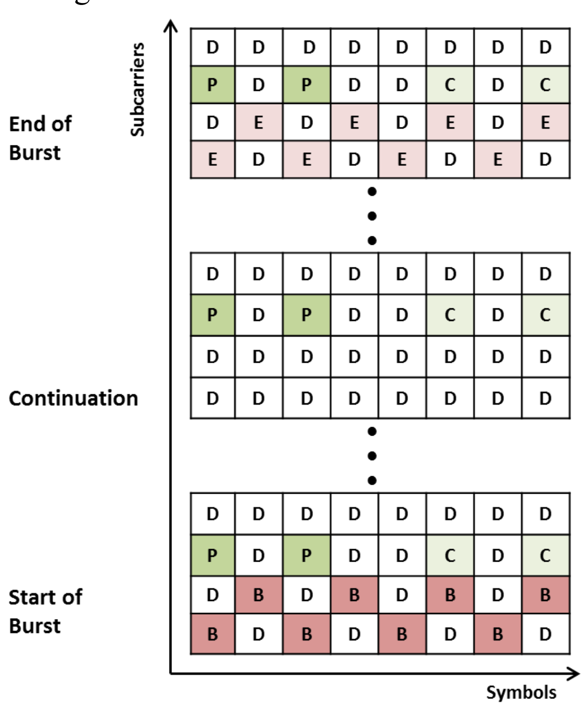


Figure 6. Example Resource Block Use

Bursts from different CNUs will be separated by one or more RBs where zero energy is transmitted. This forms a guard time.

OFDM/A RB framing with interleaving will require buffering as all bits from the PCS that are allocated within the OFDM/A frame must be processed by the Symbol Mapper so that the Interleaver operation is complete before the first OFDM/A symbol is transferred to the IFFT.

Up to 10 Gbit/s

One P802.3bn Task Force objective states: a data rate higher than the 1Gbit/s baseline data rate and up to 10 Gbit/s when transmitting in assigned spectrum and in permitted channel conditions. Another decision of the Task Force is: *the standard shall support the ability for higher capacity by combining multiple 192 MHz OFDM channels* [3]. While one 192 MHz OFDM channel will achieve at least 1.6 Gbits/s MAC/PLS data rate, exploring methods to achieve the 10 Gbits/s objective have been discussed in the Task Force. No technical selections have been made as of this time. One of the methods discussed is presented here as an example.

IEEE 802 standards have different approaches to increasing data rate between stations. Following 802.1 Bridging, multiple MAC/PHY entities can be combined using the 802.1AX Link Aggregation (LAG) standard [9]. For relevance to P802.3bn, this would mean a vendor would combine two or more distinct EPoC MAC and PHY client entities. This approach is already covered under 802 standards and is outside the scope of the 802.3 standard. Within scope, 802.3 standards increase the PHY link data rate by multiplexing PCS data over multiple channels that are referred to as lanes. Up to four lanes are typically combined by a multiplexing function in the PCS, with individual paths through the PMA and PMD. As P802.3bn is already multiplexing bits to subcarriers in the PMA, it makes sense to leverage this architecture and avoid layering another PCS multiplexing layer on top of PMA symbol mapper multiplexing.

Recall that the Symbol Mapper of Figure 4 performs multiplexing bits received from the PCS to a plurality of subcarriers in an OFDM channel under the direction of configuration information that includes bit per subcarrier loading. By extending the number of subcarriers supported and a straightforward approach to mapping subcarriers to individual OFDM Channels, several channels can be run in parallel. Figure 7 shows an example with four OFDM channels.

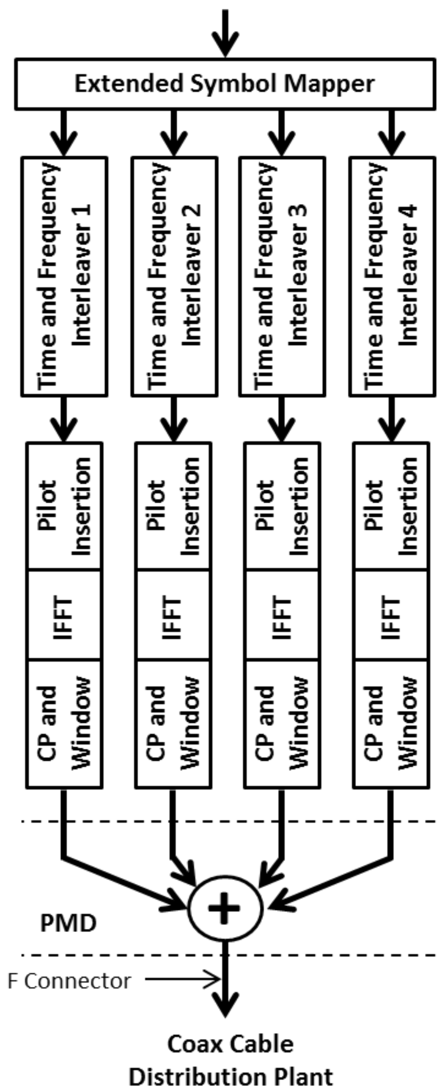


Figure 7. Multiplexed OFDM Channels

There are some assumptions with this approach. There may not be enough available RF spectrum to support all channels remaining operational. As individual subcarriers can be excluded, it would make sense for individual OFDM channels to be set as enabled or disabled until needed by the cable operator. For processing synchronization, all OFDM channels will need to have the same CP and Windowing values so the extended OFDM symbol is of the same size. The outputs of the parallel OFDM channels can be combined in the PMD. Specific combination techniques and the regulation of per channel power, and so forth, can be left up to each individual vendor.

DOCSIS 3.1 Alignment

There has been an ongoing effort to align EPoC and DOCSIS 3.1 OFDM numerology insofar as it is practical. This includes the Task Force consideration of adopting common component architectures for the industry. From an 802 process standpoint, individuals may bring forward alignment proposals, and the Task Force considers these as part of its normal socialization and technical selection process. There has been initial alignment of an OFDM FFT size, 204.8 MHz sample rate, CP and Window sizes, and upstream LDPC FEC coding and rates. In addition, the electrical input and output requirements of the PMD should be well aligned with DOCSIS 3.1 PHY OFDM channels, permitting consistent behavior and expectations when operating on a cable operator’s coaxial network. It is also expected that P802.3bn will add similar sensors and measurements to support Proactive Network Management (PNM).

SUMMARY

An overview and highlights of the IEEE P802.3bn EPoC PHY Task Force have been presented. The EPoC effort is in progress. Final architecture and operational details will be determined by consensus of the Task Force and the ballot approval process. The resulting PHY standard will enable several deployment models, increasing the number of choices for cable operators.

Ethernet is constantly evolving. EPoC is a novel solution for extending Ethernet as found in EPON services to cable operators in support of their multi-gigabit service offerings and for extending the life of the coax network.

REFERENCES

[1] IEEE 802.3 Working Group home page: http://www.ieee802.org/3/

[2] IEEE Standards Association http://standards.ieee.org/findstds/standard/802.3-2012.html

[3] IEEE P802.3bn Task Force home page: http://www.ieee802.org/3/bn

[4] CableLabs®, DOCSIS 3.1 Physical Layer Specification, CM-SP-PHYv3.1-I01-131029, http://www.cablelabs.com/specs/.

[5] Metro Ethernet Forum home page: http://metroethernetforum.org/. For more information on MEF Carrier Class Ethernet *Class of Service* requirements, look under Information Center and then MEF Technical Specifications for document MEF 23.1.

[6] DOCSIS® and DPoE™ are trademarks of CableLabs, Inc. More information DOCSIS and DPoE can be found at the CableLabs specifications web site at: http://www.cablelabs.com/specifications/

[7] IEEE P1904.1 Working Group Standard for Service Interoperability in Ethernet Passive Optical Networks (SIEPON). http://grouper.ieee.org/groups/1904/1/.

[8] Richard Prodan, *FEC Codeword Size Selection Based on Burst Size*, IEEE P802.3bn Task Force, November, 2013.

[9] IEEE Standards Association, http://standards.ieee.org/about/get/802/802.1.html

IEEE DISCLAIMER

This paper presents an overview of an IEEE Standards Association activity. The views presented in this paper are personal views of the author and do not represent the formal position of the IEEE.

For more information, refer to: http://standards.ieee.org/ipr/disclaimers.html

AUTHOR

Mark Laubach is the Chair of the IEEE P802.3bn Task Force standards project. At Broadcom, his work covers broadband management, architecture, and standards. Mark has more than two decades of experience spanning broadband, gigabit, and cable access network technologies including being CEO and President at Broadband Physics, VP, CTO, and co-founder at Com21. He is an inventor on several U.S. patents on cable modem technology. He is an IEEE Senior Member and an SCTE Senior Member. He is the principle co-author of a cable modem technology book B*reaking the Access Barrier: Delivering Internet Connections Over Cable* (John Wiley Press, 2000). He holds a BSEE and MSCS from the University of Delaware under Professor David J. Farber.

EVOLVING THE TELECOM STACK AND HOW WEBRTC PLAYS A ROLE

Chris Wendt, Principal Architect, IP Communications and Services
Comcast

Abstract

This paper provides an overview and rationale for a vision of the future evolution of telecommunications services and how real-time communications can be provided as a more flexible web oriented service. WebRTC, a web browser focused framework for real-time communications, provides standardized client-side "hooks" which can be adapted for mobile, set-top box and other popular embedded clients. The larger question is how these client technologies can be used as part of a modern, web-friendly service framework. We at Comcast Labs are proposing a framework and perspective beyond traditional telephony services, defining how real-time communications can be customized and integrated for the variety of service contexts and devices we use in our personal and professional lives. Examples include: enhanced customer service, enhanced home monitoring and device control, interactive collaboration or shared experiences. This paper will propose and discuss a framework and architecture and provide the motivation why WebRTC plays an important role, and discuss why past attempts may not have been as successful, and why based on technology and network evolution, and device capabilities, this time might be different.

INTRODUCTION

The way humans communicate in just the past five or so years has evolved at a lightning pace. With the now nearly ubiquitous availability of bandwidth and communications over wired and wireless networks and the exponential improvement in CPU and GPU capabilities of devices, the world has changed for how the human/computer interface has evolved, how we interact with our devices, how we buy and consume content and applications. It's almost unthinkable how quickly humans have adopted these new paradigms without looking back. For these reasons and others, our communications habits have been similarly transformed to an always-on multi-tasking, asynchronous method of communications, easily switching between e-mail, text, voice, video between multiple people, often simultaneously depending on our communication needs or location or connectivity. Our communications has also gotten much more contextual, often adopting fine-grained preferences for how we communicate with a particular person in a particular social context.

There is no question traditional views of communications services are quickly being challenged because of this new world. These new capabilities along with the evolution of mobile and web applications and new application frameworks and technologies like HTML5 have opened the flood gates both for application developers and end users.

What does that mean for the communications service provider? What should a telephony product look like in the next 5-10 years and how does the communications architecture evolve to support it? What role can/should the service provider play in this new world of expanding communications capabilities?

Dreaming of the Next Generation Network

IP telecommunications has evolved to become mainstream rapidly in just the past 10-15 years since it was adopted widely. The PSTN or POTS network that VoIP has been employed to carry, in many respects, looks very much the same as how it looked over 30-40 years ago. Other than the conversion from

SS7/TDM to IP protocols like SIP, the core telephone service has not changed. There have been many attempts in various industry and standards bodies at extending and evolving the telecommunications network by defining a "next generation network". To a large extent, these efforts have not pushed the ball forward. The evolution of the Internet and interconnected IP network has allowed for delivery of voice and video in better, converged, and more efficient and cost effective ways. The building blocks for adding new and web connected "telephony features" has evolved as well. However, the fundamental black phone and dial pad interface, while important for basic communications capabilities, has been somewhat of a boat-anchor for the evolution of how the products around communications are defined.

Voice telephony is still THE primary service and still the basis of a multi-billion dollar industry; who's to argue for changing the formula?

WebRTC, What’s new?

Enter WebRTC. By itself, some may argue that WebRTC is purely an API to access a camera/microphone, create a real-time channel, and display media on a remote client. What's new? There is a lot of push back from the telephony industry saying, WebRTC is just a new client in a crowded space of VoIP protocols that have existed for years. The potential for a revolution may be subtle depending on the viewpoint, but profound if put in the right context. When you view the capabilities of WebRTC in the context of a browser, all of the power and complexity of VoIP all of the sudden becomes a few lines of JavaScript code, a tiny component in the mix of a lot of other application and multimedia capabilities. It looks to be a stark new reality for the telephony product manager, but is it really a much larger opportunity in disguise?

Real-Time Application Architecture

WebRTC provides a flexible API to wrap real-time communications into new and different applications in a standard way. Interestingly, and often confusingly, WebRTC makes no assumption about signaling protocols or any definition around any specific application or service. WebRTC, at it's core, only concerns itself with the capture and display of media and the mechanics of setting up an IP channel to transmit the real-time media data.

To some, having no specified signaling or application protocol leaves WebRTC so open ended, it can be hard to wrap one's head around building any common framework to handle communications services, particularly in a generic enough way that a service provider can play any specific role. Put in a new context, however, if we embrace the limited scope of media session establishment, the world of all communications related applications, regardless of signaling protocol or functionality becomes accessible and adaptable. And isn't that really the better long term approach anyway?

Application Models

There are two classes of applications WebRTC was designed to support. As defined in [1], they are the triangular and trapezoidal models. The triangular model, shown in Figure 1, defines the case where clients that want to establish a real-time channel between each other, talk to a common central server that coordinates passing IP address contact information between each client so they can establish a media channel directly between them.

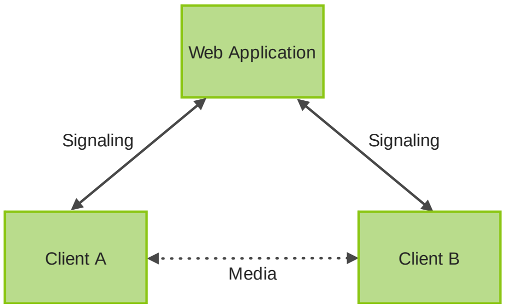


Figure 1: Triangle Application Model

The trapezoidal model, shown in Figure 2, defines the model where a client talks to a server associated with his particular application server, a remote client talks to her server, and there is an agreed upon signaling protocol and IP channel between the servers that passes the IP address contact information.

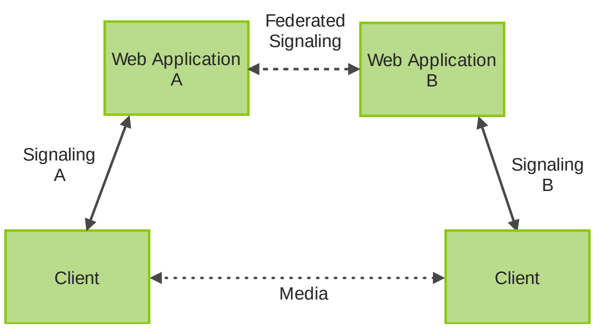


Figure 2: Trapezoid Application Model

This can be looked at as the federated model. "A" has a communications application with a particular signaling protocol, "B" has another communications application with another signaling protocol. So for a client of server "A" to talk to a client of server "B", there has to be a federation protocol that is agreed upon, to make the end-to-end communications work. There are explicitly 3 distinct legs of signaling in the trapezoid model. Both the triangular and trapezoidal models define virtually the entire universe of models of real-time communications applications.

These application models are important to help guide a general architecture that can flexibly support both models depending on the needs of the end application utilizing a real-time communications service framework. But these signaling models alone only go so far in defining what is needed to make an end-to-end service architecture real.

END-TO-END SERVICE ARCHITECTURE

Applications and services comprise components that must be delivered, secured, billed, managed, authenticated, and authorized. These requirements are common across services and can be abstracted into a generic service architecture. One such service architecture currently deployed in both cable and mobile networks is the IP Multimedia Subsystem (IMS), shown in Figure 3. The 3GPP defines the IMS specifications [2] for mobile providers; CableLabs, in turn, adopted and enhanced those specifications as the core of its PacketCable 2.0 initiative [3]. IMS defines an industry standard way for managing, billing, authentication, as well as the protocols and components used from an end-to-end perspective. IMS utilizes SIP and DIAMETER and other IETF defined protocols to deliver telephony services. SIP and IMS were born out of a specific signaling and service delivery model to support telephony services and the application and feature servers that can extend those services.

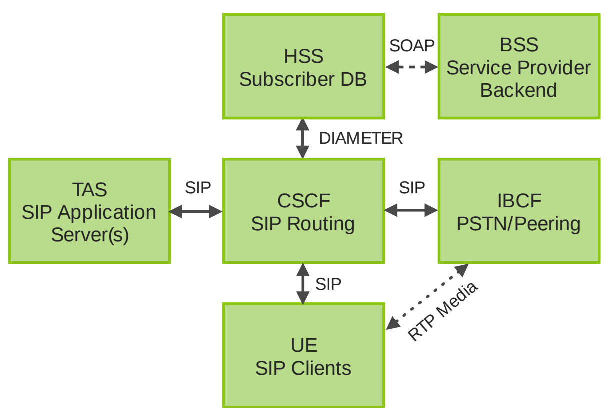


Figure 3: IMS High-Level Architecture

Comcast, as an example, has moved from distributed soft switch architecture as defined in PacketCable 1.0 and 1.5 [4] to the IMS/PacketCable 2.0 architecture. We have achieved many of the cost, operational, reliability benefits of moving to the centralized and highly scalable model that IMS offers. This is a clear win from that perspective.

However, as web-based services became more focused around building cross-platform and cross-service applications, and as newer web authentication and mobile applications models evolved, even as recently as in the past 1-2 years, there developed a clear need for a better service framework layer.

Because IMS derives its application model from the SIP family of standards and derives its service delivery foundation from SIP based deployment and management practices, there is an implied model that is hard to morph to support a general application framework. In addition, its heritage as part of the 3GPP and focus on a mobile vision of a single primary device per customer presented challenges, as will be discussed more in depth later. Extending beyond the concept of a managed device and managed primary singular identity was challenging. Incorporating non-telephony devices and services and even bring-your-own-device models, which can be supported for some cases, became burdensome to support in a general way.

These issues became the main motivation for the work presented in this paper. We propose a new telecom stack, one with the flexibility to support both traditional and non-traditional telephony servicesm and with a focus on the fundamental establishment of real-time communications channels through a flexible, web API-centric model. Just as WebRTC is designed without any specific signaling protocol, we believe being non-prescriptive to any specific service or application architecture, device management model or identity or set of identities is an approach that provides maximum flexibility in most contexts.

<u>Identity</u>

Many service providers, like Comcast, are not only telephony service providers, but also a provider of many services such as Internet and television and associated web services to supplement these services. They generally provide two main types of identities that can be associated with a particular account holder and user. These include an e-mail identity and a telephone number. Most web applications and services require an email address and user-generated password for account authentication, typically enforced by a centralized, application independent single-sign-on system (SSO). In the past few years, it is becoming increasingly popular, and for some services mandatory, required to use OAuth style token-based authentication. OAuth 2.0 [5], supports a common framework for authentication of user ids from multiple client environments, such as, mobile, browser, embedded device to name a few. This allows a single set of credentials to be managed by end users to access all of their service provider services on many different platforms with a single set of credentials managed from a central secure service.

Additionally, it is increasingly common for application and service providers to be agnostic about supporting both identities owned directly by the provider as well as those originating from third parties. These third-party identities, can be supported either by direct association of a user defined password to that third-party identity, or via what's commonly referred to as a three-legged authentication. Here, the application or service provider trusts the authentication token provided by what is usually an OAuth supporting third party identity provider (IdP). Popular examples include Google, Facebook, Twitter Auth or more general industry initiatives like OpenID. This idea, in the context of WebRTC is detailed in [6], but this framework is now very commonly used in web applications in general.

As is common for many service frameworks born before these flexible identity authentication frameworks became commonly used, IMS and SIP generally assume an identity model where a particular identity type, a SIP URI or TEL URI, is used. Many SIP networks employ a specific set of credentials for SIP services, separate and distinct from any web based login credentials.

For authentication, IMS currently specifies either a SIM mechanism for the mobile terminal world, where a physical card with an embedded certificate/public identity is provisioned for an account, or in the case of many fixed line service devices like Cable E-DVA's and enterprise PBXs, SIP Digest based username/password authentication is used. The identity or telephone number, known in IMS as the public identity, is associated with the device authentication credentials, known as a private identity, are both stored in the subscriber database, HSS.

From many perspectives, having multiple credentials associated with a user is both difficult to manage, inconvenient for the end user and can be a security risk.

Service/Device/Identity Abstraction

How do we provide the flexibility needed by the web and provide a service that can be managed in a reasonable way? The web model fundamentally a distributed, abstracted model for services, applications, devices, and identity. A service provider wants a common way to offer services to other services, applications, devices and using an identity of the users choice. The question is how do we move the telecom stack to adopt these principles in a way that is consistent with this level of abstraction.

Our proposed solution conceptually started as an extension of the routing model that is core to IMS and SIP services and incorporated a similar user/device based model and architecture with a web services style abstraction. While this work started before WebRTC was born, it quickly became quite intentional to adopt many of the ideas from the WebRTC and general web framework. These concepts include:

- Identity - support the ability for end applications to use any identity model, either self managed, third party managed, or service provider managed.

- Device - allow for the ability to support multiple devices, either simultaneously, independently, and dynamically without any dependency on provisioning or management interface

- Application - support either the triangular or trapezoidal application models with or without the dependency on service provider routing services (i.e. SIP/PSTN/RCS/federated routing models)

The model was defined as a single API to support both internal and external applications, with a primary focus on the ability to setup of real-time media channels. The resulting core service became very concise and clear, a media session establishment API, with hooks to specific traditional routing services and additionally some value added media functions as well (e.g. mixing, transcoding, etc.). This API additionally, depending on calling identity and domain, can be used to establish sessions between end clients directly, over PSTN, over SIP peering, or other federated models that may appear in the future.

In the future, this framework can be extended to new models of real-time communications services that don't exist today.

PROPOSED ARCHITECTURE

The service framework we propose in this paper began as an extended SIP framework, supporting mobile applications as an extention to primary-line IMS services. Shortly after WebRTC was first proposed in W3C and IETF the framework was quickly adapted and extended to incorporate a fully HTTP based flow utilizing technologies such as Websockets and OAuth to provide a flexible solution that can exist entirely in the HTTP world and as part of a hybrid HTTP/SIP model. It has developed both from the needs of our product evolution and new product requirements including the grander motivation to address all the issues stated earlier as a first class design criteria.

The high level architecture is shown in the figure. There are a few major architecture components, many existing and common network elements not specific to this architecture and some new components developed specifically to support general real-time communications and media session establishment aligned with WebRTC and rtcweb protocols.

The existing infrastructure utilized includes:

- IMS/SIP infrastructure - the main SIP based PSTN and federated service routing component, extendable to other SIP services including RCS and VoLTE interworking and other future interconnected federated services

- Notification Services - a common platform for events and notifications, supporting Comcast managed devices such as set-top boxes as well as federated notifications including iOS and Android notification services

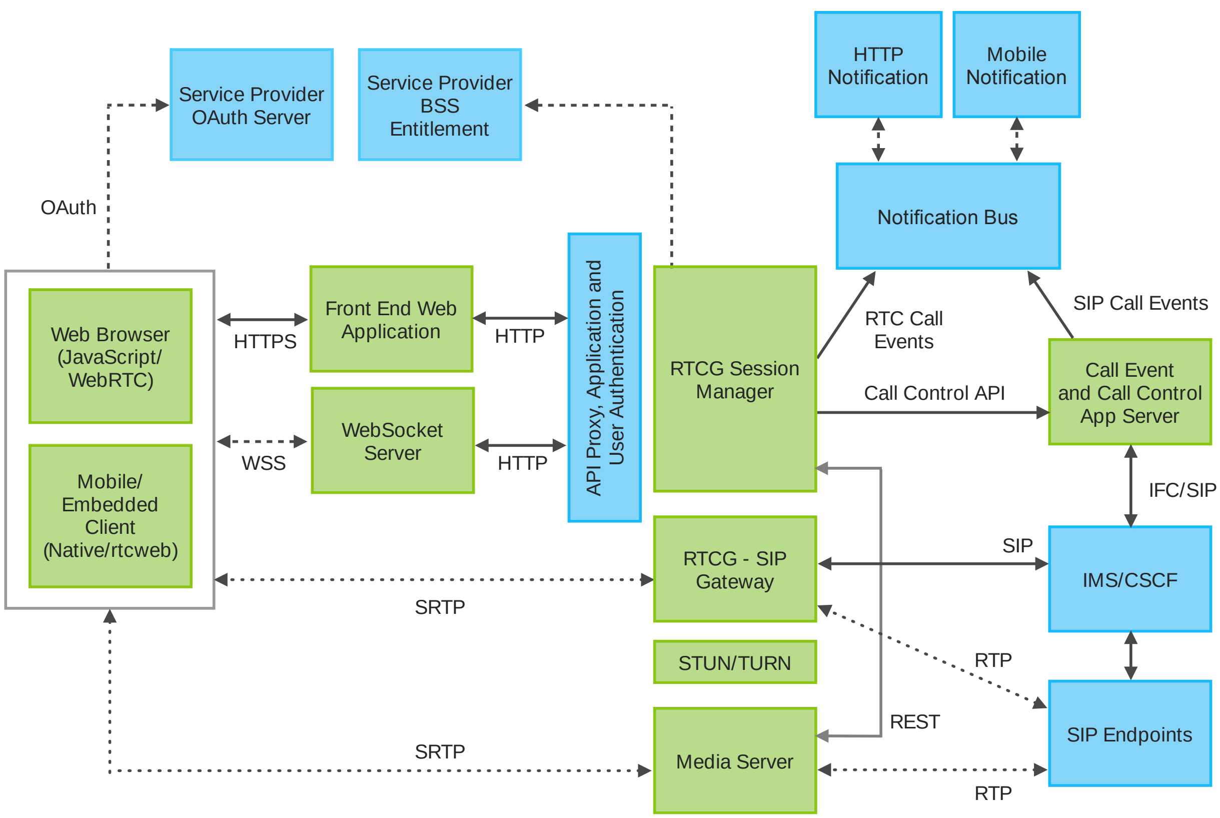


Figure 4: Proposed RTC Service Architecture

- API proxy and security layer - a common API proxy layer that incorporates application and user level authentication and typical API exposure and security functions

- SSO/OAuth Server - a common authentication service supporting OAuth token authentication for user IDs

The additional components that will be discussed in more detail include:

- WebRTC clients - including JavaScript based, browser and browser-like clients as well as native mobile SDK based clients supporting rtcweb protocols.

- Real-time Communications Gateway (RTCG) - includes a session manager, user manager, and SIP signaling and media gateway component to provide interworking with IMS/SIP and RTP media streams and codecs supported in the federated networks

- Application and WebSocket Server - represents the multitude of applications that can utilize RTCG services and provide a signaling path that is either application specific or conforms to a reference signaling model

- Call Event and Call Control - this component allows the bridging of incoming calls from federated networks to and from the RTCG

WebRTC clients

WebRTC is part of the W3C specifications and what some refer to as the HTML5 set of browser capabilities. It defines a set of Javascript APIs, including primarily getUserMedia and peerConnection. These APIs define both how to access the camera and microphone of the underlying OS platform as well as the establishment of a particular RTP based media channel between two peers. The API is defined in the W3C specifications, and the specific protocols are defined in IETF under the rtcweb working group.

We focus on two classes of clients, but aren't necessarily limited to these. The first is a JavaScript based client using the W3C defined WebRTC APIs that either supports a traditional third party browser application, or a more integrated device that provides a browser enabled environment in order to support "embedded" HTML5 applications. A JavaScript reference SDK is provided to support a particular signaling protocol based on websockets and incorporates the specifics around authentication of user and/or application. The other client model is a native approach, where an SDK native to the device is provided with either C or Java based APIs similar to the W3C APIs. These specific clients are typically in scope of most WebRTC based services today, but as stated, is not limited to these device or application models.

Real-time Communications Gateway

The RTCG is the network component handling basic registrar and media routing logic. It exposes a RESTful HTTP based API for establishing media sessions. It acts as a dynamic registrar and routing proxy and is for the most part agnostic to identity with optional configurable routing rules based on specific identity and domain. We very explicitly wanted a model that was dynamically able to support any unique identity in the context of a particular unique application id. Additionally, it was important to support this with minimal provisioning or configuration, if at all.

The main federation interface supported is SIP. This can be extended to other gateways supporting protocols or even federating to other RTCG supporting service provider networks. In the case of SIP and IMS, a media session initiation is translated into a SIP INVITE to the IMS SIP network, via Mw interface to CSCF. The SIP REGISTER method is specifically not supported. In the terminating case, an INVITE toward RTCG SIP Gateway is translated into a media session on the RTCG. Otherwise, both signaling and media are handled very similar to traditional signaling and media gateway components. In the case of SIP to WebRTC, media is likely converted between RTP and SRTP/DTLS and if transcoding is necessary, it can be incorporated. ICE and TURN procedures are followed for WebRTC clients and the exchange of credentials for TURN is handled in the API.

Another important change to note in the architecture is around the use of notifications as the primary mechanism to signal the initiation of a potential session. In the traditional SIP architecture, there is an explicit REGISTER method that provides a persistent connection to a SIP registrar that allows an INVITE to be sent to the associated contact address. We have moved away from this idea for a number of reasons:

- The number of potential registered devices is growing exponentially multiplied by the potentially exponentially growing number of applications

- Because of different network topologies and NAT and firewalls, holding persistent connections to devices can be a challenge

- Because many mobile devices are powered by batteries and persistent connections to networks can be very expensive in terms of power, it is an often discouraged practice

- Most important, the modern application interaction model has changed, web pages are not persistent, mobile applications can be persistent, but we interact with them in short sessions and based on notifications, rather than having an application always in the forground

That said, the dedicated telephone device model can be supported with a persistent registration model if needed, but we see this model less and less relevant as time progresses.

Application and WebSocket Server

One of the key design criteria for the proposed architecture is the separation of application and routing. The application interface should be a convenient API that hides the details of routing and session management from the application layer. The architecture shouldn't impose any assumptions or constraints around the application developer; there should be clear separation between application and RTCG. There also isn't any assumption around application environment. Any modern HTTP supporting server side development environment can be supported.

The minimum requirement is the application developer only needs an application key to authorize access to the RTCG APIs. Some applications that may require authenticated access to service provider services. For Comcast, as an example, this might include PSTN calling from the Comcast customer TN or access to specific paid services that require Comcast user credentials in the form of an OAuth token provided by Comcast SSO system. In this case, the application specific credentials along with the Comcast user token can be passed in the HTTP requests to RTCG and an entitlement check is performed in the RTCG to validate the association with the account specific authorized services. Another example application might need the access to a media stream from a managed secure device such as an in-home security camera. The three-legged auth model can also potentially be employed to support both end-user directed auth and revoking of a third-party application to access to these types of media stream services.

Additionally, there is a provided reference client SDK, application server and websocket server to support a particular signaling model that can be used by application developers that don't want to build their own signaling.

Call Event and Call Control

Because of the notification-based mechanism that is imposed by the architecture, for incoming or terminating calls to a federated network like IMS, there needs to be a mechanism to send a common set of call events and provide an interface for third-party call control. The importance of notifications was discussed above, but the mechanism to report incoming INVITEs and other call state details allow the end WebRTC client to interact with the call in the federated network. The call control interface provides the mechanism for when the client wants to pull or push a call to/from the federated network in the same style third party call control works in SIP today.

CONCLUSION AND NEXT STEPS

Much of this work was part of an architectural evolution born out of the necessity of supporting a more web-oriented approach to communications. From the beginning, we intentionally kept our view very broad, from the basic ability to support OTT telephone soft clients to the ability and flexibility to support the potential universe of non-traditional real-time communications applications. The guidance of the fundamental principles of WebRTC combined with a fundamentally web-centric approach to integration into the service provider common services sets the stage for a truly new approach to the integration of real-time media services.

To the casual observer, media streaming over the Internet seems like a solved problem. It is common to see HD resolution video streamed over IP with generally high quality and latency. Of course, the important distinction of telephony types of real-time communications is that minimizing end-to-end delay is an important requirement. With stored or buffered live streaming timing constraints are very much relaxed, often in the order of seconds or 10s of seconds. Delays in the order of low hundreds of milliseconds or lower are critical to delivering a quality experience. Even today, this continues to be a sometimes difficult challenge over varying network conditions and topologies. Though aggregate network speeds have improved immensely, there are still existing bottlenecks that have plagued real-time communications over IP networks from early on. There are many efforts in the IETF, as an example, to specifically tackle these issues and look at minimizing congestion over and above traditional congestion control techniques and priority packet marking techniques. As a service provider, building a standard framework that enables a more predictable experience for its subscribers across all of the applications they use can be an interesting benefit to employing standard APIs for media stream management. There have been various attempts at this in the past, but perhaps WebRTC and the proposed architectural framework can be the technology to rally around to deliver it in a consistent way.

Extending the Framework

For those familiar with WebRTC, one perhaps glaring omission from this proposal is regarding non-media related real-time streams such as the WebRTC data-channel or even websocket channels. It doesn't take much imagination to recognize that a very similar framework can be employed to support certain real-time data classes of applications including gaming, messaging, machine-to-machine or Internet-of-things types of use cases.

We would like to get industry and community feedback around this proposed framework. Not unlike IMS, we believe there is a large opportunity to evolve real-time communications as a consistent framework for either application integration or federation of services, even beyond PSTN type services. With a specific focus on a flexible web based API for media session establishment alone without any application signaling assumptions, this framework is much better positioned to evolve the telecom network to a new generation of applications and services.

ACKNOWLEDGEMENTS

I would personally like to acknowledge and thank Bryan Paluch for his collaboration and work in taking this from early concept to implementation and his valuable contributions to refine and extend this proposal to the state it is today. I'd like to thank, Erami Botchway, John Hart, and Bryan Paluch for their valuable comments and review. I'd also like to thank our product and business teams that had the vision to see the potential in being a bit disruptive to forge new ground and test the waters of new opportunities in next generation communications.

REFERENCES

[1] draft-ietf-rtcweb-overview
[2] 3GPP TS 23.228 IP Multimedia Subsystem Stage 2
[3] CableLabs PacketCable Architecture Framework PKT-TR-ARCH-FRM
[4] CableLabs PacketCable 1.5 Architecture Framework Technical Report PKT-TR-ARCH1.5
[5] draft-ietf-oauth-v2
[6] draft-ietf-rtcweb-security-arch

Field Measurements of Nonlinear Distortion in Digital Cable Plants

Tom Williams and Belal Hamzeh, CableLabs
Ron Hranac, Cisco Systems

Abstract

As cable networks transition from the carriage of analog TV signals to digital signals, the second- and third-order nonlinear distortion products, known as composite second order (CSO) and composite triple beat (CTB), are still present but now appear on conventional spectrum analyzers to be similar to random noise. Thus, to technicians the distortion energy is indistinguishable from random noise. A new digital signal processing (DSP) technique has been developed to process the spectral "noise" in a vacant band to determine if the energy is random, or correlated to a full-band nonlinear distortion signal that could have created it. This paper discusses the processing details of the new test method and presents lab, cascade, and field test results. These data demonstrate the operational benefits of making nonlinear distortion measurements in mostly- or all-digital cable plants.

DISTORTIONS IN CABLE NETWORKS

Active devices through which a cable network's signals pass compensate for various losses in the transmission path, and extend the physical reach of the network. Unfortunately, those same active devices also degrade the quality of the signals to some extent, with the amount of degradation largely related to the active circuit type (e.g., single-ended, push-pull), the active device's noise figure, number of signals carried in the network, the number and types of cascaded active devices through which the signals pass, the active devices' dynamic range, and the active devices' operating RF levels. The degradation is characterized in terms of parameters such as carrier-to-noise ratio (CNR) and carrier-to-nonlinear distortion ratio. This paper focuses on nonlinear distortions, in particular the measurement of nonlinear distortions in mostly- or all-digital networks.

Nonlinear distortions

The amplifiers used in cable networks are not perfectly linear devices, that is, their transfer function is not constant. If one were to plot an amplifier's instantaneous output voltage versus its instantaneous input voltage on a graph (see Figure 1), the resulting line would not be straight.

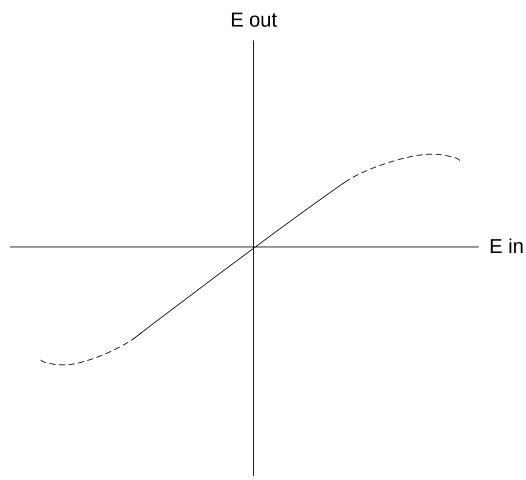


Figure 1. Example plot of amplifier output versus input voltage.

This behavior can be modeled mathematically using a Taylor Series expansion given by the equation

$$f(x) = Ax + Bx^2 + Cx^3 \ldots$$

where x is a time-varying input signal, and *f()* is a nonlinear operator such as an overdriven amplifier (or a cascade of overdriven amplifiers). The A term is linear gain, B is second-order distortion, and C is third order distortion. Higher terms, such as D, E, and F, also may be significant.

In typical operating conditions, the amplifier operates in the linear region (solid line in Figure). In non-ideal operating conditions, such as amplifier failure or overdriving the amplifier into its saturation region, the amplifier starts operating in the nonlinear region (dashed line in Figure 1).

Nonlinear behavior in an amplifier is related in part to small-signal nonlinearities in the amplifier’s semiconductor devices, but mostly to signal compression that occurs as the amplifier is operated near its saturation point.

Nonlinear distortions get their name from the nonlinear operation of the amplifier or other active device that creates the distortions. One notable differentiator between linear distortion, such as group delay, and nonlinear distortion is that linear distortion cannot create distortion energy at new frequencies, but nonlinear distortion can. Likewise, as an amplifier’s operating signal levels are increased, the relative output ratios of desired carrier levels to distortion levels will be maintained for linear distortion, but will not be maintained for nonlinear distortions.

A graphical illustration of second order nonlinear distortion caused by two signals, F_1 (75 MHz) and F_2 (100 MHz), is shown in Figure 2. Note the discrete distortions or beats at the absolute value of the difference F_1 - F_2 (25 MHz), the second harmonic $2F_1$ (150 MHz), the sum of the two frequencies $F_1 + F_2$ (175 MHz), and the second harmonic $2F_2$ (200 MHz).

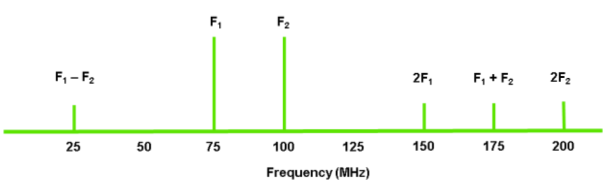


Figure 2. Simplified graphic illustrating discrete second order distortion.

Second and third order distortions

A quick operational test to determine the order of nonlinear distortion in the spectrum is to elevate the amplifier’s input signal. If the distortion energy increases 3 dB for a 1 dB step increase of the input signal, the nonlinear distortion is probably third order. If the undesired energy increases 2 dB for a 1 dB step increase of the input signal, the nonlinear distortion is probably second order.

Second order distortion in cable networks should be substantially suppressed relative to third order distortion because cable networks use balanced push-pull amplifiers. Push-pull amplifiers cancel even order (second, fourth, sixth, etc.) distortions. Expected potential sources of abnormal second order distortions are imperfect analog downstream linear lasers; damaged or unbalanced push-pull amplifiers; and distortion caused by diode-like junctions created by corrosion in the plant.

Third order distortion is the dominant nonlinear distortion in cable networks. High-output amplifiers are used to provide needed dynamic range. Cable systems are operated with up-tilt (higher frequencies operated at higher RF output levels than lower frequencies) to provide more uniform distortion over the downstream band. The potential sources mentioned previously for second order distortion also can contribute to third order distortion.

History

In the days of 12-channel operation in North American cable networks, the active devices in the outside plant were single-ended. The nonlinear distortion of concern to most operators was for the most part cross-modulation (XMOD), which, if severe enough, produced a windshield wiper effect in the pictures of analog TV signals. XMOD is a third order distortion.

Since 12-channel cable systems used the same frequencies and channel plan as over-the-air VHF broadcast television (54 MHz to 88 MHz for channels 2-6, and 174 MHz to 216 MHz for channels 7-13), the discrete second- and some of the third-order nonlinear distortions produced by amplifiers fell in parts of the spectrum where TV signals were not carried.

The introduction of push-pull amplifier technology accommodated more usable channels beyond the 12 VHF broadcast equivalents. The new push-pull actives provided a significant improvement in second order nonlinear distortion performance, so the VHF midband and superband frequency ranges could now be used. That meant an increase in cable network channel capacity.

As the downstream spectrum's upper frequency limit continued to increase with improvements in active and passive device technology over time, so did the number of channels provided by cable operators. Along with increased channel counts were an increase in the numbers of discrete nonlinear distortion products or beats, eventually to the point where it became difficult or impossible to discern individual beats.

<u>Composite distortions</u>

Because of the almost "comb-like" regular frequency spacing of analog TV signals in a cable network using the CEA-542-D channel plan, the resulting distortions cluster at various and predictable frequencies in the spectrum. The individual distortions can be so numerous within each cluster that the descriptor *composite distortion* arguably is more appropriate than *discrete distortion*. CTB distortions of concern cluster in a narrow bandwidth under each visual carrier, and CSO distortion clusters appear ±0.75 and ±1.25 MHz relative to the visual carriers.

Nonlinear distortions such as CTB, CSO, and common path distortion (CPD) don't go away in an all-digital network. Rather than clusters of discrete distortions that occur in a network carrying large numbers of analog TV signals, the nonlinear distortions in a mostly- or all-digital network are noise-like. Those noise-like nonlinear distortion products are known as composite intermodulation noise (CIN), composite intermodulation distortion (CID) or intermodulation noise (IMN) – which should not be confused with thermal noise. (Note: The term composite intermodulation noise and its abbreviation are used in this paper.)

Confusion does occur, though. It is widely-known that raising RF levels in the plant improves the CNR, where "N" is thermal noise. But in a system with a lot of digital signals, raising signal levels improves CNR to a point, then the noise floor starts to *increase* and the CNR appears to get worse.

That seems counterintuitive, but the now-elevated noise floor no longer is just thermal noise. It is a combination of thermal noise and the previously mentioned noise-like nonlinear distortions. When characterizing plant performance in the presence of CIN, the term "carrier-to-composite noise (CCN) ratio" commonly is used.

Indeed, CCN is a much more appropriate measurement metric than is CNR under these circumstances, because there is no practical way to differentiate thermal noise from CIN on a spectrum analyzer display. One could turn off all downstream signals except for automatic gain control/automatic level control (AGC/ALC) pilot(s); the noise-like nonlinear distortions would go away, leaving just thermal noise. Unfortunately, this method is service disruptive, and as such is not a practical way to characterize noise versus noise-like nonlinear distortions.

How, then, can nonlinear distortions in a mostly- or all-digital network be measured in an operating network?

Characterizing noise-like nonlinear distortion

Noise-like nonlinear distortion energy is not random, and can be quantified with knowledge of the input signal that created it. It is particularly easy to quantify in vacant test bands, such as a roll-off region. If a vacant band is not available, one can be created by demodulating the RF signal occupying the band, and then subtracting it mathematically.

There are a few possible ways that detection of signal distortion can be done. One method is to capture the same signal twice: One copy of which is a clean undistorted signal at the headend or hub site, and the other copy is captured at a test point in the field. This method has the added complexity of requiring synchronized capture and the transfer of data to a central processing point, in addition to removing linear distortion differences between the nonlinearly distorted signal and the pristine headend signal.

Another method involves capturing the signal at the input and the output test points of an amplifier, and then determining how much additional distortion was added by the amplifier. The linear distortion of the amplifier, including diplex filters response, tilt, and equalization makes this method non-trivial.

The next section of this paper describes a test methodology that requires only a signal capture at one location, where the single full band signal is captured in the field. The captured vacant band signal is stored as a “measured” signal, and processed with a “manufactured” signal. The level of match between the measured and manufactured signals determines how much nonlinear distortion was present in the captured signal’s vacant band.

ONE-SIGNAL NONLINEAR DISTORTION TEST METHOD

The following steps can be used to facilitate a one-signal nonlinear distortion measurement, using just the signal captured in the field.

1) Capture a full-band downstream signal (the entire downstream RF spectrum with all signals and vacant bands) using a digital oscilloscope that has a sampling rate of at least twice the bandwidth of the downstream band, and 10-12 bits of analog-to-digital (A-D) resolution for at least 32,768 samples. In our example, the downstream bandwidth is less than 1,250 MHz, thus the downstream signal is digitized at a rate of at least $2.5*10^9$ Hz. If necessary, a low-distortion, low-noise preamplifier can be used to boost the full band downstream signal prior to capture (digital oscilloscopes generally have a poor noise figure). A low pass filter should be used to remove any energy above 1250MHz, to prevent aliasing. Figure 3 illustrates a time domain display of a signal captured by a digital oscilloscope operating at 2.5 giga-samples per second and 12 bits of A-D resolution. The downstream signal processing requires a vacant band, which may be a roll-off band. This example 54 MHz to 860 MHz

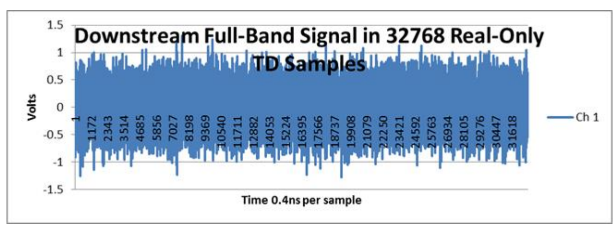


Figure 3. Captured time domain signal comprising 32768 samples, and with a duration of 16.384 μs.

signal comprised mostly digital signals, plus a few continuous wave (CW) carriers used as pilots and alignment aids. The captured signal contains a vacant band between about 770 MHz and 860 MHz, which is not evident in the time domain trace.

2) Convert the time domain signal of Figure 3 into the frequency domain with a fast Fourier transform (FFT). The resulting frequency domain plot is illustrated in Figure 4. In the frequency domain, the vacant band energy values between 770 MHz and 860 MHz are

cut and stored. These frequency domain

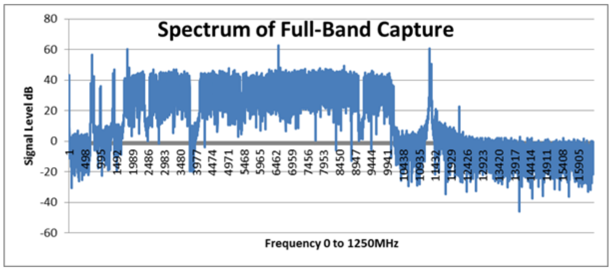


Figure 4. Full bandwidth downstream frequency domain signal obtained by performing an FFT on the time samples shown in Figure 3.

samples are called the "measured" vacant band distortion signal. Next, replace the vacant band energy in the frequency domain signal between 770 MHz and 860 MHz with zeroes. This spectral plot is illustrated in Figure 5.

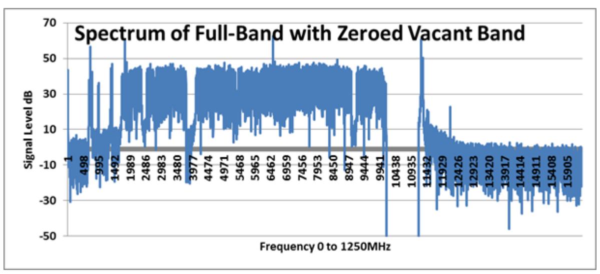


Figure 5. Full bandwidth downstream display with vacant band zeroed out.

3) Next, convert the 54 MHz to 860 MHz signal of Figure 5 with the newly-vacated band back into the time domain with an inverse fast Fourier transform (IFFT), and distort a resulting time sequence with a second and third order nonlinear distortion.

The latter is accomplished by squaring and cubing each term in the time sequence. This creates a second order "manufactured" signal and a third order "manufactured" signal. This distortion manufacturing method gives a good approximate estimate because the nonlinear distortion components are small in operating cable networks. That is:

$$f(x) = Ax + Bx^2 + Cx^2 \sim Ax$$

4) Convert the "manufactured" signals back into the frequency domain and store only the distortion components in the vacant band (770 MHz to 860 MHz in our example).

5) Process the vacant band "measured" signal with the vacant band "manufactured" signals. One processing method that has worked well is frequency domain division of the "manufactured" samples, illustrated in Figure 8, by the complex conjugate of the same frequency "measured" samples, illustrated in Figure 9, to produce frequency domain quotients.

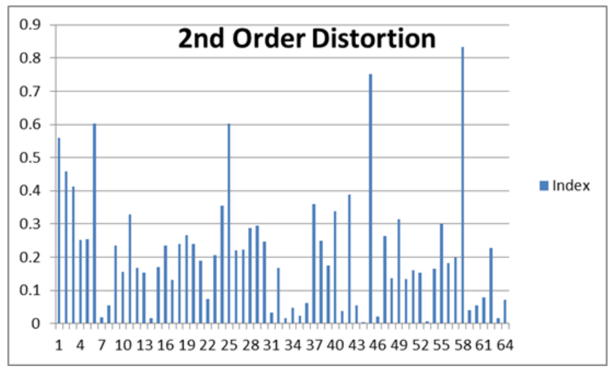


Figure 6. Time domain plot of quotient showing average first term relative to other terms, indicating low second order distortion.

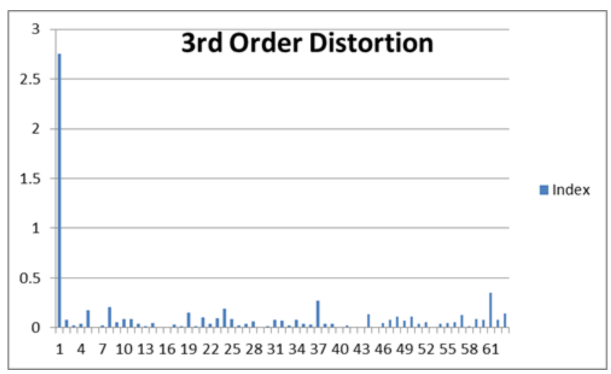


Figure 7. Time domain plot of quotient showing large first term relative to other terms, indicating third order distortion.

6) Convert the 1024 frequency domain quotients into the time domain. This is illustrated in Figure 6 for second order distortion and Figure 7 for third order distortion. Energy in the first (DC) term

indicates a match of the “measured” signal with the “manufactured” signal.

7) If necessary, averaging may be used to better discern the DC term relative to the other terms. Note that the DC terms are correlated vectors that will add, but the other terms are uncorrelated.

8) Repeat the previous steps for other orders of distortion that might be present.

The plots of Figure 6 and Figure 7 are complex time series, and only 64 sample

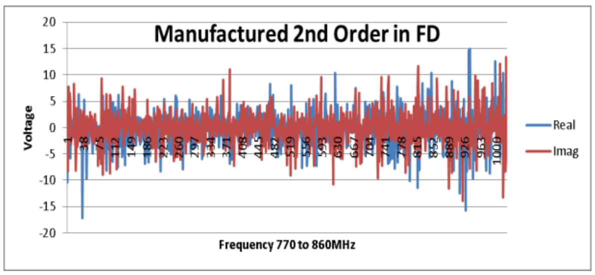


Figure 8. Sample points of manufactured second order distortion from the vacant band.

points are illustrated. As the number of averages increases, the noisy components associated with using a noise-like downstream test signal are reduced. Another improvement to reduce noise in the plots is to use a larger percentage of vacant bandwidth relative to the occupied bandwidth. There is generally a angle to the distortion, and in most observed tests on distorted cable amplifiers, the first term (t = 0) contains most of the energy. As the amplifier’s input drive level increases, both the level of nonlinear distortion and the

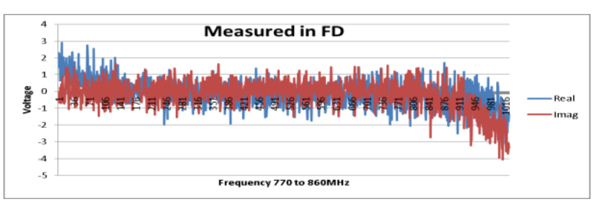


Figure 9. Sample points of measured distortion from the vacant band.

angle of the DC terms change.

OUTSIDE PLANT TESTING

Test Results Obtained from Testing Inside Houses

Figure 10 is a composite plot from seven different locations for second order distortion test results, and Figure 1 contains third order distortion results. Fifteen separate tests were conducted to determine if the measurements were repeatable, and 10 averages used for each point. Locations 6 and 7 had downstream high pass filters to block all but the data-only traffic. Note the top trace on the second order distortion Figure 10, which was location 3. This particular location was later diagnosed also have high linear distortion, as determined by ripples in the upstream equalization response.

The .750 trunk cable feeding this location appeared to have been damaged when it was installed many years ago. In the photograph of Figure 12, one can make out the conduit which may have been fractured by a boring machine. The cable was severely kinked several feet below ground and corroding as one would expect. After repair both the second and third order distortion improved significantly.

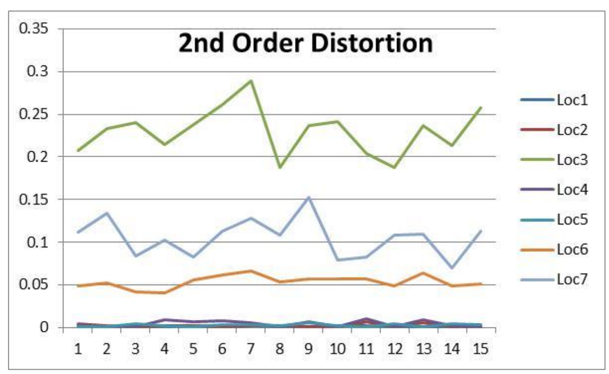


Figure 10. Second order results for seven locations (15 tests each location, 10 averages)

The highest third order distortion came from a home with known ingress issues.

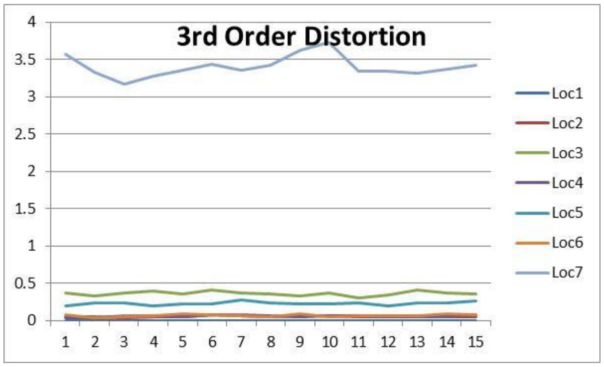


Figure 11. Third order results for seven locations (15 tests each location, 10 averages)

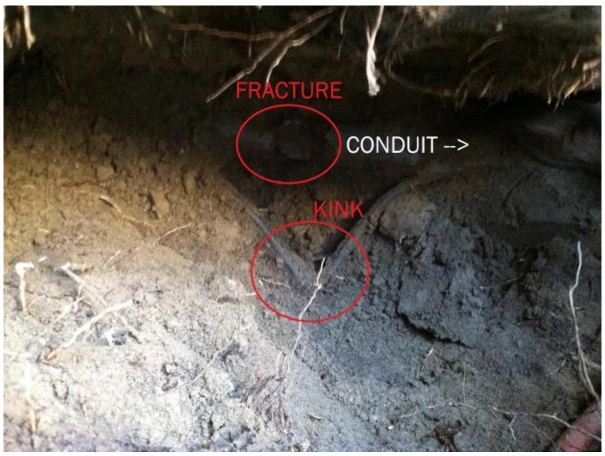


Figure 12. The kinked corroding cable causing the high second order distortion for location 3.

Test Results Obtained from Testing a Node Plus 6 Amplifier Cascade

Tests were performed on a cascade of node plus six push pull amplifiers. The amplifiers were capable of 860 MHz but were only loaded to 770 MHz, so testing was done in the vacant band between 770 MHz and 860 MHz, as described previously. Loading was with digital carriers.

Figure 13 shows the test results for second and third order distortion. Note that the third order nonlinear distortion increased through the third amplifier as expected, but then began to decrease. At amplifiers 6 and 7 the nonlinear third order distortion took a precipitous drop. Second order distortion was not significant. Note, too, that the worst second order distortion was observed at the fiber node.

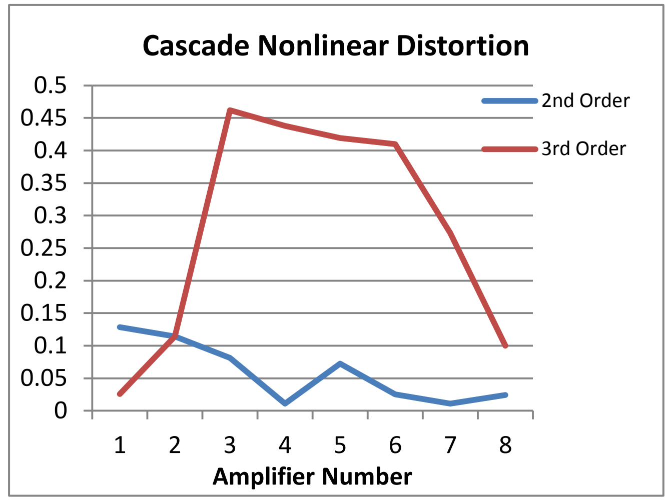


Figure 13. Distortion changes in a cascade of node plus seven amplifiers

This observed third order behavior is not understood at this time. One possible explanation is distortion cancellation by inadvertent addition out-of-phase distortion (a.k.a. predistortion). This method is used commercially in high power amplifiers to extend their dynamic ranges, but the authors are not aware of this being done in cable TV-type distribution amplifiers. Testing additional cascades will hopefully be done in the future to determine whether what was seen here was an anomaly.

LAB TESTING

Lab-based testing was conducted to validate the nonlinear distortion measurement concept discussed in this paper. A single-ended amplifier (Agilent 8447D) and a 54 MHz to 860 MHz cable push-pull hybrid amplifier (ATX model QDAXU) were used in the lab tests.

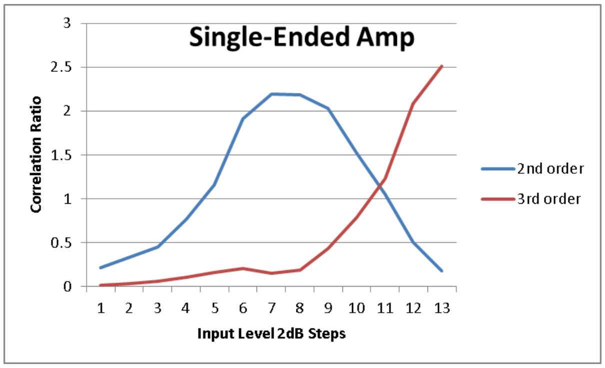


Figure 14. Single-ended amplifier performance.

The input signal was a broadband random noise source (NoiseCom UFX 7109) with its RF output filtered using a 600 MHz low pass filter (Mini-Circuits model BLP-600-75) to create a vacant test band around 800 MHz. The ratio of the energy in the DC term of the correlation plot to all other correlation coefficients was plotted versus input drive level as the input drive level was increased in 2 dB steps. Figure 14 shows that the single-ended amplifier first exhibited an increase in 2nd order distortion, followed by an increase in 3rd order distortion at higher input drive levels. Of interest was the second order distortion decreasing with heavy saturation levels.

Figure 15 shows the same plots for the push-pull cable downstream amplifier. The second order distortion remains low for all drive levels, while the third order distortion increases with higher drive levels.

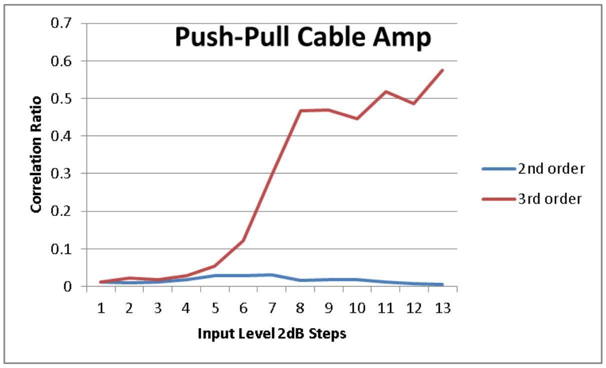


Figure 15. Push-pull amplifier performance.

RETURN PATH TESTING

The nonlinear distortion measurement method described in this section was demonstrated at the CableLabs 2014 Winter Conference. This new method was developed to analyze RF signals to determine the presence of nonlinear distortion using digital signal processing techniques. The method involves multiple transmissions of a burst (or a known variant of the burst), and performing a full-band signal capture containing a vacant band having noise and/or distortion for each transmission. As described here, it is possible to use this measurement method to evaluate a cable network’s return spectrum, although it also could be used for other applications, such as wireless transmitters.

The energy in the vacant band from the first burst transmission in Figure 18 is processed with the energy in the vacant band from the subsequent transmissions in Figure 18 to see the correlation between the two energies. Figure 16 shows the mid-split 5 MHz to 85 MHz upstream burst test signals

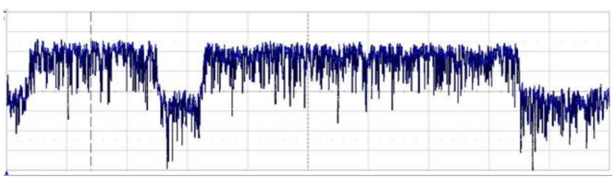


Figure 16. Full band capture for upstream signal with a vacant band.

with a vacant band 25 MHz to 35 MHz, where the energies were processed and analyzed. In situations where there is no nonlinear distortion occurring in the network, the signal in the vacant band is expected to be white Gaussian noise, and thus the energy captures in the vacant band from the multiple transmissions are expected to be uncorrelated. When nonlinear distortion does occur in the network, the signal in the vacant band is expected to contain distortion components from the transmitted signal; the energy levels of the distortion components in the vacant bands are directly related to the amount of distortion occurring in the network. In these situations, the energy captures in the vacant band from the multiple transmissions are expected to have a certain level of correlation.

By correlating the energy captures in the vacant band, detection of non-linear distortion is possible. Figure 17 shows the result from correlating (using frequency domain processing) the vacant band energy captures for two identical time-separated bursts sent

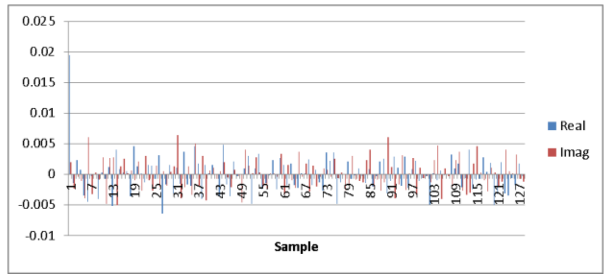


Figure 17. Correlation plot of vacant band energy from two identical bursts.

through a network suffering from nonlinear distortion. The presence of a large DC term in the correlation plot is indicative of the presence of nonlinear distortion.

As mentioned previously, the transmitted bursts are sent in sequence, and are either identical, or the subsequent bursts could be a variant of the original burst. One option would be to transmit a second identical burst and the third burst as an inverted version of the original burst as is shown in Figure 18.

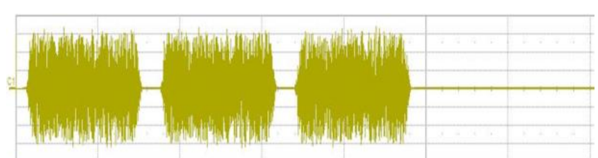


Figure 18. Multiple burst transmission in time domain: identical first and second bursts, identical but inverted third burst.

Transmission of the third inverted burst enables the identification of odd-ordered distortion in the network such as CTB.

To detect CTB (or odd-ordered distortion), the correlation results of the vacant band energy of the first burst and the second burst are compared to the correlation results of the vacant band energy of the first (or second) burst and the third burst. If the correlation results from the inverted burst go negative, it is indicative of odd-ordered distortion.

These test methods can potentially be incorporated into digital terminals and CMTSs, so that testing can be performed remotely without the need for a truck roll.
Patent pending.

SUMMARY

This paper discussed nonlinear distortions in cable networks and methods to discriminate nonlinear distortion from thermal noise in all-digital networks. One method is appropriate for downstream signals which are continuous transmissions, and another method is for upstream transmissions, which are generally bursty. With these new test methods, cable operators can potentially move to new levels of plant fault detection.

REFERENCES

Broadband Cable Access Networks, by David Large and James Farmer (© 2009, Morgan Kaufmann Publishers, ISBN 978-0-12-374401-2)

Code of Federal Regulations, Title 47, Part 76

Modern Cable Television Technology, 2nd Edition, by Walter Ciciora, James Farmer, David Large, and Michael Adams (© 2004, Morgan Kaufmann Publishers, ISBN 1-55860-828-1)

SCTE Measurement Recommended Practices for Cable Systems, Fourth Edition, (© 2012, Society of Cable Telecommunications Engineers, TS46)

ACKNOWLEDGMENTS*:*

The authors wish to express their gratitude for access and support to the cable operators that supported research for this paper.

Growth Architectures: Built to Last, Built to Launch

Dr. Robert L. Howald
Vice President of Network Architecture
Comcast

Abstract

Since the introduction of DOCSIS in the late 90's, cable operators have embarked on an aggressive phase of enhancing existing services and adding new ones. The triple play and the tipping point of mass HD soon followed DOCSIS, and bigger and better service offerings continue today with the addition of whole-home services, multi-screen video, vast on-demand libraries, Wi-Fi access points, and cloud DVR (cDVR). Ironically, while the transparent nature of better services is testimony to the tremendous flexibility of the HFC architecture, it can be perceived by customers as a lack of attention to investment in network upgrades – i.e. out-of-sight, out-of-mind – whereby other would-be service providers announce network installations often with extensive media fanfare. Of course, this perception of cable system evolution is far from the reality.

Referencing the launch of DOCSIS means that this renewed phase of investment has been going on for about 15 years. In many cases, upgrades are required simply to keep pace with uninterrupted traffic growth. While today's incremental upgrade approach has been effective, the trend of needing to deliver even more and at an accelerated pace makes this approach less cost effective and less practical going forward. In essence, the necessary pace of service evolution exceeds that of conventional network evolution. Instead, an even more aggressive response aligned with the pace of technology change and service demand is required – and it must take place with at least the same transparency to the end user achieved today.

Operators have cost effectively evolved HFC since its inception, relying on a proven, robust, and flexible architecture able to adroitly match architecture and technology to service evolution and are evaluating the avenues and timing for the next phase of network investment. A key objective going forward is to address evolution comprehensively, synergistically, and proactively in anticipation of next-generation service and customer expectations. Alignment across all impacted areas – Headend/Hub, access network, CPE, and cloud/software – will maximize ROI, create agility and service velocity, and optimize the customer experience. There are many simultaneous and interdependent parts to assess given the pace of consumer demand and technology change. In this paper, operator guidance reflecting key areas around services, technology, architecture and system engineering will be discussed. Operators have important bets to place to be prepared for the future, and will team with key industry partners to help drive the continuous improvement of the customer experience enabled by sound investment strategies. This paper will outline operator thinking around future network evolution paths, and offer insight to solution partners in order to fulfill this mission.

INTRODUCTION

Cable operators have seen downstream IP traffic sustain a Compound Annual Growth Rate (CAGR) of 40-50% for many consecutive years since introducing data services. Upstream has grown as well, although on a more irregular trajectory. These trends are a useful foundation to consider when evaluating long term capacity, investment, and drive strategic decisions on architecture and technology. Reasonable debates can take place over the long-term growth trends of media consumption [2,8,9], although to-be-seen future applications may emerge beyond media consumption. Placing strategic bets based on anticipating a slowdown in bandwidth growth is not prudent.

Regardless of how future long-term trends actually play out, the prior decade plus represents a "long term" of periodic investment to support service growth. This has particularly been the case since high-speed data (HSD) and HDTV became service cornerstones. Operators are now evaluating potential avenues and timing for the next phase of network investment. Operators have consistently kept ahead of the "need for speed," and in providing the aggregate network capacity to meet demand. Cable's uniquely nimble and cost-effective HFC architecture has delivered more and better services over an explosive 20 year stretch of media innovation that has included the rise to ubiquity of consumer Internet. It is now poised to deliver even more with strategic evolution and the integration of key architecture and technology components.

In this paper, we will quantify the service growth challenges, define the architecture and technology steps that deliver a cost-effective sustainable evolution playbook, and describe how the right architecture investments and strategy create a touch-once network migration path. We will show how the right choices install a long term capacity runway that supports the pace of service evolution and positions the network for a continuing bright future in this era of HFC evolution and beyond.

CAPACITY MANAGEMENT TIMELINE APPROACH

To quantify the challenge of service growth and capacity management, we can turn to a Capacity Management Timeline style analysis [5,8,9]. A sample analysis is shown in Figure 1. This tool, a one-page snapshot capturing all aspects of services, traffic growth, service group splitting, available spectrum, and technology implications, is invaluable to operator planning. The logarithmic scale enables the plot to cover the wide range of Gbps inherent in compounding long-term bandwidth, which can be otherwise be difficult to capture given a foundation of exponential growth.

Figure 1 shows various horizontal capacity boundaries for four different possible thresholds. Thresholds shown include allocations of 24 and 32 DOCSIS QAM carriers of 256-QAM, as well as 2 Gbps of capacity. The latter threshold could be, for example, 32 DOCSIS 256-QAM carriers joined by approximately 96 MHz of DOCSIS 3.1.

In general, cable operators tend to use their entire spectrum to maximize services for their customers. It takes proper planning and investment to increase the spectrum allocated to DOCSIS channels, and thus the awareness of when there will be a need to act, such as Figure 1 provides, is extremely valuable.

A fourth threshold is drawn as a reference to compare what may be available today for IP data (such as 24 DOCSIS QAMs) and what the network is capable of providing in terms of total Gbps. The example shown is the threshold labeled 750 MHz of QAM (for 750 MHz systems, obviously), in this case referencing only DOCSIS 3.0 capability, which is 256-QAM.

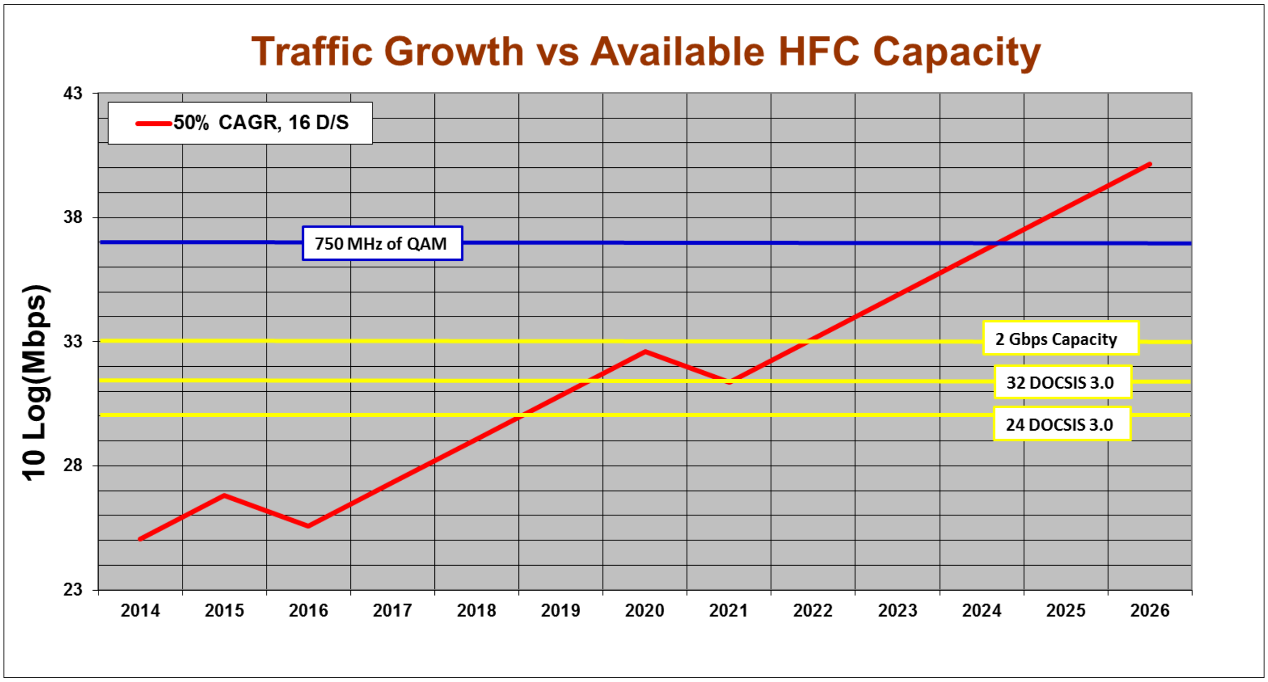


Figure 1 – A Capacity Management Timeline Guides Service and Architecture Evolution

Interpreting the rest of Figure 1, the growth of IP data (DOCSIS) is shown by the red trajectory trending upward with a slope that represents a 50% CAGR. It begins with 16 downstream channels deployed – the 2014 objective – utilized at inception at 50% at peak busy hour. The upward trajectory is broken twice along the way, representing service group splits. These splits (node segmentation or physical node splitting) represent a common, straightforward "business-as-usual" (BAU) approach to capacity management. In splitting service groups, operators manage capacity growth by sharing the existing bandwidth over a smaller number of subscribers, adding the narrowcast ports to support the new service groups, and in so doing (ideally) double the average bandwidth per subscriber.

IP video, or cloud TV (cTV), is assumed part of the engine that keeps the 50% CAGR going. There is not a separate "set-aside" of DOCSIS bandwidth for IP video. It's effect is embedded in the 50% CAGR itself the way over-the-top (OTT) video has been for many years. Various prior analysis consider an alternate approach of a steady 50% CAGR with IP video services introduced as an additive block of bandwidth [3]. In the long-term, such as over the time scale of Figure 1, the difference becomes very small as an offset of fixed channels is overwhelmed by the persistently aggressive compounding growth assumption itself. The assumption used thus only plays a role only in the details of managing near-term simulcast.

Figure 1 shows continued aggressive CAGR exceeding thresholds set by the shown DOCSIS channel allocations about 4.5-5.5 years down the road with one node split assumed. Using the two node splits shown provides 10 years of life on a 750 MHz plant of 256-QAM (DOCSIS 3.0) if all of the bandwidth were available for IP growth by that time. Not all of this bandwidth is available until the IP transition is 100% complete. In fact, this analysis approach is precisely valuable for its ability to assess the timing recommended for providing more DOCSIS channels over time, and ultimately the timing it suggests for completing the all-IP transition.

Basic conclusions from Figure 1 are that existing capacity and service growth for 750 MHz networks may present long term challenges under the assumption of persistently aggressive downstream CAGR. Also, current BAU evolution practices offer a solid runway of time to analyze trends and, as we are describing in this paper, develop a plan that overcomes the challenges and continues to deliver more and better services for the long term.

Upstream Lifespan Perspective

Similar to the calculations in the downstream, we can project upstream lifespan under various growth and service scenarios. CAGR for the upstream tends to be less predictable – spiking when Napster and YouTube were introduced for example, and lagging during other periods of time, including recent years.

With a varying CAGR range, the format shown in Figure 2 handily provides a sensitivity analysis around upstream CAGR. The chart is very straightforward to interpret – for the given scenario identified for each curve on the plot, the lifespan of the upstream is shown on the vertical axis against an average year-on-year (YoY) CAGR on the horizontal axis.

An underlying assumption for lifespan calculations is that four 64-QAM carriers (DOCSIS 3.0) of 5.12 Msps/6.4 MHz represent a full 5−42 MHz upstream. That is, there no DOCSIS carriers centered below 18 MHz and S-CDMA is not turned on. This represents a typical band usage assumption today.

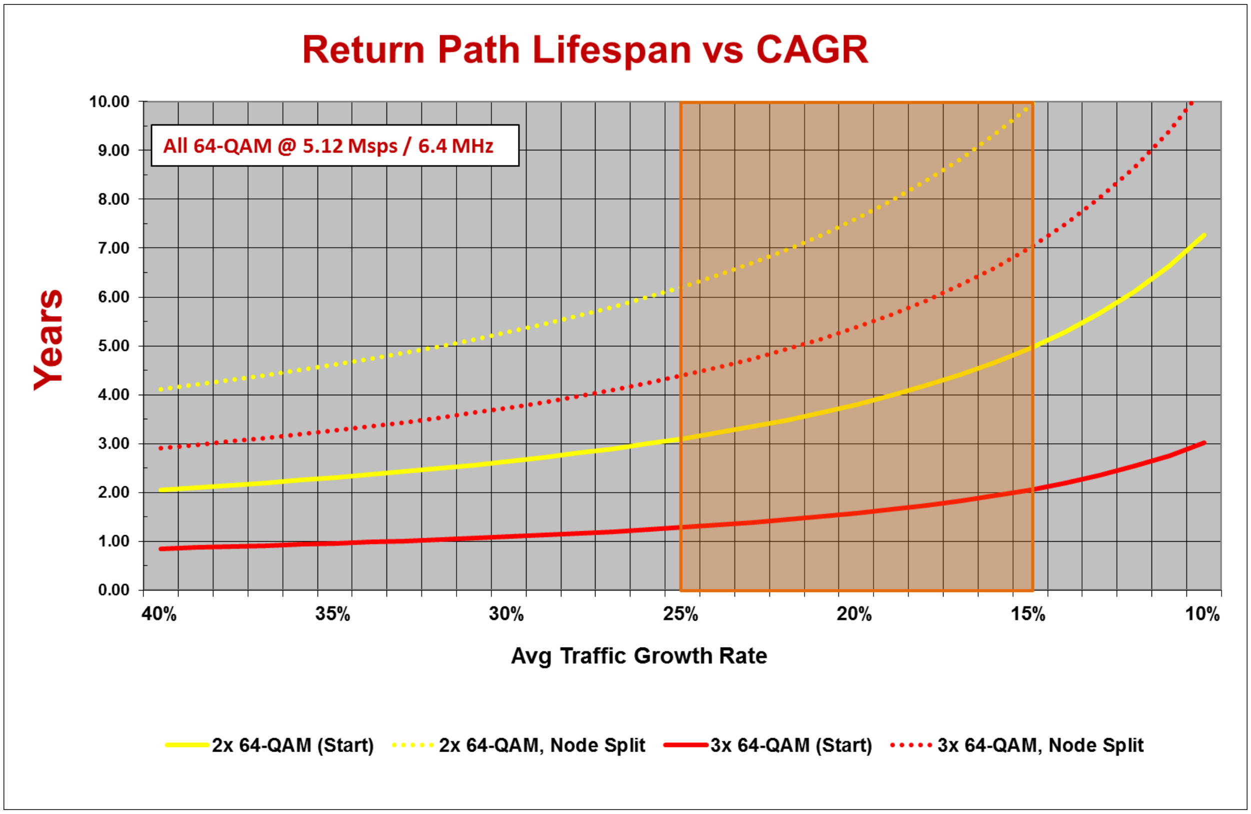


Figure 2 – Quantifying Upstream Lifespan

From Figure 2, we can conclude that where there are three 64-QAM carriers deployed today at full load, then for an upstream CAGR of 15-25%, there are 1-2 years of lifespan left before the upstream is fully utilized. As with the downstream, fresh runway is created by a BAU node split, which buys at least 3 years. Also, there is inherent margin in the years of lifespan where utilization is not 100%, which is typically the case.

For example, if the channel utilization is 20% below the threshold of congestion that would trigger a node split (i.e. it is 80% full), then a 25% CAGR means there is one extra year of margin to what the chart shows. In the 20% underutilization case with 25% CAGR and three 64-QAMs deployed, there are therefore about 2.5 years of life before another node split is required. Including this second node split there are about 5.5 years. Again, this runway provides for a comfortable time window for planning. The timing is right now to set this direction. Since areas of high traffic usage tend to be the ones with 3 or 4 upstream carriers, it is these cases that are of most interest for crafting strategy and projecting investment timing.

Note that perhaps as important as aggregate capacity of 5−42 MHz return is the limit placed on peak burst rate. In addition to delivering higher and higher upload speeds to customers, this is important for most efficiently enabling Gbps downstream speeds.

CAPACITY OPTIMIZATION LEVERS

Theoretical capacity is straightforward and based on two variables – bandwidth (B) and Signal-to-Noise Ratio (SNR). The well-known Shannon Capacity limit is the maximum error-free rate that can be obtained in an additive white Gaussian noise (AWGN) channel, and given by

$$C = [B] \, \mathrm{Log}_2 \, [1+SNR] \qquad (1)$$

This is further simplified in cable, in particular for the downstream, by using high SNR assumptions typical of cable networks. In high SNR conditions, capacity is closely proportional directly to bandwidth, B, and SNR *expressed in decibels* (dB):

$$C \approx [B] \, [SNR \, (dB)] \, / \, 3 \qquad (2)$$

The simple message of (2) is that increasing the available capacity involves increasing spectrum, increasing SNR, or both. Architecture evolution should thus aim for these goals.

Let’s first consider SNR.

QAMazing

Improving link SNR enables more bandwidth efficient modulation formats. Forward Error Correction (FEC) has an important role in new capacity in how it relates to SNR, and therefore capacity. While nothing about FEC shows up in (2), better FEC enables a given M-QAM format to operate at a lower SNR. Simply put, the best FEC makes for the most efficient use of the “SNR” in (1) and (2).

As an alternative to a given M-QAM operating at a lower SNR, we can also say that for a given SNR, better FEC enables M-QAM formats with higher bandwidth efficiency. DOCSIS 3.1 takes advantage of this, and is therefore a critical technology component to the evolution path forward. Regardless of improvements forthcoming in end-of-line (EOL) SNR, DOCSIS 3.0 (and below) limits the downstream to 256-QAM, or 8 bits/symbol. However, HFC commonly delivers performance higher than what is needed to support 256-QAM. What is wasted system margin today is taken advantage of in DOCSIS 3.1 by allowing higher order M-QAM profiles. An architecture evolution path should take advantage of these DOCSIS 3.1 innovations by delivering better SNR to most efficiently use spectrum.

Figure 3 shows the higher modulation formats associated with DOCSIS 3.1 including two of the key DOCSIS 3.1 additions – 1024-QAM and 4096-QAM. Each is shown for an *equivalent uncoded BER* of ~1e-8. The 46 dB SNR for uncoded 4096-QAM highlights both the expectation of an architecture with improved network performance, as well as the power of the modern FEC technology being adopted.

Compared to 256-QAM, the bandwidth efficiency increase of 1024-QAM, 2048-QAM, and 4096-QAM is 25%, 37.5%, and 50%, respectively.

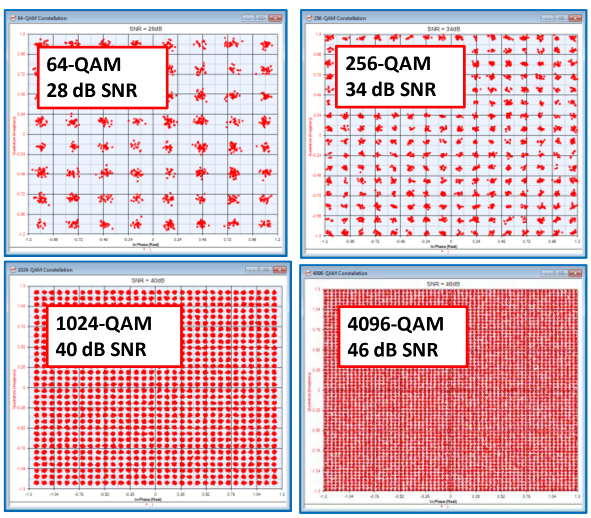


Figure 3 – DOCSIS 3.1 M-QAM Format Comparison

<u>Shannon Nuances</u>

Not obvious in the capacity equations in (1) and (2) is the assumption of a fixed SNR over the bandwidth, B. In real systems, variations in SNR are likely across the band and across an HFC physical footprint. DOCSIS 3.1 also aims to take advantage of this by providing modulation format flexibility, as well as having modulation formats targeted to the performance of CMs on a "binned" basis – a tool known as Multiple Modulation Formats (MMP). Figure 4 illustrates this concept [3].

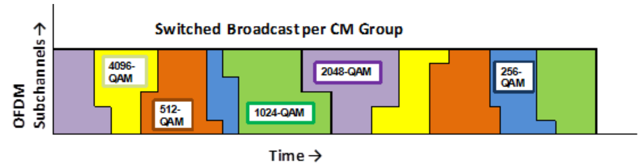


Figure 4 – DOCSIS 3.1 Includes Use of Multiple Modulation Profiles (MMP)

Fielded cable modems show a wide enough range of received SNR that to enforce the same M-QAM format on every user would be to deliver lowest common denominator performance, leaving many Mbps on the table. Figure 5 shows the reported SNR of a large sample of CMs [11]. It shows the majority of CMs are spread across three QAM profiles, while at least five profiles are observable when considering the non-zero percentage of CMs at the top and bottom end of the distribution. Figure 5 is a clear justification for the use of MMP to optimize capacity while preserving the basic simplified broadcast structure of the downstream.

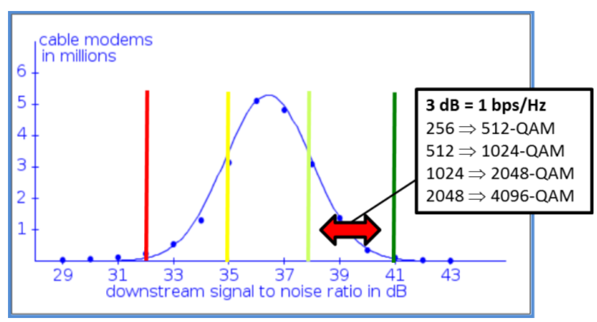


Figure 5 – SNRs Reported by Fielded CMs

Role of OFDM

To deal with non-constant SNR across the band as it relates to the "B" in the capacity equations, DOCSIS 3.1 introduces to cable the use of Orthogonal Frequency Division Multiplexing (OFDM) as the waveform on the wire. OFDM is notable for its ability to maximize capacity on a channel, and in particular on unknown or complex channels.

Why is this important to cable? An example of spectrum variability is the new bandwidth above 1 GHz, support for which is included in DOCSIS 3.1. Figure 6 [1], shows insertion loss characteristics of various Tap models above 1 GHz for "1 GHz" specified taps. A cascade of Taps will have wider variation.

Waveforms such as OFDM, with its narrow subcarriers that can be fitted to the channel characteristics, allow maximum capacity extraction from such channel performance variations. The DOCSIS 3.1 OFDM parameters (subcarrier spacing, cyclic prefix values, FFT size) were chosen based on HFC channel characteristics developed as part of a renewed channel modeling exercise [12, 13]. Both OFDM and MMP are tools selected for DOCSIS 3.1 to deal with the practical aspects of maximizing capacity in ways that are not accounted for in the simple capacity calculation of (2).

Summarizing, per capacity equation (2), higher SNR is an important component of increased HFC capacity. DOCSIS 3.1 has a primary objective of optimizing capacity by providing bandwidth efficient modulation tuned to the available SNR through:

1) Higher QAM formats
2) Better FEC
3) Optimal waveform (OFDM)

Note that while more capacity is indeed available with a higher SNR, it is with logarithmic proportionality. For example, either 50% more spectrum (B), or 50% more SNR (in dB) yield 50% more capacity.

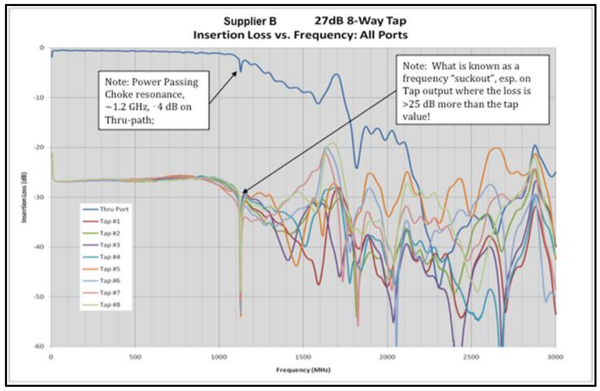


Figure 6 – Varying Tap Responses Above 1 GHz [1]

However, turning a 40 dB network SNR into a 60 dB SNR is a very tall if not impossible order, practically speaking. The implication of this is that, while deploying architectures with higher SNR is a significant objective, ensuring more spectrum offers more bang for the buck, and should be high priority.

We've identified both more SNR and more bandwidth as clear avenues that a network architecture should pursue to deliver new capacity, and DOCSIS 3.1 as the technology tool to take advantage of the higher SNR. We now discuss an evolution path that delivers on these objectives.

FIBER DEEPEST

We've emphasized the important capacity levers to target for architecture evolution. "More" capacity is of course better, but it is also critical that the investment lead to a quantifiably long-term capacity runway such as we can project via analysis such as Figures 1 and 2, as opposed to temporary gains that lead to repeated investment. As always, efficiency of capital expense is important.

For "brownfield" HFC, a logical conclusion that aligns with each of the objectives emphasized is to drive fiber directly to the coaxial last mile for carriage into the home. This is referred to as a "Fiber Deep" (FD) architecture. The foundation of the FD approach is an N+0 network architecture.

We can identify some of the fundamental objectives of the FD approach:

1) Leverage the high capacity coaxial last mile for cost effective growth
2) Provide a one-time touch sufficient for long-term bandwidth needs
3) Improve EOL performance through the elimination of the RF amplifier cascade
4) Deliver all current and emerging new services including Gbps speeds, 4kHD, Cloud TV (cTV), near-ubiquitous HSI via WiFi hotspot access, and is enabling of FTTP and DOCSIS-based business services
5) Significantly reduced operations and maintenance costs and enhanced customer QoE
6) Serve as the launch architecture for FTTP alternatives for long-term migration as needed

Executing a Fiber Deep network evolution is the sought after one final outside plant (OSP) touch operation, providing long term sustainable bandwidth, while delivering operational efficiencies and savings for as long as we can reasonably expect to project "long term" for a business built on technology.

Fiber Deep System Benefits

Figure 7 shows phased BAU service group segmentation via deeper fiber penetration as it might take place over an actual serving area footprint. It illustrates how a natural architectural end-state of HFC in a BAU sense would culminate in an N+0 system with an all-passive coaxial last mile through repeated node splits. In Figure 7, a single node serving area (top left) is ultimately broken into serving group sizes, which may be as small as (approximately) 50 hhp. Note that physical node boundaries of the map do not capture virtual segmentation of the nodes, of which configurations up to 4x4 are typical.

A fiber deep strategy instead bypasses the incremental phases and migrates directly to Fiber Deep (N+0), understanding that the persistency of CAGR inevitably drives the network to this optimal end state. Of course, a FD evolution, just like today's capacity upgrades, is incrementally introduced market -by-market based on service demand. The history of capacity upgrades suggests a rippling through the footprint over the course of many years.

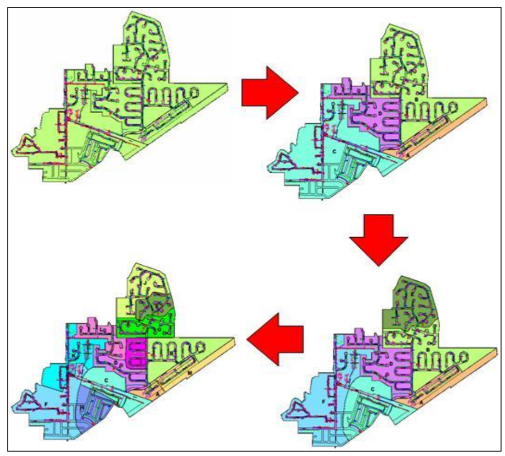


Figure 7 – Segmentation Phases of BAU Fiber Deeper Migration

The approach delivers the sought-after components of more capacity per equations (1) and (2):

1) Improved EOL SNR
2) Enabling of New Spectrum, B

What does N+0 mean to EOL performance? The performance of HFC networks in the downstream is well understood from decades of achieving fidelity acceptable for analog video. RF cascade reduction improves EOL SNR due to the reduced impact of the noise accumulation associated with active components. Physical node splitting shrinks average cascade depths. Note that service group splitting may still occur "virtually" though a segmentable node, but in this case there is no effect on the amplifier cascade depth.

On the upstream, SNR performance tends to be dominated by the optical link, with minor contributions from aggregated return path RF amplifier noise. However, the shrinking service group size has the significant effect of decreasing external interference funneling.

Quantifying the downstream performance, we can determine the effect when all RF degradations beyond the node are removed. The increased EOL SNR means an increased likelihood of supporting higher modulation formats.

Downstream performance is also largely set by the AM optical link. However, the subsequent amplifiers each noticeably nick away at SNR performance down the cascade. Using M-QAM SNR requirements, we can derive what QAM format can be supported over a range of HFC and home architecture variables. This is shown in Figure 8 [3,7].

Shown on the x-axis is expected EOL Composite Carrier-to-Noise using typical link length 1310 nm AM optics of 42 dB (N+6), 44 dB (N+3) and 47 dB (N+0), each labeled via the pink vertical lines. CCN is effectively the same as SNR, just comprised of AWGN *and* digital noise-like distortions), Note CCN is shown from *left to right* as highest to lowest.

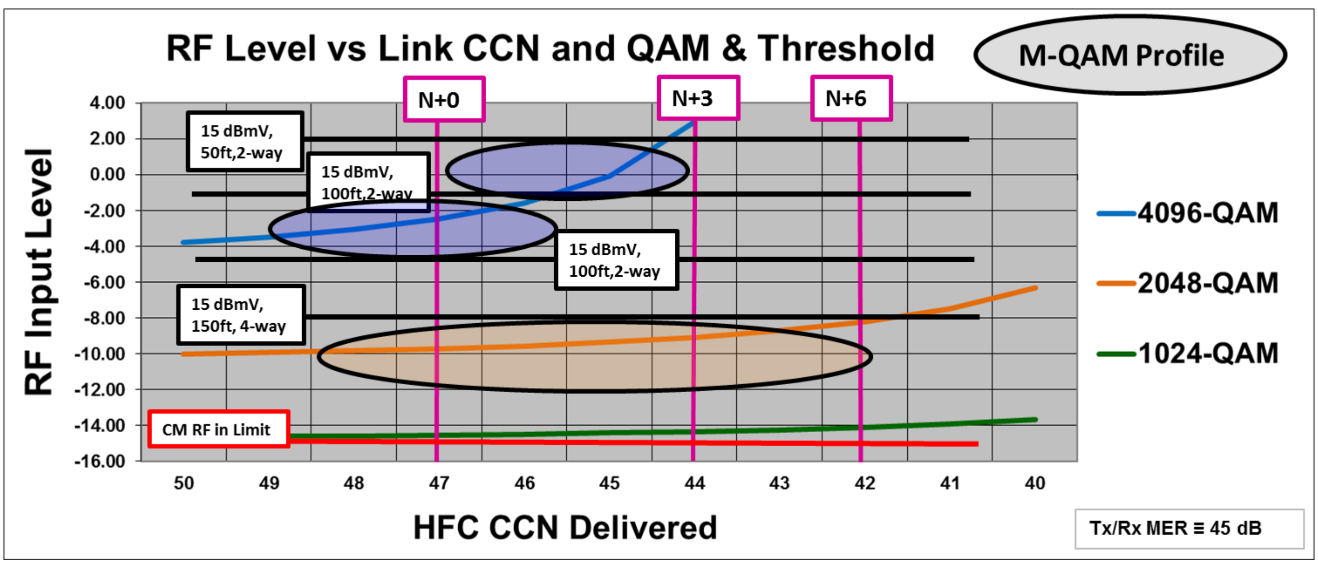


Figure 8 – HFC Cascade and Home Architectures vs. M-QAM Supported

In general, we can observe from Figure 8 that the higher the order of M-QAM format, the higher on the chart the base of its curve is and the further to the left that its upward trajectory begins. The upward trend indicates the CCN point below which it begins losing the ability to support the format. In other words, unsurprisingly, the higher CCNs of shorter cascades more readily support the most bandwidth efficient M-QAM formats, and we can predict under what conditions.

Specifically focusing on Fiber Deep, we can zero in on the "N+0" vertical line in estimating QAM expectations. We note that, for example, that a Fiber Deep (N+0) system comfortably supports 4096-QAM over a range of input levels, while for N+3, the CM must see about 2 dBmV or above at its input, and for N+6 we cannot support 4096-QAM at all for the range of scenarios shown.

Additionally, for N+6, the input level to the CM must exceed -8 dBmV in order to use 2048-QAM. Otherwise, only 1024-QAM can be supported. This chart and variants of it [3,7] quantify the N+0 advantages in EOL SNR performance in terms of bandwidth efficiency.

Use of Digital Optics

While Figure 8 is encouraging with respect to the ability to extract more bandwidth in Fiber Deep systems, another potential avenue for future Fiber Deep migrations is to employ a digital optical solution in place of classical AM optics.

The example of Figure 8 assumes classic 1310 nm point-to-point optics. Meanwhile, digital narrowcast service growth, continued node splitting, and site consolidation has driven operators to use of WDM technologies to maximize fiber usage. The trade-off parameter space for AM optics includes link length, wavelength band, number of wavelengths, and MER performance. These are carefully balanced for optimal deployment, creating architecture or performance constraints.

A digital optical link largely eliminates optical link length issues, increases the number of wavelength per fiber, and delivers a fixed (DAC limited) CCN performance at the node output. The node output SNR performance matches the performance of DRFI-compliant equipment installed in Headends, such as the output port of a CMTS or EQAM. Therefore, when digital optics is combined with the removal of RF amplifiers, nearly all the EOL performance variations are removed and link performance becomes extremely consistent and predictable. As such, it will enable consistent use of the most bandwidth-efficient DOCSIS 3.1 modulation profiles.

There are multiple solution types that offer a digital-to-the-node solution, and the industry is still vetting these options. In general, the available options are less about technology, and more about overall architecture, interfaces, and long-term network strategy. Fiber Deep merely creates the unique opportunity to define and implement a modular node solution enabling of a digital optical approach. Furthermore, the transition to digital optics is a natural and necessary step to a long-term plan that enables FTTP.

Downstream Spectrum: 1.2 GHz

In addition to the improvement in system performance as amplifiers are removed, the passive last mile provides a unique opportunity for considering changes to spectrum allocations, which are now much more efficiently available than with a cascade of RF amplifiers. In the downstream, with the defined extension in the DOCSIS 3.1 standard to 1.218 GHz and optionally to 1.7 GHz, Fiber Deep creates that unique opportunity to implement such an extension. (Note we will use “1.2 GHz” throughout the rest of this paper to represent 1.218 GHz).

One of the significant instantaneous advantages of Fiber Deep is the immediate access to 1 GHz of available bandwidth across most areas using current products. With plants more likely to be limited by RF actives than Taps, the bypassing of the amplifier cascade for an optimally placed N+0 node frees up approximately 130-250 MHz of heretofore unused spectrum in 750 MHz and 870 MHz systems, because a standard new node product today will be at least 1 GHz. As we shall see, this new bandwidth is offset somewhat by an extension to 85 MHz that will take place in the upstream, but this significant new bandwidth adds valuable shelf space for capacity growth.

While the bandwidth added by 1 GHz is quite powerful, there are nonetheless several key reasons to take advantage of the DOCSIS 3.1 extension of bandwidth beyond defined to 1.2 GHz. Among these are:

- Guarantees that bandwidth is set aside for DOCSIS 3.1 regardless of services added below 1 GHz
- Ensures a full 192 MHz block (and Gbps services) of DOCSIS 3.1 can be turned up in fully utilized 870 MHz systems
- Enables the potential for 10 Gbps of total downstream capacity via DOCSIS 3.1
- Offers more bandwidth flexibility for known service additions:
    - Time flexibility to complete all-IP transition
    - Addition of 4kHD video
    - Move to Cloud-based DVR (cDVR)
    - Implement Multi-Gbps service rates
    - Growing footprint of WiFi APs
- OSP equipment housings are already bandwidth-capable and the necessary RF actives available in CY14.
- Fiber deep (N+0) offers the uniquely efficient opportunity to make spectrum adjustments
- Taps that are in place are suitable as is in some cases, or need a faceplate change only
- 1.2 GHz keeps MoCA™ available (through channels defined from 1300-1675 MHz) as an HLAN option

MoCA™ has been widely deployed as Home LAN (HLAN) technology in recent years. In using 1.2 GHz for downstream services, most of the defined MoCA™ band is not overlapped. However, this would not be the case extending to the 1.7 GHz DOCSIS 3.1 option.

In addition to MoCA™ management, other potential issues were evaluated as part of the assessment for moving to 1.2 GHz. Two of the more important ones were powering and SNR loss.

1) Power requirements of the RF actives will increase roughly proportionally to the RF loading increase. However, in the larger picture of a power-saving Fiber Deep architecture, the effect is expected to be a wash, if not a net gain (less power) overall
2) There is slight SNR loss due to RF loading (using AM optics), but again a net gain in EOL performance from Fiber Deep is expected overall.
3) There are unknowns of optical nonlinearity performance at the extended bandwidth. However, the risk of major issues is considered low.

Note that we have seen in Figure 5 that bandwidth to 1.2 GHz in Taps can vary. Nonetheless, some families of deployed Taps are actually quite well-behaved above 1 GHz, and virtually all major remaining Tap vendors have developed technology that enables the existing fielded Taps to be expanded to up to 1.7 GHz with a simple change of the faceplate. Others have built such faceplates for a different vendor's Taps.

Example Tap response showing "excess bandwidth" on a 1 GHz Tap, and the faceplate-style bandwidth extension for it to extend to 1.7 GHz, are shown in Figures 9a and 9b.

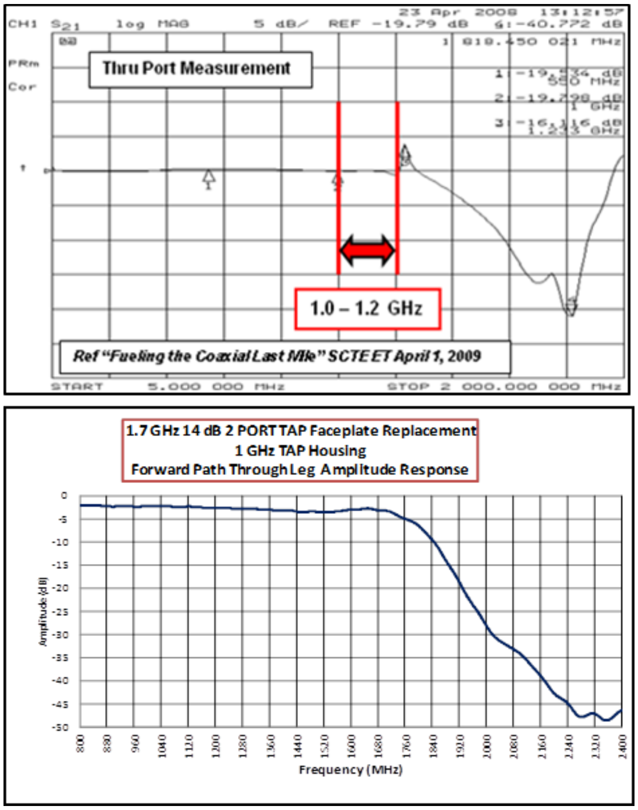


Figure 9 – a) (Top) Some 1 GHz Tap Models Have Quality "Excess Bandwidth" b) (Bottom) Tap Bandwidth Can Sometimes Be Easily Expanded with new Faceplates

<u>No Free Launch</u>

Now let's take a closer look at 1.2 GHz actives. For an N+0 architecture, of course, this means only the node. For the optics, because the transmitter loading is flat, the SNR cost is only about 0.6 dB relative to 1 GHz loading. Again, however, the small delta is more than offset by the SNR gains associated with the elimination of amplifiers altogether.

More significantly, however, the RF spectrum launched onto the coax is up-tilted. Thus, adding RF loading at the top of the spectrum disproportionately increases total power. Figures 10 demonstrates this effect.

In Figure 10, the extra 200 MHz to 1.2 GHz slightly more than doubles – 4 dB – the total RF power load using a particular higher N+0 tilt design. Fortunately, modern Gallium Nitride (GaN) RF technology is trending towards this higher total output power at equivalent distortion performance. It is expected that this technology will enable 1.2 GHz of bandwidth at, or very close to, equivalent 1 GHz performance and levels, including extending the tilt line all the way to 1.2 GHz. Or, conversely, these devices will deliver identical performance at or very close to the full 1.2 GHz bandwidth.

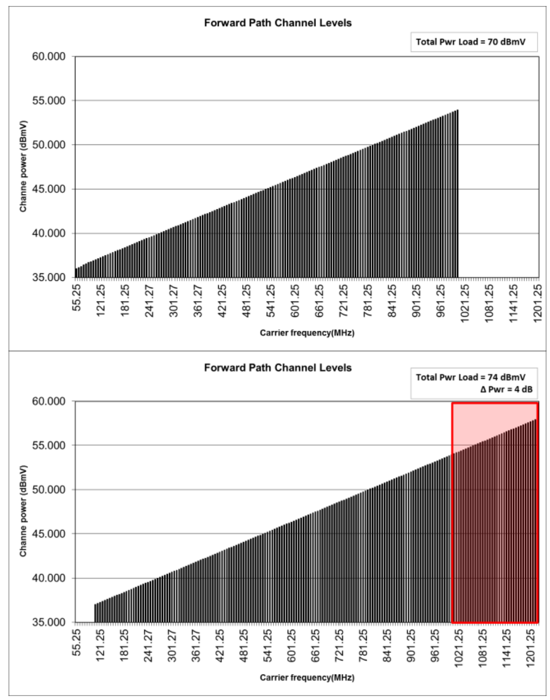


Figure 10 – Tilted RF Outputs Increase Total Power Disproportionately with Increasing Frequency

Now consider Figure 11, where the bandwidth is extended to 1.7 GHz. In this case, loading the spectrum at the identical tilt and equivalent relative PSD requires 9 dB more RF power, *nearly an 8x increase*. The impact on powering of nodes and plant powering overall to drive this additional RF power is substantial. It is not practical within the constraints of HFC as we know it, or can imagine reasonably evolving it cost-effectively in a useful timeframe.

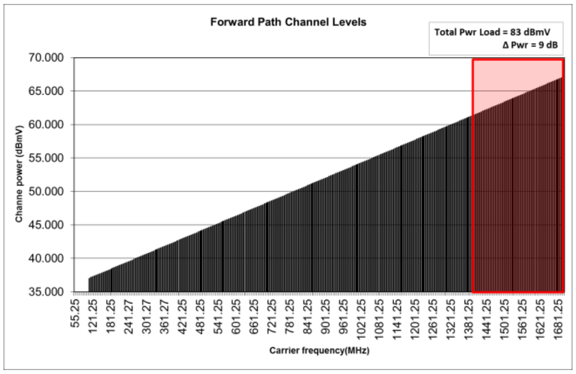


Figure 11 – Downstream Bandwidth Extension to 1.7 GHz Creates an RF (and AC) Power Dilemma

Other considerations that finalize the case against a 1.7 GHz extension are:

1) Optical loading SNR loss for the 1.7 GHz case would be significant (~2.5 dB) should it be carried over AM optics. Other optical nonlinear effects are unknown for what is now almost a full new octave. Only an overlay (parallel) architecture would make sense.
2) The OSP housings and RF interfaces themselves do not easily extend to 1.7 GHz, requiring substantial redesign of node platforms and possibly of fundamental materials used such as circuit board dielectric.
3) When extending to 1.7 GHz, the entire MoCA™ band is overlapped. A point-of-entry demarcation architecture is a requirement in this case. By restricting the downstream to 1.2 GHz, most of the MoCA™ band remains accessible without this requirement.
4) Sufficiency of 1.2 GHz plus FD serving group size reduction to support aggressive long term CAGR.

In summary, the addition of 1.2 GHz of bandwidth is a manageable extension using current techniques for optical loading, RF distribution, and powering, including tilt line extensions of tilts used today. It supports projected bandwidth needs for the long term, and technology availability is just around the corner. This is not the case for enabling up to 1.7 GHz. The recommended evolution path is therefore 1.2 GHz.

Upstream Spectrum: 85 MHz

A key component to the long-term enabling of HFC is more upstream capacity. The capacity bounds of 37 MHz, and considering that the bottom spectral portion can be less capable, drives the need to pursue a wider spectrum allocation to manage continued, albeit slower, traffic growth. Multiple upstream carriers are deployed in most markets today, and each new carrier consumes more of the total available bandwidth. Thus, looking ahead to project the timing of introducing more spectrum, while factoring in what DOCSIS 3.1 has to offer to 5-42 MHz, is important. The actual implementation of spectral re-allocation requires planning and coordination.

DOCSIS 3.0 already defines the 85 MHz mid-split for a wider upstream spectrum. Fortunately, technology supporting this band is already mature. In Figure 12, typical performances of an upstream DFB at nominal link length and Digital Return over the 85 MHz mid-split bandwidth are shown [3]. M-QAM performance thresholds, with typical margin associated with ensuring robustness in the less predictable upstream, is built-in using DOCSIS 3.1 QAM profiles.

Digital returns have a distinct advantage of not degrading NPR as a function of RF loading, and not being sensitive to the link length for performance or receiver output level setting. Performance is completely determined by the effective number of bits (ENOB) of the A/D converters.

Figure 12 shows that with new DOCSIS 3.1 FEC, the link is 1024-QAM capable using high performance DFB optical links and digital return systems available today over the full 85 MHz and with the dynamic range of NPR typically required. In fact, the NPR performance looks capable of supporting 2048-QAM in theory (3 dB higher than the 1024-QAM threshold). However, in practice, the margin that exists between the estimated 2048-QAM threshold and the NPR curve itself is less than what is typically seen as robust in the upstream.

Nonetheless, Figure 12 indicates why all of the M-QAM formats in Figure 2 are worth considering for the upstream as well as the downstream. Technology and FD architecture variables are falling into place to make these possible in the plant, shifting the performance burden to the complex task of high fidelity burst transmitters and high sensitivity burst receivers.

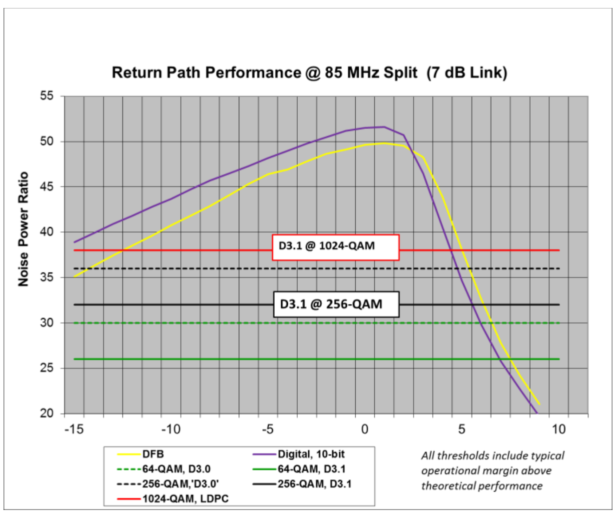


Figure 12 – Mature 85 MHz Technology Supports DOCSIS 3.1 QAM Profiles

New Capacity = New Spectrum

The highest spectral efficiency gain to expect in the downstream is 50% (256-QAM @ 8 bps/Hz to 4096-QAM @ 12 bps/Hz). In the upstream, the increase is now up to 67% using 1024-QAM. These are the "SNR" related capacity gain components in (1) and (2).

We now turn to the "Bandwidth" component of new capacity in (1) and (2). In the upstream, the relatively large (>2x) bandwidth extension to 85 MHz is particularly attractive. In the downstream, extending to 1.2 GHz is extending a bit into the unknown and a more modest addition to the current maximum forward band definition of 1 GHz. By contrast, in the upstream we are instead extending a partially troubled channel into an area where it will typically be better behaved. The upstream generally becomes increasingly cleaner with frequency above about 15 MHz in North America. And, as we have shown, the 85 MHz mid-split is available in current DOCSIS 3.0 and HFC technology.

Perhaps most importantly, as we shall quantify, an 85 MHz upstream combined with FD segmentation offers a long lifespan window. In addition, the 85 MHz mid-split offers the opportunity for Nx100 Mbps service rates, whereas 100 Mbps can be a challenge in DOCSIS 3.0 systems with 5-42 MHz of spectrum.

Similar to the downstream spectrum extension recommendation, the upstream "optional" edge defined in DOCSIS 3.1 to 212 MHz is *not* the recommended path, doe several reasons:

1) Unacceptable SNR degradation using DFB return optics, due to increased RF loading sharing a fixed total power, incurring about a 7 dB penalty over 5-42 MHz
2) Costly digital return A/D conversion and optics for high-fidelity sampling of very wideband spectrum
3) Significant implications to return path set-up, alignment, and technology implementation due to the increase in frequency dependent cable losses
4) Perhaps most importantly, removes significant downstream spectrum, where the higher traffic growth exists

The recommended return path architecture is therefore an 85 MHz mid-split.

KEY ARCHITECTURE PRINCIPLES

The major defining principles of HFC architecture migration are therefore:

- Implement Fiber Deep (N+0) for better SNR, efficient access to spectrum reallocation, reduction of service groups, and savings in maintenance, MTBF, and Opex.
- Target the N+0 design to be service group sized for long-term bandwidth runway as well as alignment with future potential architectures.
- Take advantage of cost-effectively accessible additional spectrum for the efficient enabling of new capacity.
- Use all spectrum most efficiently by taking advantage of key advances in M-QAM technology, Forward Error Correction techniques, and optimal waveform design (i.e. DOCSIS 3.1).
- Architect the FD investment – the "last" fiber node – and define the OSP equipment requirements as a template that positions the network for FTTP extensions if or when a fiber last mile is required on either a targeted or large scale basis.
- Consider the FD migration as an efficient path to transition the network to digital optics.

Applying the above tools – Fiber Deep, 85 MHz Mid-Split upstream, expansion to 1.2 GHz downstream, DOCSIS 3.1, a modular Fiber Deep node template, plus a migration plan to all-IP – creates a one-touch sustainable HFC architecture for many, many years to come. Key architecture principles above are visualized in Figure 13.

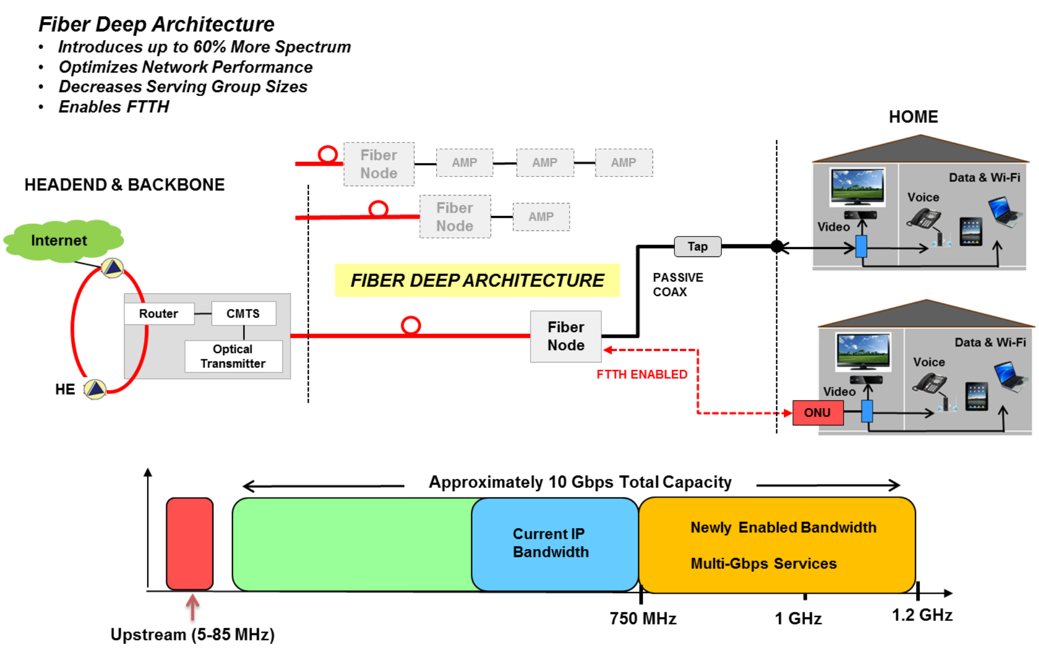


Figure 13 – Fiber Deep: Architecture Principles and Spectrum Allocations

A FINAL STEP

Using the above principles, an updated version of the downstream Capacity Management Timeline first demonstrated in Figure 1 is shown in Figure 14. It quantifies the fundamental case for one-touch long-term network lifespan based on a Fiber Deep architecture foundation. As the figure shows, the one-step FD migration path combined with the technology refresh components mentioned herein and designed to deliver long term capacity, does just that. As always, trying to predict service evolution, technology breakthroughs, and network evolution options ten years down the road, or even five, is difficult if not impossible.

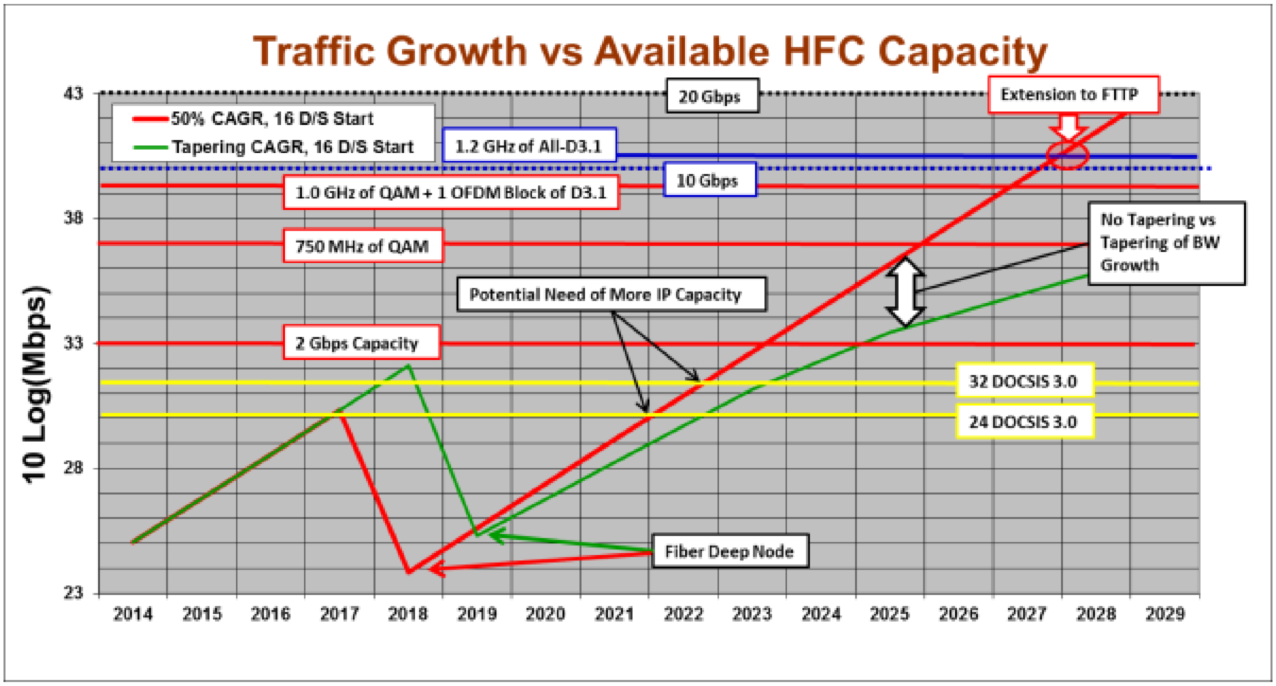


Figure 14 – Long-Term Downstream Capacity Enabled by the Fiber Deep Architecture

Looking at Figure 14, there are two trajectories – a constant 50% CAGR (red) with a Fiber Deep step taking place in 2017, and a trajectory with a built-in tapering of bandwidth growth with time (green). This is not to suggest the expectation of tapering, although this argument can be made [9]. More so, it is to show the power of compounding – or in this case, the power of *not* compounding as aggressively over time. It can result in what is essentially a "forever" network.

Note that the long runway of 15 years, even accounting for steady, aggressive CAGR, is comforting in that there is a solid window of time to observe trends and plan or re-plan accordingly. Even with this perspective of growth runway, it is important to point out that introducing Fiber Deep or FTTP throughout the footprint would inherently follow the common approach of targeting investments as driven by local service demand, as is done with any capacity upgrades, and thus the spreading out the capital investment over many years,

Examining Figure 14 further, the thresholds of 24 and 32 slots of DOCSIS 3.0 are again shown. The IP bandwidth growth trajectories reveal when these thresholds would be breached with continued bandwidth growth. As discussed, operators generally utilize their entire spectrum to provide the fullest set of services for customers. Various tools are used to manage the evolution of services, technology, and spectrum, such as for adding IP bandwidth to optimize the service mix to customers. Adding DOCSIS channels and splitting service groups are BAU operations used whenever the need arises.

The 24 and 32-channel thresholds shown in the figure show the modest impact in terms of years that incremental increases in IP bandwidth alone provide in the face of exponential growth. Without Fiber Deep or a node split, these thresholds are exceeded after 3-4 of years.

New thresholds have been added to Figure 1 that demonstrate the benefits of a Fiber Deep strategy and DOCSIS 3.1. As described previously, ensuring the full bandwidth efficiency benefits of DOCSIS 3.1 may also include a digital optical transition as part of the migration plan for Fiber Deep, and several solutions are under consideration for this path forward. Observing the persistently aggressive 50% CAGR IP growth curve and the implementation of the FD and technology strategy, it is not until 2028 (red circle on Figure 14) until there is a potential need to consider more capacity – 14 years away.

This conclusion is obviously important, but the intervening thresholds along the way inform us for setting the guidelines around the pace of the all-IP transition.

Summarizing, we note the following as key components that together achieve long term, one-touch network sustainability:

- Fiber Deep migration
- Downstream BW to 1.2 GHz
- DOCSIS 3.1
- Transition (implicit) to all-IP

Forever CAGR?

Perhaps the more intriguing long term case is if there is in fact a "Tapering CAGR." As previously described, it would not be prudent to base a strategy on the expectation or requirement that consumer bandwidth growth will slow or end. History does not support this as a logical assumption. However, it would be negligent not to evaluate the possibility and recognize the implications.

The tapering trajectory is based on the understanding that streaming video has been the recent driver of persistently aggressive CAGR. Furthermore, a quantifiable maximum video bit rate and concurrency of subscribers at peak-busy-hour can be calculated. The tapering example here also assumes 4kHD plays a role in driving continued aggressive CAGR, but that formats beyond this are not significant contributors in scale for a number of practical reasons [2]. The conclusion of these assumptions suggest that a long term HFC capacity of 10 Gbps to a Fiber Deep sized serving group may be all that is necessary – *ever* – to satisfy consumer broadband as it relates to media consumption-centric applications.

As always, yet-to-be uncovered non-media applications may replace video as the CAGR growth engine. This is a key reason why keen attention to service growth trends over the next decade or so is critically important to determining future architectural evolution and timing required.

Every strategic plan, of course, is essentially a living document. Coarse corrections are constantly being made. However the actual growth of consumer bandwidth plays out, the lifespan enabled using the architecture principles described here provides a very comfortable window of time to assess the trends in services and technology, and develop appropriate responses. Based on our understanding of the key variables and driving forces, the strategies outlined represent the best plan of attack today.

Upstream

Figure 15 updates Figure 2 with the lifespan extension implications of the 85 MHz mid-split, Fiber Deep migration, and the eventual implementation of DOCSIS 3.1, which increase spectral efficiency up to 67% (64-QAM to 1024-QAM).

With a standard node split on a 42 MHz architecture, the 25% CAGR case managed about four years with a 3x 64-QAM (red dashed) starting point – again ignoring margin offset due to any potential peak busy hour underutilization on Day 1. With an 85 MHz mid-split, the lifespan is extended to last over 8 years (blue dashed).

Implementing Fiber Deep when combined with the 85 MHz mid-split architecture extends the lifespan to over 12 years. The 12-year range now closely synchronizes upstream lifespan to the downstream. It is this long-term capacity, the compatibility with legacy STB out-of-band signal frequency range (130 MHz maximum), the support of Nx100 Mbps services, and the opportunity that FD migration efficiently presents, that makes 85 MHz the right approach.

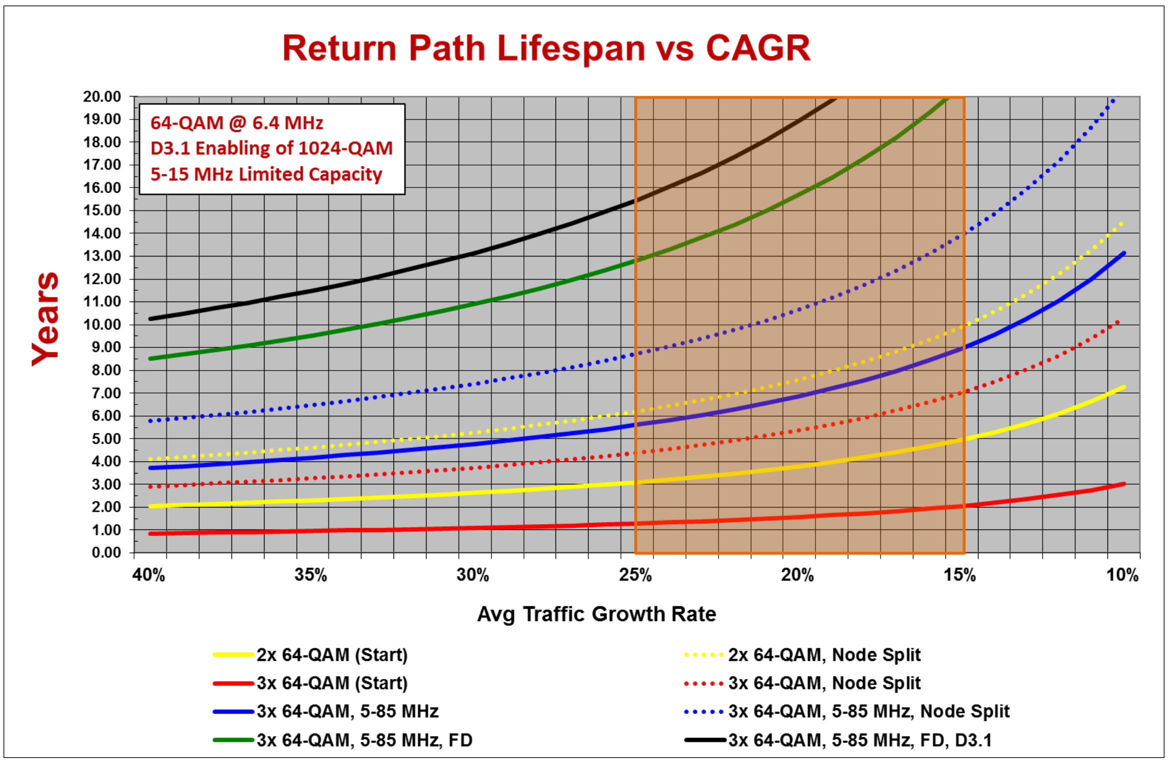


Figure 15 – Long-Term Upstream Capacity Enabled by the Fiber Deep Architecture

We now factor in that DOCSIS 3.1 is around the corner, so even more capacity potential can be baked into the calculation. This extends (black curve) the upstream expiration date out 15 years from an aggregate capacity perspective. These projections quantify and confirm why an 85 MHz upstream is the recommended spectrum option for Fiber Deep migration.

Lastly, one more aspect is worth looking at that further solidifies the upstream recommendation. As with the downstream, it will be important to keep a watchful eye on nominal CAGRs. We've discussed the upstream growth mostly around the high end of its recent CAGR range. However, with the more dynamic (relative to downstream) CAGR history of upstream in mind, note that the 85 MHz mid-split and a DOCSIS 3.1 Fiber Deep upstream still delivers *10 years* of life at a 40% CAGR – a CAGR range associated with actual *downstream* YOY growth in the past several years. In other words, if upstream CAGR does begin to resemble the downstream and see a large increase for a period of time, then the lifespan provided by implementing the above set of upstream strategies still delivers on a 10-year lifespan.

BEYOND FIBER DEEP?

In the N+0 FD architecture, the physical path from the Headend or Hub to the home is extended to now be (approximately, in physical distance) 98-99% fiber, with only the last 1-2% coaxial cable. The infrastructure itself has been heavily based on fiber optic technology and has taken advantage of the continued technology breakthroughs since all-coaxial networks first became HFC, and this will obviously continue.

Beyond maintaining the fundamental premise of a coax-to-the-home last mile, a core FD architecture strategy includes the concept of enabling the final launch of a fiber last mile if and when services drive this. In this sense – a network adaptable to the demands of customers – it is exactly how cable operators have been evolving their networks and services since they first built them. The cable industry, which pioneered AM optics, has also been deploying FTTP digital optical technologies for many years in very demanding business services and cellular backhaul applications. Thousands of FTTP Gigabit connections, readily available through years of investment in a rich fiber infrastructure and optical technology, exist in cable architectures. These FTTP systems, among other mature FTTP technology solutions, are natural candidates for use in residential services where it makes sense to do so, and enabling them is therefore considered a basic core requirement of a Fiber Deep node.

Summarizing, then, cable services can be run over traditional HFC, or Fiber Deep, or an all-fiber infrastructure. Furthermore, the shift to all-IP is erasing historical differences of service delivery and access, and IP is agnostic to whether it is an RF or fiber medium. The optimal solution will vary by situation, re-emphasizing the importance of a cost effective architecture based on flexibility and modularity, which cable operators have always embraced. Fiber Deep, and in particular the definition of the Fiber Deep node requirement itself, will continue this approach, and RF and fiber last mile access technologies will be options available.

Lastly, while architectures and engineering options dominate the thought processes in engineering departments, it is important to not lose sight of the fact that the choice of service provider by consumers is largely based on the services provided themselves and the servicing of them.

SUMMARY

Cable networks have demonstrated a remarkable adaptability to service demand, as well as the integration of the advanced technology required to support this demand. In recent years, operators have settled into a steady cadence of capacity upgrades that have delivered enhancements to best-in-class video, high-speed data, and voice, and have introduced a range of new services. These investments have been very effective, but the periodicity suggests a more comprehensive, "big picture" approach to network architecture and evolution may be more efficient.

We aimed to describe such a path in this paper, mapping out where the services and customer requirements are headed, how architecture and technology alternatives align to deliver on these requirements, the logic behind some of the key decisions, and the approach to positioning the network on a long runway of sustainability. The Fiber Deep approach infused with targeted architecture and technology refreshes produces a sound path for the next phase of network investment, and one that shall not need to be revisited for many, many years to come.

REFERENCES

[1] Chapman, John, Mike Emmendorfer, Dr. Robert Howald, Shaul Shulman, *Mission Is Possible: An Evolutionary Approach to Gigabit-Class DOCSIS*, Cable Show Spring Technical Forum, Boston, MA, May 21-23, 2012.

[2] Howald, Dr. Robert L, *Boundaries of Consumption for the Infinite Content World*, SCTE Cable-Tec Expo, New Orleans, LA, October 20-22, 2010.

[3] Howald, Dr. Robert L, *Breathing New Lifespan into HFC: Tools, Techniques, and Optimizations*, Cable Show Spring Technical Forum, Washington, DC, June 10-12, 2013.

[4] Howald, Dr. Robert L, *Fueling the Coaxial Last Mile*, SCTE Conference on Emerging Technologies, Washington DC, April 2, 2009.

[5] Howald, Dr. Robert L, *Looking to the Future: Service Growth, HFC Capacity, and Network Migration*, SCTE Cable-Tec Expo Capacity Management Seminar, sponsored by the Society for Cable Telecommunications Engineers (SCTE), Atlanta, Ga, November 14, 2011.

[6] Howald, Dr. Robert, *The Grown-Up Potential of a Teenage PHY*, Cable Show Spring Technical Forum, Boston, MA, May 21-23, 2012.

[7] Howald, Dr. Robert L, *Time on Our Side: Using the Past to Pave the Way Forward*, SCTE Cable-Tec Expo, Orlando, FL, October 17-19, 2012.

[8] Howald, Dr. Robert L, Dr. Sean McCarthy, *Accounting for Techies: Taking it to the Ultra*, Cable Show Spring Technical Forum, Washington, DC, June 10-12, 2013.

[9] Howald, Dr. Robert L, Dr. Sean McCarthy, *Bits, Big Screens and Biology*, Cable Show Spring Technical Forum, Boston, MA, May 21-23, 2012.

[10] Howald, Dr. Robert L. and John Ulm, *Delivering Media Mania: HFC Evolution Planning*, SCTE Canadian Summit, Toronto, ON, Canada, March 27-28, 2012.

[11] http://www.ieee802.org/3/bn/public/adhoc_mmp/adhoc_mmp_urban_01.pdf

[12] http://www.ieee802.org/3/bn/public/mar13/howald_3bn_01_0313.pdf

[13] http://www.ieee802.org/3/bn/public/mar13/howald_3bn_02_0313.pdf

Implement Closed-Loop Network Decisioning Now with Big Data Analytics and Fuel Future-State SDN Use Cases Through a Common Platform Deployment

Brennen Lynch and Anukool Lakhina
Guavus, Inc.

Abstract

Communication Service Providers (CSPs) are finding an increased reliance on their IP networks by consumers from a service perspective. As consumers become more dependent on IP services and thus bandwidth demand grows, CSPs require dynamic, stateful tools to optimize network operation/QoS and monetize the streaming data generated by the network. CSPs also have the opportunity to incorporate potential consumer-based revenue streams by targeting elastic, innovative, and potentially disruptive services and technologies.

Operational virtualization can occur with systems that are present within the network now. CSPs have the opportunity to become early adopters of these closed-loop use cases and roll out incrementally, starting with solutions that can be supported by current network conditions and placing their business in a strategic position to exploit SDN in a timely manner. CSPs that implement a standardized data ingest/processing/fusion platform can fuel use cases that are supported by current network conditions (such as real-time routing decisions). As the industry and backbone/access networks move towards complete virtualization through SDN and NFV techniques, the common platform can ingest additional, relevant data sources (such as hypervisor logs) to power coherent virtualized, automated, closed-loop decision processes.

This paper explores platform integration and techniques that can be implemented in the network now, which place operators in a prime position to exploit fully virtualized control systems as they become incorporated. Use cases from an operator's network perspective (traffic optimization) as well as a consumer's perspective (dynamic service delivery) are explored and demonstrated through experimental investigation of both current network and fully SDN-enabled test-environment scenarios.

Transition towards dynamic resource allocation

Code-line mediation of network infrastructure generally requires physical intervention at the hardware site. This highly inflexible way of mediating the network is not congruent with the way that current services are becoming virtualized. No longer does the ISP or content provider connect directly to the subscriber to deliver media. Intermediary content delivery networks (CDNs), placed near to the subscriber, cache and relay content to reduce latency and increase the quality experienced by the subscriber. Datacenter operations are in many cases have transient requirements; significant hardware/CPU optimization can occur if these operations are virtualized and implemented, when and where they are required.

Network analytics can extract highly valuable information and insights. However, there must be integration within the network to allow for these insights to drive operations. Devices currently make stateless decisions based on the traffic they carry, and not based on the network's entire state. Resource allocation is based on fixed presets, rather than network conditions and demands. Large and complex networks simply compound the problem of manual intervention within network settings – a tedious and error-prone process.

Virtualization of networks is at odds with centralized functions/services, yet does centralize control functions. As physical servers are virtualized and operate under dynamic provisioning, the innovation bottleneck is now placed within network provisioning. New protocols and services require tedious manual changes of the software within network devices. Enabling a centralized control plane and providing separation from the forwarding functions offers the ability to dynamically program the behavior of the network using well-defined interfaces. This controller can be deployed as a cluster for high availability and scalability. These are the core facets of Software-Defined Networking (SDN) implementation from a functional perspective. Network Function Virtualization (NFV), developed by service providers, is a complementary concept allowing for relocation of network functions (such as NAT, firewalling, DNS, intrusion detection, caching, etc.) from dedicated appliances to generic servers, enabling dynamic resource allocation.

Analytics is the brain behind a highly optimized and efficient network. Traditional 'store-then-analyze' paradigms are simply too latent to fuel real-time decisioning in the face of petabyte-scale streaming data. By implementing a distributed-compute model, where known end-actions are translated into pre-processing algorithms at the collector state, an operator can effectively power real-time processes that are dependent on such fusion. Effective integration of SDN will require a specialized, *big data fabric* to provide stateful decision-making, as these inveterate & transient functions are conditional on real-time processing of network data (as well as fusion to appropriate reference datasets).

Architecture & State of Integration

<u>Software Defined Networking</u>

At its core, SDN provides a construct for centralizing the control plane (typically layer 2/3 functions). This controller provides a logical representation of the network by abstracting the network topology and the network elements. It 'replaces element-based, box-at-a-time configuration with network control; that is, instead of hop-by-hop control, it provides end-to-end control' [17]. By providing a centralized stateful controller, applications can then manipulate control functions dynamically via northbound APIs. Through such orchestration applications, SDN can automate service chaining for service functions in higher-order layers (3+) – a boon for minimizing deployment disruptions when implementing SDN functions.

Southbound protocols (such as standards-based OpenFlow) provide the controller access to the data plane devices - whether these devices are physical or virtual. By standardizing the control protocols, different vendor devices can be easily rendered as interoperable. This ease of integration is hugely supportive of industry innovation, as the control plane can ensure interoperability on a virtual basis – no longer do entire software stacks need to be updated in support of a new service (imagine if every time you wanted to download another app on your smartphone, a new operating system update was required!).

A non-disruptive integration of SDN towards a wholly device-based implementation (requiring SDN-compatible devices) must support virtual switches to communicate with devices, as well as provide support for bare-metal hosts such as datacenter servers. The figure below details this hybrid architecture.

(consuming a fixed amount of resources at a fixed location). These traditional implementations could span from centralized-service data centers, or highly distributed appliances (such as video caching). Virtualizing these functions with standardized hardware (i.e. x86) transitions from the traditionally unique (& costly) implementations and supports one single, highly efficient investment of an operational model. NFV breaks the traditional linkage between IP location and identity. This is a very similar model to how modern data centers operate, and is highly complementary to SDN techniques that abstract the control plane from the data plane.

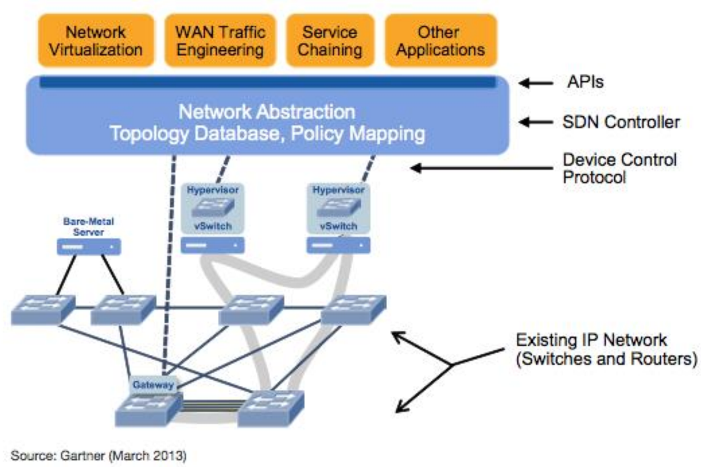


Figure 1: hybrid architecture of SDN + bare metal [17]

<u>Network Function Virtualization</u>
Since standard hardware – multicore processers – have made exponential gains in their processing power, network functions can be embedded as software agents in these standard processors, as opposed to the traditional implementation of a unique appliance

<u>State of Integration within Service Providers</u>
Hybrid architectures will be implemented in order to maintain a seamless transition over time to a fully SDN architecture. Traffic forwarding must be differentiated into centralized controller forwarding (i.e. via

OpenFlow) or forwarding via conventional means. The Converged Cable Access Platform (CCAP) provides a transitional model towards virtualized control, by combining the eQAM for digital video with the CMTS for data into one device (or logical entity, if virtualized). The key CCAP feature towards SDN techniques is that it provides a control/management plane (for IPDR, DOCSIS, RF, QAM MIBS PCMM, & edge resource management). Decentralizing control from the headend in a virtualized CCAP environment would allow these compute processes to occur in the most efficient location as determined by the controller. The flow table can communicate with the controller to install flow entries based on load balancing, failover, traffic control, service prioritization, etc. Implementing real-time data analysis to support this controller communication with the forwarding plane will ultimately enable dynamic resource allocation and drive key network optimization techniques.

Big Data Fabric; Real-time Processing

Use cases supporting SDN techniques will require an underlying data fabric that supports real-time analysis and supports closed-loop, automated processes. A loosely coupled, component based architecture is the basis to this platform. Furthermore, the collection architecture should be intelligently structured with edge-based processing components. At the petabyte-scale of high velocity network data, collection must have the ability to pull data from the edge through the backbone network in real time without taxing mission-critical infrastructure. Furthermore, the network data needs to be dynamically fused and correlated with static & reference data sets to produce key causal relations. Thus, a departure from the traditional 'store-then-analyze' approach must be explored in order to move from reporting-based BI platforms to one that supports real time, closed loop processes. Although this paper is not a detailed investigation of the specific platform components, a high level view will be presented of the overall structure.

Basing the platform on open-source architecture is important to maintain interoperability standards. However, radical extensions of traditional open-source platforms are required to optimize for specific use cases (low latency, high cardinality, compute requirements, etc.). A highly available, 'compute-first' architecture that is based on the open source Apache Spark project is the cornerstone of this *new data fabric* that supports this paper's use case explorations. By integrating batch analytics, real-time analytics, and iterative machine learning into a single stack, significant latency gains can be made over the Apache Hadoop's Map-Reduce implementation (also significant development primitives over map-reduce). At the point of collection, data can be normalized and fused by using agent-based software components that are optimized for low-latency, high-throughput transactions.

Using Apache Spark as the base, extensions can be incorporated into a single stack to support high cardinality use cases, complex event processing and machine learning iterative algorithms, as well as fast query processing (outside of structure workflows) and mutable data-store integration. Furthermore, some of these platform components *must be* *virtualized* in order to effectively support SDN use cases. The compute engine supports stream processing via a rules engine & analytics based on state variations, batch processing through cubes representation, & iterative processes through machine learning engines. Apache Spark has great support for machine learning through its Mlib library (which leverages Scalanlp's Breeze and JBlas), and graph analysis through its GraphX library. Shark, a distributed-query engine implemented over Spark, can support low-latency discovery use cases over data that resides in Hadoop's Distributed File System. Through cluster management (ie YARN), multiple applications and workloads can run on a single cluster, optimizing resource allocation and providing cluster resilience. The processed data can then be presented for visualization via in-memory caching & disk storage using columnar compression, or can be used as an outbound action trigger through a message broker framework with standardized adaptors. This could easily communicate with any variety of SDN controllers such as the open-source OpenDaylight, or vendor-specific such as Juniper's Contrail. The virtualization of this platform to support efficient SDN use cases could leverage QEMU/Openstack. [20]

Use Cases for Fixed-line Networks

<u>Access Network</u>

By deploying a platform with the ability to provide real-time correlated outputs, an operator can begin to drive virtual functions that support closed-loop use cases around optimization, routing/caching, security, and more. This also allows for continual visibility of VNF (virtual network function) & VRF (virtual routing & forwarding) implementations.

The data sources required vary by use case, but can be segregated into the network-based data, and contextual datasets.

Network data:

- *IPDR/SP*: IP detail record resolving traffic flow, QoS metrics to network topology (cable modem & CMTS). IPDR is already enabled throughout the entire network through DOCSIS.
- *DPI*: deep packet inspection involving probe-based implementations at the router to process each packet & output relevant data, such as the user agent to define device, or the content/application accessed.
- *Flow (i.e. OpenFlow / NetFlow)*: traffic flow structured by the matching packets to the traditional 5-tuple of ingress interface, source/destination IP, IP protocol, source/destination port for TCP/UDP, IP Type of Service
- *DNS*: domain name system provides IP-to-URL translation and, with a given IP, provides visibility into content & application.
- *Router configs/SNMP*: gives interface level details for capacity, verification, etc.
- *Hypervisors*: such virtual topology information is contained within the hypervisor logs, such as virtual-to-physical mapping, routing paths, and bandwidth reservation state. This could be in a distributed environment to achieve scalability [3].

– Other configuration data: application IPs/ports, site IP ranges, etc.

QoE-driven Optimization; Content Routing & Caching; CDN

As consumers are increasingly reliant upon IP delivery of content, QoE-driven mechanisms are important to ensure a high quality to the end user, as well as optimize the use of the access network. Content caching through CDNs have been implemented for some time now as a method to reduce core traffic and to increase the quality of service to the consumer. Traditional engineering would build fixed-appliance caches on expected peaks, which ultimately means that much of the caching capacity would be unused for a majority of the time. These typical deployments are in-line in order to capture and respond to http requests (thereby forced to examine all traffic, whether it can be served by the cached server or not). SDN techniques allow CSPs to deploy servers out-of-line, directing specific flows to the caching servers. This results in an increase of the overall scalability and reliability of the solution. In addition, specific targeting of sites based on caching can occur.

When services can be deployed close to the end user, latency and overall network efficiency benefit. Stateful decisions must be made; in other words, considering not just distance, but also latency, jitter, and available bandwidth. By correlating streaming inputs of link performance, latency of services, delay, & jitter, an SDN controller function can spin up or spin down services by directing flows to the caching servers. Although this can be structured on the bit-level, it is much more valuable to resolve these flows into content, in order to aggregate and optimize by content or service. This level of granularity can be obtained in two ways: (1) through some level of packet discovery via DPI, which can either be full-census or via intelligent extrapolation of flows based on sample-census DPI, or (2) by resolving NetFlow logs to DNS queries on a real-time basis (requiring low-latency compute operations). Controllers can then obtain latency information from the routers to ultimately build quantitative analytics and even predictive; to ultimately program the network itself in response to routing in certain conditions or *anticipated* conditions. Although the mechanisms of

Figure 2: example visualization of services usage, classified by content category. This gives visibility towards any caching VNFs implemented on a per-content basis, and parameters such as by utilization or WAN cost allow for operational oversight. Furthermore the sites/site-pairs could be exposed to view top demand generators for each service.

controller-based VRF is out of scope of this paper, the controller could act at the router level by dynamically implementing distributed, virtual publish/subscribe systems in the Ternary Content-Addressable Memory (TCAM) thus dynamically programming the access control lists (ACLs) [2]. Financial information can also be incorporated to bring peering costs via leased links/routers into context [12]. An MSO could use GSLB (global server load balancing) to direct traffic according to these criteria, and interface with PCMM PS (the PacketCable Multimedia Policy Server) to provide some level of DOCSIS QoS assurance. From a video optimization standpoint, business-contextual data around weather, large events i.e. Superbowl, etc. can be integrated into the dynamic caching/routing algorithm. The operational metrics associated with demand generators and VNF functions are shown in figure 2 through an example user interface.

Traffic Optimization

Current implementations of traffic forwarding (in DOCSIS networks) see the CMTS performing routing functions. By learning the network topology through routing protocols and by providing a default subscriber gateway, the CMTS routes traffic from subscriber devices to aggregation routers, which in turn route traffic across core network to the Internet or CSP servers. With the converged cable access platform (CCAP) providing the transitional bridge from physical hard-wired appliances to virtualized network functions, controller interfaces that rely upon real-time analysis of data can perform optimization-based functions. The CMTS/CCAP can serve as a layer 2/3 device by interfacing its Flow Table with the controller, providing dynamic forwarding function as well as aggregate flow processing (matching multiple traffic types and/or destinations). Once these distributed-compute algorithms are virtually implemented, they can operate on a closed-loop basis to support such VRF functions (see figure 3). Implementation of local DNS caching can also take advantage of this analytics engine to drive traffic optimization, reducing response time to the subscriber as well as reducing traffic on the core network (minimizing centralized server queries).

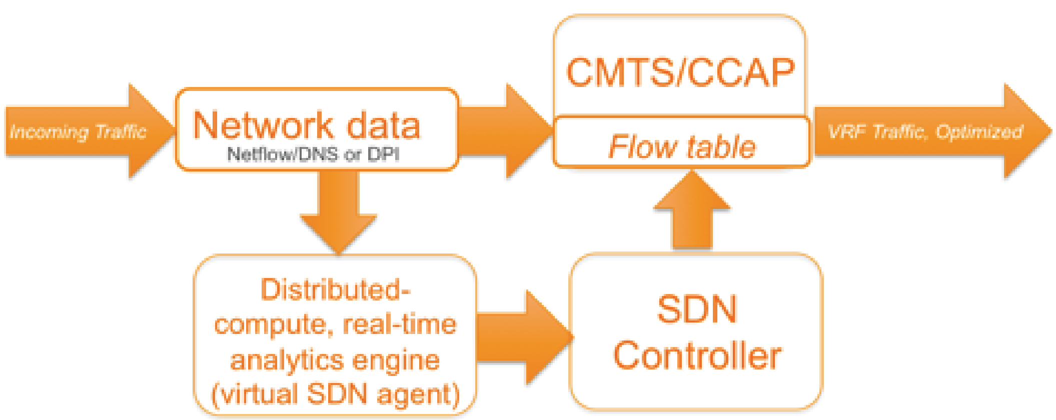


Figure 3: example representation of closed-loop VRF driven by a distributed-compute, real-time analytics engine

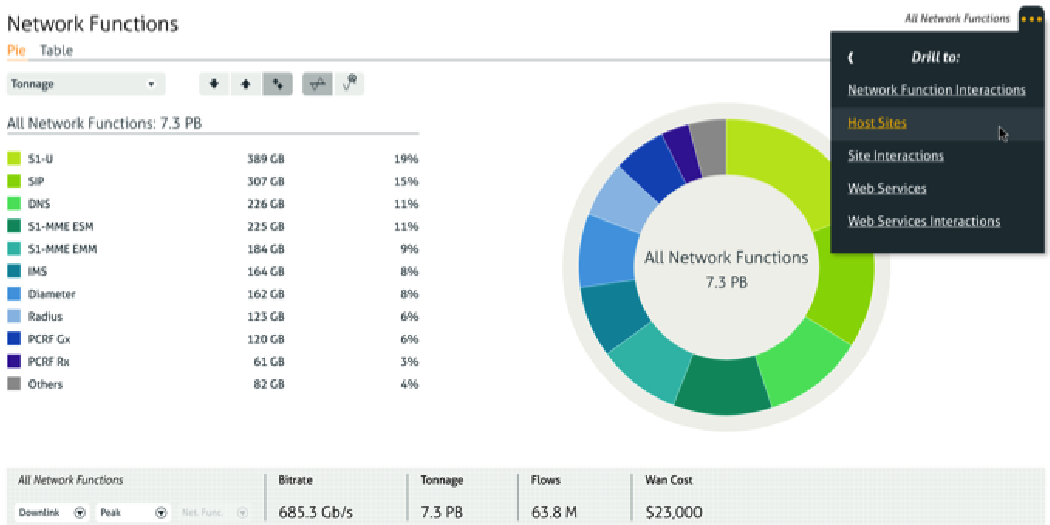


Figure 4: example of a centralized UI that provides visibility towards how network functions are operating. Executive oversight is enabled through key drill downs to host sites, site interactions, & web services/interactions.

Network Function Visibility; SLA Monitoring

Visibility towards how VNFs are operating is a key executive use case to allow for ongoing monitoring of their interactions with the network (see figure 4 for example UI). A site-to-site breakdown of traffic supports site visibility. For network functions, one could centralize in a UI the app usage for the network by data center (to satellite site,..), the applications by bitrate/tonnage/flows, WAN cost, & utilization. Furthermore one could drill down to view top consuming sites for a given application, or drill down to sites and site-pairs to see top demand generators. SLA monitoring can occur by visibility towards performance metrics; views by links (by delay/packet loss/jitter), or drill into top applications and cloud services on impacted links. Alerts can be configured via application performance to inform the stakeholder of any anomalistic behavior. This type of executive oversight enables operational assurances.

Framework for other use cases

As diverse as software functions are, there are many use cases that can be enabled through the real-time analysis of data. This paper will mention a few more use cases at high level, but this is not meant to be exhaustive. Security use cases can support Lawful Intercept (virtualized vs. at the CMTS), and enable stateful firewall decisions that can interact with packet flow through the Flow Table and can be instantiated on a virtual basis. Or, support for distributed CG-NAT or dynamic encapsulation of VPNs directly from the cable modem [5]. Furthermore, targeted billing services can be implemented on a virtual basis with a high level of accuracy.

Datacenter Operations

Datacenters are lacking an integrated view across infrastructure, applications, and users – such an integrated view is required on a continuous basis to trigger timely operational decisions. User traffic will generally span many datacenters & network functions. Traditionally these must be constructed via prediction – calculating the expected traffic at peak and engineering to support that. This leads to inefficiencies during non-peak hours, as well as potential bottlenecks if traffic is exceedingly high. By understanding how these network functions are operating in real-time, a provider can route traffic and optimize these functions via a software-based controller. This capability is highly dependent on analyzing these network logs and fusing them with contextual reference data (for example, fusing with graph-based representations of topology via hypervisor logs), and having the ability to do so in real-time.

The interface of the physical and virtual domains is key to survivable, scalable, operable data centers. In today's datacenters, physical and virtual assets are largely silo-ed and very cumbersome to fuse even on a reporting/historical basis. By integrating a seamless end-to-end view, previously un-targeted use cases can be exploited such as real-time security or root cause analysis, as well as power VNFs around dynamic resource allocation. Closed-loop signaling is used to automate critical response to issues that are discovered. From an executive standpoint, business leaders can implement assurances around security

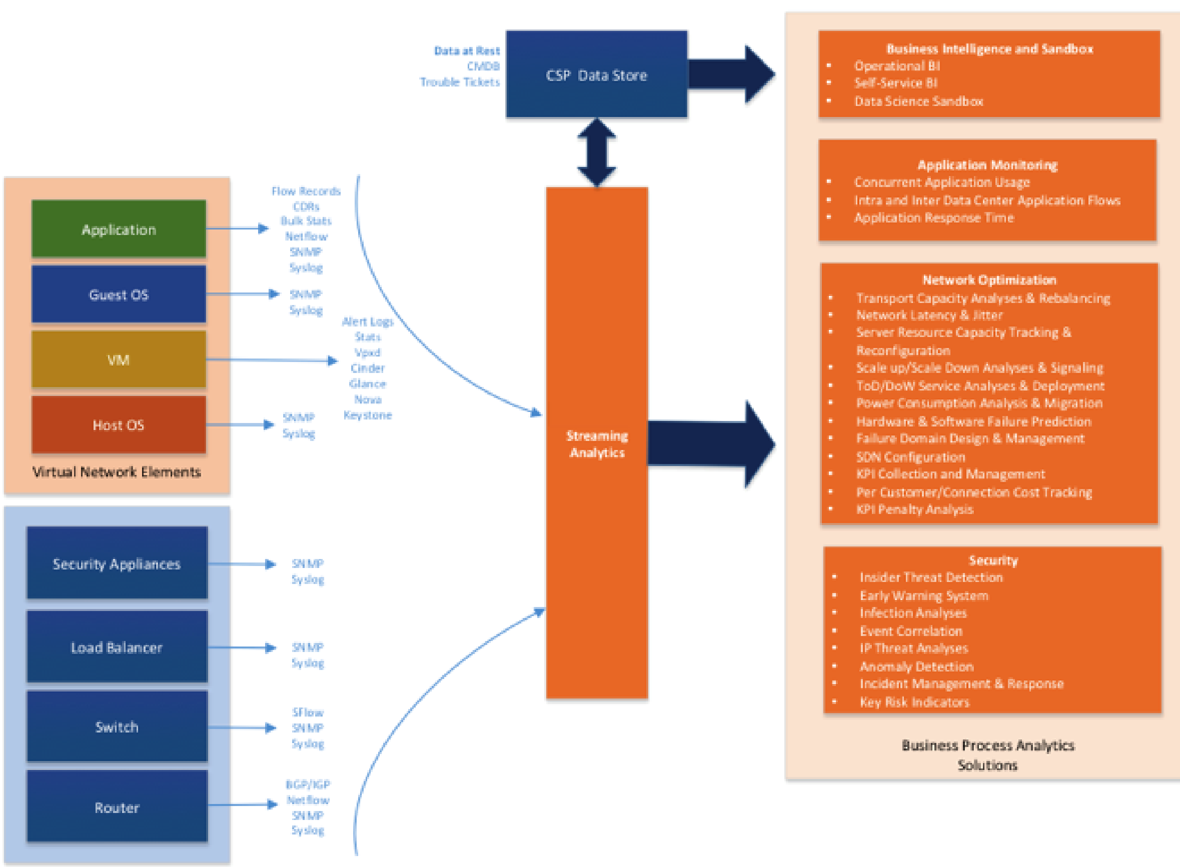


Figure 5: reference architecture for DCOI. Logs are collected directly; VM data pulled directly from orchestration systems. Flow, BGP collected from transport infrastructure.

and operational optimization, as well as identifying shadow IT.

By integrating closed-loop architecture, intelligent edge processing and fusion supports highly valuable use cases dependent on real-time analysis. A sample architecture would collect the necessary data from their points of origin; illustrated in figure 5. Network elements forward traffic & polling data such as syslog from the application or OS; SNMP, and flow. VM data is collected from the elements and orchestration via SDKs (software development kits; API plug-in libraries), and server data is collected from the management systems (iLO, SEL,..). Static data such as customer information and semi-static such as DNS and trouble ticketing is collected. As additional data centers are brought online, they can be easily integrated within the analytics framework to create a single, unified view.

Resource Utilization - Capacity Planning; Reduce WAN Costs
Business users are always searching for ways to optimize capacity planning, such as reducing WAN costs, deferring data center link upgrades, etc. By integrating NFV, the optimal data center can be identified and the service deployed in a virtual instance. Data center sites can be identified through ingesting data provided by SDKs; in a virtualized instance. Fusing this utilization-based data with the context of WAN costs, along with link-based metrics from the access network (latency, hop count, etc.), will drive a closed-loop optimization application (a VNF). This specialized VNF can furthermore integrate with forwarding & routing decisions to optimize the access network flows dynamically (via software controller interacting directly with the Flow Table in the CMTS/CCAP). This VNF could be monitored by centralizing the real-time operational metrics into a user interface (see figure 6). Redeployment of services reduces the WAN costs. Rehoming of traffic also defers link upgrades and application licenses could be reduced.

The ability to forecast resource consumption as well as respond in real-time to utilization conditions will drive

Figure 6: example visualization of site-based traffic. Site-to-site interactions give visibility towards how DCOI-based resource allocation VNFs are operating, such as those driving WAN-based network optimization.

efficiency-based processes. By collecting flow data from the physical domain that fuses with server, VM, & storage data, one can create support resource optimization VNFs on a real-time basis and present a seamless view of resource utilization from the data center gateway to the application (see figure 7). Furthermore, data center resources can be forecasted per application based on historical trends and rules-based engines (operating via key indicators). By including financial-based data around energy, VMs can be instantiated based on minimizing cost as well as delivered performance [1]. This allocation algorithm interacts with the centralized controller to manage the servers and switches across the logically virtualized data centers [3]. Ultimately this will reduce the network downtime, radically improve root cause analysis, and improve end-user QoS.

Application Optimization
Similar to resource optimization, but extended to manage L3 and above application traffic. This could involve applying application-based policies (including QoS assurance), as well as defining the deployment, user, and server contexts. For high-value applications, specific SLAs around QoS might be in place, involving a cost model of both revenue and penalty [21]. KPIs around CPU/memory utilization, disk i/o, storage latency, switch/firewall/link/load balancer performance & utilization, etc. can be implemented to alert or trigger automated actions once they are surpassed. Or, based on application on-demand (i.e. data backup or replication scenario), network bandwidth can be dynamically allocated for backhaul connectivity or inter-datacenter links.

Efficient utilization of resources to support these applications must occur under a fluctuating demand and unpredictable failures. This requires a dynamic adaptive mechanism in the form of a VNF that is driven by real-time analysis of the physical- and virtual-domain data.

Cognitive Computing
With growing data-center demands at over 60% per year, the infrastructure needs to adapt dynamically. Real-time granular detail on response times, arrival rates, & concurrency for capacity planning prevents performance-sapping measurement code in core applications.

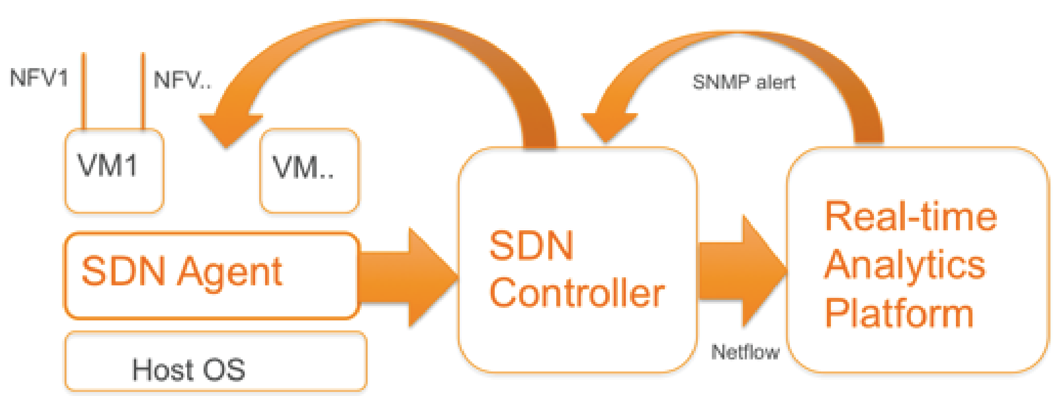


Figure 7: dynamic resource allocation allows for resources and functions to be optimized across data centers

This will maximize server utilization while maintaining user performance. From a business standpoint, an analytics solution could accurately predict scale points to manage investment.

Business-powered use cases (closed-loop trigger or visual discovery) around assessment of performance-related issues. Real-time browsing and purchasing data can be streamed into analytics and forecasting applications. As business & performance conditions change, the IT systems must follow suit & adapt; interacting with the network functions to eliminate inadvertent but costly errors that are driven by static (poor) system configuration. Or, from a product standpoint, applications with different QoS requirements can be sold with differentiated pricing – bulk applications that are not critical to revenue can be transported when the network is not being heavily utilized.

Security & Root Cause

Data center security is traditionally a linear firewall implementation. The current service-deployment implementation strategies allow for large vulnerabilities, as security is generally added on after deployment, as an afterthought. The solution is to distribute security systems that are built in from the beginning - systems that carry out regular automated tests to track compliance.

By tracking and characterizing malware threats / fraudulent transactions etc. in real-time (for neutralization), machine learning algorithms can be implemented to effectively characterize new as-yet unknown threats. By providing a seamless correlation between the physical and virtual domains, obscure patterns (such as those that span datacenters) can be identified to characterize and neutralize such threats. A virtualized firewall instance can be then implemented, or the centralized controller can interact directly with the routing/switching mechanism to drop offending packets. Service chaining based on such virtual instances could occur – having a flow pass through a network monitoring VNF, followed by a load balancing VNF, to ultimately a firewall VNF. Furthermore, the underlying analytics platform could provide predictions based on network behavior (using data science techniques such as application of Granger causality analysis).

This data must be structured in an appropriate visual/closed-loop architecture to target the described use cases. For example, a Security solution might tackle its use case set through the following three modules:

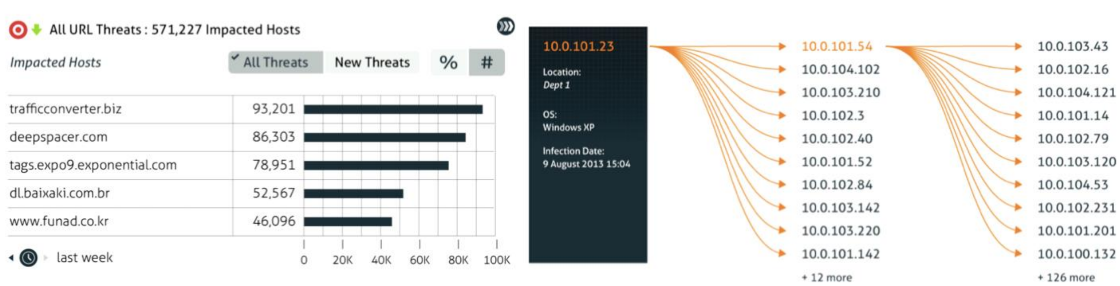


Figure 8: example user interface of threat identification and propagation analysis

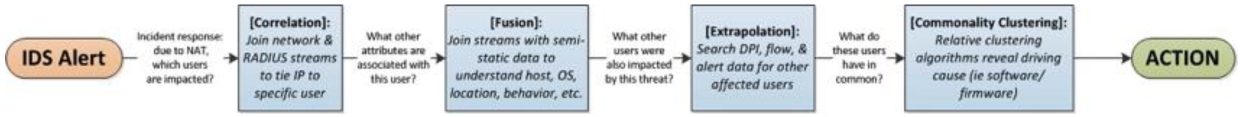


Figure 9: flow chart detailing security use case from Alert, Correlation, Fusion, Extrapolation, Root Cause Analysis, & Action

Security dashboards: by exposing the top threats, correlating the events to the user, and overlaying the context (device, location, OS etc.), a security stakeholder can implement assurances and ongoing monitoring.

Security investigations: once events have been correlated, their attributes and network behavior can be extrapolated to understand threat propagation and threat impact (see figure 8). This provides required insight in to the threat proliferation and monitors the results of any incident containment effort.

Advanced threat detection: to target the driver of any threat, a compute module can be implemented that not only detects anomalies, but also defines the root cause. Through relative commonality calculations, attributes of the subscribers who visit malicious URLs/IP addresses can be correlated as a driving cause. Furthermore, baseline users exhibit various behavior features (connection ratios, destination entropy, fan-out, etc.), from which a risk score can be calculated based on anomalous behavior (that has been previously shown to precede a threat). To understand impact (now & potential), the Most Likely Origin system (ground zero) can associate with the subscriber attributes to drive a proliferation algorithm that extrapolates the potential impact across the entire subscriber base and systems architecture. For an example of the threat investigation flow, see figure 9. Once the threat is identified, an increase in sampling rate could be triggered to increase operational awareness. Further, SDN techniques to interrupt the packet flow (as previously-described) can be implemented or the element could be instructed to configure the span port to collect all packets and decode to PCAP.

Conclusion

Implementing real-time datacenter operational intelligence is key to an effective hybrid datacenter strategy. Holistic intelligence of physical, virtual or cloud functions (across multiple data centers) powers the type of scalable analytics that drives automation through structured, end-to-end use cases. The network can be dynamically allocated via functions, storage, compute assets that require deeply business contextual data. Spend can be optimized through data-driven resource allocation. From a product standpoint, deployment of new services can be radically accelerated and SLAs can be more fortified.

The access network can be targeted with a number of data-driven use cases. As CCAP becomes integrated the distribution of flows can be optimized via demand, latency, priority, or other metrics. Content can be dynamically cached as CDN networks move from traditionally purpose-build appliances towards a software-controlled cache on standard hardware. Other virtualized functions such as security, CG-NAT, VPN, or more can leverage the same underlying data and processing capabilities inherent to an operational intelligence platform.

A distributed-compute, real-time analytics platform is the key enabler to these closed-loop use cases around data center & access network control virtualization. Key virtual network functions (such as VNF) are dependent on compute components that are fed pre-processed data on an extremely timely basis. With many modern datacenters already operated under virtual control, access networks are following suit with a slow transition towards a fully virtualized and centralized control scheme. With this in mind, implementing the underlying analytics (operational intelligence) platform will position an operator with the ability to implement virtualized control when and where it becomes available. This platform inherently operates on a 'collect-once, analyze-many' paradigm that supports extremely quick deployment of specialized VNFs.

REFERENCES

[1] Beloglazov, Anton, and Rajkumar Buyya. "Energy Efficient Resource Management in Virtualized Cloud Data Centers." *Cloud Computing & Distributed Systems (CLOUDS) Laboratory, Univ. Melbourne* (2010): n. pag. Print.

[2] Bhowmik, Sukanya. "Distributed Control Algorithms for Adapting Publish/Subscribe in Software Defined Networks." *University of Stuttgart* (2013): n. pag. Print.

[3] Chuanxiong, Guo, Guohan Lu, Helen Wang, Shuang Yang, Chao Kong, Peng Sun, Wenfei Wu, and Yongguang Zhang. "SecondNet: A Data Center Network Virtualization Architecture with Bandwidth Guarantees." *ACM CoNext* (2010): n. pag. Print.

[4] Dong, Yaozu, Xiaowei Yang, Jianhui Li, Guangdeng Liao, Kun Tian, and Haibing Guan. "High Performance Network Virtualization with SR-IOV." *Journal of Parallel and Distributed Computing* 72.11 (2012): 1471-480. Print.

[5] Donley / CableLabs, Chris. "Leveraging Openflow in DOCSIS Networks." *NCTA Spring Technical 2013* (2013): 59-76. Print.

[6] Durr, Frank. "Towards Cloud-assisted Software-defined Networking." *University of Stuttgart / Institute of Parallel & Distributed Systems* (2012): n. pag. Print.

[7] Haider / NICT Tokyo, Aun, Richard Potter / NICT Tokyo, and Akihiro Nakao / NICT Tokyo. "Challenges in Resource Allocation in Network Virtualization." *ITC Specialist Seminar* (2009): n. pag. Print.

[8] Jain, Raj, and Subharthi Paul. "Network Virtualization and Software Defined Networking for Cloud Computing: A Survey." *IEEE Communications Magazine* 51.11 (2013): 24-31. Print.

[9] Koldehofe, Boris, Frank Durr, Muhammad Tariq, and Kurt Rothermel. "The Power of Software-defined Networking: Line-rate Content-based Routing Using OpenFlow." *University of Stuttgart / Institute of Parallel & Distributed Systems* (2012): n. pag. Print.

[10] Kulkhani, Shridhar. "Software Defined Networking (SDN) for Cable Access Networks." *SCTE Technical Forum 2013* (2013): n. pag. Print.

[11] Liou / Infinera, Chris. "Optimizing Multi-Layer Networks with Transport SDN." *SCTE Technical Forum 2013* (2013): n. pag. Print.

[12] Lively / Cisco, Dave. "Software-Defined Networking and Cloud – Enabling Greater Flexibility for Cable Operators." *NCTA Spring Technical 2013* (2013): 77-84. Print.

[13] Nandiraju / Arris, Nagesh, and Sebnem Ozer / Arris. "Applying the Software Defined Networking Paradigm to MSO Commercial Networks." *NCTA Spring Technical 2013* (2013): 85-98. Print.

[14] "OpenFlow-enabled SDN and Network Functions Virtualization." *Open Networking*

Foundation, ONF Solutions Brief (2014): n. pag. Print.

[15] Rahman, M. R., and R. Boutaba. "SVNE: Survivable Virtual Network Embedding Algorithms for Network Virtualization." *IEEE Transactions on Network and Service Management* 10.2 (2013): 105-18. Print.

[16] Ram / Rice University, Kaushik, Jose Santos / HP Labs, Yoshio Turner / HP Labs, Alan Cox / Rice University, and Scott Rixner / Rice University. "Achieving 10 Gb/s Using Safe and Transparent Network Interface Virtualization." *VEE* (2009): n. pag. Print.

[17] Sharma, Akshay, Deborah Kish, Sree Chadalavada, Laura Flowerree, and Sean Moore. "Forecast Overview: SDN and NFV in Carrier Infrastructure, Worldwide, 2013." *Gartner* (2013): n. pag. Print.
Skorupa, Joe, Mark Fabbi, and Akshay Sharma. "Ending the Confusion About Software-Defined Networking: A Taxonomy." *Gartner* (2013): n. pag. Print.

[18] Smith / Juniper Networks, Andrew, Colby Barth / Juniper Networks, and Saif Rahman / Comcast. "Enhancing DOCSIS Services through Network Functions Virtualization." *NCTA Spring Technical 2013* (2013): 99-105. Print.

[19] Somohano / Juniper Networks, Pilar, Brian Mutcherson / Comcast, and Boon Thau Loo / University of Pennsylvania. "Unleashing Network Potential through SDN in Virtual Networking Lab Environments." *SCTE Technical Forum 2013* (2013): n. pag. Print.

[20] "Spark Documentation & Overview." *Spark 0.9.0 Documentation*. Apache, 2014. Web. <http://spark.apache.org/docs/latest/>.

[21] Wang, Xiaoying, Zhihui Du, Yinong Chen, and Sanli Li. "Virtualization-based Autonomic Resource Management for Multi-tier Web Applications in Shared Data Center." *Journal of Systems and Software* 81.9 (2008): 1591-608. Print.

[22] White / Arris, Gerry. "Can DOCSIS Networks Leverage SDN?" *NCTA Spring Technical 2013* (2013): 105-25. Print.

KEY ACRONYMS

- ACL: access control list
- BGP: border gateway protocol
- CCAP: converged cable access platform
- CDN: content delivery network
- CMTS: cable modem termination system
- CSP: communications service provider
- CG-NAT: carrier-grade network address translation
- DCOI: data center operational intelligence
- DNS: domain name system
- DOCSIS: data over cable service interface specification
- DPI: deep packet inspection
- eQAM: edge QAM
- GSLB: global server load balancer
- iLO: HP Integrated Lights-Out
- IPDR: internet protocol detail record
- ISP: internet service provider
- MIB: management information base
- MLlib: Spark machine learning library
- MSO: multiple systems operator
- NF: network function
- NFV: network function virtualization
- PCAP: Packet capture
- PCMM: PacketCable Multimedia
- PS: policy server
- QAM: quadrature amplitude modulation
- QEMU: Quick EMUlator
- QoS/QoE: quality of service/experience
- RF: radio frequency
- SDK: software development kit

- SDN: software defined networking
- SLA: service-level agreement
- SNMP: secure network monitoring protocol
- UI: user interface
- VNF: virtual network function
- VNR: virtual network resource
- VPN: virtual private network
- VRF: virtual routing & forwarding
- WAN: wireless access network

Leveraging external applications for DOCSIS 3.1 HFC plant optimization.

Pawel Sowinski
Cisco Systems Inc.
psowinsk@cisco.com

Abstract

DOCSIS 3.1 offers a robust upgrade to the physical layer transmission methods by taking advantage of the state-of-art technologies such as OFDM and LDPC. These new techniques allow the DOCSIS networks to better utilize spectrum of HFC network, effectively pushing the efficiency towards the Shannon limits. The improvements come at a price; very complex configuration and the need for constant monitoring and probing processes which generate vast amounts of test metrics and maintenance information.

This information needs to be analyzed and acted upon by the CCAP software in order to adapt to changing network condition in a timely manners. The results are in the form of customized transmission profiles matched to groups of cable modems. However, the analysis of complex and voluminous test data on the CCAP has its limitations which we will explain in the presentation.

This paper presents a case for opening components of the CCAP OFDM phy maintenance sub-system to rely on an ecosystem of external applications and services. In particular, the paper reviews the requirements and compares options for definition of Cable Unique API for increasing the value proposition of DOCSIS. Such external APIs will offload CCAP from HFC optimization tasks and enable implementation of these functions in third party tools and applications. As result, CCAP configuration and operational complexity will be reduced. The set of compared external interface options includes an SDN controller based approach. The paper examines the relevance of such an approach to accelerate software development as well as the economic value for customers who desire to extend the CCAP functions with third party services and applications. Finally, the paper explores the need for multivendor interoperability and prospects for standardization of APIs.

Introduction

The ink is still drying on the recently issued DOCSIS 3.1 specifications. They offer vast improvements to the physical layer techniques and network scaling. The improvements come at a price; a complex channel structure, much higher PHY processing requirements (approx. 3-fold increase in the number of required silicon gates), multiplication in scale of operational and configuration data as well as the need for constant monitoring and probing processes which generate vast amounts of test metrics and maintenance information.

The intended audience

The intended reader of this paper is assumed to have a rudimentary familiarity with DOCSIS 3.1 technology.

What are DOCSIS 3.1 OFDM profiles?

This section provides a brief introduction to OFDM Profiles including a discussion about how a CMTS manages them.

DOCSIS 3.1 introduces Downstream and Upstream channels based on the OFDM technology, which enables effective deployment of higher level modulation orders. An OFDM channel has a complex structure; it consists of large collection of individually modulated sub-carriers, which are processed by FFT and can be individually modulated. DOCSIS 3.1 supports downstream modulation orders up to 4096-QAM (in the future up to 168384-QAM and higher) and upstream modulation of 1024-QAM (in the future up to 4096-QAM and higher).

After much discussion, DOCSIS 3.1 adopted the concept of modulation profiles for OFDM and OFDMA channels. We will refer to them as OFDM Profiles. An OFDM Profile defines modulation order for all active data sub-carriers of a downstream OFDM channel and for the data symbols of a minislot of an upstream OFDMA channel. OFDM Profiles allow for effective optimization of the transmission path to or from the CMs that can tolerate a higher modulation.

Sub-carrier modulation orders in a profile may vary across the frequency range of a channel. For example, the active sub-carriers in the range 100-188 may be configured for 64-QAM modulation; the active sub-carriers in the range 200-399 use 256-QAM modulation and the remaining active sub-carriers of a channel use 1024-QAM modulation order.

DOCSIS protocol primitives allow for unlimited variability of modulation orders where each of the eight thousand sub-carriers of OFDM channel may have a different modulation order from its neighboring sub-carriers. In most cases, the sub-carrier modulation order fluctuation will be more limited; in many cases sets of thousands of adjacent sub-carriers will share the same modulation order.

DOCSIS specifications require that a CM must provide support for 4+1 (four active and one test) profiles for each downstream OFDM channel and two profiles for each upstream OFDMA channel. The CMTS may provide up to 16 profiles for downstream OFDM channels and up to 7 profiles for upstream OFDMA channels.

In the control plane, the CMTS communicates downstream profile configuration to the CM via Downstream Profile Descriptor (DPD) messages. The upstream profile information is sent in UCD messages.

In the dataplane, downstream packets belonging to the same profile are organized into FEC codewords and FEC codewords are mapped by the NCP signaling channel. In the upstream an OFDM profile becomes synonymous with data IUCs; their allocations are signaled by MAP messages.

<u>OFDM Profile Management</u>

CMTS vendors will undoubtedly apply a various techniques and algorithms to effectively manage OFDM Profiles. The paper does not prescribe a specific method for this purpose. However, for background information, in order to appreciate the complexity of OFDM profile management functions we feel we need to examine those functions at high level.

The CMTS profile management involves two distinct, yet interdependent tasks. The first task is the determination of the OFDM profile parameters. For the second task, the CMTS assigns OFDM Profiles to groups of cable modems which receive downstream signal with similar fidelity or from whom the CMTS receives upstream signals with similar fidelity.

In its simplest form, the CMTS could create OFDM profiles based on a static device configuration. By "static device configuration" we understand a mode of operation where the CMTS is provisioned with a persistent set of configuration parameters for OFDM Profiles.

The static method for configuration of OFDM Profiles has a number of inherent limitations. Static profile configuration is difficult to manage and the profile settings are less efficient than dynamically created profiles. A typical CMTS/CCAP can house hundreds of OFDM channels, each with thousands of sub-carriers. The large number of sub-carriers makes the static configuration a difficult task, even if the configuration can be automated and involve only machine-to-machine interactions. Unless the OFDM Profile settings are greatly simplified, with very little variability in modulation levels across the channel's frequency range, any direct human involvement in OFDM profile configuration management tasks will overwhelm even the most patient operators. Furthermore, statically configured OFDM profiles are … static. They cannot be flexibly adapted to changing network conditions, such as occurrence of ingress noise, cable failures, component degradation, etc. Because the profiles are static they must be provisioned with larger error margin therefore have less then optimal efficiency. Static OFDM Profile configuration does not represent an interesting case for an external application and will not be considered further.

Alternatively, the CMTS may create OFDM profiles dynamically. At a high level, the CMTS algorithm includes collection and analysis of signal quality measurements from receivers, sorting of CMs with similar signal quality measurements into groups and determination of per sub-carrier modulation order for the groups. However, after diving into the details, it quickly becomes obvious that the problem has many dimensions.

The volume of signal quality measurement samples can quickly grow to considerable proportions. CMTS algorithms must scrutinize various MAC and PHY level error reports and incorporate elements of root cause analysis. Error reports that communicate the potential for direct impact on the user data must be acted on immediately. In such cases the CMTS will reduce the profile modulation or switch the set of affected CMs to a different, less strenuous profile. The frequency in which the CMTS collects the signal quality measurements has a direct impact on the ability to promptly detect and resolve issues that may be shared by groups of CM or by entire channels. Reduction in the intervals for gathering and analyzing the signal quality metrics helps with the response time, but it also increases the system processing overhead. The goal to maximize the spectral efficiency has to be balanced with other performance criteria, such as the need to minimize latency and overhead. To ensure minimal impact on the users, the newly created OFDM profiles candidates have to be tested before they are enabled to carry user data.

Dynamic OFDM Profile generation functions have the potential to become a complicated and a CPU intensive task because they need to fulfill many, often contradicting requirements and because they operate on large volumes of signal quality and diagnostic data.

Naturally, a hybrid of these two approaches to OFDM Profile management is feasible. For example, the CMTS could use static profile configuration as a starting point and further refine profile parameters through a dynamic process.

A more detailed description of OFDM profiles and a discussion of their management can be found in [01] as well in [03].

Problem Definition

Next, we will examine basic scaling and CPU performance requirements for dynamic management of OFDM profiles. As mentioned earlier, the CMTS periodically collects signal quality metrics from OFDM receivers and based on the analysis of the collected data can progressively build OFDM Profiles. How large is the volume of the signal quality metrics that a CMTS needs to collect and comb through to determine optimal OFDM Profile settings?

For this purpose, we decided to examine a performance metric which has per CM and per sub-carrier scaling multipliers. DOCSIS 3.1 specifications define standard methods for the measurement of a receiver's ability to receive modulated signal known as Modulation Error Ratio (MER). MER values can be effectively represented as 8-bit values using a logarithmic scale (dB). Cable modems measure MER for each active downstream OFDM sub-carrier based on pilot signals inserted in the channel.

The CMTS gathers MER information for each active sub-carrier of upstream and downstream channels. The CMTS requests MER measurements from CMs and collects MER statistics reported by CMs via newly added DOCSIS OFDM Downstream Spectrum (ODS) messages. In the reverse path, the upstream receiver embedded in the CMTS measures upstream sub-carrier MER from OFDMA probe signals.

The MER statistics database size examples have been calculated by taking into account the following parameters. Each downstream OFDM channel may include up to 7680 active sub-carriers. Each upstream OFDMA channel can consist of up to 3840 active sub-carriers. Both upstream and downstream MER stats for each sub-carrier are maintained as 8-bit values. Table 1 displays the MER database size estimates for a few combinations of CMTS Scale and service group compositions.

	SG Composition (OFDM Channels)	
CMTS Scale	2 DS + 1 US	5 DS + 2 US
3 000 CMs	**58 MB**	**138 MB**
10 000 CMs	**196 MB**	**470 MB**
30 000 CMs	**575 MB**	**1.4 GB**
60 000 CMs	**1.2 GB**	**2.8 GB**

Table 1 MER Database Size Examples

These values have been calculated considering the worst case scenario, for systems operating with 25 kHz sub-carrier spacing. The MER database size estimates for systems with 50 kHz sub-carrier spacing should be reduced by half. Nevertheless, Table 1 demonstrates that the MER statistics database size can grow quickly with the CMTS scale and with the increase in spectrum dedicated to OFDM. Considering that numbers in Table 1 represent memory sizing estimates for only one of several possible signal quality measurements, the actual memory requirements for the signal quality measurement database may be much higher. The database size could grow by another factor of magnitude if the processing algorithm includes elements of trend analysis and requires access to multiple generations of MER measurements. While these numbers won't stun readers familiar with modern, general purpose computing platforms, a comparison to the limitation of existing CMTS platforms may give a better perspective. The estimates may be approaching or exceeding the total memory pool size of many currently deployed CMTSs. Jumping a few years forward with Moore's law in mind, the signal quality measurement database, even if partitioned to fit the modularity of CMTS processing components (think cable linecards) will likely consume a significant portion of available memory in currently developed DOCSIS 3.1 compliant CMTSs. Undoubtedly, removing the need to maintain signal quality database from the CMTS will lower CMTS's memory and performance requirements.

Outline of the operation

The main goal of the paper is to examine the case for separation of the majority of the OFDM Profile management functions from the embedded programming environment of a CMTS and moving them to an external application which may be operating in a virtualized environment. We will refer to such application as HFC Profile Manager, or HPM. In this section, we will describe how such distributed system could operate.

The architecture of a system incorporating HPM is shown on Figure 1.

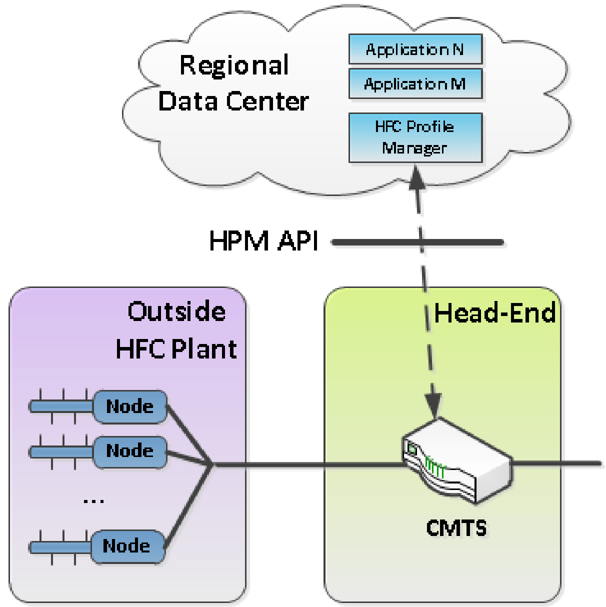


Figure 1 - Proposed architecture

CMTS and HPM

The authors assert that the bulk of non-real time OFDM Profile management functions can be implemented in the HPM. In the proposed functional split the CMTS responsibilities include:

- Periodic and on demand collection of signal quality measurements and error reports from CMs and the CMTS's upstream receiver
- Real-term evaluation of error reports and necessary profile adjustments to address urgent issues, such rapid profile downgrades.
- Initiation of on demand profile test procedures with CMs and collection of test results
- All protocol interaction with Cable Modems

HPM responsibilities include:

- Implementation of complex and CPU intensive functions to analyze signal quality measurement data

- Determination of the optimal set of OFDM profiles, backup OFDM profiles, and the most common denominator profile, referred to a profile “A” in DOCSIS 3.1
- Evaluation of error reports for the purpose of long term profile adjustments

The CMTS and HPM interact through an abstract application programming interface (API). We will refer to this interface as HPM API. The HPM API can be divided into several functional components:

- **The channel registration** component, through which the CMTS registers and unregisters its OFDM channels and their attributes with HPM. This is a process roughly analogous to the resource registration process described in Edge Resource Management Interface [4] (ERMI) specification to manage QAM channels. For each OFDM channel the CMTS communicates to the HPM the channel parameters, current OFDM Profile settings, and dynamic changes to those parameters as well as the channel’s unique identifier and HFC topology information.
- **The device registration** component through which the CMTS informs the HPM about the CMs which are using registered channels and their current OFDM profile assignments.
- **The signal quality analytics** component through which the HPM can request the CMTS to deliver a variety of diagnostic and performance information which may be useful in evaluation of OFDM Profiles. The set of analytical data includes performance metrics defined in DOCSIS specification such as RxMER collected from Cable Modem or CMTS upstream receiver as well as other performance indicators, for example LDPC performance statistics or upstream pre-equalizer coefficient settings. The data flowing through the signal quality analytics interface constitutes the bulk of information exchanged between the HPM and the CMTS. The throughput of the exchanged data can reach the levels of many megabits per second for each CMTS.
- **The test and command** component through which the HPM communicates to the CMTS OFDM Profile candidate parameters, requests from the CMTS to test OFDM Profile candidates and through which the CMTS delivers the results of requested profile tests.

The HPM API with logical partitioning is shown on Figure 2.

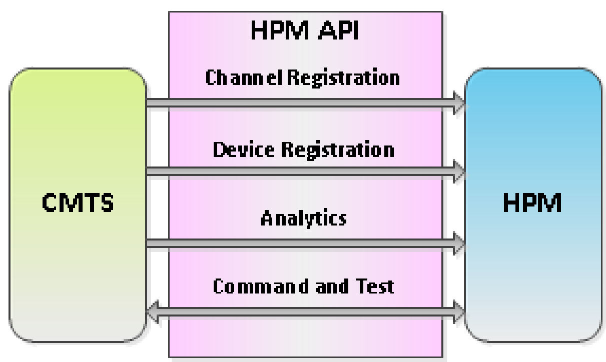


Figure 2 – HPM API

<u>Possible Extensions</u>

HPM application, initially developed for traditional CCAP/CMTS could become a building block of a future, virtualized CMTS. HPM could be integrated and operate in concert with other HFC management applications, such as Proactive Network Maintenance PNM server, explained in [2].

<u>Is HPM suitable for an SDN application?</u>

Next, we'll evaluate whether the OFDM Profile management functionality meets rudimentary criteria for a SDN application. We believe that OFDM Profile management as well as the presented HPM concept fit well into the mold of SDN application.

- OFDM Profile management can be broadly categorized as custom control plane functionality.
- As we discussed throughout the paper, OFDM Profile management involves complex, highly customized SW.
- The HPM application can be efficiently isolated or abstracted from other components of the CMTS system as we have demonstrated earlier, by outlining HPMI.
- There are few real-time processing constraints on OFDM profile management.

Standardization of HPMI

The benefits of standardization of network interfaces cannot be overstated. Virtually all interfaces between components of a modern network, including external interfaces of a CCAP/CMTS are based on industry standards. Open standards cover not only the external interfaces but also selected internal interfaces between components of a CMTS. The majority of Cable Operators networks deploy equipment, including CMTSs from multiple vendors. Interface standardization is rudimentary in enabling multivendor interoperability and reducing deployment costs. Undoubtedly, if HPM is to be developed and adopted as a decoupled cloud application, its interface to the CMTS will be formally defined. HPMI standardization will benefit CMTS vendors, the prospective vendors of HPM application software and ultimately, the Cable Operators.

The benefits of the proposed idea

Finally, let us review the benefits of the proposed approach.

Moving OFDM Profile processing from the embedded environment of a CMTS to the data center provides benefits generic to SDN and virtualization; those include the acceleration of software development and improved feature velocity, shorter test cycles, fewer memory constraints and scalable processing power. Once the data gets into the cloud it is generally easier to manage it, for example to archive it or perform historical analysis on it.

The application specific benefits include elements of CapEx and OpEx reduction:

- The removal of the bulk of OFDM Profile processing functions lessens CMTS processing burden and lowers CMTS memory requirements, resulting in lower equipment cost.
- It leverages the more sophisticated application development environment (eg commercial data bases) and much lower cost of generic processing and storage and available in the virtualized data center.
- Operations can be simplified because HPM as a cloud application makes dynamic OFDM Profile management possible, thus eliminating the need for complex and error prone OFDM Profile configuration settings in the CMTS.
- A decoupled, centralized HPM application will execute a single, consistent set of OFDM Profile processing algorithms and offer a single set of configuration knobs to control them even when serving CMTSs from different vendors. Centralized configuration and unified processing algorithms further help in operational simplification.

- HPM as a cloud application can be directly integrated with other HFC management applications, for example the PNM servers, becoming an integral part of the HFC plant management and service monitoring ecosystem.
- HPM application developed initially for a traditional CMTS/CCAP can be reused as a building block of a future, virtualized CMTS.

Conclusion

In recent years the Cable Industry has embraced two areas for innovation and significant investments: DOCSIS 3.1 and SDN. These areas appear to be completely unrelated. DOCSIS 3.1 aims at the physical network capacity optimization and scaling. DOCSIS 3.1 goals have been accomplished at the cost in increased network complexity. SDN intends to simplify the network by allowing operators to abstract control plane functionality from physical network nodes and implement them in a virtualized environment. Increased complexity of DOCSIS 3.1 represents a new opportunity for application of SDN concepts. The paper presents a cogent case for decoupling one of DOCSIS 3.1 control plane functions, OFDM Profile management and for implementation as an SDN application.

References

[1] DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.1-I02-140320, CableLabs, 2014

[2] DOCSIS 3.1 PHY Physical Layer Specification, CM-SP-PHYv3.1-I02-140320, CableLabs

[3] The Power of DOCSIS 3.1 Downstream Profiles, John T. Chapman, NCTA 2013

[4] DOCSIS Edge Resource Manager Interface Specification, CM-SP-ERMI-I04-110623, CableLabs

List of Acronyms

CM – Cable Modem
CMTS – Cable Modem Termination System
CCAP – Converged Cable Access Platform
DPD – Downstream Profile Descriptor
HPM – HFC Profile Manager
HPMI – HPM Interface
IUC – Interval Usage Code
NCP – Next Codeword Pointer
ODS – OFDM Downstream Spectrum
OFDM – Orthogonal Frequency Division Multiplexing
OFDMA– Orthogonal Frequency Division Multiplexing with Multiple Access
OUDP – OFDM Upstream Data Profile
PNM – Proactive Network Maintenance
SDN – Software Defined Networking
UCD – Upstream Channel Descriptor

Leveraging Wideband and Full Spectrum Receiver Capabilities to Develop and Utilize Software-based Tools for Remote Spectrum Analysis and Troubleshooting

K. Scott Helms
ZCorum
Jim Koutras
MaxLinear

Abstract

Spectrum analysis tools have been in use in the cable industry for troubleshooting and design purposes for a long time. Traditionally this meant buying and maintaining large numbers of meters, spectrum analyzers, and other specialized tools. This increased expenses and training time, and limited the number of people in any organization who could leverage spectrum analysis.

Today, using new software in conjunction with newer components in modems and set-top boxes, remote spectrum analysis is possible without purchasing specific test equipment. This allows spectrum captures to be performed from anywhere and the exact downstream characteristics in a customer premise can be seen and analyzed without having to drive to that location or the need to gain access to the in-home wiring.

NEW CAPABILITIES IN REMOTE SPECTRUM ANALYSIS

New DOCSIS 3.0 modems using wideband or Full-Spectrum Capture™ receivers have the ability to capture and report local (end-user) cable-plant spectra, providing Cable and Multi-System Operators (MSO) new freedom in remotely monitoring, managing, and troubleshooting their networks. Embedded remote spectrum monitoring and diagnostic data is delivered to a software application for performance analysis. This remote spectrum monitoring technology allows rapid capture of frequencies from 54 MHz to 1GHz using dedicated on-chip FFT hardware to transform time-sampled input signals to the frequency domain for subsequent user analysis and interpretation.

When the remote spectrum monitoring function executes, the tuner inside the cable modem receives the spectrum information and is commanded to fill an on-chip buffer with time-series data. Various windowing functions can be applied to time-series data on either side of the buffer to smooth discontinuities at the window boundaries (thus reducing spectral leakage) and an FFT is performed. The frequency-domain data read from the device are then interpolated as necessary to achieve the requested frequency step size and finally converted to units of dBuV.

This means that not only can we see the spectrum wherever we have a modem that employs wideband or Full-Spectrum Capture tuners; we can also look at multiple sources at the same time. This allows us to compare multiple devices in the same home to isolate an in-home problem or compare modems in the neighborhood to determine if impairments are specific to one drop or are more systemic.

In terms of data flow, the spectrum data from the receiver is sent to the DOCSIS 3.0 modem SoC or MPEG decoder SoC depending on the application. In the case of a DOCSIS modem or STB with an embedded DOCSIS modem, the data can be transferred to a CMTS/CCAP via a DOCSIS 3.0 MIB. In the case of a STB without DOCSIS, standard protocols like TR69 and SNMP can be used to transmit the data to a headend/operations center.

BENEFITS OF REMOTE SPECTRUM ANALYSIS

There are significant benefits in developing or purchasing a remote spectrum analysis tool.

Reduce travel time.

Reducing the amount of time, especially travel time, needed to capture spectrum information at specific locations in the cable plant eliminates one of the major inconveniences of using traditional, equipment-based spectrum analysis tools. The burden of carrying spectrum analysis equipment to the location where you want to make the measurements will no longer be necessary.

Measuring levels in a combining plant can be relatively straight forward and quick, but in many cases we need to measure a forward path at specific locations. We generally have two choices in order to accomplish this. We can either drive the tools to that site or we can install probes ahead of time and hope that we have enough coverage so that we can use our probes for most forward path testing.

By leveraging wideband or Full-Spectrum Capture tuners in modems and set-top boxes, we basically turn all of our compatible Customer Premises Equipment (CPE) into probes for spectrum analysis without having to purchase dedicated hardware. We also eliminate the costs involved in deploying multiple dedicated probes.

Reduce the cost of spectrum analysis tools.

Reducing the costs associated with spectrum analysis tools also eliminates the need to replace dedicated hardware as DOCSIS technology changes. This is especially important as we look forward to DOCSIS 3.1. Today our spectrum analysis tools have to incorporate a large amount of custom circuits and electronic components in order to perform the capture of the spectrum data and present it in a usable format. This causes the costs of even basic meters to be hundreds of dollars while more advanced devices are substantially more expensive.

By leveraging the existing RF tuners and demodulators in our set top boxes and cable modems for spectrum analysis we can now move the display of that information to consumer grade tablets and smart phones. This lets us take advantage of the competition in consumer electronics, which greatly lowers the costs of the devices our technicians have to carry and reduces the number of separate devices they have to carry with them. Because we are now using the RF components already in our modems and set top boxes we no longer have to replace or upgrade meters to account for changes in DOCSIS protocols. As we deploy new modems and set tops in DOCSIS 3.1 we have the ability to leverage those devices as probes without having to purchase new hardware.

Increase the number of technicians who can leverage spectrum analysis.

By increasing the number of people who can effectively use spectrum analysis in troubleshooting, we can decrease training time for new technicians and leverage the comfort level that younger employees have with tablets and smart phones. This serves a two-fold purpose and is driven by two different factors. One of the issues for many MSOs is that the cost of spectrum analysis tools limits how many meters and analyzers are purchased, which of course limits the number of people who can use those tools at any one time.

The other challenge is that few new technicians are comfortable using a spectrum analyzer and many don't use them effectively even when they have access to a meter or analyzer. Training time and costs, especially with contractors in the mix, frequently prevent good tools from being used to their full potential.

One of the advantages of moving the technician interface to tablets and smartphones and away from dedicated hardware is the level of comfort that most new technicians will have with the interface. These current interfaces are much more intuitive than most meters and will make it easier to use and incorporate more features to help the user identify problems.

PREREQUISITES BEFORE YOU BUILD

If your plan is to develop your own remote spectrum analysis tool, the following prerequisites need to be in place.

A method to handle modem communication (SNMP).

You can manage the modem SNMP communication as well as the communication to the spectrum client utilizing a custom-built centralized server that can run in the head-end or in a MSO's datacenter. The server handles discovering modems that support spectrum capture as well as authentication and authorization of clients that should be able to run spectrum analysis. The server can be available on the public Internet (our server supports TLS/SSL encryption) or behind a corporate firewall so that a VPN must be used to reach it. Spectrum data can also be sent to a device in the home network. For instance, a tablet running spectrum analysis software connected via Wi-Fi allows a technician to analyze the spectrum while troubleshooting in the home.

A method to discover modems that support docsIf3SpectrumAnalysis OIDs.

A method to poll through the modem pools looking for devices that correctly respond to a SNMP GET request on docsIf3CmSpectrumAnalysisCtrlCmdEnable OID (Object Identifier). Figure 3, on page 5 below, shows an example of a server-side admin panel that is displaying four modems in a lab test network that have responded to the SNMP GET request. This data can be refreshed as frequently as the MSO would like by allowing for a setting in the server's configuration file.

A client to display the information.

By implementing a client-server approach with the client being apps for iOS and Android, we're able to run them on standard tablet and smartphone platforms. This allows technicians to simply connect to the server from anywhere in the world and then request the server to run the spectrum analysis functions and relay the results to the client for display.

Applications utilizing the spectrum data can include but are not limited to:

- Channel Loading/Response
- Tilt Measurement
- Ingress Detection
- Remote Network Monitoring
- In-home Spectrum Analyzers

A client distribution method in place

The most popular sites for distribution and follow-on updates are the Google Play Store and iTunes App Store. This will also be where the client subscription is managed. One of the biggest challenges for anyone involved in software distribution is dealing with updates. Keeping software up to date on the various mobile platforms is even more difficult, which is why choosing to tie into the most popular app stores makes sense.

In addition to smart phone and tablet clients, offering desktop clients via the iTunes market for Macs and the Windows Store for Windows 8 and later PCs will be beneficial for in-office use.

Modems in place that currently support docsIf3CmSpectrumAnalysis

Remote spectrum capture capability is supported on the hardware side by modems that utilize the Intel Puma 5 and 6 chipsets with MaxLinear wideband and Full-Spectrum Capture receivers such as the MxL261, MxL265 and MxL267.

Note: While the hardware may support wideband or Full-Spectrum Capture tuners, the firmware on the cable modem might require updating to a version that supports docsIf3SpectrumAnalysis OID.

EFFECTIVENESS OF REMOTE SPECTRUM ANALYSIS

In testing we found that remote spectrum analysis was usually more effective in detecting and responding to impairments than traditional approaches of debugging network problems. The ease of running an analysis made it an earlier part of the troubleshooting process. The ability to compare multiple sources of data at the same time greatly improved technicians' ability to discern between common problems and channel tilt, ingress, suck outs, and other distortions of the forward path (see Figures 1 and 2). Technicians' ease of use and quicker adoption will rely on developing user interfaces that allow them to quickly find RF issues.

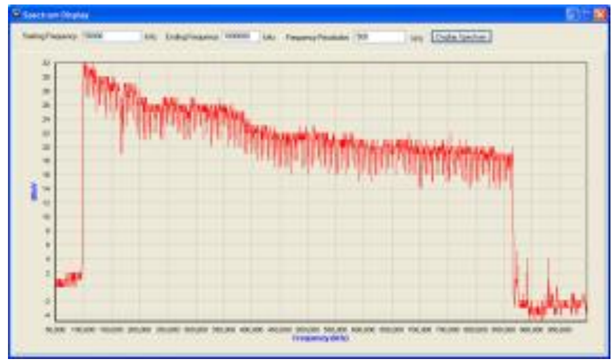


Figure 1 Course View - Channel Tilt

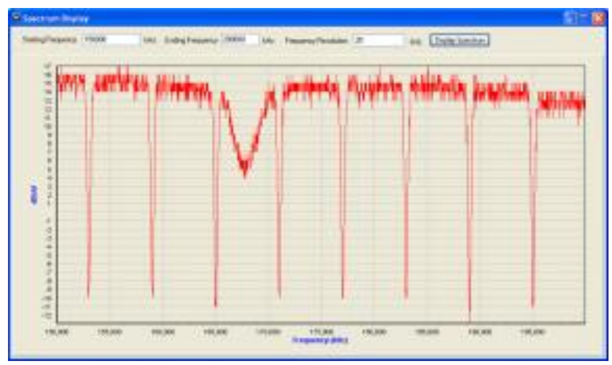


Figure 2 Zoomed View - Suck out

REMAINING CHALLENGES

The largest remaining challenge is deploying more modems in the field that have both the hardware and firmware to support remote spectrum monitoring through wideband or Full-Spectrum Capture tuners. Many tens of millions of existing modems already have the needed hardware but deploying the firmware upgrade to support docsIf3SpectrumAnalysis OIDs is an ongoing exercise.

As this technology is more widely adopted we expect that this problem will largely disappear with all of the major modem vendors pushing to get the firmware in place as rapidly as possible.

FUTURE APPLICATIONS

We have only scratched the surface of what's possible with remote spectrum monitoring and wideband or Full-Spectrum Capture receivers. We expect that future applications will be able to proactively detect problems on the forward path so that the tools can become more proactive in nature and we no longer have to wait for a customer in a neighborhood to tune to a specific channel, notice the impairment, and call to report it. We also expect to see Full-Spectrum Capture receivers implemented in other kinds of equipment including terrestrial and other technologies. Another area for future development is giving technicians more guidance about problems to shorten the troubleshooting process.

Figure 3 Example of a server-side admin panel

AUTHORS' CONTACT INFORMATION

Scott Helms, VP of Technology
ZCorum

4501 North Point Parkway, Suite 125
Alpharetta, GA. 30022
Tel 678-507-5000 Ext 5551
khelms@ZCorum.com
www.zcorum.com

Jim Koutras, Director of Product Marketing
MaxLinear

2051 Palomar Road, Suite 100
Carlsbad, CA 92011
Tel 760-517-1303
jkoutras@maxlinear.com
www.maxlinear.com

LOW-LATENCY IPTV NOTIFICATIONS WITH MINIMAL SERVER IMPACT

Gary Horton
Time Warner Cable

Abstract

IPTV platforms provide value-add with notifications such as Caller ID and EAS alerts. An efficient architecture provides timely delivery of these messaging events without unnecessarily binding resources in their absence. The technology choice to support this is influenced by message frequency, latency requirements and client population size.

Among the contexts combining those variables, low latency delivery of infrequent messages becomes more challenging for the server as client size increases. A system is described here targeting that scenario, scaling gently with increases in client size and latency constraints. The system impacts server resources only when a message is available.

INTRODUCTION

Background

With HTTP as the application protocol, notification across the Internet relies on a client-initiated request. This presents the problem of how to convey event information from a service to a client, for example an EAS service delivering a flash flood alert to an HTML5 browser.

Periodic client requests provide opportunities for the server to send the EAS alert. However, with such "short poll" solutions, the client can suffer from high latency and untimely message delivery if polling frequency is too low; or server capacity is challenged when frequency is high. In particular, the aggregate cost of setup, maintenance and teardown of TCP connections[1] increases as a function of the polling frequency and client size.

So-called COMET[2] or "long poll" solutions have been developed to reduce the TCP activity. With these approaches, a connection is established and held by the server, either until a message is available or with periodic timeouts. Streaming servers can hold the connections open persistently and deliver ongoing messages as information becomes available. These techniques not only reduce TCP setup and teardown cost, but also support the requirement for low-latency message delivery. However, while setup and teardown cost is lessened, maintenance remains: system provisioning requires concurrent connection capacity as a function of client size. This incurs cost in RAM[1] and brings other considerations[3].

The benefit of the RAM cost is a function of messaging frequency. That is, in a context with frequent message availability such as a stock exchange, the capacity is well utilized. However, that use declines with decreasing messaging frequency. Persistent connections are increasingly idle, becoming an unused commitment.

Examples calling for low-latency delivery of infrequent messages include EAS, Caller ID and location-aware advertising.

LightWeight Polling

The traffic profile combining timely delivery of infrequent messaging events is the context for this discussion. Targeting that context, "LightWeight Polling" (LWP) is a communication protocol described here, with performance test results suggesting it provides low-latency messaging while moderating many shortcomings of traditional short poll systems. In particular, significant savings in network and CPU activity have been measured and will be discussed in detail. In short, LWP accomplishes this by opening its destination port for clients only when a message is available. Resources are used primarily for those events while terminating any TCP connection attempt immediately in their absence.

Acronyms

For purposes of this discussion, the following acronyms are used:

CPU	Central Processing Unit
EAS	Emergency Alert System
HTTP	HyperText Transfer Protocol
JVM	Java Virtual Machine
LP	Long Poll
LWP	LightWeight Poll
RAM	Random Access Memory
SP	Short Poll (traditional polling)
TCP	Transmission Control Protocol
TTL	Time To Live

ARCHITECTURAL CONSIDERATIONS

Use Cases

Testing done here is intended to simulate broadcast messages that occur infrequently but call for timely delivery.

The point-to-point class of use cases, for example Caller ID and soft remote, imply private instead of broadcast messages. This is not an intended use for LWP as it is currently implemented. The current prototype relies on a *port-per-topic* protocol, using a single destination port to host a given topic. For example, the "topic" for Caller ID is the individual phone number, and this requires N ports for N phone numbers. Supportable Caller ID client size then becomes a function of available ports[9]. While larger values of N appear to be a variation of the unused commitment problem, LWP capacity is not dedicated in the absence of messaging since ports are not opened. Given this, LWP applied with point-to-point may show benefits, but experiments remain to-date undone.

Latency requirements that are more relaxed are not examined here since this is not a particularly challenging context.

The class of use cases calling for frequent messaging, for example a stock exchange, are also not an intended application of LWP systems. As frequency of messaging increases, LWP systems increasingly resemble standard polling, and long poll systems become increasingly favorable.

System Requirements

An example set of requirements where an LWP system would be suitable include the following:

1. **one-second-latency:** The system must support latency requirements of one second or better after message availability.
2. **120s-messaging:** System capacity requirements must grow linearly or better as a function of messaging frequency with a period of 120 seconds or more.
3. **graceful-scalability:** System capacity requirements must grow in sublinear fashion relative to client population size.
4. **dynamic-capacity:** The system must release CPU and network resources when messaging work is not required.

Candidate Approaches

Given the target use cases and system requirements, there are various candidate solutions included in testing done here. The range of candidates was not intended to be exhaustive, for example omitting Web Sockets[4]. The intention was to determine if short polling with an LWP approach is a viable alternative to long poll.

In terms of the system requirements, the **120s-messaging** context is not challenging taken alone. However, low-frequency messaging undermines the **dynamic-capacity** requirement with a long poll system, which by design is unwilling to release capacity in the absence of messages. The **one-second latency** requirement is problematic with a traditional short poll system in terms of support for **graceful-scalability**[5]. This is true regardless of messaging period, due to impact on RAM and network. Increasing client sizes with a long poll system would require corresponding growth in concurrent connection capacity, failing to support **graceful-scalability**.

LIGHTWEIGHT POLLING: DETAILS

Overview

The “LightWeight Polling” mechanism described here is not sensitive to high rates of client polling, with the destination port opening only when messages appear. Without messaging, TCP connection attempts are terminated immediately. This is done with a connection reset issued from the network stack on the LWP host[6], incurring modest impact solely on that stack.

At message time, TCP communication in LWP flows as usual[7]. When messaging events are infrequent, the cost of TCP setup, maintenance and teardown are reduced relative to traditional short poll. With lessened demand on server, network capacity and CPU, these resources are available for other activities when no messages are available. LWP capacity needs are largely determined by messaging frequency.

The LWP protocol side-steps the provisioning needs of long poll solutions, where capacity for “always-open” connections must grow as a function of population size. These connections are idle a majority of the time when messages are infrequent.

The combination of “strict latency” (timely delivery) and infrequent messages is the most suitable context for an LWP system. The remainder of this discussion assumes strict-infrequent as the context.

Communication Flow

The sequence diagram in Figure 1 illustrates the typical LWP flows. The timeline is captured to the left of the Client, with system State changes to the left of the Publisher. To elaborate on the numbered steps:

1: A web Client “subscribes” for “notification” when an EAS alert occurs. The notion of EAS alerts is regarded as a “topic”, and the client request is referred to as a “subscription”. Port 8088 is used for the subscription, with this port remaining open for the duration of LWP execution.

The LWP service responds to the Client subscription with information on the port for polling and a TTL (Time To Live) – that is, how long any EAS alert will be cached for availability. The TTL is the maximum recommended polling interval for this particular topic.

In this example, the port to be used for polling is 8091 and the TTL is 10 seconds. If the Client polls more frequently, it is less likely to miss any messages and may receive messages with less delay. If polling is done less frequently, the Client may miss messages.

2a-2c: The client is motivated to get any EAS alerts with little delay, polling every second. Note from the current State (No Alert, Alert, …) that each of steps 2a, 2b and 2c are executed when no alert is available.

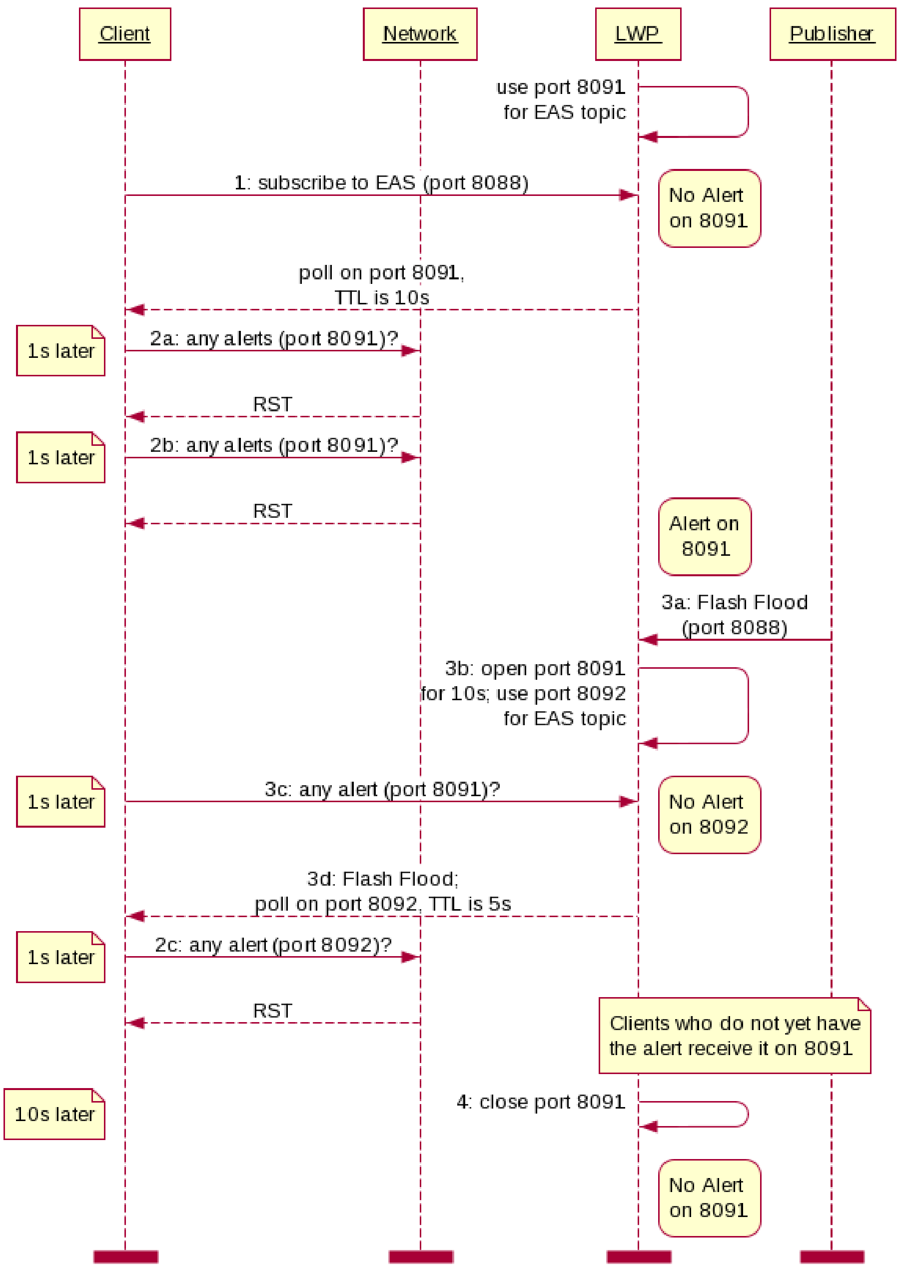


Figure 1: Typical LWP Communication Flows

The impact from one-second polling is quite low as quantified in the TEST RESULTS section. When no message is available, LWP is not aware of the polling frequency, since the Network terminates the connection attempt immediately with a TCP reset.

3a: The Publisher provides an EAS message to LWP at 3a using port 8088. LWP responds at 3b by opening port 8091 with the message and establishing port 8092 for subsequent messages. Note that port 8092 is not opened until that next message arrives. The next time any Client requests an EAS alert on 8091 (3c), the message is delivered (3d), and LWP informs that client to poll for subsequent messages on port 8092 with a TTL of 5s.

Note that the port and TTL have changed. LWP changes the port so that this first client is not repeatedly asking for the same message on port 8091, which remains open for other clients during the TTL. This would undermine the LWP goal of “no unneeded work” and result in duplicate messaging to the first client. Other clients that have not yet received the message will find it at port 8091 for the next 10 seconds. Meanwhile the first client will now poll against 8092.

This *port-per-message* protocol requires a client to resubscribe if it has not contacted LWP within the TTL period. If a message was delivered during that TTL, the port will have changed for the EAS topic. Alternately, clients can resubscribe on port 8088 on a regular basis as insurance against changed port numbers, and a message numbering scheme can be used to deliver potentially missed messages to clients.

Note that proof-of-concept for the port-per-message protocol remains undone. Tests done here use a *port-per-topic* implementation of LWP, and this did result in duplicate messages for a given client and additional impact on LWP during the TTL period. A port-per-topic approach is sufficient for point-to-point, since this is equivalent to port-per-client and no redundant requests for the same message will happen (since the port is closed immediately after message delivery).

Finally, as illustrated in the sequence, LWP optionally changes the TTL to adaptively manage its resources as needed. TTL adjustments are also used to encourage higher polling frequencies. For example, an initial EAS alert is likely to soon be followed by more.

TEST CONSTRUCTION

Test Setup

Test results provide some insight into why traditional short poll systems are considered inefficient and how LWP addresses the problems. An LWP server prototype was constructed using the netty (Java-based) framework[8]. In the interest of reducing the variations, SP and LP server prototypes were also constructed with this framework. The SP implementation represents a traditional short poll system for purposes of testing and subsequent analysis in this discussion.

Each server was load tested in separate executions on a physical hardware machine dedicated to the tests. The machine has two Intel E6750 CPUs running at 2.66GHz. The JVM was configured at 256MB minimum and 2GB maximum heap size.

SP and LWP Test Phases

The SP and LWP systems were executed in various *test phases* as described below, and as annotated with the timeline in Figure 2 in the TEST RESULTS section.

The test phases **Q**, **P**, **S**, and **M** are, briefly:

Q: the system is quiescent
P: polling is underway
S: server is running, handling polling only
M: messages are being delivered

The test phase details are as follows:

Q: The system is in quiescent state for at least 180 seconds to establish a system baseline.

P: Polling is initiated for 300 seconds to establish a polling-only baseline. For each test, either 3K or 30K client threads are initiated in sets of 3K at a time. For the 30K population, a 6-second delay is added between each 3K set, with the intent of adding some temporal spread to the client request traffic. This results in a “bumpy” rather than a uniform distribution of client request traffic, due to staggering and resource contention delays with thread startup.

S: At the first of two “S” phases, a given type of server (SP or LWP) is started and opens port 8088 to accept publisher messages. SP opens port 8091 to accept client requests. LWP opens that port only when a message appears, and then only for the TTL period.

Polling against SP and LWP continues without any message delivery for 600 seconds to establish a no-messaging baseline.

M: Messages are delivered five times, once every 120 seconds. Once a message appears, the SP and LWP servers make it available at port 8091 for the next 10 seconds.

Clients of the SP server receive an "HTTP 204 No-Content" response in the absence of messages. Clients of the LWP server receive a connection reset in the absence of messages. In each case, the client receives an "HTTP 200 OK" response if a message is available, and the client sends a new request one second after either a reset or response.

S: At the second "S" phase, the server continues running for 120 seconds after the last message, with polling against SP and LWP, but no message delivery. The server is finally halted to allow the system to return to quiescent state before the next test execution is started.

These executions were done first using SP and LWP against a 3K population, and then using SP and LWP against a 30K population. This was the approach for all tests and associated graphs that follow, except where otherwise indicated.

<u>LP Test</u>

A single test correlating RAM with concurrent connections was done for the LP server with the 30K population. For the LP server, all clients connect immediately on startup and remain connected until the first message arrives. The LP server delivers the message immediately with "HTTP 200 OK" to these clients, which reconnect one second later to wait for the next message.

<u>SP and LWP Graphical Output</u>

Each graph of test results displays a particular metric for a given population size. Each graph includes both SP and LWP executions:

1. 3K clients, SP and then LWP
2. 30K clients, SP and then LWP

As annotated with Figure 2, the red markers in each graph indicate the startup times of SP and subsequently LWP servers, with messaging events indicated by the groupings of light-blue markers. While both populations are shown for the bandwidth metric in Figures 2 and 3, only the 30K population is shown for any metrics with similar results for both populations.

The selected set of metrics captured includes the following:

1. Network activity, to include bandwidth, socket use, TIME-WAIT and connection reset levels
2. CPU idle time
3. RAM and heap size usage

Sampling of the metrics was done at five-second intervals for finer-grained analysis. Graphs were displayed using minimum or maximum depending on the metric, with 20-second resolution for simpler visualization.

To facilitate understanding of overall test results and graphical presentations, reference the annotations in Figure 2. This can be consulted as a template for subsequent graphs in the TEST RESULTS section.

TEST RESULTS

Network Impact

Maximum kilobytes of input and output bandwidth are shown in Figure 2 for 3K clients and Figure 3 for 30K clients, with SP and LWP executions in each graph. Figure 2 provides detailed annotations of test phases.

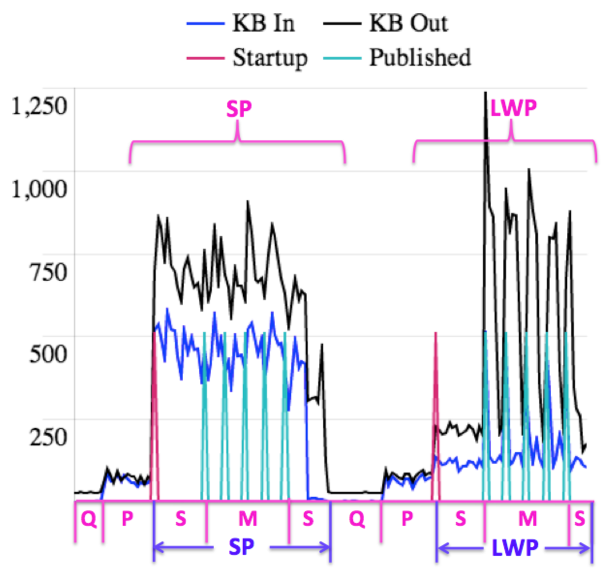


Figure 2: Bandwidth Impacts, 3K clients for SP and LWP executions

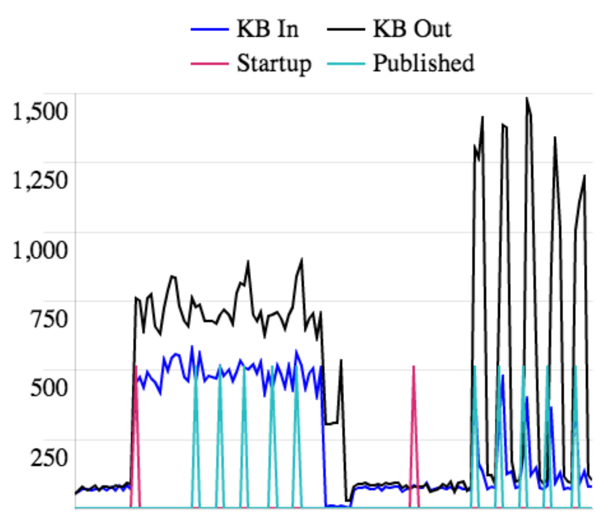


Figure 3: Bandwidth Impacts, 30K clients for SP and LWP executions

Variability in output peaks are likely due to the bumpy profile of client traffic. If capacity is to be provided for these infrequent peaks, it would be useful to reclaim capacity that is unused when there is no messaging. As illustrated in Figures 2 and 3, an SP environment exhibits bandwidth usage that varies in a narrow range near its highest levels, regardless of messaging traffic. This leaves less bandwidth margin to be reclaimed.

Alternately, bandwidth consumption for LWP stabilizes at about 40% of the SP levels in the absence of messaging traffic, engaging network capacity only when messages arrive. Additionally, the LWP bandwidth profile tapers shortly after the message event. This profile offers margin that can be applied to alternate capacity allocation. This margin increases as messaging frequency decreases. With an LWP system, this inverse correlation is independent of client population or polling rates.

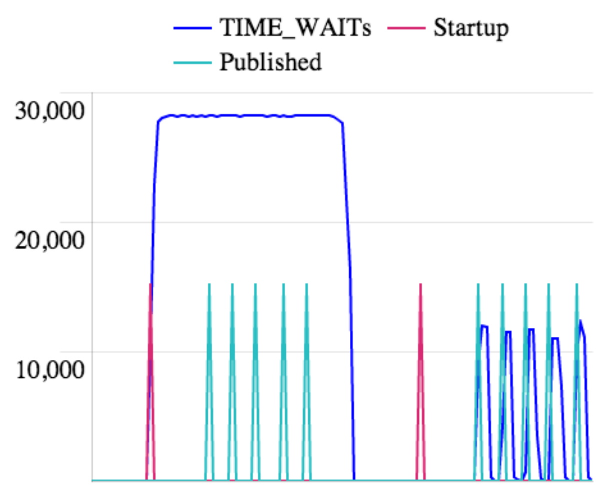


Figure 4: TIME-WAIT Activity, 30K clients for SP and LWP executions

The correlation with TIME-WAIT behavior can be seen in Figure 4. Because of configuration on the load generator machine, the limit on source IP-port tuples (which are stored in TIME-WAIT structures) was about 28K[13]. The chart shows that SP systems max out the potential TIME-WAIT levels for those connected source IP-ports immediately, keeping those levels high for its entire execution. However, LWP establishes fewer connections since the destination port is only open for brief periods, exhibiting the same reduced peak and post-message tapering for TIME-WAIT as seen with bandwidth. TIME-WAIT activity can have impact on scalability and performance in various ways, as discussed in the FOOTNOTES section[10].

Figures 5: Total Sockets Profile, 3K clients for SP and LWP executions

Figure 6: Total Sockets Profile, 30K clients for SP and LWP executions

Maximum levels of sockets used by SP and LWP are shown in Figure 5 for 3K clients and Figure 6 for 30K clients. As with bandwidth, these peaks drive capacity planning. SP peaks are higher relative to LWP for both population sizes. LWP use of sockets remains slightly elevated above baseline independent of population size, peaking at higher levels for the 30K population. The difference in absolute levels of socket use with the 30K population is about 50% greater with SP vs. LWP; this suggests that LWP would scale more gradually (providing spare capacity in terms of sockets) as the population increases. More full-featured load generation would be needed for higher confidence.

The engagement of TCP connection reset activity using an LWP system accounts for the reductions in bandwidth, TIME-WAIT and socket levels. This is illustrated with Figure 7.

Figure 7: Connection Reset Activity, 30K clients for SP and LWP executions

A TCP connection reset can be triggered by various conditions[11]; in this case it is due to an unopened destination port. This happens immediately at test startup, as seen in Figure 7 preceding both red startup markers; this is because tests are begun with the polling-only baselines without the server running. Immediately on SP startup, the reset rate drops effectively to zero since SP has opened port 8091 and is returning “HTTP 204 No-Content” with no messages available. However, the resets continue after LWP startup since LWP waits until a message is available to open the destination port.

CPU Impact

The timeline in Figure 8 illustrates the minimum levels provided by SP and LWP for CPU idle percentage.

Figure 8: CPU Impact, 30K clients for SP and LWP executions

Available CPU declines immediately at SP startup time, with no benefit to the delivery of messages. Idle CPU remains below the 40% mark for the duration of the SP run for both populations. The CPU exhibits no noticeable reaction at LWP startup, remains high in the absence of messaging and has minimums higher than the SP maximums with messaging events. LWP engages the CPU only when delivery of a message is needed, and unlike an SP system, that CPU usage is transient with a quick recovery after the message.

It would seem LWP should drop to the same levels of idle CPU during message events as seen for SP, but it remains significantly higher. This is due to thousands of No-Content responses from SP to clients continuing to poll at one-second intervals. While SP engages the CPU to tell the client that there is no message, LWP systems avoid this wasteful activity.

The differences in CPU availability are shown with a frequency distribution of idle percentages, for SP and LWP in Figures 9 and 10 respectively. LWP never uses more than 50% of CPU, while SP sometimes consumes 100%.

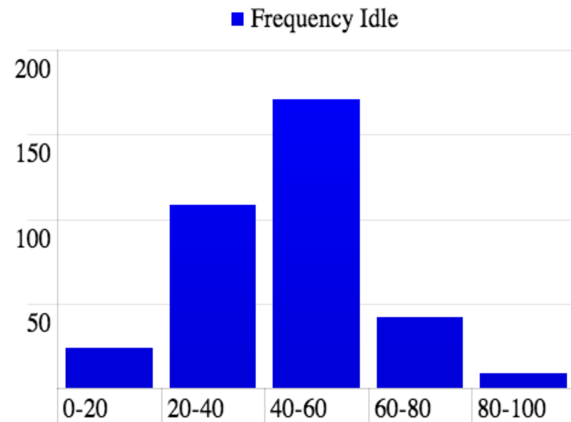


Figure 9: CPU Idle, Frequency Distributions, 30K clients for SP

With LWP there are many fewer 204 responses during messaging periods, since LWP closes its port at TTL expiration. This raises the question of why *any* 204 responses are seen, instead of all requests receiving either 200 OK or a connection reset. The reason is that the LWP implementation does not apply locks to the processing, allowing some requests into the request-handling pipeline while the port is open. By the time the last handful of these requests are processed by the handling logic, the TTL has expired and the message has been removed. In this case, as with the SP system, the response handler sends out the 204 responses.

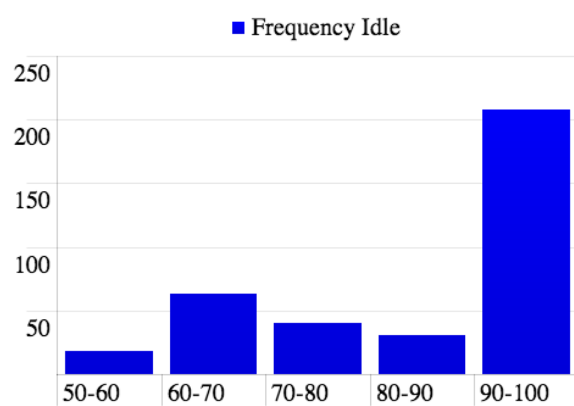


Figure 10: CPU Idle, Frequency Distributions, 30K clients for LWP

RAM Impact

There is only a slight difference in maximum RAM usage between SP and LWP systems, regardless of client size. Both plateau after all 30K clients have been engaged, as per Figure 11. The absolute values of RAM used by the SP and LWP systems are about 140MB and 170MB respectively.

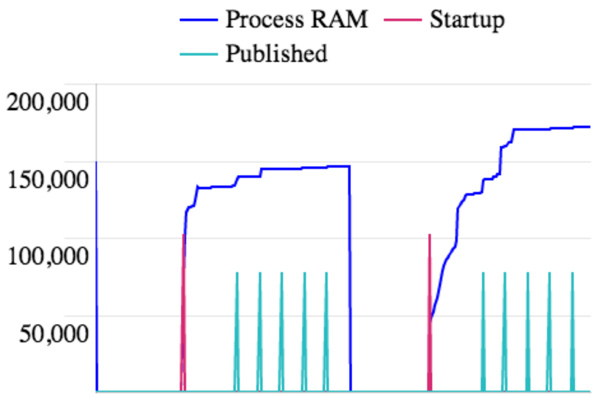


Figure 11: Process Resident Memory, 30K clients for SP and LWP executions

Figure 12 shows the RAM levels achieved by the LP implementation for the same 30K population, displayed with the incremental increase in connections as clients are started in 3K sets. Connection levels here are multiplied by eight to more readily visualize the correlation. The peak RAM level approaches 30% higher than LWP (220MB). Note that the LP execution did not include any messaging events; it was done solely to determine the influence of additional connections on growth in RAM.

Figure 13 shows the incremental changes in RAM size as connections increase. The RAM cost for each connection averages about 9K. For all 30K connections the increase totals about 275MB, which is primarily unused capacity in an infrequent-messaging context. Note the cost of RAM per connection found here should not be regarded as typical, since implementations, system tuning and other factors will vary[1].

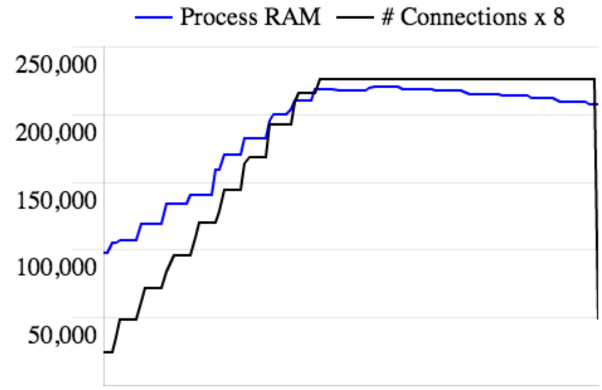


Figure 12: RAM Growth with Increases in # of Connections, 30K clients for LP only

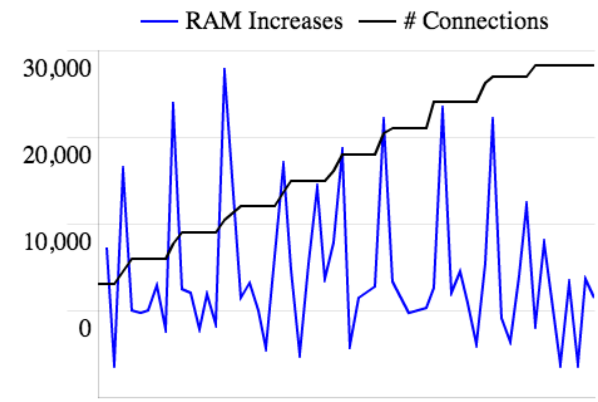


Figure 13: RAM Delta with Increases in # of Connections, 30K clients for LP only

Heap Size Impact

Java heap activity is shown using SP in Figure 14 and LWP in Figure 15, against the 30K population only.

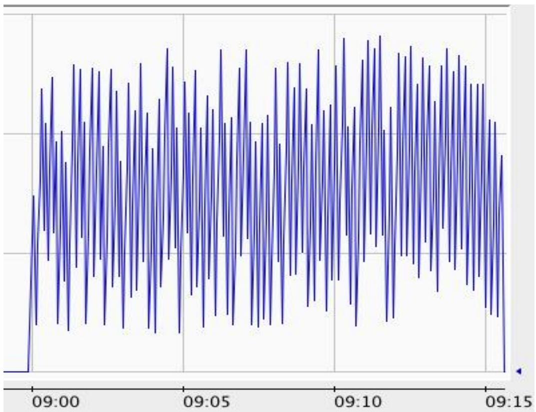


Figure 14: JVM Heap, 30K clients for SP

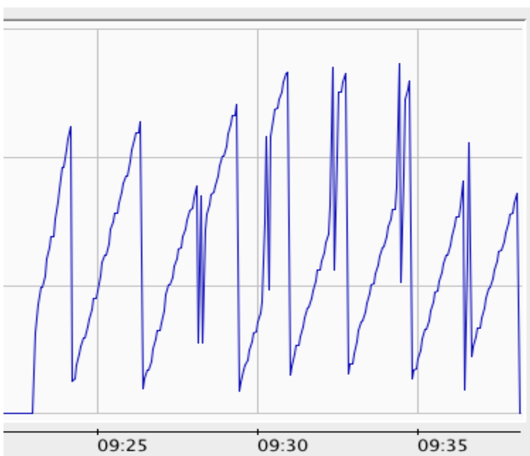


Figure 15: JVM Heap, 30K clients for LWP

Both heaps achieved approximately the same size, but this is a function of JVM and garbage collection tuning parameters. The variations in heap size appear to be a result of garbage collection activity, which is known to have measurable impacts on system response[12]. There is considerably less of that activity in the LWP system.

Another test execution over a longer period compares heap activity for SP in the first half, followed by LWP in Figure 16. SP has a great deal more activity but is losing ground over time, while LWP shows more desirable behavior. Additional testing would be useful here to capture not only heap size, but also generational activity and pause times for a more insightful analysis.

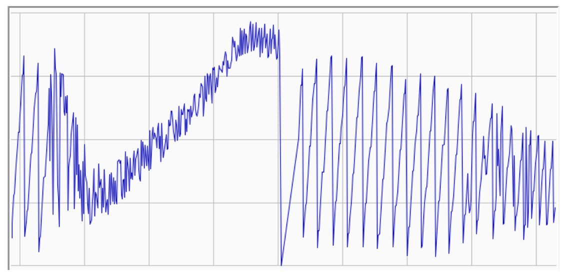


Figure 16: JVM Heap Size, 30K clients for SP and LWP executions

CONSIDERATIONS

All tests were done with HTTP (no SSL).

LWP Tradeoffs

LWP has many considerations, to include at least the following:

1. A specialized server is required.
2. This is prototype technology, barely emerging from proof-of-concept stages.
3. Clients must be aware that they will receive frequent connection resets.
4. Clients cannot determine if the service is down since connection resets are the norm. That said, clients – including load balancing systems that manage failover – can execute health checks against LWP at port 8088, which is open permanently. See the Class Diagram in the APPENDIX for details.
5. If SSL is used and the load balancer handles SSL offload, then there is an undesirable impact against that network component. This undermines the original purpose of LWP.
6. If SSL is used but the load balancer does not handle SSL offload, then load balancer

caching cannot be leveraged. This would call for SSL offload at LWP and a shared cache instead, assuming horizontal scale for LWP servers.

7. If SSL is used but the load balancer does not handle SSL offload, then routing decisions, session management based on headers, and any other application-layer functionality provided by the load balancer cannot be leveraged.
8. Messages must be cached for a TTL of at least the expected latency tolerance of the clients. However, LWP is intended only for “strict latency” environments, so that TTL is expected to be short.
9. If the message TTL is longer than the polling interval, the *port-per-message* protocol (which is to-date undeveloped) should be used. Otherwise, clients polling at high frequency will unnecessarily burden LWP, and these clients must distinguish duplicates.

The points in items 5-7 raise a question of whether LWP-type logic should live in a proxy server acting as SSL offloader, load balancer and centralized cache.

CONCLUSIONS

The architectural goal of LWP is to use resources only when needed to deliver messages, releasing that capacity for other purposes in their absence. Tests results show that LWP makes low-impact demands on network and CPU, moderating various undesirable characteristics of traditional short poll systems, and that it is well suited for strict latency, infrequent messaging contexts.

LWP also addresses the unused commitment problem (dedicated but idle capacity) of long poll systems in the strict-infrequent context. This justifies a second look at LWP-based short polling as an alternative to long poll for use cases calling for that context.

In short, LWP accomplishes this by opening a TCP port only when the message is available. With this approach, data center provisioning is now primarily driven by messaging frequency rather than client size or latency requirements. The resources supporting an LWP system are increasingly available for other activity as messaging frequency decreases.

Conversely, with increase in messaging frequency, LWP performance degrades to more closely resemble traditional short poll. The increasingly leveraged capacity of long poll provides more favorable cost-benefit in this context.

As with many technology solutions, LWP brings trade-offs and considerations. The most suitable use is narrow: strict latency requirements with infrequent messaging. The prototyping done to-date has moved beyond proof-of-concept to demonstrate promising performance results, but LWP remains for now a research project with numerous open questions to be resolved.

APPENDIX

Class Diagram

For reference at the implementation level, an overview class diagram describing responsibilities and collaborations is given here.

The *ContentListener* is the component that opens a port permanently (8088 as tested), listening for both Client subscriptions and messages from the Publisher. The *ContentListener* is responsible for creating one *TopicHost* for each supported topic.

The *TopicHost* component opens a port (8091 as tested) for its assigned topic only when a message is available (and only for a brief period as specified in the TTL). *TopicHost* delegates to the *MessageSender* to handle incoming client requests for that message.

Since the *ContentListener* always listening, its port should be used for health checks instead of the *TopicHost* port, which is usually closed.

The *port-per-message* protocol has not to-date been implemented; only *port-per-topic* is shown here.

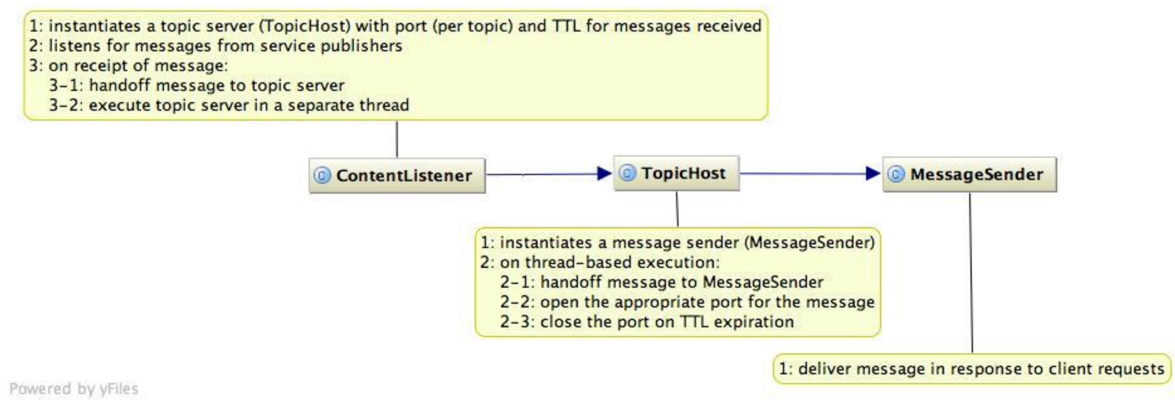


Figure 17: Class Diagram for netty Implementation of LWP

FOOTNOTES AND REFERENCES

[1] TCP Cost

http://en.wikipedia.org/wiki/Transmission_Control_Protocol#Resource_usage
http://stackoverflow.com/questions/7669293/performance-implication-of-creating-new-tcp-connection-per-message
http://stackoverflow.com/questions/4139379/http-keep-alive-in-the-modern-age
http://stackoverflow.com/questions/4840116/general-overhead-of-creating-a-tcp-connection
https://tools.ietf.org/html/rfc955
http://en.wikipedia.org/wiki/Transmission_Control_Protocol#Resource_usage
http://stackoverflow.com/questions/3173720/keeping-1000-tcp-connections-open-inspite-of-very-few10-20-actual-communicatio

TCP setup involves a three-way handshake, with teardown requiring a four-way exchange, adding latency to each connection. The size of each transfer is typically small enough to fit into one network packet. While network latency and impact on network stacks (CPU and kernel) comprise the majority of cost, bandwidth use is proportional to client population size. Estimates from others, as seen in some of the discussions referenced above, suggest varying levels of RAM cost per connection.

[2] COMET

http://en.wikipedia.org/wiki/Comet_(programming)

[3] Long-poll and Streaming Considerations

http://tools.ietf.org/html/draft-loreto-http-bidirectional-07#section-2.2

[4] WebSockets

http://en.wikipedia.org/wiki/WebSocket
http://www.websocket.org/
http://www.whatwg.org/specs/web-apps/current-work/multipage/network.html

[5] Options to Support “graceful-scalability”

http://www.w3.org/Protocols/rfc2616/rfc2616-sec8.html#sec8.1
http://web.archive.org/web/20100813132504/http://www.io.com/~maus/HttpKeepAlive.html
http://users.cis.fiu.edu/~downeyt/cgs4854/timeout
http://tools.ietf.org/id/draft-thomson-hybi-http-timeout-01.html
http://stackoverflow.com/questions/4139379/http-keep-alive-in-the-modern-age

Connection Keep-Alive is used by default in HTTP 1.1, but the customized clients and servers used here would require additional steps to configure it. Those steps were not done for the tests done here since the use case under test assumes infrequent messages. Given that assumption:

- Applying Keep-Alive to SP would blur the distinction with LP and its unused capacity problem;
- Applying Keep-Alive to LP would moderate reconnect cost, but was left undone since only the impact of persistent connections on RAM (vs. latency and CPU usage from setup and teardown) was measured for the LP implementation; and
- Keep-Alive is irrelevant for LWP by virtue of that system’s overall strategy (no connections are made without messages)

HTTP Pipelining is an interesting option that can be revisited in follow-ups, but was omitted here to reduce the number of test permutations.

[6] TCP Reset Sequence

http://blog.creativeitp.com/posts-and-articles/networking/exploring-idle-scanzombie-scan/
http://en.wikipedia.org/wiki/Transmission_Control_Protocol

When a port is not open, as done with LWP by design, the attempted setup consists of a SYN from the source peer, followed immediately by RST from the destination peer. This terminates connection setup immediately.

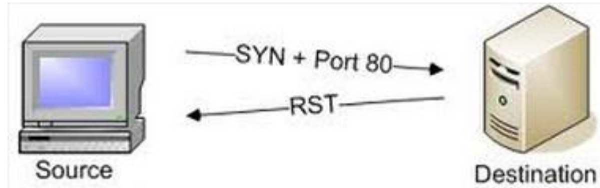


Figure 18: TCP Connection RST Sequence

[7] TCP Setup/Teardown Sequences

http://blog.creativeitp.com/posts-and-articles/networking/exploring-idle-scanzombie-scan/
http://en.wikipedia.org/wiki/Transmission_Control_Protocol

TCP setup resulting in a connection consists of three exchanges: the client sends a SYN to start the dialog, the server responds with a SYN, ACK, and finally the client sends an ACK. From that point the connection is in the ESTABLISHED state. Teardown uses four exchanges, with FIN from terminating peer, ACK and FIN from the remote peer, followed by an ACK from the terminating peer.

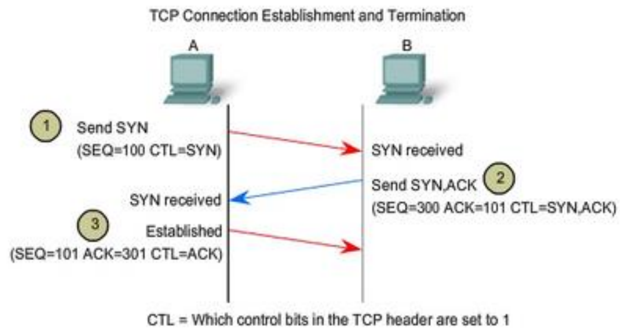


Figure 19: TCP Connection Setup Sequence

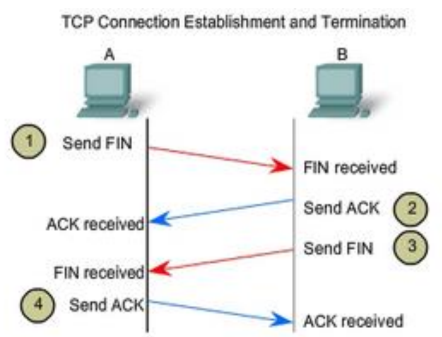


Figure 20: TCP Teardown Sequence

[8] netty

http://netty.io/
http://en.wikipedia.org/wiki/Netty_(software)

[9] Port Capacity on Linux

http://www.ncftp.com/ncftpd/doc/misc/ephemeral_ports.html
http://www.nateware.com/linux-network-tuning-for-2013.html

Linux systems are configured to allow ports with numbers ranging from 32768 to 61000 by default, whether ephemeral (outbound from clients) or assigned to applications (inbound to servers), as managed using `sysctl` with `net.ipv4.ip_local_port_range`. This range can be increased up to 65535, and possibly as low as 1024 depending on ports reserved for system services.

[10] TIME-WAIT

http://www.ncftp.com/ncftpd/doc/misc/ephemeral_ports.html
http://www.lognormal.com/blog/2012/09/27/linux-tcpip-tuning/
http://www.nateware.com/linux-network-tuning-for-2013.html
http://www.hjp.at/doc/rfc/rfc3102.html#sec_6.1
http://www.isi.edu/touch/pubs/infocomm99/infocomm99-web/
http://www.serverframework.com/asynchronousevents/2011/01/time-wait-and-its-design-implications-for-protocols-and-scalable-servers.html

The duration of TIME-WAIT (typically between 60-240 seconds by default) prevents the same client (source IP) from connecting to the same service (destination IP-port) using the same ephemeral port (source port). IP stacks will typically allocate different ephemeral ports for the next connection request from that client, but as the ephemeral port range decreases and the client polling frequency increases, it becomes more possible to run out of ephemeral ports and to be unable to establish a connection.

TIME-WAIT on Linux systems can be managed using `sysctl` with the parameter `net.ipv4.tcp_fin_timeout`. On the system under test in experiments done here, the TIME-WAIT timeout was 60 seconds.

[11] Causes of TCP Resets

http://stackoverflow.com/questions/251243/what-causes-a-tcp-ip-reset-rst-flag-to-be-sent
http://myaccount.flukenetworks.com/fnet/en-us/supportAndDownloads/KB/IT+Networking/protocol+expert/What_are_TCP_RST_Packets_-_Protocol_Expert
http://blogs.technet.com/b/networking/archive/2009/08/12/where-do-resets-come-from-no-the-stork-does-not-bring-them.aspx

[12] Java Heap and Garbage Collection

http://javabook.compuware.com/content/memory/analyzing-java-memory.aspx

[13] Socket and Connection Capacity on Linux

http://www.cyberciti.biz/faq/linux-increase-the-maximum-number-of-open-files
http://www.lognormal.com/blog/2012/09/27/linux-tcpip-tuning/
http://www.nateware.com/linux-network-tuning-for-2013.html

On Linux systems, maximum sockets and connections can be configured. The relevant parameters for connections are under `net.netfilter` and include the current count `nf_conntrack_count` and the maximum `nf_conntrack_max`, as managed by `sysctl`. The maximum number of connections configured for the system under test in these experiments is about 64K.

As these limits increase, kernel memory usage also increases. Sockets are also limited by number of maximum open file handles, as managed with `fs.file-max`.

While the server could have accepted 64K connections, the number of (outbound) sockets for the load generator in these experiments was limited by its ephemeral port range of 28K (as per the default range of 32768 to 61000). This can be revisited in subsequent testing to generate more client threads, but more useful effort would go towards establishing multiple source IP addresses instead.

LTE-UNLICENSED: AUGMENTING MOBILE DATA CAPACITY, BUT COEXISTENCE NEEDS CONSIDERATION

Bernie McKibben, Kyung Mun, Rob Alderfer
Cable Television Laboratories, Inc.

Abstract

Cable Operators (MSOs) are exploring new service opportunities based on wireless voice and data. There have been several interesting options proposed from wholesale small cell backhaul for licensed mobile operators to retail mobile services such as Wi-Fi First MVNO to widely deployed Wi-Fi for broadband services or widely deployed LTE neighborhood small cells in homes similar to community Wi-Fi. An interesting opportunity with significant strategic implications is emerging based on a nascent LTE technology.

Key players in the mobile ecosystem have recently put forth a set of proposals to the 3GPP standards body to enable LTE in unlicensed spectrum bands to effectively aggregate licensed and unlicensed spectrum use under a single radio technology ("LTE-Unlicensed" or LTE-U). This report highlights three different modes of operation of LTE-U and its likely deployment scenario in small cells. The report suggests potential opportunities that LTE-U may afford to cable operators and highlights challenges that lie ahead, mainly the ability of LTE to equitably coexist with Wi-Fi and other technologies in unlicensed spectrum. Timing of likely implementation for standardized and pre-standard releases are explored, and technical differences between LTE and Wi-Fi are detailed to highlight the further development of LTE-U that is required to enable coexistence necessary for continued open innovation in unlicensed spectrum.

WHAT IS LTE-UNLICENSED (LTE-U)?

The massive growth in data traffic on both mobile and wireline networks and the proliferation of smartphones and other connected devices continue to put pressure on network operators to increase capacity. For mobile network operators, spectrum is a fundamental resource in this pursuit. However, the licensed spectrum, especially the valuable low-frequency bands with low propagation loss traits, is limited, and is rapidly being exhausted by a dense and growing subscriber base. With a significant amount of unlicensed spectrum globally available in the 5GHz band,[1] the mobile operators and vendors are looking to use unlicensed spectrum to augment the capacity of licensed frequency carriers. In a 3GPP Radio Access Network (RAN) plenary standards meeting in December 2013, the proponents, including Qualcomm, Ericsson, Verizon, China Mobile, Huawei, and others, formally proposed "LTE-Unlicensed" (LTE-U) to utilize unlicensed spectrum to carry data traffic for mobile services with the initial focus on the 5725-5850 MHz band for this use. [2]

As the name implies, the goal of LTE-U is to extend LTE to unlicensed spectrum. LTE-U proponents seek to leverage unlicensed spectrum as a complement to licensed spectrum to offload best-effort traffic through the carrier aggregation framework that has already been defined in

LTE Advanced. In this so-called "licensed-assisted" access, a primary cell carries critical control signaling, mobility, and user data that demand high quality of service on licensed spectrum while less-demanding, best-effort traffic is carried on a secondary cell on unlicensed spectrum (Figure 1)[3]. In this framework, the use of unlicensed spectrum is always accompanied by a primary carrier on licensed spectrum.

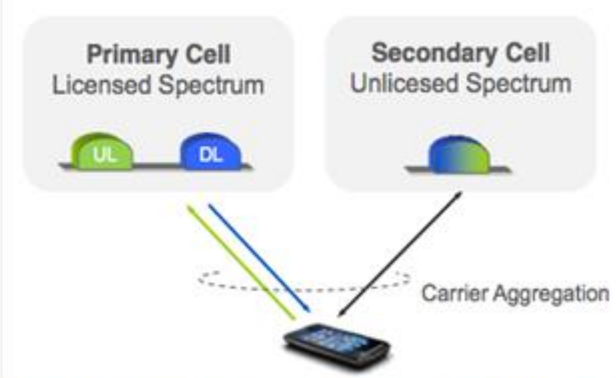


Figure 1 – "Licensed-Assisted" Access in LTE-U (Ericsson 3GPP contribution RP-140260)

There are two main deployment options for aggregating unlicensed spectrum to a licensed carrier to augment capacity, and a third option to run LTE on unlicensed spectrum standalone, without the primary cell on licensed spectrum. First, in the Supplemental Downlink (SDL) mode of operation, the unlicensed spectrum is utilized only for the downlink to augment capacity and increase data rates in a heavily trafficked downlink, which is typical in today's network use. Secondly, the Carrier Aggregation (CA) mode of operation allows use of unlicensed spectrum in both the downlink and uplink, just like typical LTE TDD systems. A key advantage of the CA mode is the flexibility of adjusting the amount of unlicensed spectrum resource that can be allocated for uplink or downlink.

In both the SDL and CA modes, unlicensed spectrum is used only for the data plane, and all the control plane traffic is handled through licensed spectrum in the licensed-assisted manner as depicted in Figure 1, to maintain operator control of both licensed and unlicensed spectrum resources. A primary focus of the LTE-U deployment scenario envisioned by the key proponents in 3GPP currently is the SDL mode as it affords simplicity of deployment and operation to mobile operators and less complexity on devices. A third option that has not been formally proposed in 3GPP is the Standalone (SA) mode, which offers the possibility of higher spectrum efficiency of LTE over unlicensed spectrum. In the SA mode, both control plane and data plane traffic are carried over unlicensed spectrum such that operators without licensed spectrum can potentially take advantage of high efficiency and seamless mobility handling in interference-limited scenarios that is a hallmark of LTE technology. In essence, LTE-U as currently envisioned in 3GPP, without an SA option, will tie use of the technology to wireless carriers with licensed spectrum, precluding use by those who do not have spectrum licenses.

Regulations typically limit unlicensed operations to a maximum transmit power of 1 watt or less[4]; therefore, LTE-U will be deployed as small cells in outdoor venues or indoor. While the LTE-U standard development progresses, it is likely that certain mobile operators are economically motivated to deploy pre-standard LTE-U small cells as early as end of 2015. These pre-standard solutions may leverage the LTE R10/R11/R12 carrier aggregation features and the 5GHz RF front end, (primarily from 802.11ac), which are already in place in vendor product plans. These underlying technologies will allow pre-standard LTE-U to ramp up to scale. To facilitate possible coexistence with Wi-Fi, the initial pre-standard[5] LTE-U small cells will likely be integrated with Wi-Fi in jurisdictions that do not require explicit Listen Before Talk (LBT) regulations, such as the United States, China, Korea, and a few other countries.

A key motivation for mobile operators in considering the use of LTE-U is that it helps to make more spectrum available under a single radio access technology. Today, some mobile operators have "carrier" Wi-Fi networks to offload traffic to ease the load on their primary mobile networks. By combining the use of both licensed and unlicensed spectrum under one (LTE) radio technology, the operators are looking to simplify the network management with a tighter integration under a single RAN and gain additional control over unlicensed spectrum resources. In essence, LTE-U enables tighter integration of both core LTE network infrastructure, and established management and security capabilities that it affords, along with a single radio technology over both licensed and unlicensed spectrum for better control of user experience.

<u>Opportunities for Cable Operators with LTE-Unlicensed</u>

LTE-U offers several interesting ways for cable operators without licensed spectrum holdings to partake in the mobile value chain. One possibility is the "LTE-U small cell as a service" (SCaaS) model in which a cable operator can build out LTE-U small cells in select locations and offer SCaaS as a neutral host to multiple mobile network operators on a wholesale basis. Another possibility is for cable operators to parlay owned LTE-U small cells along with Mobile Virtual Network Operator (MVNO) agreements to provide retail mobile services. A third model is to extend the "Cable WiFi" concept with LTE-U Standalone mode to take advantage of possible efficiency gains. In all cases, LTE's built-in attributes for high efficiency are leveraged in unlicensed spectrum use as described below.

LTE has been designed to offer high efficiency in interference-limited mobile scenarios. Several vendor studies on LTE-U performance relative to Wi-Fi show about 3-5x improvement of LTE over Wi-Fi in unlicensed spectrum.[6] Qualcomm claims that LTE-U provides better RF coverage and offload as compared to Wi-Fi, and that the same capacity can be provided with fewer nodes with LTE-U vs. Wi-Fi. For an "inside-out" deployment scenario wherein small cells, deployed indoors, provide indoor as well as outdoor coverage and capacity, Qualcomm states that Wi-Fi requires 5x more Access Points (AP's) to provide the same capacity as LTE-U.[7] Separately, in the Nokia research, for a given system bandwidth and transmission power, the average user throughput on LTE is reported to be about 4x higher than Wi-Fi in both "sparse" and "dense" scenarios.[8] As indicated in both the Qualcomm and Nokia studies, it is generally accepted that LTE provides a better link performance over Wi-Fi due to the centrally coordinated and managed nature of LTE, which consequently provides more reliable and predictable performance over Wi-Fi. By comparison, Wi-Fi supports collision avoidance features to operate in unlicensed spectrum, and as network load increases, the collision avoidance becomes more burdensome and reduces throughput, with link performance negatively impacted.[9]

These studies of efficiency rely on modeling of the respective air interfaces of LTE and Wi-Fi. However, unlicensed spectrum enables participation of many network operators, and the inherent sharing properties of Wi-Fi enable this coexistence. Today, Wi-Fi carries the majority of Internet traffic as a function of its coexistence properties, making its implementation extremely efficient given the comparative amount of spectrum available for use with Wi-Fi.[10] LTE, however, is not built for a contended RF environment, and the impact of concurrent LTE operations in unlicensed spectrum is as yet unclear. If LTE-U standards require an equitable Wi-Fi coexistence technology we expect some of its efficiencies to be reduced.

In the "LTE-U Small Cell as a Service (SCaaS)" model, cable operators can leverage their network assets for backhaul and pole attachment rights to offer operator-neutral LTE-U small cell as a service on a wholesale basis. Cable operators can potentially offer a complete facility-owned, offload solution to mobile operators, including the (unlicensed) spectrum, with no concerns about who owns what and which spectrum is being used. This opportunity may allow cable operators to offer coverage and capacity solutions to mobile operators without making a significant investment in the Enhanced Packet Core (EPC) infrastructure, which in this wholesale model would be furnished by mobile network operators. LTE-U can be a better technology than Wi-Fi (on a single operator basis) especially for dense deployment scenarios where higher overall capacity can be achieved with better efficiency gain promised by LTE and improved coverage with deployment of newer LTE techniques such as enhanced Inter-Cell Interference Coordination (eICIC) and Coordinated Multipoint (CoMP).

Another interesting opportunity for cable operators is to leverage owned LTE-U small cells along with MVNO arrangements to offer retail wireless service. By leveraging owned small cells for offload, and the MVNO licensed spectrum control channel with lower traffic, cable operators have the potential to offer mobile services at lower cost as offload traffic can be delivered at lower cost than the wholesale rate at which cable operators would "rent" mobile capacity from mobile operators. With a bulk of MVNO service cost tied to network expenses, the MVNO business case can yield higher operating profitability as LTE-U small cells are deployed to where subscribers dwell and consume most of their traffic. A CableLabs' internal analysis of "Wi-Fi first" business model has shown that the service operating margin, excluding handset sales and costs, can be increased more than 13% for incremental 10% offload to Wi-Fi.

In the "Cable LTE-U SA" model, cable operators without licensed spectrum can potentially leverage LTE-U SA mode as an alternative way to offer end-to-end services to their home subscribers. Whether the "better" performance through LTE as delineated above can be achieved in the LTE-U SA mode is unclear as the key concept of "anchoring" on licensed spectrum, for crucial control signaling to ensure quality of service, is obviously not possible in this particular method. In addition, the coexistence properties of LTE-U are unknown, and multiple operators using the same unlicensed band may significantly diminish the efficiency benefits noted above. Assuming that LTE-U SA can somehow achieve a similar performance gain as LTE-U SDL, or even moderately better than today's Wi-Fi network for a given network load, then it may be advantageous for a cable operator to deploy LTE-U instead of Wi-Fi (assuming that cost differential of deploying Wi-Fi vs. LTE-U small cells are "reasonable") for utilization of unlicensed spectrum as LTE-U could conceptually provide more reliable and predictable wireless services.

Wi-Fi Coexistence Challenges and Risks

LTE-U in the context of the existing Wi-Fi ecosystem is not without challenges. The expected ramp of 802.11ac in the targeted 5GHz band for LTE-U could raise challenges and risks if the LTE-U does not take proper measures to equitably coexist with Wi-Fi in a "fair" manner. In general, the lack of coordination and management of mutual interference is the main challenge in the coexistence of different wireless technologies in unlicensed spectrum. Although most technologies are designed to handle interference management between systems of the same kind, it becomes especially challenging in heterogeneous systems with different time slots, scheduling modes, for example, time division multiple access (TDMA) vs. carrier sense multiple access (CSMA), and other media access control mechanisms. As the possibility of two dominant wireless access technologies, LTE and Wi-Fi, sharing common unlicensed spectrum bands becomes more likely with the active push for LTE-U by the mobile ecosystem, it is imperative for the respective standards organizations, 3GPP (for LTE) and IEEE (for Wi-Fi) to carefully study the LTE-U/Wi-Fi coexistence problem and define standard mechanisms to fairly utilize unlicensed spectrum without detriment to performance and user experience of the respective services. These coexistence strategies should consider both the existing base of access points and user devices as well as the future base of (especially 802.11ac) devices to come. (As depicted in the "green" region in Figure 2, a bulk of 802.11ac chipsets and subsequent Wi-Fi use of 5GHz is still to come.) Without a mutually agreed upon standard for coexistence that defines a trusted "arbitration" of fairness, LTE and Wi-Fi ecosystems risk the "tragedy of the commons," whereby each ecosystem may try to rationally allocate the unlicensed spectrum for its use, but may deplete efficient use of that spectrum for all by not coordinating effectively across technologies.

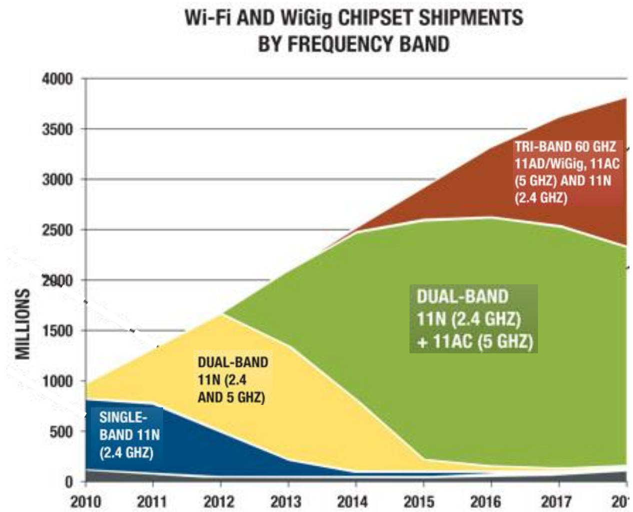


Figure 2 – Wi-Fi/WiGig Global Chipset Shipment Forecast (source: ABI Research, 2013)

Wi-Fi is designed to be asynchronous and decentralized in nature as it is intended for unlicensed spectrum use where multiple radio technologies can potentially contend for the spectrum resource. Wi-Fi uses channel access known as Carrier Sensing Multiple Access with Collision Avoidance (CSMA/CA). In this contention-based channel access protocol, the Wi-Fi node (AP or device) first "listens" to a channel before transmitting. Only when a channel is deemed "empty" (i.e., observed interference level is below a certain threshold) is the node allowed to transmit. If the channel is deemed "occupied," then the node defers transmissions for a random time period ("backoff") to avoid collision. For this reason, Wi-Fi is generally not efficient in highly dense environments with lots of devices competing for the common unlicensed spectrum resource.

Unlike Wi-Fi, resource allocation in LTE is much more efficient. LTE is synchronous and centralized in nature as it is designed for licensed spectrum where exclusive use of spectrum is guaranteed. LTE uses the Orthogonal Frequency Division Multiple Access (OFDMA) channel access technique that allows simultaneous transmissions with optimized allocation of frequency and time. LTE does not typically perform carrier-sensing detection (as exclusive use of licensed spectrum is assumed in the technology) and schedules channels optimally based on control and management signaling. With its scheduled nature and without carrier-sensing detection, LTE tends to be more "aggressive" in allocation of spectrum resource, while Wi-Fi tends to be more "polite" in its usage of spectrum resource.

In the absence of standardized LTE/Wi-Fi coexistence, LTE-U, as a simple re-band of LTE without any coexistence mechanisms, will crowd out Wi-Fi. A recent Nokia research paper[8] highlights this risk clearly. In the research, the team evaluated system level performance of coexistence of LTE and Wi-Fi systems based on simulations of "sparse" and "dense" environments.[11] The simulation results show that LTE performance is nearly unchanged in the presence of Wi-Fi use of the same band (less than 4% in most cases) while Wi-Fi performance degrades drastically in the presence of LTE (~70% in "sparse" deployment and over 90% in "dense" deployment). The primary cause is that Wi-Fi usage is often blocked by LTE co-channel interference, making Wi-Fi stay in "listen" mode most of the time, which directly impacts user throughput. With certain coexistence mechanisms, such as "smart" channel selection based on Wi-Fi and LTE channel measurements for example, it is plausible that the Wi-Fi performance degradation as observed in the Nokia simulations can be reduced. However, it really depends on how "aggressive" or "friendly" LTE-U will be in various coexistence scenarios, and those details have not been fully defined or reviewed by the wider body of ecosystem participants. Therefore, an early launch of pre-standard products in 2015 may expose coexistence concerns in deployed networks.

Unlicensed Spectrum Policy and Impact on Deployment Timing

Unlicensed spectrum policy, as determined years ago, generally did not anticipate the growth of wireless broadband or how integral Wi-Fi would become to the broadband ecosystem. As a result, in the U.S., China, Korea, and elsewhere, regulations governing the use of unlicensed spectrum primarily set governing power limits to protect adjacent band and co-band primary users, without specific requirements that facilitate coexistence among unlicensed users.

In this environment, coexistence in unlicensed spectrum has come about through the cooperation of all relevant stakeholders toward common goals. Therefore, there is no discernible legal barrier to entry of LTE-U in the U.S. and other jurisdictions with similar regulatory frameworks, and the features that will determine coexistence (both among LTE-U operators and between LTE-U and Wi-Fi) are unknown.

However, in Europe, Japan, and India, the regulatory framework is different, and specific coexistence protocols are mandated in unlicensed spectrum. These protocols, known as "listen before talk", generally replicate the CSMA/CA operation of Wi-Fi, thus enshrining Wi-Fi "politeness" in rule. In these jurisdictions, additional features will be required of LTE-U to ensure compliance and to achieve a true global scale.

Divergent regulatory requirements are likely to impact the rollout of LTE-U. In areas without mandated coexistence, LTE-U, in theory, can be implemented "as is", meaning with unspecified or proprietary and configurable coexistence features. In areas with specific coexistence parameters enshrined in rule, additional development and standardization will be required of LTE-U, which is likely to push out the timeline for initial implementation for such jurisdictions.

In subsequent sections of this paper, we explore the challenges posed by LTE-U to equitable coexistence in unlicensed spectrum, as well as the anticipated timeline and impact of the standards process.

LTE-U Timeline for Specifications and Products

LTE-U specification development remains in the planning phase within 3GPP. The timeline illustrated below is CableLabs' estimated projection of the LTE-U timeline based on recent 3GPP activity and conversations with key stakeholders.

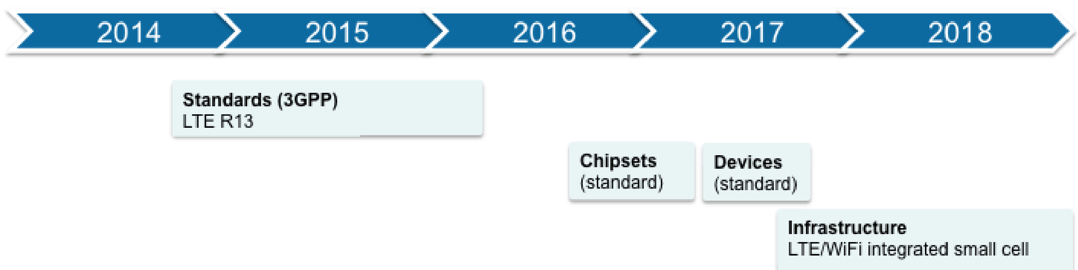


Figure 3 – LTE-U Timeline for Specifications and Products

A workshop dedicated to LTE-U is being held at 3GPP-ETSI headquarters in June of this year. The purpose of the workshop is to form a consensus among 3GPP companies for the scope of LTE-U technical studies and specification work. The consensus will provide the basis for study item and potentially specification work item approval at the RAN plenary in September 2014. At least fourteen companies are advocating a brief study so that specification development can commence at the beginning of 2015 and be completed by the end of 2015. A more conservative estimate targets specifications to be completed in early 2016. Interest in LTE-U is sufficiently high that significant contributions for the anticipated study item are already being submitted into 3GPP in advance of study item approval. With standards being completed in 2016, it is reasonable to expect compliant products in 2017. But it is important to note that pre-standards products can be introduced well in advance of this date, since regulations in U.S., China and other regions do not prevent proprietary solutions within unlicensed spectrum.

The technical scope of the LTE-U standards will likely be determined at the RAN September meeting when the study items and potential specification work items are scheduled for approval. The emerging consensus among key stakeholders in 3GPP suggests that LTE-U will be specified for the 5GHz unlicensed band. Control signaling will remain in licensed spectrum, and the unlicensed spectrum will be used for traffic channels. Effort will be placed upon non-interference among LTE-U systems and enough coexistence features with Wi-Fi to satisfy global regulations of the 5GHz band. The "Listen Before Talk" requirements in Europe are one example of these regulatory requirements, and it is reasonable to expect that standards will enable implementation in a majority of global jurisdictions. However, pre-standard coexistence features are unknown and will likely be proprietary. In jurisdictions such as the U.S. that are likely to see pre-standard deployments, it is not clear that 3GPP coexistence features will be subsequently adopted, and divergent proprietary approaches may persist for some time.

In summary, it is reasonable to expect that LTE-U standard products capable of operating in the globally ubiquitous 5GHz band and meeting international regulatory requirements will be available in 2017. It is also reasonable to expect that the vendors will offer pre-standard products well in advance of 2017 that may not reflect the coexistence features mandated in many jurisdictions.

Wi-Fi MAC and PHY Primer

This section takes a closer look at the Wi-Fi Medium Access Control (MAC) and physical (PHY) layers in order to provide a reference for the discussion of LTE-U and Wi-Fi coexistence described in subsequent sections. Wi-Fi frequency use, channel structure and medium access mechanisms are described.

Wi-Fi leverages Orthogonal Frequency Division Multiplexing (OFDM). The same frequency channels are used for uplink and downlink traffic. Channel bandwidths include 20MHz, 40MHz for 802.11n. 802.11ac adds 80MHz and the potential for 160MHz channel bandwidths. Sixty-four subcarriers are spaced every 312.5 kHz within each 20MHz channel. Fixed location pilots are placed in every modulated symbol. Modulation rates are up to 64 QAM for 802.11n and 256QAM for 802.11ac.

Devices discover and associate with APs using MAC layer signaling and procedures. The Wi-Fi AP can identify itself and its capabilities by broadcasting MAC layer beacons, or it can respond to network discovery queries from clients. Beacons on APs can be disabled such that the AP remains hidden until responding to a client discovery request.

Wi-Fi is designed such that multiple Wi-Fi networks can coexist in the same unlicensed frequency band through the use of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) procedures as explained below. These procedures also allow for Wi-Fi's coexistence with other technologies such as Bluetooth and Zigbee. Since Bluetooth and Zigbee implement their own versions of coexistence procedures, all three technologies operate together within unlicensed spectrum.

Wi-Fi devices and APs attempt to transmit by first sensing if traffic is already on the medium from another Wi-Fi source, or non-Wi-Fi source. The channel must be free of energy for at least 34 microseconds (μs) before transmission is allowed. If the medium is free, the Wi-Fi device or AP applies traffic to the medium. Burst durations on Wi-Fi can range from ~13 µs to 65 milliseconds (ms). For example, a single 1518 Byte Ethernet frame may be transmitted within ~110 µs to 1.8 ms depending upon the modulation rate used for the transmission. The procedure to sense before applying traffic is also commonly known as "Listen Before Talk."

If Wi-Fi devices and APs sense traffic or interference on the medium, they will back off for a specified period of time, and then sense the channel again to determine if a transmission is possible. This is depicted in Figure 4 below. Device B senses that the medium is busy. Device B backs off a random period of time and then senses the channel again. The figure below illustrates the case where it is determined that the channel is free after the initial back off, so that Device B applies traffic to the channel. Had Device B determined that the channel is still busy, it would have backed off for a longer period of time before attempting to apply traffic again. All Wi-Fi devices use these procedures when they attempt to apply traffic to the channel. Collisions are possible when two devices correctly sense the medium is free, and then apply traffic at the same time. Wi-Fi devices will back off a random period of time before sensing the channel again whenever a collision is detected.

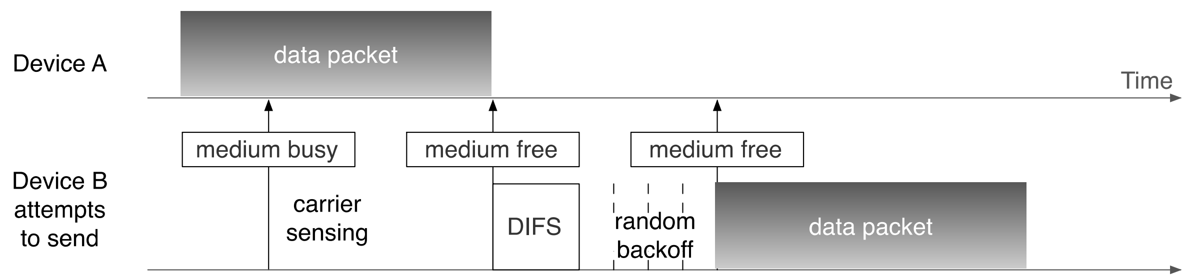


Figure 4 – Procedures to Apply Traffic to a Wi-Fi Channel

While the CSMA/CA procedures allow multiple Wi-Fi devices and network to coexist, they also introduce inefficiency as more devices attempt to use the channel and create collisions or back off for increasing periods of time. As more devices attempt to use the medium, the total aggregate traffic delivered by an AP decreases and service eventually becomes seriously degraded.

<u>LTE MAC and PHY Primer</u>

This section takes a closer look at the LTE MAC and PHY layer in order to provide a reference for the discussion of LTE-U and Wi-Fi coexistence described in subsequent sections. LTE frequency use, channel structure and medium access mechanisms are described.

LTE can be operated as either a Frequency Division Duplex (FDD) or Time Division Duplex (TDD) air interface. As with Wi-Fi, OFDM is applied on the downlink. Single-Carrier Frequency-Division Multiple Access (SC-FDMA), a very similar technique to OFDM, is applied to the uplink. Channel bandwidths of 1.4 MHz, 3 MHz, 5 MHz, 15 MHz and 20 MHz are standardized. Subcarriers are spaced every 15kHz and pilot locations vary from symbol to symbol. Traffic is scheduled in 10ms frames. Modulation rates are up to 64 QAM. Because the physical layers between LTE and Wi-Fi are different, they cannot coexist in the same spectrum at the same time without interference.

LTE is currently designed for use in exclusive licensed spectrum, meaning only one LTE network can exist in a given frequency band. LTE is not designed to coexist with other networks. The LTE network periodically broadcasts training and network identification information. Devices need to scan all exclusive license bands in the area in order to select the correct network. Once the proper network is identified, the mobile device will authenticate and attach using LTE access control channels. LTE control channel signals from the network are periodically transmitted. For example, primary and secondary synchronization signals are sent every 5 ms among other signaling, which results in very minimal "quiet time" in the channel.

When an application requests services or a data session, the network schedules a dedicated traffic channel for the device. Traffic is periodically transmitted in the scheduled traffic channel until all data is delivered. The transfer of traffic is carried within one or more Resource Blocks (RBs). As shown in Figure 5 below, mobile devices or User Equipment (UEs) are allocated subcarriers in time. A total traffic transferred to a mobile device may take several RBs. In periods of heavy traffic, the LTE network constantly places RBs on the channel in an organized schedule.

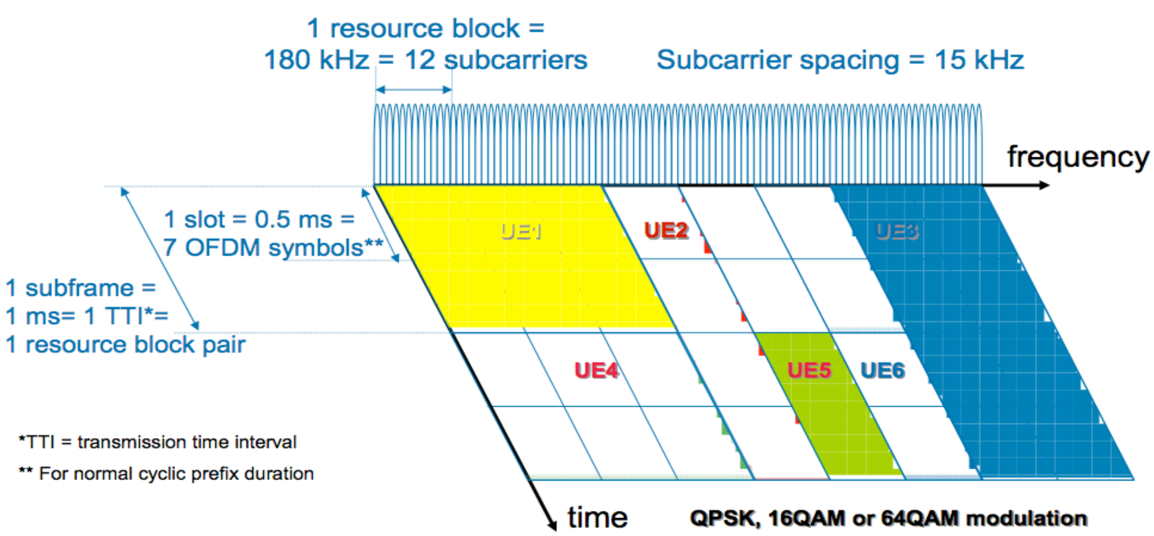


Figure 5 – LTE Resource Block for Traffic Transmission

As seen in Figure 6, the LTE network schedules periodic transmission of resource blocks that support multiple mobile devices with dedicated traffic channels. This centralized scheduling of the air interface eliminates the possibility of traffic collisions among mobile devices. Since only one LTE operates in a single frequency band, there is no possibility of collisions among LTE network providers.

As noted above, the LTE system supports both FDD and TDD operations. TDD operation requires that the frequency channel is divided between uplink and downlink transmissions. 10 ms frames are divided between uplink and downlink 1-ms subframes. The number of subframes assigned to the uplink and downlink is configured by the network operator. Guard time periods are placed between uplink and downlink transmissions. There are no network transmissions during the guard time.

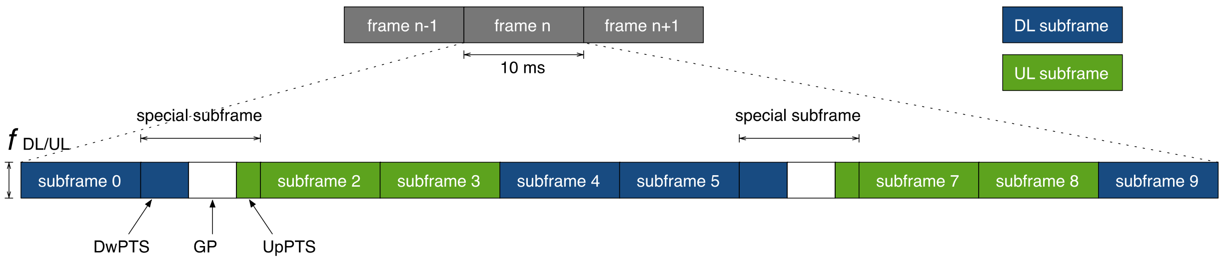


Figure 6 – LTE-TDD Transmission

In summary, the LTE air interface is centrally scheduled by the LTE network, which has exclusive use of the licensed spectrum. This avoids collisions among LTE devices and LTE networks. It also means that more of the total LTE cell data traffic capacity is still available for use as the number of devices in the cell increases. The centrally scheduled nature of the LTE air interface helps make it a more efficient method to transmit data compared to Wi-Fi.

Wi-Fi / LTE-U Coexistence Challenges

Wi-Fi and LTE-U data transmissions will interfere with each other if transmitted simultaneously. As explained in the Wi-Fi and LTE primers sections above, Wi-Fi includes coexistence procedures designed to allow multiple Wi-Fi systems to coexist, whereas LTE is designed with the assumption that one operator has exclusive control of a given spectrum. LTE traffic channels are designed to very efficiently continuously transmit when delivering traffic such that Wi-Fi will have little chance to sense the channel unoccupied and suitable for transmission. LTE also transmits periodic control and synchronization signaling even when no traffic is delivered to devices. So unless the LTE-U traffic channels are redesigned differently than LTE traffic channels in licensed spectrum, LTE-U will apply continuous traffic to devices in a periodic fashion. This raises the possibility that LTE-U may essentially control the unlicensed spectrum at the expense of Wi-Fi devices and other technologies in times of congestion. A research paper from Nokia indicates that LTE-U interference may degrade Wi-Fi performance over 90% during heavy traffic times.[12] Therefore, the LTE-U MAC layer will need to be designed to coexist with Wi-Fi if Wi-Fi is to be afforded a useful portion of the unlicensed spectrum. But how best to design coexistence into LTE-U without substantially degrading the data throughput efficiency of LTE-U remains an open question.

Ideally, coexistence requirements and solutions should provide a level playing field for each network and technology while accounting for local regulatory requirements. Air time fairness and data throughput efficiency are important considerations. The U.S. and China do not mandate specific coexistence requirements for 5 GHz unlicensed spectrum. Europe, however, does mandate the coexistence requirements as summarized below:

For Frame Based Equipment:

- Clear Channel Assessment time: equal to or greater than 20 µs
- Channel Occupancy Time: between 1 ms and 10 ms
- Minimum Idle Period: greater than 5% of the Channel Occupancy Time

For Load Based Equipment

- Clear Channel Assessment time: equal to or greater than 20 µs
- Time back off: N*Channel Occupancy Time, where N~[1,q]. q=4 or 32
- Channel Occupancy Time: less than (13/32)/q ms

Coexistence mechanisms should ideally provide each network an equal opportunity for airtime fairness. Specifically, each network needs to be able to utilize equivalent portions of spectrum over time as traffic conditions meet or exceed the data throughput capacity of the air interface. For example, if 10 networks attempt to utilize 100 MHz of unlicensed spectrum in UNII-3 band in the U.S., each network should be afforded an average of 10 MHz of spectrum over the time of the high traffic period. This does not necessarily provide each device in the network the same average data rate, which is dependent upon a number of factors. Air time fairness shares equivalent megahertz portions of spectrum equally among participants.

Coexistence mechanisms should also strive for data rate efficiency. But a range of coexistence techniques to help ensure air time fairness may present costs to data rate efficiency. For example, excessive clear channel assessment times, long back off periods, and short time periods where a device can apply traffic before being forced to yield the channel can all decrease the data rate efficiency delivered to devices. MAC layer signaling designed to communicate or negotiate air frequency use will increase the overhead of system. While the European regulations described above may bound some of these characteristics, in other areas of the world, such as the U.S. and China, the etiquette is unclear.

The legacy coexistence mechanisms in Wi-Fi have been accepted as sufficient in the wireless industry and are currently in use by billions of Wi-Fi devices across the globe. LTE-U is the new entrant technology targeted for the unlicensed band where Wi-Fi and other technologies currently successfully coexist. LTE-U needs to be designed in light of existing Wi-Fi systems in order to support coexistence.

One design direction for LTE-U would follow the Wi-Fi model for coexistence. LTE-U would first listen to a channel to ensure it is idle before applying traffic. Traffic would be applied for a specific maximum length of time. LTE-U would then release the channel for a specific back off period before starting the process again. Parameter values for these LTE-U procedures could be specified in order to ensure air time fairness with Wi-Fi. While this may provide an equitable solution for all technologies and networks in the band, it may come at a price for LTE-U. Proponents for LTE-U claim it can achieve up to 5x the data throughput efficiency compared to Wi-Fi if it is designed as a simple re-band of LTE with minimal coexistence techniques. But a significant portion of the LTE-U efficiency is due to the centralized and continuously scheduled nature of its air interface. If LTE-U is subject to the inefficiencies of the Wi-Fi's "listen before talk" procedures, it would lose much of the benefit of its scheduled air interface. The efficiencies of LTE-U could approach those of Wi-Fi if Wi-Fi like coexistence procedures are applied to LTE-U. While coexistence parameters may cause LTE-U to lose some efficiency relative to its theoretical maximum, means of coexistence are necessary not only to prevent interference with Wi-Fi, but also interference between multiple LTE-U operators using the same frequency band.

An alternative design direction may be for LTE-U to release the channel temporarily using currently designed scheduling mechanisms. Specifically, TDD LTE-U could be designed to intentionally not transit data for X frames during the period of every Y total frames. This duty cycle approach to coexistence allows LTE-U to maintain the efficiencies it enjoys due to the scheduled nature of the LTE air interface. This design direction would also afford other technologies to transmit for a portion of the LTE-U data transmit duty cycle in order to help ensure that LTE-U does not consume the entire spectrum as shown in the Nokia paper referenced above. But it should be noted that this approach leaves LTE-U firmly in control of the unlicensed spectrum. What kind of duty cycle should be specified to ensure air time fairness? Should this LTE-U duty cycle be designed to give up the channel 50% of the time when other systems are attempting to use the interface? Should the duty cycle be adaptive to take into account how many other Wi-Fi systems are in the area? How should the duty cycle be enforced to ensure air time fairness? And by whom?

In summary, air time fairness for all technologies and networks may come at a cost to the projected data throughput efficiencies of LTE-U. Therefore, a consensus of cross industry stakeholders is needed to ensure LTE-U is properly designed. A combined effort between 3GPP and organizations responsible for Wi-Fi, such as the Institute of Electrical and Electronics Engineers (IEEE) and Wi-Fi Alliance (WFA), may be needed to reach a standardized conclusion to these coexistence design tradeoffs. Coexistence solutions that are vendor proprietary or defined by a single set of stakeholders may come at the expense of air time fairness, and may persist in the marketplace even after standards are developed.

STEPS TOWARD LTE-U'S FULL POTENTIAL

LTE-U may provide higher data throughput to users in the increasingly crowded unlicensed frequency bands. However, the features that make LTE an efficient technology also make it a challenge to equitably coexistence with other technologies that use unlicensed spectrum, such as Wi-Fi. The U.S., China, and Korea do not mandate specific coexistence mechanisms, which both paves the way for expeditious implementation of LTE-U and raises additional questions about the impact to unlicensed users. Additional development is required to ensure that LTE-U is implemented equitably within unlicensed bands, and that it is available for all to use, regardless of licensed spectrum holdings. Doing so will not only preserve open innovation in unlicensed spectrum, but it will also increase the scope and scale efficiencies of LTE-U.

ENDNOTES

[1] About 500 MHz of unlicensed spectrum is available globally in the 5GHz band.

[2] With the recent FCC rule change in the 5150-5250 MHz band to allow outdoor use with a higher regulated power limit, the scope of unlicensed spectrum utilization can readily be extended to this UNII-1 band, effectively doubling the most useful spectrum for outdoor small cell deployments to 200 MHz in the U.S. Another 355 MHz is available in the 5GHz band for unlicensed use, with additional limits on transmit power and interference avoidance. And LTE-U may be implemented in other unlicensed bands, such as 2.4 GHz.

[3] 'Study on Licensed-Assisted Access using LTE: Motivation', Ericsson, Qualcomm, Huawei, 3GPP contribution RP-140260, March 2014. See http://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_63/Docs/

[4] In the U.S., maximum regulated power is 1 watt in 5725-5850 MHz ("UNII-3") and, following recent FCC action, also in the 5150-5250 MHz ("UNII-1") bands. For more information on this recent FCC action to expand unlicensed use in the UNII-1 band, see: "FCC Votes to Expand Wireless Spectrum: A Win for Wi-Fi", Rob Alderfer, CableLabs, March 31, 2014, available at: http://www.cablelabs.com/fcc-votes-to-expand-wireless-spectrum-a-win-for-wi-fi/.

[5] LTE-U will likely be part of Release 13 study/work item based on latest 3GPP standards activity.

[6] Most of the Wi-Fi vs. LTE-U performance comparisons that we have looked at are based on licensed LTE coordination and control, low-level coexistence features and do not account for 802.11ac. It is unlikely that LTE-U with listen-before-talk or other coexistence features would perform as well.

[7] "Extending LTE Advanced to unlicensed spectrum" white paper published January 17, 2014. See http://www.qualcomm.com/media/documents/white-paper-extending-lte-advanced-unlicensed-spectrum

[8] Andre Cavalcante et. al., Performance evaluation of LTE and Wi-Fi coexistence in unlicensed bands, IEEE Vehicular Technology, Spring 2013.

[9] The link performance in LTE system also gets impacted as network load increases, but its impact is less severe as resource allocation is tightly managed through central coordination in LTE, unlike the "ad hoc" nature of Wi-Fi networks.

[10] According to the Cisco's 2013 Visual Networking Index, Wi-Fi carried 49% of all Internet Protocol traffic globally in 2012, compared to 48% for the fixed network and 3% for mobile. Accounting for the amount of spectrum dedicated to mobile versus Wi-Fi, Wi-Fi is therefore approximately 30 times more efficient than mobile.

[11] The Nokia paper showed Wi-Fi performance may be degraded over 90% under certain traffic load environments. In the Nokia simulation, the "sparse" environment consisted of 4 Wi-Fi AP's

and 4 LTE AP's and the "dense" environment consisted of 10 Wi-Fi AP's and 10 LTE AP's. In both scenarios, up to 25 Wi-Fi and LTE user devices each communicated with respective AP's. Both scenarios show substantial impact to Wi-Fi.

[12] Andre Cavalcante et. al., Performance evaluation of LTE and Wi-Fi coexistence in unlicensed bands, IEEE Vehicular Technology, Spring 2013.

Moving CCAP To The Cloud

Alon Bernstein
Cisco Systems

Abstract

Network function virtualization (NfV) is gaining traction as a viable method for implementing network appliances on generic compute resources (e.g. the same intel processors used in laptops) instead of custom-built networking hardware.

What would be the benefits of fully virtualizing a Converged Cable Access Platform (CCAP) as an NfV appliance? The presentation will outline the new ways in which a virtual CCAP can solve challenges of scaling, performance, availability, qualification, test and other operational issues in a cost effective way.

Moving CCAP To The Cloud

This paper does not prescribe a specific functional division between the various components that make a virtual CCAP solution. The paper will outline the benefits of moving both the control plane and data plane of a CCAP solution to the cloud (a.k.a. NfV) and the overall network architecture surrounding it.
CCAP owns the physical access to the plant and that part is not ready for virtualization yet, therefor this paper assumes that the virtual CCAP relies on an external solution for HFC access.

Network Function Virtualization Overview

Up to the mid 80's there were not many dedicated network devices. Most routers where implemented on general-purpose servers that had multiple interface cards. However, those general-purpose servers were fairly expensive and as demands for networking increased several companies began building cheaper devices for the sole purpose of routing/switching Internet traffic. Custom built network devices ruled the earth for a couple of decades but as general purpose CPUs became more powerful and better integrated with Ethernet input/output (IO) it was demonstrated that they can be used for forwarding millions of packets per second making them a viable alternative to custom built network devices – in a way completing a circle from server to appliance and back to a server implementation again.
Anything that can run natively on a CPU (a.k.a "bare-metal") can run on top of a hypervisor, opening the door to place these networking devices in the cloud alongside other cloud applications. As a result of these observations ESTI (ref [1]) created an NfV working group to study and make recommendations on an end-to-end framework, including provisioning, testing, monitoring and scaling for NfV solutions.

Running Over a Hypervisor

There is a "virtualization tax" in terms of performance when running over a hypervisor, but there are clearly benefits as well. What are the benefits of running a network function on top of a hypervisor? Clearly there is the original benefit of virtualization: having a single binary software run on multiple types of physical hardware, but there is more. The hypervisor allows multiple virtual CPUs (vCPU) to run over a single physical CPU. The hypervisor can also move a function that was running on one server fairly seamlessly to another server. What these two capabilities provide generally fall into these two categories:

- Availability
- Efficient resource usage

In the following sections we will look in more details to what these capabilities mean in the NfV world.

NfV Benefits

When looking at NfV from a pure engineering point of view of power/performance/space it may not look that appealing. After all it's pretty intuitive that custom-built hardware would be more efficient then a general purpose solution and that hardware optimization would be more efficient then software optimization. However, when taking the operator point of view and especially from a CAPEX point of view a different picture emerges:

- Power efficiency: We need to consider the peak-to-average power. A big router sitting idle at 4:00AM draws more power then an NfV appliance sitting idle or turned off completely. Some of general purpose CPUs are so optimized to save power (mobile devices and such) that security experts claim that they can tell what a CPU does based on how much power it draws.
It's also worth noting that when a CPU company rates a CPU at a certain value it's a worst-case estimate with the floating point processor and other features running at once, which is typically not the case for NfV. On top of all that we have the ability to shrink and expand the number of NfV instances based on demand so that CPUs can be tuned off if not used. Because of the above the average consumption per-hour of an NfV appliance may end up being attractive.
- Granular scaling: similar to the point above but from a different angle: An NfV solution has very fine grained scaling, and so some power/space efficacy is derived from the fact that an operator can use exactly (!) the right amount of resources. Based on the monetization model this may mean that the operator pays exactly for the resources used and that translates to a cost benefit.
- Reuse of resources: A general-purpose server can perform other services at off-times when NfV functions are not needed. For example, it can do packet processing by day and payroll processing by night. The power/cost/space advantages come from the fact that the compute resources are optimally utilized.
- Availability: the ability to quickly move virtual machines from one server to the other is playing a key role in providing high availability. If a virtual machine fails it can be re-instantiated on a different server. Geographical redundancy is also possible - a server at a remote location can take over in case of a catastrophic even if the local data center is down.
- Feature velocity: network appliances are embedded systems that require embedded systems disciplines. Development on general purpose servers is easier because most software developer are familiar with that environment and the software development tools associated with it, therefor the time to develop and test software is shorter.
- Fast qualifications: because of the granularity of NfV services its possible to test new code drops for a very small population and slowly increase the deployment of the new code drop as it proves itself. This can reduce qualification times by being more aggressive in moving to a production environment because the risk of global failure is greatly reduced.

Scale out vs. Scale up

The discussion on scaling benefits with NfV warrants a separate discussion. The key message is the following:

NfV is not about performance. Its about scaling

What does the above statement mean?
The traditional approach in physical network appliances is to deal with increasing

bandwidth demands and feature requests by “scaling up”, i.e. building a family of small/medium/large appliances with a larger number of linecards and faster network processors as demand grows. In contrast, the data center approach is to “scale out” which means to build a basic service instance that gets replicated (dynamically) with resource demands. So, for example if a virtual CCAP reaches its maximal packet forwarding capacity the operator – or more precisely an orchestration system - can instantiate a new virtual CCAP to handle the extra load.

The cloud world offers a feature called “cloud bursting”; when the local cloud runs our of resources the system can use another cloud, such as a public cloud offered by several companies, to take over. This is helpful if the peak conditions are relatively rare. For virtual CCAP there is another type of “cloud burst” that is possible; one where the average demand is handled by a physical device and only peak conditions by the cloud. This allows building a very cost optimized solution making the best of both physical and virtual versions of CCAP.

<u>NfV Network Architecture</u>

The basic physical building block of an NfV solution is a server. The server is typically a collection of multi-core CPUs, one or more network interfaces and storage. The multi-core CPUs can host multiple virtual machines, which in our case might be multiple instances of a virtual network function. But how do all these machines connect to the Ethernet and how can they exchange information between each other? The entity that helps with that is the vSwitch (Figure 1):

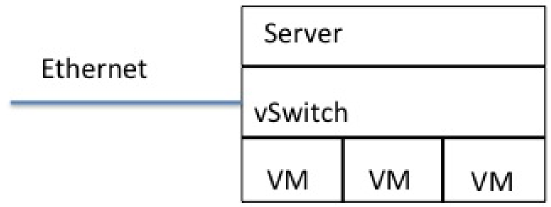


Figure 1 Server and vSwitch

With the magic of virtualization each one of the VMs “thinks” it owns the Ethernet to the outside world. The vSwitch, which typically uses one of the CPU cores, switches traffic between the VMs and traffic in/out of the physical Ethernet port connection.

The next level of how an NfV assisted cable network may be connected is depicted in Figure 2. Traffic from the Internet is directed at the front end of the data center that handles

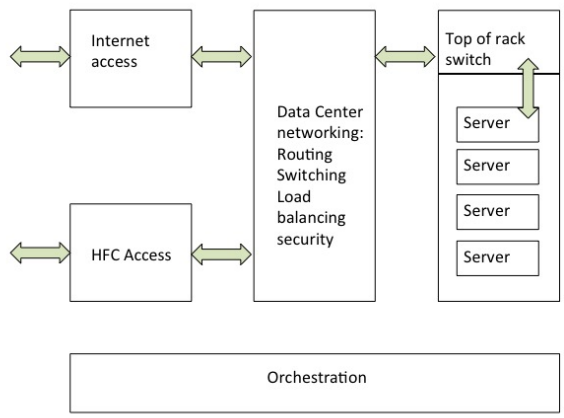


Figure 2 NfV packet processing

basic functions such as load balancing traffic between servers and first level sanitization of the traffic - for example denial of service protection. The next stop is typically the “top of rack switch” that sends the packet stream to the right server (in a data center the servers are stacked in a rack and connected to a top of rack switch). Once the packet stream is intercepted by a server it gets switched to the right VM via a vSwitch. The access to the HFC is essentially a mirror image of the process described above.
Considering how dynamic the data center is and how VMs and functions can move

around, the need for automation in this system is obvious. This automation is referred to as "orchestration". The following user story helps outline the role of orchestration in the workflow of creating a virtual CCAP.

<u>A User Story: Virtual CCAP instance workflow</u>

The following section will outline how a CCAP instance can be created:

1. A cable operator notices an increase in traffic demand and the need to add a virtual CCAP instance.
2. The operator can look in a "service catalogue" for a virtual CCAP function (as we cover in the next section other services may be firewall, deep packet inspection, parental controls etc.).
3. Assuming that an HFC access is already available then all the cable operator has to do is add the virtual CCAP function to the HFC segment that is under load.
4. The Orchestration system takes care of creating the right path between the Internet, HFC access, data center switching/routing, fiber interconnect and vSwitch.
5. Service up and running. No need to locate new linecards or CCAP chassis, no need for an install or hook up any equipment.

<u>Virtual CCAP Q&A</u>

The following sections are answers to frequently asked questions about virtual CCAP.

<u>Is the virtual environment stable enough for packet processing?</u>

The virtual environment has a reputation of being somewhat unpredictable, however, this reputation is not justified. Issue with variable latency and packet drops start showing up when we over subscribe (or close to oversubscribe) the server – not that different then issues with oversubscribing bandwidth on a physical appliance. So when provisioning NfV one should be careful to allocate enough resources and possibly set up VM scheduling for reliable operation of the NfV appliance as well as take into account other users of CPU in the sever, in particular the vSwitch and the hypervisor itself. If provisioned correctly the virtual environment can be as reliable as a "bare-metal" environment where the OS rides directly over the hardware.

<u>How is a virtual CCAP related to SDN?</u>

SDN, in a nutshell, is about separating the control/management plane from the packet forwarding in the data plane. The key to NfV is packet processing in a virtualized environment. NfV appliances can be controlled by an SDN control plane or can operate independently so that an NfV appliance may or may not be part of an SDN solution. In other words NfV and SDN are two independent technologies that are designed to solve different problems. It so happens that both can run in the cloud (aka "virtualize") but that's about all that's common between them.

<u>Is the vCCAP a CCAP replacement?</u>

One common question about the vCCAP is if it's going to replace the physical CCAP. The short answer is "no". The vCCAP should be viewed as another packaging option for a CCAP that is cost effective in certain cases, most likely the areas where a small scale CCAP fits in today. Another use case of the vCCAP is to handle peak usage, i.e. a physical CCAP can handle average traffic loads scenarios but the vCCAP can kick in with extra capacity is needed.

<u>What if the data center is not in the normal data path?</u>

Different operators may have different designs for their data center. The two base options are to (a) build a massively centralized data center a-la Google/amazon (b) build a distributed data center.
The first option could be a challenge for NfV. If the data center is built around video distribution and is physically remote from Internet peering points then packets would have to loop in the network in order to provide service through a virtual CCAP. Even in that environment a virtual CCAP can be used for certain applications (help with test/qualification, geographical redundancy etc.) but it would be harder to deploy at scale. A distributed data center with servers closer to the end users lands itself more easily to a scale deployment of virtual CCAP since the packet flows are similar to those with a physical CCAP network architecture.

How is CCAP Related to Service chaining?

NfV appliances handle dedicated functions. For example, one NfV appliance can be dedicated to deep packet inspection, another to a firewall and a third to a virtual CPE. Service chaining is the ability to set a pre-determined path that is per-application (or subscriber or service) along that path, e.g. one application can go through parental control and deep packet inspection and another application can be directed at only a parental control service. A CCAP appliance, either physical or virtual, is very likely to be both at the entry and the exit of a service chain. At the entry the service flow classification can assist with the selection of a service chain, and at the exit of the service chain the application of Quality of Service (QoS) is the last feature to be applied.

How are virtual CPE and virtual CCAP related?

The CPE can be swept along with CCAP in the virtualization wave. Once we have both CCAP and the CPE in the cloud they start appearing as elements in a service chain, where CCAP might take care of some aggregate policies and shared resources while the CPE is tailored around a specific subscriber needs. In other words, the traditional breakup of functions between a CPE and a CCAP device will be challenged once both move to the cloud.

Conclusion

The following table would help place a boundary around what virtual CCAP is and is not:

Virtual CCAP Is	Virtual CCAP Is not
Scaling optimized	Performance optimized
Connected to node	Integrated CCAP
Dependent on the data center	Standalone device
Dependent on orchestration	Self managed
Data plane processing in the cloud	Separation of only control plane

Network function virtualization is a viable implementation and deployment option in certain use cases and a virtual CCAP can fit well inside an NFV ecosystem. The operational benefits that made virtualization a success in storage and compute apply to networking in general and more specifically to the virtual CCAP where it can be the right tool to address certain operational challenges in a cost effective and scalable way.

References

1. http://www.etsi.org/technologies-clusters/technologies/nfv

MULTICAST AS A MANDATORY STEPPING STONE FOR AN IP VIDEO SERVICE TO THE BIG SCREEN

Eshet, Amit   Ulm, John   Cohen, Uzi   Ansley, Carol
ARRIS

Abstract

With the availability of new technology, namely home gateways with 16 or more DS tuners and CCAP platforms, the transition to a managed IP video service is now becoming feasible and cost-effective. The biggest challenge remaining is the bandwidth; especially with the continual growth of high-speed data service leading to constant contention among the services for the limited spectrum resources available.

This paper provides a window into the viewing habits in 2013. We show that linear content is still the number one service with the number of peak time viewers at over 75%. We compare the efficiencies of switched multicast, and unicast (native for ABR delivery) approaches. Furthermore, to give indication of the transition period to a full IP service, we further present the significant benefits of multicast even at small groups of subscribers.

We conclude by introducing a Multicast assisted ABR architecture that offers the benefits of multicast as well as building upon a single infrastructure to enable an IP video service to all devices in and outside the home.

THE TRANSITION TO A MANAGED IP VIDEO SERVICE

An IP Video service is nothing new in the cable space. Over the past few years operators have been offering VOD and linear services to mainly secondary screens. IP Video is yet to be used as the vessel for the primary video service. The motivation to expand IP Video to the big screen is clear – first and foremost it enables CPE cost reduction, eliminating components such as the cable card, and QAM tuners. Transition to IP video also opens the path to leveraging consumer electronics like consoles and Smart TVs without relying on a STB. An additional driver is the opportunity for significant OpEx and CapEx saving by maintaining a single architecture for video service to all devices. A baseline IP Video delivery architecture is depicted in Figure 1.With this architecture, linear channels are transcoded to multiple profiles, segmented, encrypted and pushed to the Origin server. From the Origin server, the segments are distributed via a CDN and made available over HTTP to ABR clients residing on STBs, Smart TVs, consoles, mobile devices, or browsers, where the latter may consume the linear service inside or outside the home, depending on content rights. The ABR client registers with the DRM application server and accesses it to get the keys to decrypt the content.

With the availability of new technology, namely home gateways with 16 or more downstream tuners as well as CCAP platforms, the transition to IP is now becoming feasible and cost-effective. The biggest challenge remaining is the bandwidth. DOCSIS 3.1 is still several years ahead, and so is HEVC. Cost of massive node splits is still significant and proactively replacing all legacy STBs with IP STBs, needed to reclaim the spectrum of QAM Video, is many years away from being an economical solution. Moreover, high-speed data service keeps growing rapidly and is competing with IP video for the limited spectrum resources available.

TRENDS IN CONSUMPTION OF A MANAGED VIDEO SERVICE

We believe today's viewership trends must also apply to IP Video service as the subscriber should be agnostic to the actual delivery scheme. Other trends that may have an impact on viewership, such as nDVR and new, more user friendly UI are not limited to IP Video. Examples are the Cablevision QAM-based nDVR service and the new Xfinity UI launched on the X1 QAM STB.

Looking at viewership information, it is very clear that live linear (consuming linear channels in real-time) service to the STB is still the predominant service by far. Next is DVR viewership also with significant viewership. Finally, VOD, as well as multiscreen, trail with few percentiles or less of the eyeballs.

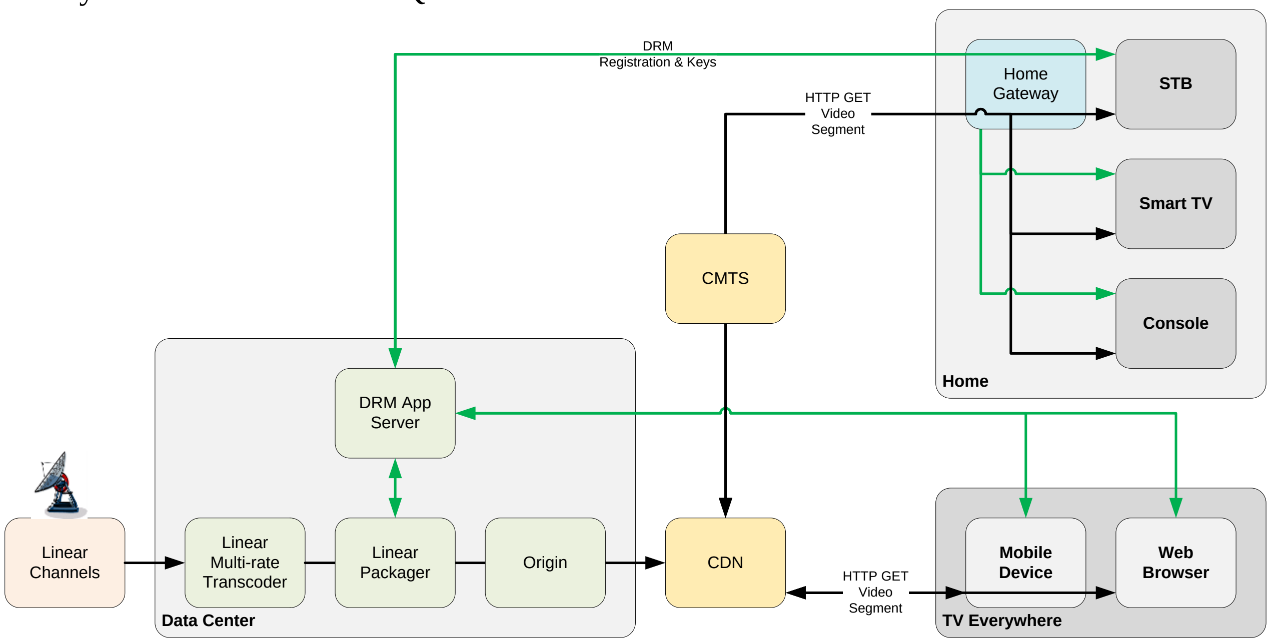


Figure 1 - Baseline IP Video Delivery Architecture

<u>Case Study – Linear vs. DVR trends</u>

Figure 2, Live TV Viewing, provides insight into the two dominant video services in the home, linear and DVR, over a period of a week. The data was collected from 1032 households with a total of 2,386 STBs. Data includes number of viewers (STBs tuned to linear content), number of channels, number of DVR recording sessions, and number of DVR playback sessions.

Multiple key insights become obvious from looking at the data:

1. At peak time (~6-8 PM) almost 60% of the STBs (1.4 STBs / Household) are consuming linear channels.
2. Number of STBs consuming linear channels changes dramatically throughout the day (300-1400), but is very similar between days at a particular time of day.
3. As expected, number of linear channels watched is dramatically lower, at ~300 channels at peak.
4. Number of DVR recording sessions is also very high, at peak time with over 1000 concurrent sessions (the gateway system used to collect this data can support up to six concurrent sessions).
5. DVR recording peaks are correlated with the linear peak with limited recording happening at off-peak hours.
6. Number of concurrent DVR play back sessions is significantly lower than the number of recording sessions and accounts to 9% of the STBs.
7. People are recording much more than they are actually watching

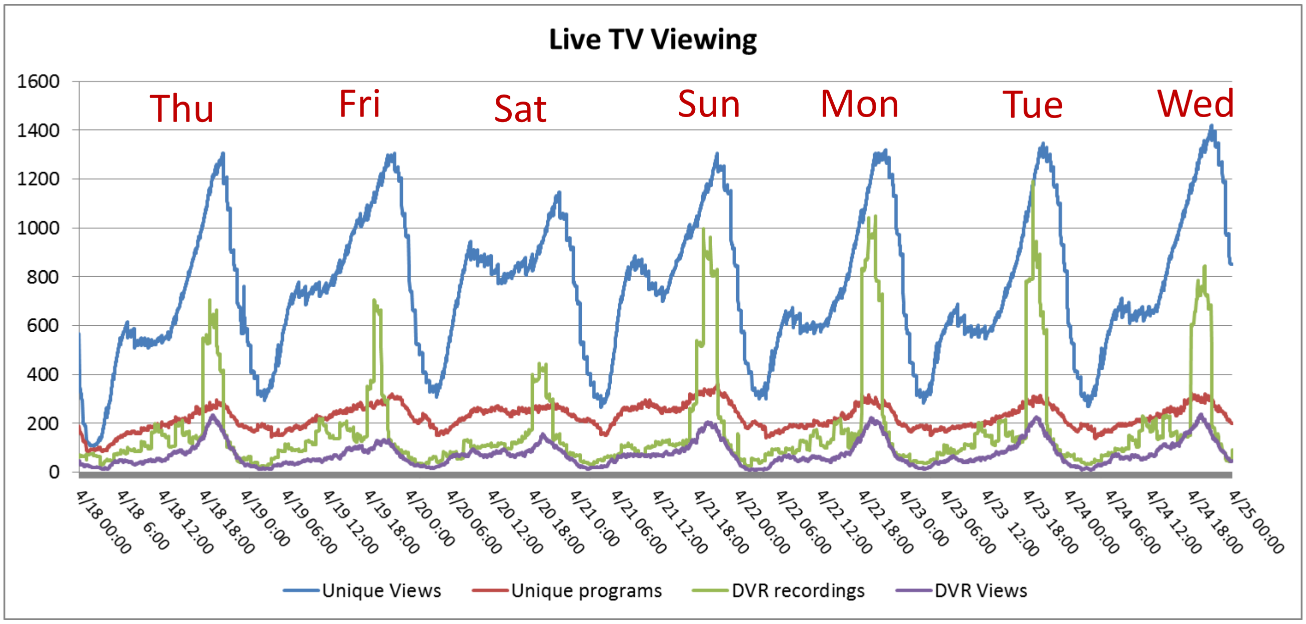


Figure 2 - Live TV Viewing trends

With all these said, the three key takeaways are:

1. Linear is still the king. No other service comes close.
2. The potential advantages of multicast are evident from comparing the unique linear views to the number of unique linear channels (~75% saving).
3. With DVR recording being a dominant factor, a unicast IP video service that doesn't include nDVR is not realistic. This becomes clear from comparing the number of unique linear views plus the number of DVR recording sessions to the number of unique linear channels (~85% saving).

Case Study – Bandwidth Requirements

At the end of the day the benefit of Multicast is to be proven at typical service group sizes. We have thus focused on three typical size, 125, 250, and 500 tuners. These model groups are equivalent to about 35, 70, and 140 subscribers per service group. The motivation was to explore the benefit in a small scale deployment as well as to see the benefit as the linear IP video service ramps up.

We have looked at a viewership data collected over a period of a month from a group of ~27,000 tuners (~20,000 STBs). The lineup consisted of 69 HD and 266 SD channels. We further assumed 6 Mbps for HD and 2 for SD (H.264).

Figure 3 compares the peak number of viewers to the peak number of channels, like the one done in Figure 2, only for smaller service groups. Note that a viewer can be either a STB tuned to a linear channel or a DVR recording a linear channel. For each service group, two data points are included, one representing the highest number of viewers, the seconds representing the largest number of linear channels over the one month period.

In all cases, during peak time, on average 50% of the tuners are tuned to linear channels, with a maximum of 61-67% depending on the service group size. In the 125 tuner service groups, less than 50 channels (40 on average) are watched at peak, less than 90 (average of 63) for the 250 tuner service groups, and less than 120 (average of 40) for the 500 tuner service groups.

Looking at this data and the graph, one can argue that the saving at the 125 tuners SGs will be marginal with an average worst case of 63 viewers and average worst case of 40 channels. This information is misleading as HD and SD channels are not weighted proportional to their respective bitrate. In order to clearly see the value of multicast a true bandwidth comparison is due. This comparison is illustrated at Figure 4, Multicast vs. Unicast Bandwidth requirement.

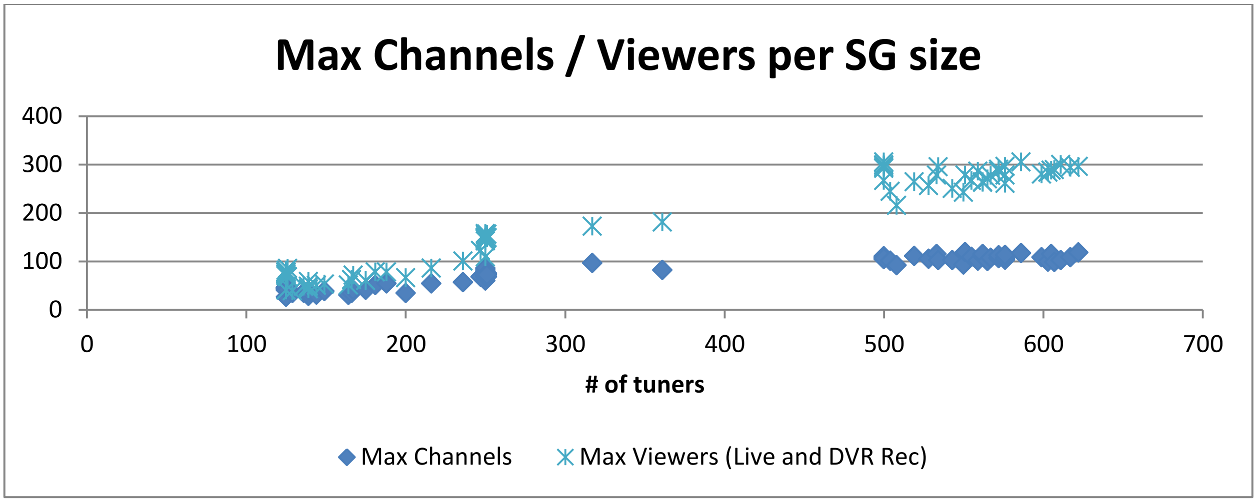


Figure 3 - Max Viewers vs. Max Channels

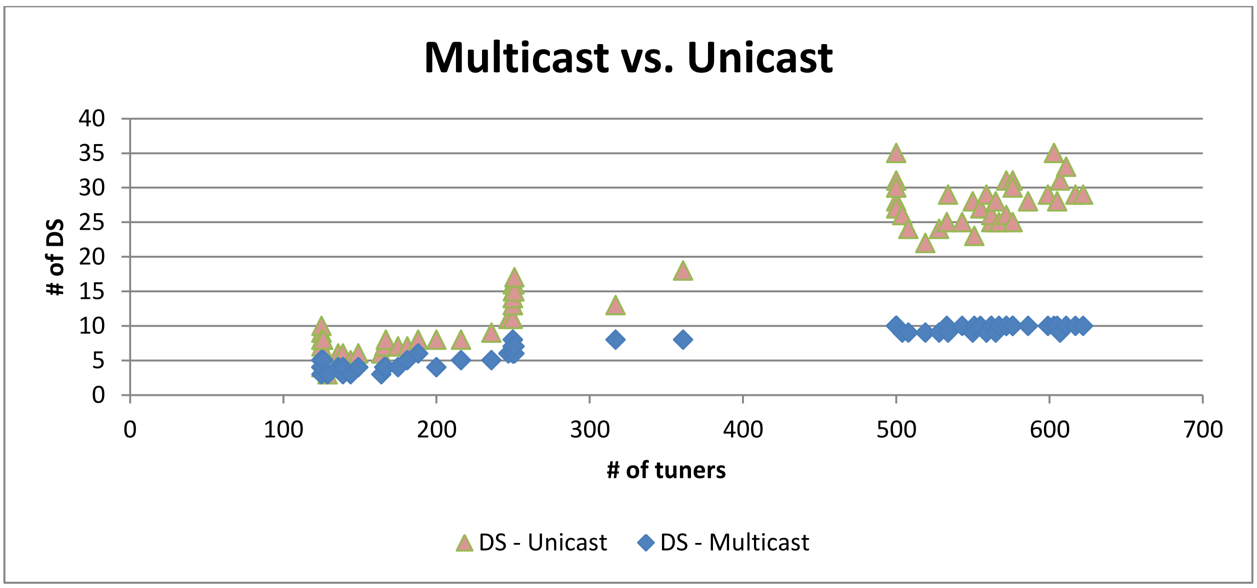


Figure 4 - Multicast vs. Unicast Bandwidth Requirements

For each service group, two data points are included, one representing the highest number of downstream DOCSIS channels needed for a multicast implementation, the second representing the highest number of downstream DOCSIS channels needed for a unicast service. The following table summarizes the worst case per SG size (125, 250 and 500 tuners). As the spectrum allocation across all service groups is the same, the worst case represents actual capacity planning.

# Tuners	Unicast (Max DS)	Multicast (Max DS)
125	10	5
250	17	8
500	35	10

Figure 5 - Linear IP Video Capacity Planning

It is clear that 50% saving on capacity can be achieved even at small service groups of 125 tuners (or ~35 subscribers). The saving grows with the service group size reaching over 70% at 500 tuners (140 subs). Moreover, this data indicates that a basic Linear IP Video service is feasible with a minimal spectrum of 5 to 10 DOCSIS downstream channels. It should be emphasized that actual capacity requirement are highly dependent on the HD take rate. In the case of the data analyzed here the number of SD channels at peak time was roughly 2.5 the number of HD channels. Moreover, other aspects of the solution, like unicast traffic for enabling fast channel change, and targeted ad insertion may dramatically increase the bandwidth requirements, if not addressed effectively.

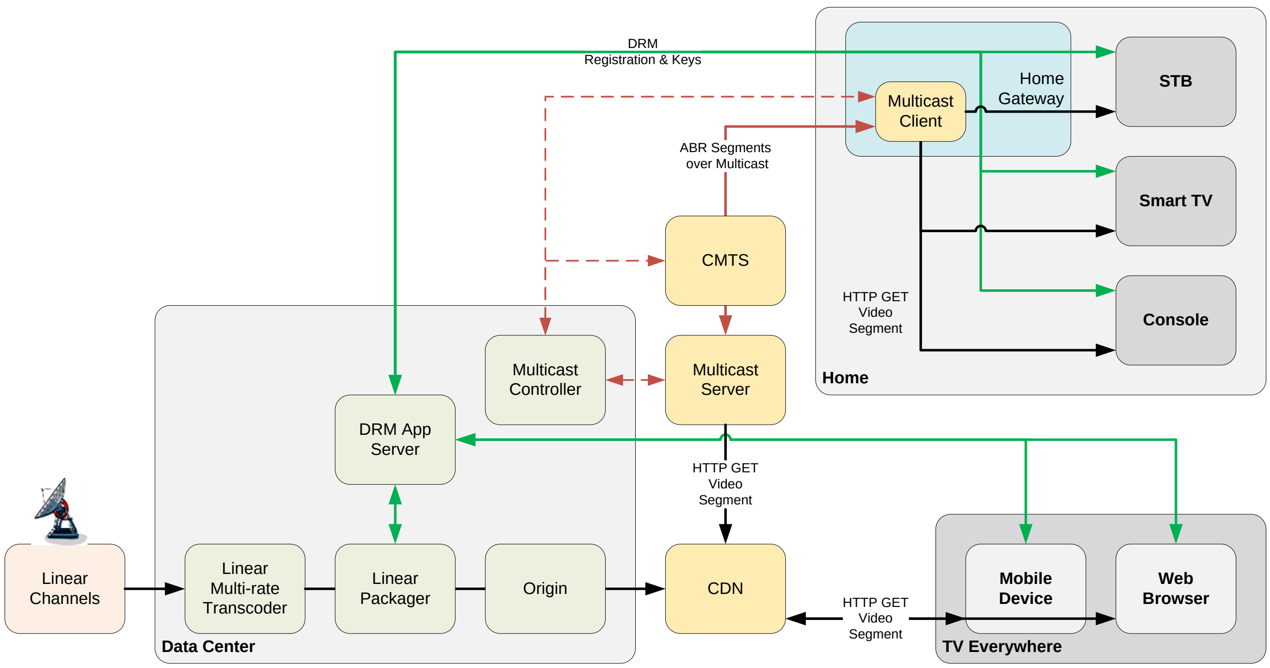


Figure 6 - Multicast Assisted ABR architecture

A MULTICAST ASSISTED ABR ARCHITECTURE

The architecture outlined in Figure 6 achieves both the goal of a unified architecture to reach all devices in and outside the home and the goal of minimizing the bandwidth requirements for a linear IP video service. Three new components are added to the baseline architecture: a Multicast Server, a Multicast Client, and a Multicast Controller. The role of the Multicast Server is to pull new Linear segments as they are made available in the CDN and deliver them over multicast. The Multicast Client serves two roles. It serves as a cache for segments arriving over multicast as well as a transparent proxy for requests coming from the ABR clients. When an ABR client requests a segment, the Multicast Client will intercept the request, check if it can be fulfilled from the cache, and if not, pass it to the CDN. A request for a linear channel segment not already cached in the ABR client can trigger an IGMP join request to the appropriate multicast in order to start filling the cache. As such, in a typical situation, the first few requests for segments will be fulfilled via unicast whereas all following requests would be met by the cache being filled by the multicast. Finally, the Multicast Controller serves multiple roles:

1. Collect viewership reporting from the Multicast Clients
2. Control the lineup being offered via multicast
3. Control the (proactive) caching on the Multicast Clients
4. Control the delivery and caching of ads in the Multicast Client.

Note that the some of the roles of the Multicast Controller are directly aimed at optimizing the Multicast assisted ABR service, and ensuring high efficiency compared to a pure unicast service.

SUMMARY

We have shown that to date linear service is still the predominant method used by subscribers to access MSO video content by far. As such, to offer a linear IP video service, operators should be implementing a Multicast architecture thus saving over 50% of the bandwidth compared to a pure unicast solution. We concluded by introducing a Multicast assisted ABR solution leveraging a common architecture for delivering video to all IP devices.

NETWORK VIRTUALIZATION IN THE HOME

Chris Donley
CableLabs

Abstract

Networks are becoming virtualized. While there has been significant focus on virtualization in core and data center networks, network virtualization will also provide benefits in the home. From reducing equipment costs to simplifying software upgrades

CableLabs has been exploring how Network function Virtualization (NfV) and Software Defined Networking (SDN) can affect cable subscribers' home networks. This paper will present a vision for future home networks, specifically:

- *A Virtualized Home Network Architecture*
- *Virtualized Home Network Functions*
- *Virtualization Benefits to MSOs and Subscribers*

INTRODUCTION

Home networks are growing more sophisticated; customers are not. As home networks become more complicated, many customers are looking to MSOs to support these more complicated networks, and MSOs need tools to support them. Network virtualization using technologies such as Software Defined Networking (SDN) and Network function Virtualization (NfV) provides such a set of tools.

Generally speaking, SDN describes an open architecture comprising a set of APIs, and control protocols such as OpenFlow that allow for dynamic, distributed provisioning and automation.

NFV decouples network functions such as firewalls, deep packet inspection, caching, etc., from proprietary hardware so that they can be run in software on generic (e.g., x86) servers.

While SDN and NFV can be implemented independently, the benefits multiply when the technologies are combined The architecture described below illustrates a combined approach for MSO subscriber networks.

HOME NETWORKS TODAY

Home networks are evolving. Most subscribers today connect to the Internet using a router. As shown in Figure 1, subscribers are connecting additional routers to their networks to extend the reach of their wifi, or to add services such as home automation and security, IP video, and sensor networks (e.g., Internet of Things). Home routers, however, typically do not run a routing protocol, and networking these routers was challenging, and usually required multiple layers of IPv4 Network Address Translation (NAT). As customers are interconnecting devices within the home for video streaming or remote printing from tablets, these multiple layers of NAT are becoming problematic and severely hamper these in-home services.

To address these problems, CableLabs, in conjunction with MSOs and technology suppliers, developed HIPNet™, a new architecture leveraging IPv6 provisioning to automatically configure home routers into a routable network without requiring NAT on interior routers. HIPNet functionality is becoming available on cable eRouters, and represents a significant improvement over previous technology. However, some challenges still remain. Service Discovery across routers (e.g., to allow Smart TVs to locate DLNA media sources) is challenging, and MSOs do not have an easy way to manage this proliferation of home routers on behalf of their subscribers. In addition, it is difficult to add new home network services, as they rely on the capabilities of the routers already deployed, and may require a new device to support new features.

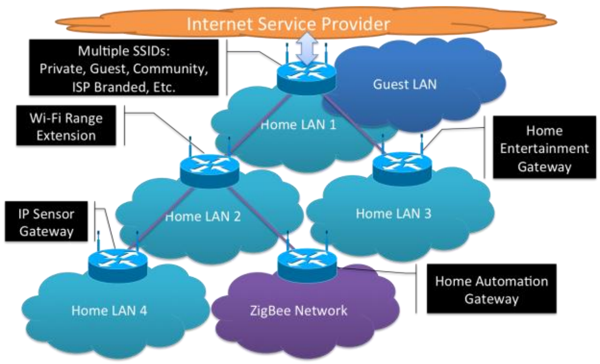


Figure 1: Evolving Home Network

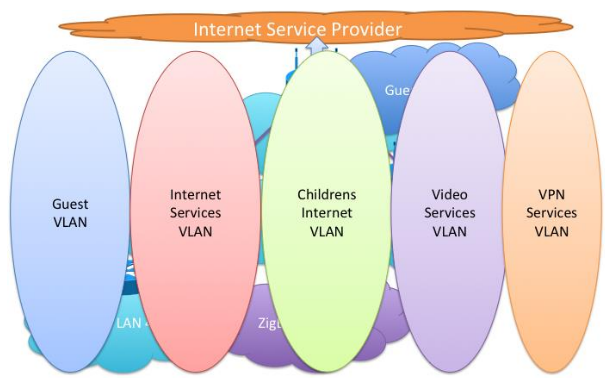
A VIRTUALIZED HOME NETWORK ARCHITECTURE

One solution to the growing complexity of subscriber home networks is to virtualize the home network so that it can be managed by the MSO (or the subscriber via a self-service portal). This allows us to move beyond the device-centric architecture we use today and consider a virtualized service-centric architecture, which offers MSOs the ability to better manage subscriber networks and to understand how customers are using them, and offers subscribers a way to tailor the network to optimize their specific use cases such as gaming or video streaming.

Routing vs. Bridging

There has been a debate among home networking experts about whether to use routing or bridging inside the home. Many problems experienced in a routed home network, such as service discovery, multiple firewalls, and multicast forwarding, become simpler in a layer 2 (bridged) network. However, existing devices typically include routers. Also, some emerging services such as Smart Grid or home automation and security require routed networks for security purposes or to satisfy regulatory requirements.

In a virtualized home network, we can have the best of both worlds. First, the home network can be separated into different logical policy domains, such as for Internet access, guest access, VPNs, or in-home video sharing. See `Figure 2`. Each zone can be assigned its own firewall and connectivity policies. Next, each zone is distributed throughout the house using encapsulation techniques such as VXLAN. Finally, hosts are assigned to one or more zones. Because devices can receive multiple IPv6 addresses, it is conceivable that they will receive unique addresses for each zone. By default, they would be assigned to the Internet or Guest zone (for a Guest WiFi network), and could be assigned to different zones, as well.



`Figure 2: Virtualized Home Network`

Within each zone, which could stretch across router boundaries, traffic is bridged. This would offer subscribers an improved quality of experience. No longer would nested internal NAT functionality interfere with printing or video streaming, and link-scoped service discovery mechanisms such as mDNS would show all the devices in a particular zone, rather than just those devices on the local subnet.

Bootstrapping

When the network first comes online, it needs a basic level of automatic configuration support plus a path to reach the MSO network controller. HIPNet, included in eRouter devices, provides this level of connectivity using DHCPv6 prefix delegation to provision routers in a tree topology and establish routes to all the devices. It is optimized for Internet connectivity, and also supports host-to-host communication, but perhaps not in an optimized manner. Once network connectivity is established, the home routers can contact the MSO network controller for optimized forwarding instructions using protocols such as OpenFlow or TR-069.

To create optimized forwarding paths, the MSO network controller needs topology information from the home network devices. Home routers can collect this topology information using Link Layer Discovery Protocol (LLDP) and communicate it to the MSO controller using OpenFlow or similar protocols. The MSO controller can then use the Dijkstra algotithm (also used in routing protocols such as OSPF and ISIS) to compute optimal forwarding paths and communicate them back to the subscriber's routers. Subscriber routers can also collect and report attached host MAC and IP addresses to help troubleshoot issues that may arise in the home and to further optimize traffic forwarding.

In the event of an Internet connectivity failure, this architecture would allow the network to use a backup connectivity mechanism such as WiFi. If that is not available, the home network will continue to operate, albeit with more basic HIPNet functionality. Thus, the MSO controller provides optimizations when the service is connected, but the home has local survivability.

While a virtualized network architecture as described above can improve a subscriber's quality of experience, it is not yet sufficient to deliver on the promise of enhanced management and customizability. For that, let's explore various home network functions and how they could be delivered to our virtualized network.

HOME NETWORK FUNCTIONS

Home networking devices typically perform a number of functions on behalf of the customer. These features can be divided into two types: control plane and data plane. Control plane features look at packet headers and enforce policy on a network, while data plane features are inserted in the traffic forwarding path and affect the payload of the traffic.

While not an exhaustive list, control plane features include:

- Network Address Translation (NAT), which provides differentiation between customer space and public space and which is used to manage IPv4 address scarcity during the transition to IPv6.
- Firewall, which enforces security policy on the network
- Routing and forwarding, which identifies the optimal paths to send traffic through the network.
- Virtual Private Networks (VPNs), which provide private connectivity to remote networks such as corporate offices.
- IPv6 transition technologies

Likewise, data plane features include:

- Dynamic Host Configuration Protocol (DHCP) and Domain Name Service (DNS), which provision devices with IP addresses and provide database lookup services to identify other hosts
- Deep Packet Inspection, which looks into packet payloads and helps with Denial of Service and Parental Controls
- Denial of Service protection, which looks for traffic anomalies and block unwanted traffic streams.
- Parental Controls, which block objectionable content.

Until now, these features have generally been offered on home routers, and configured separately on each router. This has led to a sub-optimal experience for subscribers, who have looked to the teenager down the street or commercial services such as Geek Squad to configure their routers. With network virtualization techniques, MSOs can host all of these services in their data centers and offer them to subscribers as cloud services.

In addition, customers are interested in some control plane features that are not widely available today, either because they have not been possible, or because they have been difficult to implement with existing devices, but that could be delivered in a virtualized environment:

- Bandwidth on demand, where subscribers can change bandwidth levels on the fly to accommodate large file transfers (e.g., downloading a movie before a flight).
- Priority service for video or gaming services, allowing subscribers smooth delivery of entertainment content.

Recently, many in the industry have been considering virtual CPE (vCPE), an approach for moving network functions into per-subscriber virtual machines in the cloud. Control plane features such as firewalls and parental controls are obvious candidates to move into the cloud. Both can be configured via a self-service or technician web portal, with policy pushed into the home using SDN, and enforcement performed as close to the customer device as possible. See Figure 3.

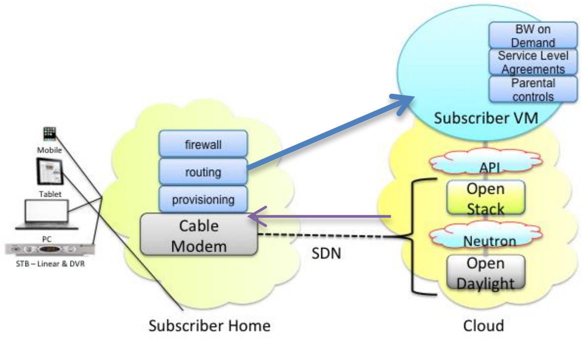


Figure 3: Virtualized CPE

Parental controls are slightly more interesting, as they also involve content filtering, and require deep packet inspection to look for objectionable content. Such a service would be far more robust than traditional DNS-based controls available on home routers. To perform these functions, the parental control function needs to be performed in-line on the data plane. Thus, the customer traffic requiring parental control will be routed through the data center and passed through parental control and other data plane network functions using service chains, a mechanism for passing traffic flows through multiple network functions on their way to the Internet or back to the home network.

Once the virtual network is in place, it allows MSOs to offer new network services such as Bandwidth on Demand or enhanced service levels for high-value content such as video streaming or gaming. Indeed, we have already taken the first steps. CableLabs has developed a PacketCable MultiMedia (PCMM) plugin for OpenDaylight that can be integrated into such a framework. OpenDaylight, combined with our PCMM plugin, provides a RESTful interface for adding, modifying, and deleting DOCSIS®

service flows. This would allow a subscriber to create a new service flow (e.g., for gaming) with defined bandwidth and DOCSIS QoS characteristics.

BENEFITS

The home network described above offers benefits for both MSOs and subscribers. MSOs benefit from reduced expenses, faster time-to-market with new services, and optimized use of deployed resources. Subscribers benefit from mass-customized services and service-centric policies (as opposed to device-centric policies today).

MSOs stand to benefit from reduced expenses, as this virtualized network architecture allows for self-service provisioning via a web portal, simplified upgrades managed by DevOps tools such as Puppet and Chef, and simplified inventory management and certification testing, as the functionality is delivered in software, rather than via specific devices. It also gives MSOs more visibility into the devices attached to the subscriber network, helping them troubleshoot and optimize the network on the subscriber's behalf. As network functions are deployed in software, this architecture offers MSOs shorter build-measure-learn development cycles that will bring new features to market faster. Finally, as virtualized network resources can be shared across multiple subscribers, it allows MSOs to optimize the use of deployed resources.

For subscribers, network virtualization offers a mass-customized Internet service. Just as we have seen with cellphone app stores, subscribers value different aspects of a service. Under this approach, they can drag and drop those features that are important to them. For example, an avid gamer might select optimized gaming service, while parents might opt for strict parental controls. As services can be tailored to individual subscriber needs, this approach offers an enhanced quality of experience over today's networks. In addition, network policies are tied to the user, and not the device. This allows subscribers to have the same Internet experience at home or on the road through Cable WiFi.

CONCLUSION

In conclusion, home networks are becoming more sophisticated, but subscribers are not. Network virtualization allows MSOs to offer subscribers a new network architecture that is mass-personalized, automated, and tailored to individual needs. This architecture includes service-(or policy-) specific overlay zones that can be extended into the MSO data center to allow delivery of MSO-managed network features. From the data center, MSO SDN controllers can push policy to individual network devices, optimizing network forwarding paths and enforcing firewall policies. These changes offer improved economics to MSOs and an improved quality of experience to subscribers.

REFERENCES

John Berg, The Future of Home Networking: Putting the 'HIP' in HIPnet, http://www.cablelabs.com/the-future-of-home-networking-putting-the-hip-in-hipnet/

Chris Donley, SDN & NFV: Moving the Netowork into the Cloud, http://www.cablelabs.com/sdn-nfv/

VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks, draft-mahalingam-dutt-dcops-vxlan-08.txt

Puppet: http://puppetlabs.com/

Chef: http://www.getchef.com/

NIELSEN'S LAW VS. NIELSEN TV VIEWERSHIP FOR NETWORK CAPACITY PLANNING

Michael J. Emmendorfer and Thomas J. Cloonan
ARRIS

Abstract

Our industry is fully aware of Nielsen's Law of the maximum Internet service tier offered to consumers, also known these days as the "Billboard Internet Speed", and that this has been growing at a 50% CAGR since 1982. To date, some MSOs have sized their network on a method of multiplying the billboard speed by either doubling (2X) or tripling (3X) the billboard speed to determine the amount of DOCSIS capacity per service group, this is sort of a Rule of Thumb method for DOCSIS Network Sizing. This method worked for the most part, but this approach will break in the future and we will show why. Additionally, ARRIS plans to unveil a new Traffic Engineering and Capacity Planning Formula to help MSOs properly size their networks to accommodate Nielsen's Law, Traffic, and Competition. This is called the Network Quality of Experience (NQoE) Formula and is a unit of measure that may be used to size any service provider network and network technology.

Our industry is also fully aware that Internet Traffic or consumer bandwidth demand has seen explosive growth from historic averages. We are also aware that this is in large part driven by over the top (OTT) video services causing explosive growth in Internet Traffic, which may range from 40% to over 100% in annual growth rates. This has moved the symmetry between downstream and upstream traffic from 2:1 or even 4:1 from a decade earlier to now over 10:1. We will examine the impact of video service as a key driver for traffic consumption and growth rates. ARRIS will also show several Internet traffic growth rate predictions that may help network planners.

High-speed Internet is only one service that utilizes spectrum and drives network investment, and this paper examines the role of other services and delivery technologies on network utilization as well. We will show that Coax to the Home (CTTH) will be able to sustain the needs of the customer through the year 2030. Obviously, forecasting until 2030 is difficult, but we want to illustrate the controls the MSOs have and the visibility that traffic may not grow at this rate forever. We will also introduce new network architecture for accommodating 1) the Billboard service tier growth rates and 2) the Internet traffic growth rates. ARRIS will also introduce a new method for network architecture that should reduce capital costs and extend the life of the CTTH network, while competing or beating FTTH networks.

Should our industry migrate to DOCSIS as the unified video delivery network supporting both MSO delivered content and for OTT to extend the life of the HFC?

The paper will unveil:

1. *New Traffic Engineering and Capacity Planning formulas*
2. *New Video Traffic Growth Rate projections*
3. *A new approach to DOCSIS Network Architecture Capacity*
4. *A new forecast for Network Capacity through 2030*

DRIVERS FOR TRAFFIC ENGINEERING AND CAPACITY PLANNING

The MSO's competitive landscape has changed rapidly in just the last 12 months especially from Over the Top (OTT) video providers such as Apple TV, Amazon, Hulu, Netflix and others entering the On-Demand video market. In many ways, the consumer electronics companies, like Apple, are becoming service providers enabling the video experience across all platforms and across any carriers' network. The OTT competition affects the MSOs in lost revenues for On-Demand services and perhaps a reduction in the subscription service. Adding to the lost revenue are increased costs to the high-speed data network due to increased consumer usage.

SHOULD OUR INDUSTRY MIGRATE TO DOCSIS AS THE UNIFIED VIDEO DELIVERY NETWORK SUPPORTING BOTH MSO DELIVERED CONTENT AND FOR OTT TO EXTEND THE LIFE OF THE HFC?

Service Providers will have to forecast their video network resource requirements as well as the Internet Service Tiers and Traffic Growth rates to properly size the data network. What if our industry did something very different? What if we began to analyze the high-speed data traffic to determine use of video services being delivered OTT, would this influence our video network planning, and even our overall network planning? Our industry may consider the migration to IP Video in an effort to have a single video delivery network, whether the consumer selects an OTT video provider or the MSO video service the video traffic will be transported across a single network. Today, consumers select video from two separate networks: the Analog and Digital Broadcast network and the unicast MPEG-TS VoD network. Additionally, the customers are switching to a different video network, the high-speed data network and this may cause wide variations in network planning to allocate network resources for essentially two different video delivery networks. The MSO may actually extend the life of the HFC by transitioning to DOCSIS to enable a single video delivery network support regardless of the origin.

NIELSEN'S LAW OF INTERNET MAXIMUM SERVICE TIER OFFERED

The network traffic estimates need to consider the downstream and upstream high-speed Internet service tier, in other words the data speed package that the MSO offers to consumers. The highest data speed offered in either direction is a determining factor for sizing the network. The High-Speed Internet service tier and traffic will grow considerably during this decade moving from perhaps four 6 MHz channels downstream, which is less than 4% of the MSO's total spectrum allocation and may grow to perhaps 40-50% in the next 10 years.

This model illustrates that Data Service Tiers offered to consumers increase at about a 50% compound annual growth rate (CAGR) and this model also is used to forecast actual consumer traffic usage which also grows at roughly a 50% CAGR. This is based on Nielsen's Law of Internet Bandwidth or Max Internet Service Tier. We have also combined Nielsen's Law with the research of Dr. Thomas J. Cloonan, CTO of ARRIS and co-author. The research is captured in the "Max Internet Data Services Tier Offering Downstream and Upstream graphs in this section. Dr. Cloonan begins with the data rate offered since 1982 and charts growth through to the present day. This data, referred to as Cloonan's Curve, also reflects the historical 50% CAGR as does Nielsen's Law. The data service portion of the model is predictable but at some point, as with Moore's Law, Nielsen's Law may not continue on this 50% CAGR trajectory for another 20 years, and it may break.

The high-speed Internet service tier offering will be a key contributor to overall bandwidth drivers. Figure 1 below shows a thirty-year history of the max bandwidth offered or available to consumers. This figure also attempts to predict the max service tier we may see in the future, if the growth trend aligns with the preceding years. The models are a combination of Cloonan's Curve, a 30 Year History of Max Service Tier Offered, and Nielsen Law of 50% CAGR.

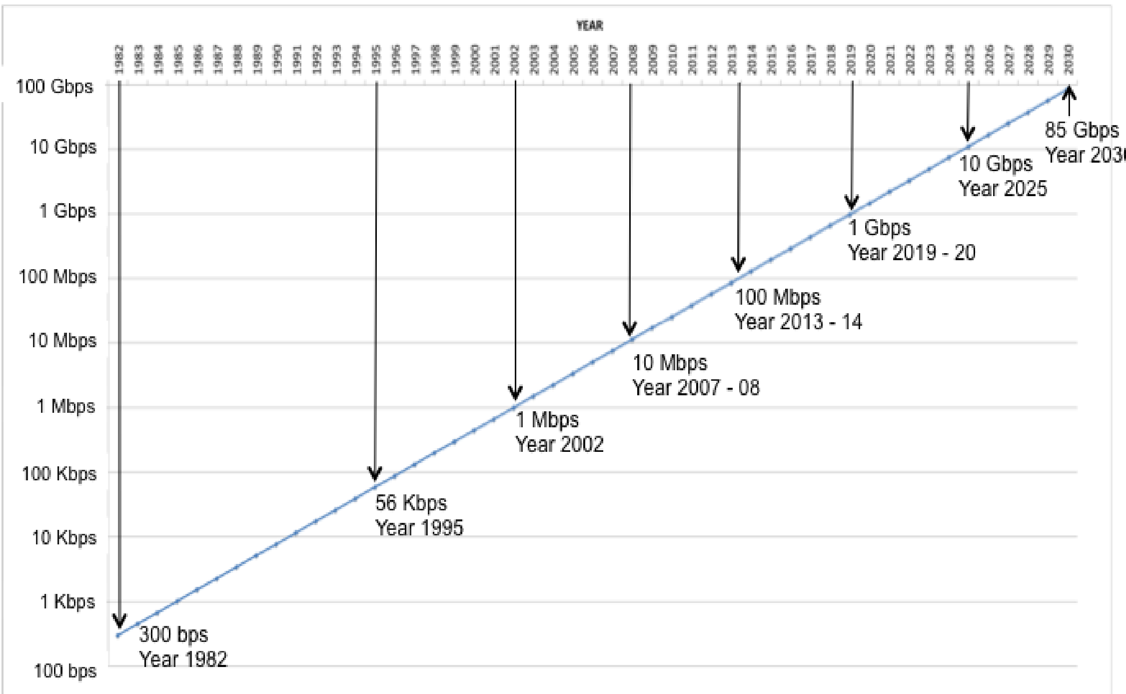


Figure 1 – Nielsen's Law Max Internet Service Tier Offered – Downstream (1)
Source: http://www.nngroup.com/articles/law-of-bandwidth/

Is it possible that a service provider will offer a residential 11 Gbps Internet service by 2025? When will Nielsen's Law Break? It is unlikely that a 50% CAGR for the Max Service Tier Offered will last forever. Moore's law broke and Nielsen's Law will too, but when and by what rate? The Internet service tier offered is ultimately a decision of the service providers and they may simply pull the lever of growth back, as this is a driver for investment that is desired for only a small percentage of their Internet customer base.

INTERNET TRAFFIC OR CONSUMER BANDWIDTH DEMAND

Measuring and Estimating Customer Traffic

In addition to the service tier offered to consumers, the actual usage of the network by the consumers is a critical factor for network planners. This is known as the bandwidth per subscriber (BW per Sub). The determination of bandwidth per sub is a measurement of the total amount of bandwidth or traffic in a serving area divided by the number of consumers in the serving area. This may be measured during busy hour(s) to drive operator traffic engineering limits. Dr. Cloonan has collected traffic data from many sources to determine the traffic per subscriber data rates as seen in Table 1 and 2. The bandwidth per subscriber is measured in the downstream and upstream direction. The downstream was measured at a 100 kbps per subscriber and the upstream at 43 kbps per subscriber in the year 2010, as illustrated in tables 2 and 3. The bandwidth per subscriber CAGR may vary, so we have used several growth rates for the downstream and the upstream. These numbers are used for planning purposes in this analysis, it is important that each operator capture their own CAGRs.

DOWNSTREAM DATA NETWORK TRAFFIC PREDICTIONS			
North Amer. & Europe	Per Subscriber Traffic (Shown in Mbps)		
Year	40% CAGR Downstream	50% CAGR Downstream	60% CAGR Downstream
2010	0.10	0.10	0.10
2011	0.14	0.15	0.16
2012	0.20	0.23	0.26
2013	0.27	0.34	0.41
2014	0.38	0.51	0.66
2015	0.54	0.76	1.05
2016	0.75	1.14	1.68
2017	1.05	1.71	2.68
2018	1.48	2.56	4.29
2019	2.07	3.84	6.87
2020	2.89	5.77	11.00
2021	4.05	8.65	17.59
2022	5.67	12.97	28.15
2023	7.94	19.46	45.04
2024	11.11	29.19	72.06
2025	15.56	43.79	115.29
2026	21.78	65.68	184.47
2027	30.49	98.53	295.15
2028	42.69	147.79	472.24
2029	59.76	221.68	755.58
2030	83.67	332.53	1,208.93

Table 1 – Downstream Bandwidth per Subscriber Table

UPSTREAM DATA NETWORK TRAFFIC PREDICTIONS				
North America	Per Subscriber Traffic (Shown in Mbps)			
Year	10% CAGR Upstream	25% CAGR Upstream	35% CAGR Upstream	50% CAGR Upstream
2010	0.04	0.04	0.04	0.04
2011	0.05	0.05	0.06	0.06
2012	0.05	0.07	0.08	0.10
2013	0.06	0.08	0.11	0.15
2014	0.06	0.10	0.14	0.22
2015	0.07	0.13	0.19	0.33
2016	0.08	0.16	0.26	0.49
2017	0.08	0.21	0.35	0.73
2018	0.09	0.26	0.47	1.10
2019	0.10	0.32	0.64	1.65
2020	0.11	0.40	0.86	2.48
2021	0.12	0.50	1.17	3.72
2022	0.13	0.63	1.58	5.58
2023	0.15	0.78	2.13	8.37
2024	0.16	0.98	2.87	12.55
2025	0.18	1.22	3.88	18.83
2026	0.20	1.53	5.23	28.24
2027	0.22	1.91	7.07	42.37
2028	0.24	2.39	9.54	63.55
2029	0.26	2.98	12.88	95.32
2030	0.29	3.73	17.38	142.99

Table 2: North America Upstream Bandwidth per Subscriber Table

The Internet Traffic CAGR is determined by the consumers, service provider speed tiers offered, and technologies that will influence network usage. The growth rate of traffic will vary widely even within a service provider, because usage patterns will be different between demographics. Traffic growth rates are very hard to forecast because there are many possible influences to drive traffic growth.

Key Questions

- When will Nielsen Law of 50% CAGR for Max Service Tier Offered Break?
- Which "Downstream Traffic" CAGR do you believe and when will it break?
- Which "Upstream Traffic" CAGR do you believe and when will it break?

Understanding the Composition of Customer Traffic

The Over the Top (OTT) video providers such as Apple TV, Amazon, Hulu, Netflix, You Tube and others entering the On-Demand video market are driving traffic and peak period consumption upward, this is called Busy Hour Busy Day (BHBD) traffic utilization. In the figure below, from the Sandvine, Global Internet Phenomena Snapshot: 2H 2013 North America, Fixed Access report this illustrates that 67% of the downstream traffic is Real-Time Entertainment from these OTT providers. The percentage of Real-Time Entertainment has gone up over the last several years and a decade ago this was a rounding error as traffic in that time frame was dominated by web browsing and file sharing. It is very important we understand those types of traffic and the volume of traffic as a percentage. We can begin to understand to separate out say Video traffic from the traffic data and estimate the amount of users watching video over the top (OTT) as they are not watching an MSO delivered video offering. This is key for network planning.

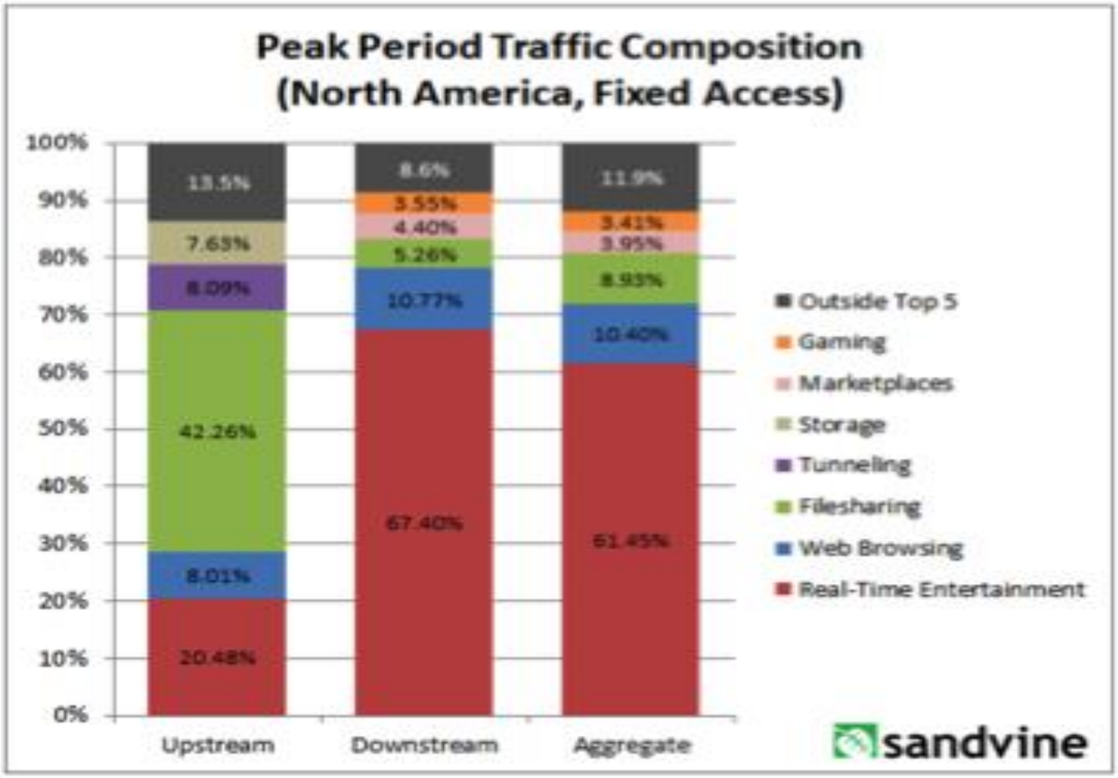


Figure 2 – Peak Period Traffic Composition Complied by Sandvine (2)

	Upstream		Downstream	
Rank	Application	Share	Application	Share
1	BitTorrent	36.35%	Netflix	31.62%
2	HTTP	6.03%	YouTube	18.69%
3	SSL	5.87%	HTTP	9.74%
4	Netflix	4.44%	BitTorrent	4.05%
5	YouTube	3.63%	iTunes	3.27%
6	Skype	2.76%	MPEG - Other	2.60%
7	QVoD	2.55%	SSL	2.05%
8	Facebook	1.54%	Amazon Video	1.61%
9	FaceTime	1.44%	Facebook	1.31%
10	Dropbox	1.39%	Hulu	1.29%
	Top 10	**66.00%**	**Top 10**	**76.23%**

Figure 3 – Traffic Application Share Upstream and Downstream Complied by Sandvine (3)

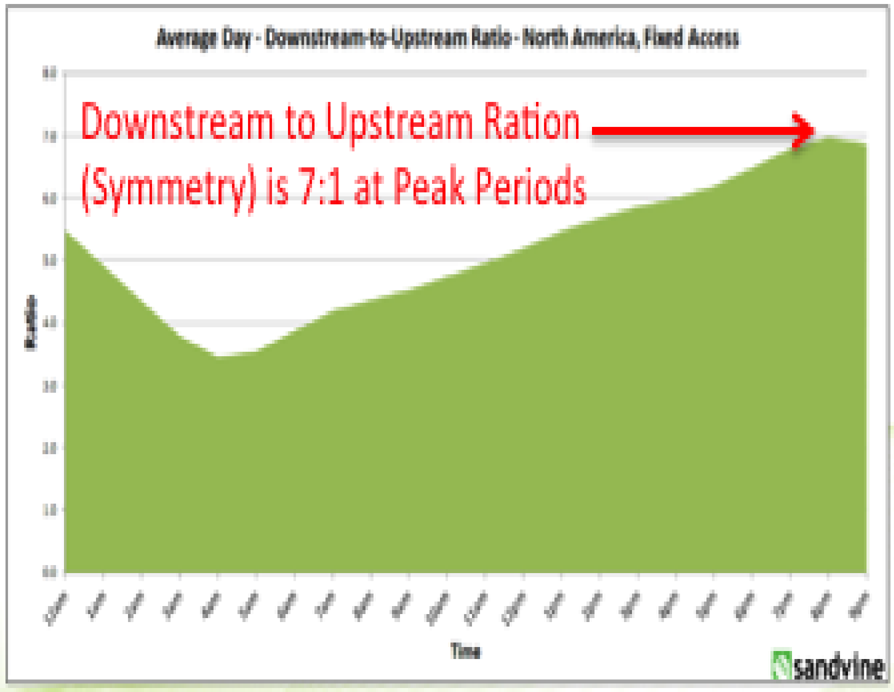


Figure 4 – Average Downstream-to-Upstream Ratio Complied by Sandvine (4)

In the figure above, the Downstream-to-Upstream Ratio is shown during an average day. This data useful to understand the ever increasing spread between the traffic directions. In the early 2000s, this number had far more symmetry during the peak of peer-to-peer some traffic assessments showed a near 1:1 ratio but more often, this ratio was 2:1. As seen in figures 2 and 3 the amount of video downstream is causing the ratio to spread dramatically. We will expect that until there is some upstream application that will consume bandwidth at a faster rate and duration during peak period this spread in the ratio of traffic will continue.

The Nielsen TV Research Company

In this section, we will examine the number of devices per home as well as the types of devices. The model will use both High-Speed Internet projections, like the Service Tier Offering and bandwidth per subscriber to predict Network Utilization and Capacity Planning. The number of TVs per household has been growing since 1975, but the rate of growth has declined in the last decade. According to the data from Nielsen, there is about 2.5 to 2.93 TVs per household according to year 2011 and 2010 data. The type of TV in each home is important for future network sizing as well as estimating the trend for the future like a transition from SD sets to HD sets and then HD set to UHD sets. Finally, according to the U.S. Census Bureau there are 2.61 persons per household between from the years 2008-2012. (7)

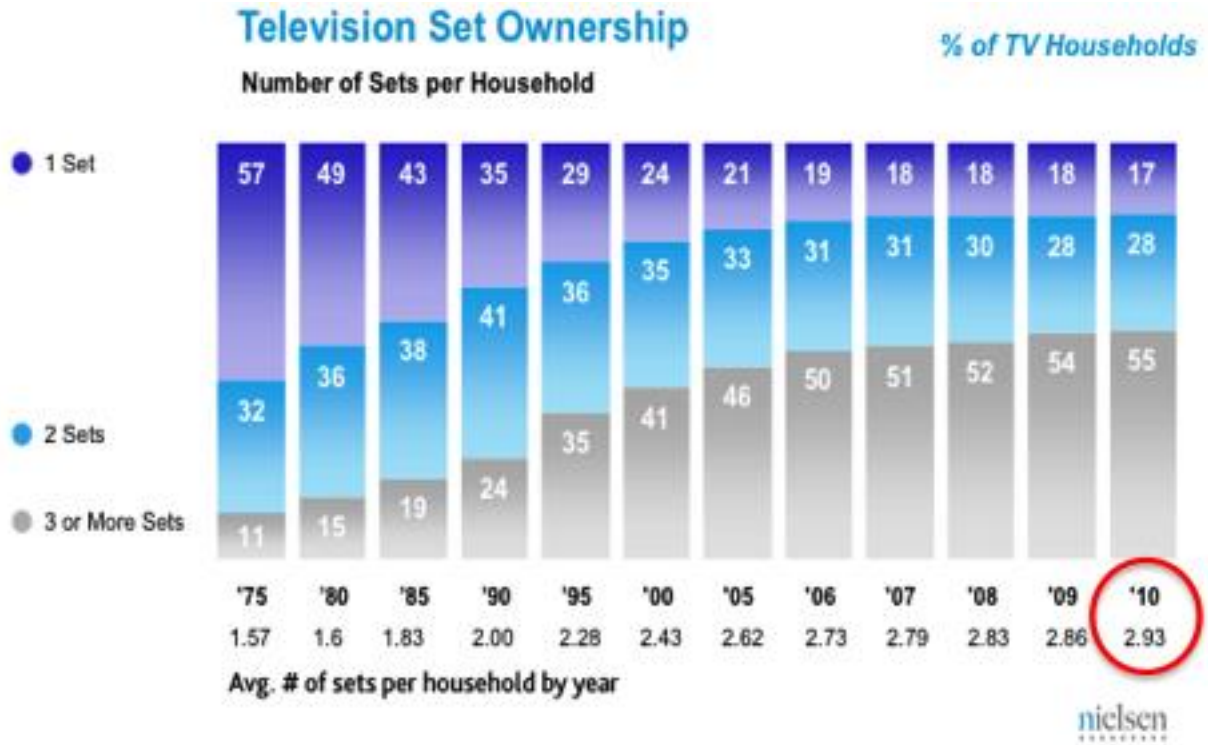


Figure 5 – Nielsen Data on TVs per Household – 35 Year Trending Model (5)

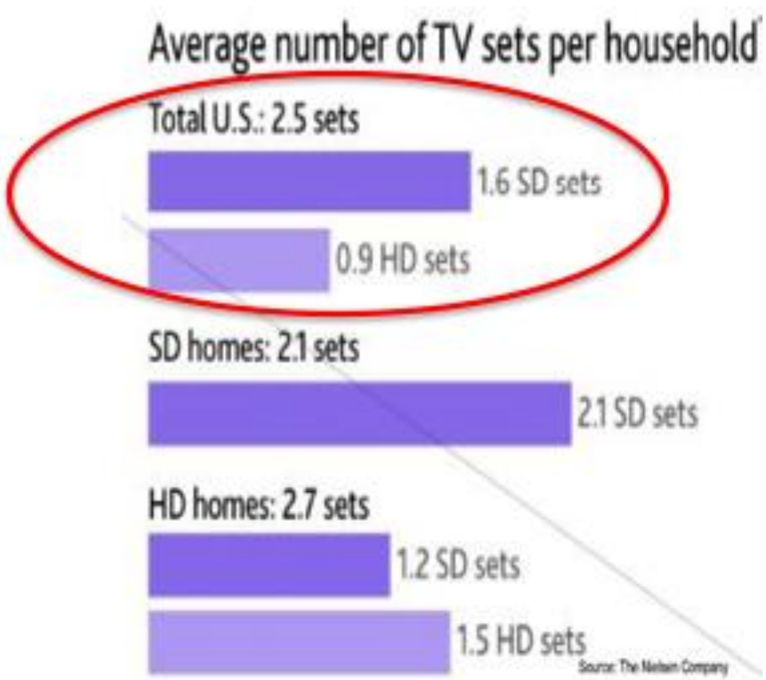


Figure 6 – 2011 Nielsen Data on TVs per Household U.S. TV Sets “Equal People per House (6)

SHOULD WE MIGRATE TO DOCSIS AS A UNIFIED VIDEO DELIVERY NETWORK FOR MSO "AND" OTT?

This question may seem a bit controversial, but if we consider some of the key points captured in this paper, then this may begin to become something that may be our target architecture. First, video is dominating the DOCSIS network and is consuming about 67% of the traffic during peak periods. We are building two highways and our customers pick which one to use. This will make capacity planning for parallel video delivery networks a challenge. What if we forecasted the traffic needs that would accommodate video traffic and whichever video content source it went over one network and that network was DOCSIS? What would be the capacity requirement if we applied some estimates based on the research collected in this paper?

According to Nielsen, the number of TVs per household is in the range of 2.5 – 2.93 sets. If we also consider the number of people on average per home which lines up with the number of TVs, as 2.61 we can start to forecast the support for video service needs. Therefore, we consider the migration of SD sets to HD sets, HD sets to UHD sets, and we use the HEVC data rates for HDTV and 4K UHDTV.

Resolution		Digital Compression Method and Bit Rate		
Resolution Terms	Frame Size / Scanning System / Frame Rate	MPEG2	H.264	HEVC
SDTV	480 / i / 30	3.7 Mbps	2 Mbps	1 Mbps
HDTV	720 / p / 30	6 Mbps	3 Mbps	1.5 Mbps
HDTV	720 / p / 60 (or 1080 / i /60)	12 Mbps	6 Mbps	3 Mbps
HDTV	1080 / p / 60	20 Mbps	10 Mbps	5 Mbps
4K UHDTV	4Kx2K / p / 60	80 Mbps	40 Mbps	20 Mbps
8K UHDTV	8Kx4K/ p / 60	320 Mbps	160 Mbps	80 Mbps

Figure 7 – Video Data Rates

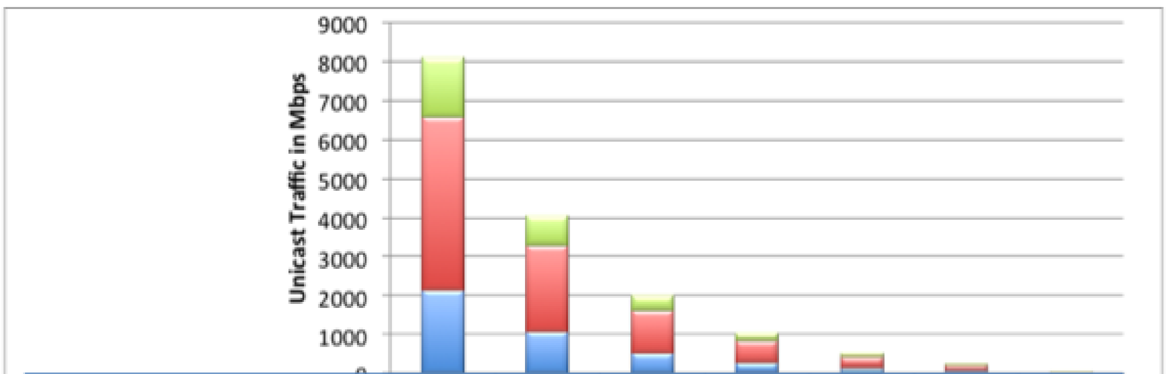
Key Observations Considered for Forecasting Capacity

- HSD traffic is growing at a faster rate than ever
- Traffic growth has not always grown at 50% CAGR
- 20-30% through most of 2000s was observed
- Symmetry Downstream-to-Upstream was not always 7:1 DS:US
- 2:1 through most of 2000s was observed
- OTT Video is the driver for the traffic growth
- Should we separate video planning from traditional HSD traffic?

- A person watching OTT video is not watching CATV service (and vice versa)
- Video Capacity Planning Factors
  - 2.61 number of people per household (2008-2012 U.S. Census)
  - 2.5 – 2.93 number of TVs per household (2010-2011 according to Nielsen)
  - 66% are SDTV and 34% are HDTV according to Nielsen
  - What if in the future 66% are HDTV and 34% are 4K UHDTV
  - What are some worst-case numbers for video planning?

Assumptions used in the figure below

- 2.61 number of people per household
- 2.61 number of TVs per household actively stream unique unicast content
- Assume the future mix of TVs with 66% are HDTV and 34% are 4K UHDTV
- Assume the use if HEVC for HDTV at 5 Mbps and 4K UHDTV at 20 Mbps
- Assuming a worst case of 100% of the TVs are receiving Unicast Video Traffic
- Year 2030 HSD Traffic Estimates Removing Video Traffic and use a 30% CAGR from 2010



Number of Customer Homes	250	125	62	32	16	8	1
Year 2030 HSD Traffic Estimates (without video)	1520	760	380	190	97	49	6
HEVC 4K UHDTV 4Kx2K / p / 60	4437	2219	1100	568	284	142	18
HEVC HDTV 1080 / p / 60	2153	1077	534	276	138	69	9

Values in Mbps

Figure 8 – Unicast Traffic Estimates for Video and Data Services

The main conclusion is that the migration to a single network for video delivery will likely extend the life of the HFC network and perhaps the service group levels. The figure above illustrates that a typical node and customer take rate could be served if we allocated capacity and spectrum to a single video delivery network. If there were two video delivery networks, MPEG TS and DOCSIS, the allocation of capacity would likely be higher, since it may be hard to predict which network a customer would use in a given evening.

HOW IS YOUR HIGH-SPEED DATA NETWORK MEASURED? SAMKNOWS

As many MSOs are aware, there are measurement services that measure the offered speed tier against the actual ability for the customers to reach those speed tiers and different test times and different test point. This figure below captures the architecture and test points of SamKnows.

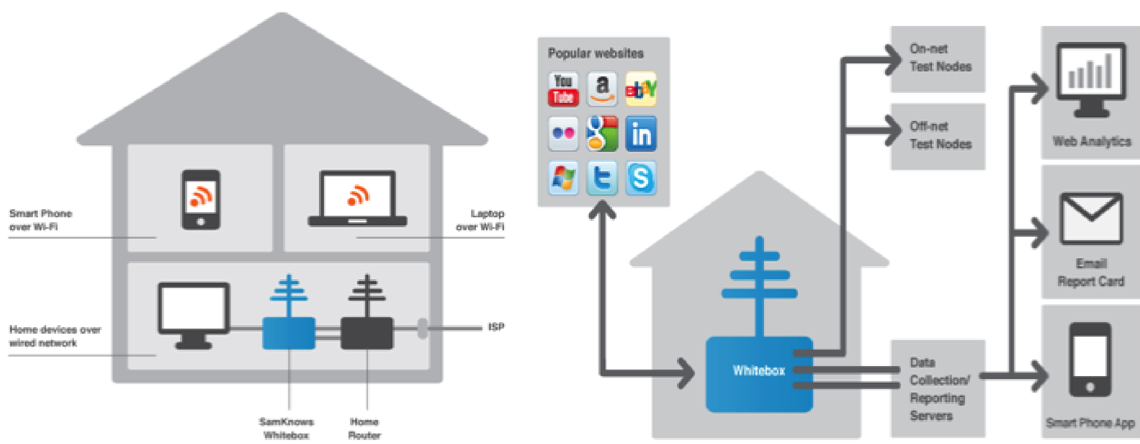


Figure 9 – SamKnows Installation Architecture and Test Points (8)

Service Providers believe it is very important to meet the service tier expectation of the customer and to do so at peak usage periods. As the service tier grows to higher data rates it is critical that we have methods to allocate the correct proportion of capacity. To that end, this next section will discuss traffic engineering and capacity-planning approaches used today and those possible in the future.

NETWORK QUALITY OF EXPERIENCE NQOE FORMULA (NQOE UNIT FORMULA)

This section will introduce a new method for sizing the service provider's data network, called the Network Quality of Experience (NQoE) Formula. The purpose of the formula is to account for traffic and service tier as well as growth rates to increase the probability of customers reaching the desired speed tier. The formula is described in more detail in the following three figures.

Rule of Thumb Approach for Network Sizing

To date some MSOs have sized their network on a method of multiplying the billboard speed by either doubling (2X) or tripling (3X) the billboard speed to determine the amount of DOCSIS capacity per service group, this is sort of a Rule of Thumb method for DOCSIS Network Sizing. If this rule of thumb approach is used as service groups get smaller, then too much capacity could be allocated.

NQoE Formula Goals

- Achieve Max Service Tier even during busy periods
- Allocate appropriate amount of network resources
- Configurable to accommodate any data network
- Accommodate estimates of Service Tier and Traffic Growth Rates
- Achieve Max Service Tier through next network capacity adjustment

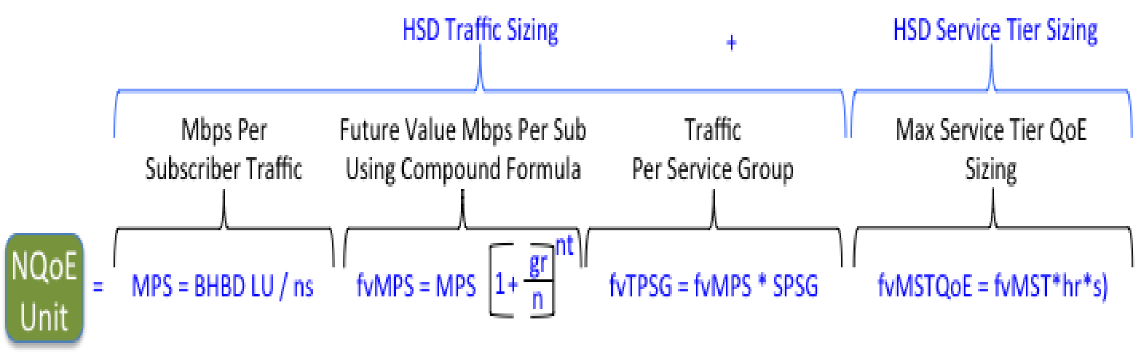


Figure 10 – Network Quality of Experience (NQoE) Formula

HSD Traffic Sizing

Mbps Per Subscriber Traffic

$$MPS = BHBD\ LU / ns$$

Where,

- **MPS =** Mbps per Subscribers (peak periods know as BHBD)
- **BHBD LU =** Busy Hour(s) Busy Day Link Utilization stated in Mbps (recommend to measure data links not at utilization thus all links measured should not exceed say 65% utilization but this depends on the data technology and even direction down/up link. Measuring links with high utilization will have inaccuracies for MPS)
- **ns =** Number of Subscribers served from the network link measured for BHBD LU

Future Value Mbps Per Sub Using Compound Formula

$$fvMPS = MPS \left[1 + \frac{gr}{n}\right]^{nt}$$

Where,

- **fvMPS =** Future Value of Mbps per Subscriber is an estimate for traffic utilization at the end of the data links planned useful life
- **MPS =** Mbps per Subscribers (initial or beginning value)
- **gr =** annual nominal growth rate (as a decimal, not in percentage) The MSOs must know the behavior of users and applications
- **n =** number of times the traffic is compounded per year
- **t =** number of years (must be expressed in years, because data growth rates are expressed annually, thus partial years are expressed as follows: 6/12=.5 or 15/12=1.25)

Traffic Per Service Group

$$fvTPSG = fvMPS * SPSG$$

Where,

- **fvTPSG =** Future Value Traffic Per Service Group
- **fvMPS =** future value of Mbps per Subscriber)
- **SPSG =** Subscribers per Service Group (Number of Subscribers served from the network link measured)

Figure 11 – Network Quality of Experience (NQoE) Formula – Traffic Sizing

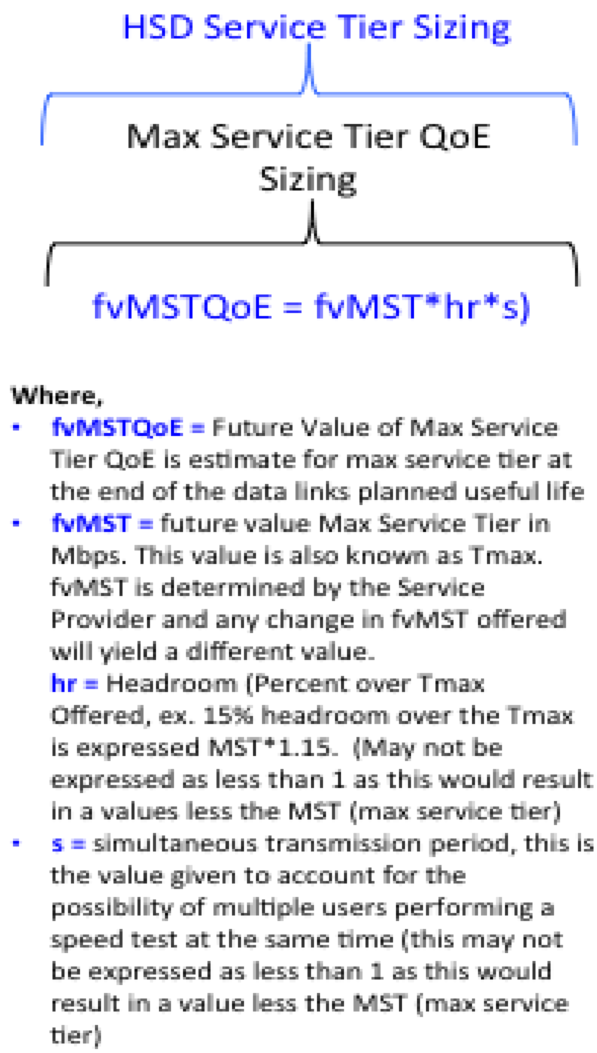


Figure 12 – Network Quality of Experience (NQoE) Formula – Service Tier

NQoE Unit is solved for Mbps in this example; however, the unit value may change if desired. NQoE Unit is a measurement for required data link throughput capacity, not link speed, and thus service providers will need to account for data link overheads and desired operational link utilization thresholds. ARRIS does have models that incorporate these factors.

NETWORK UTILIZATION AND CAPACITY PLANNING

Capacity Planning for High-Speed Internet Max Service Tier plus Data Traffic per Service Group

The downstream High-Speed Internet service tier growth from 2010 through 2030 is estimated and direction is used to forecast the date when the downstream may be at capacity see figure 13.

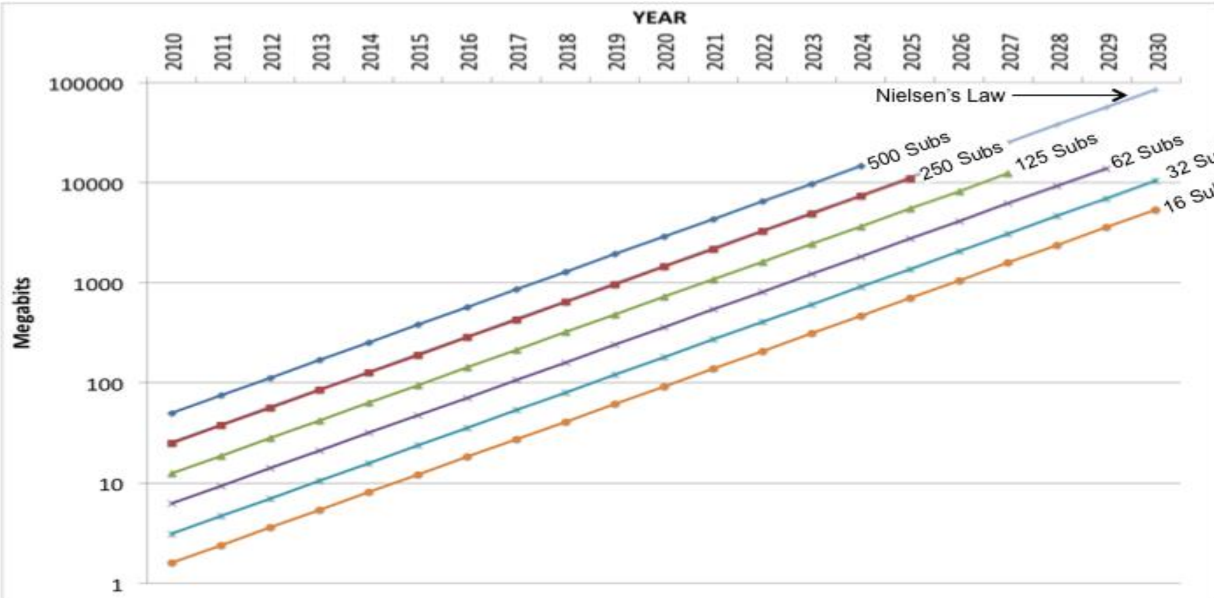


Figure 13: Nielsen's Law with Traffic per Service Group Estimates 2010 to 2030

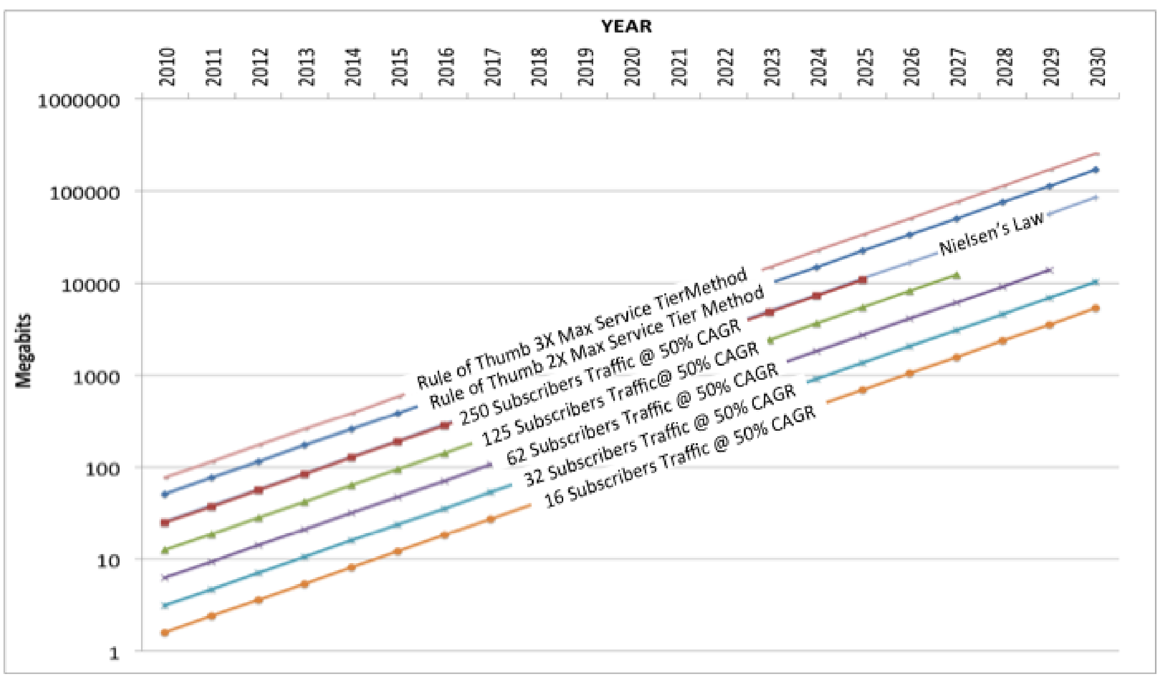


Figure 14: The Rule of Thumb Method and Traffic per Service Group 2010 to 2030

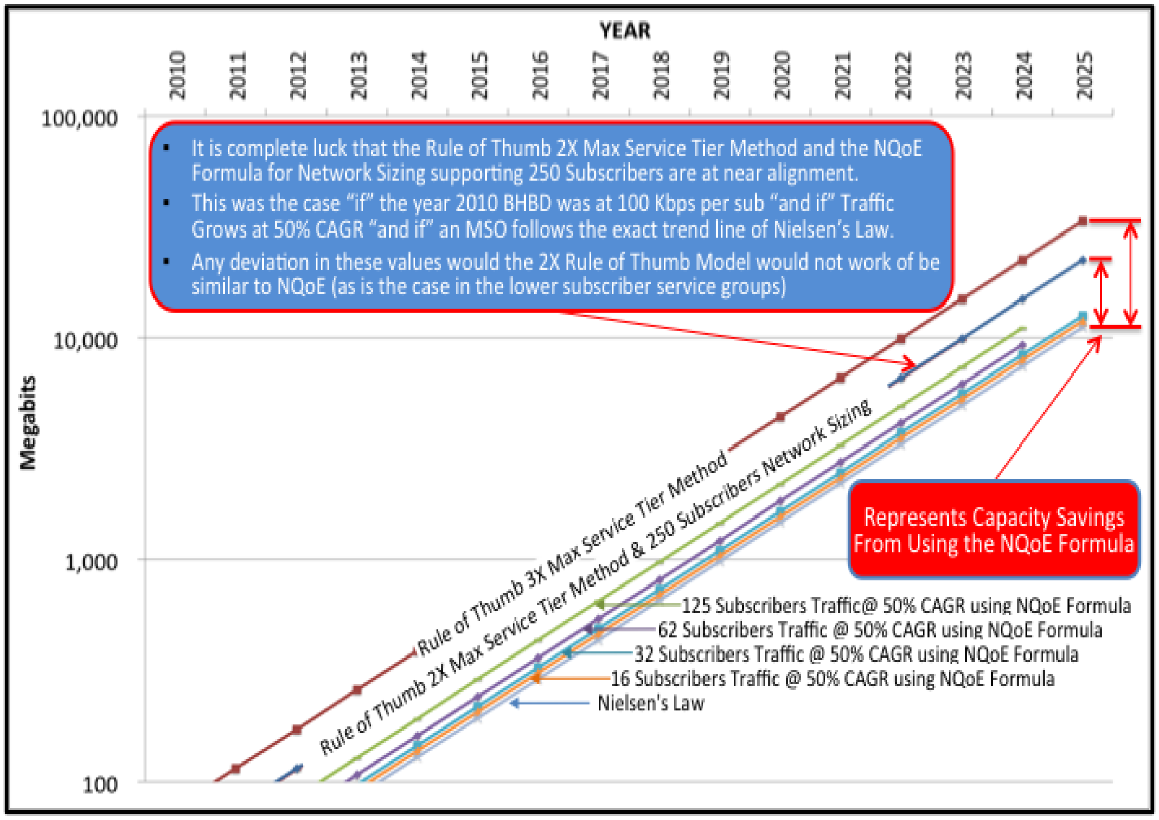


Figure 15: Rule of Thumb Method Compared to the NQoE Formula

The HFC downstream capacity assumptions will use several reference points; these include 192 MHz, 384 MHz, 576 MHz, 768 MHz and 960 MHz of usable DOCSIS downstream spectrum. These assume High-Speed Internet Max Service Tier and Traffic continues at a 50% CAGR. It again should be stated that these are predictions for the next decade or more, it is uncertain if growth for either or both will continue at this pace, see figure 16.

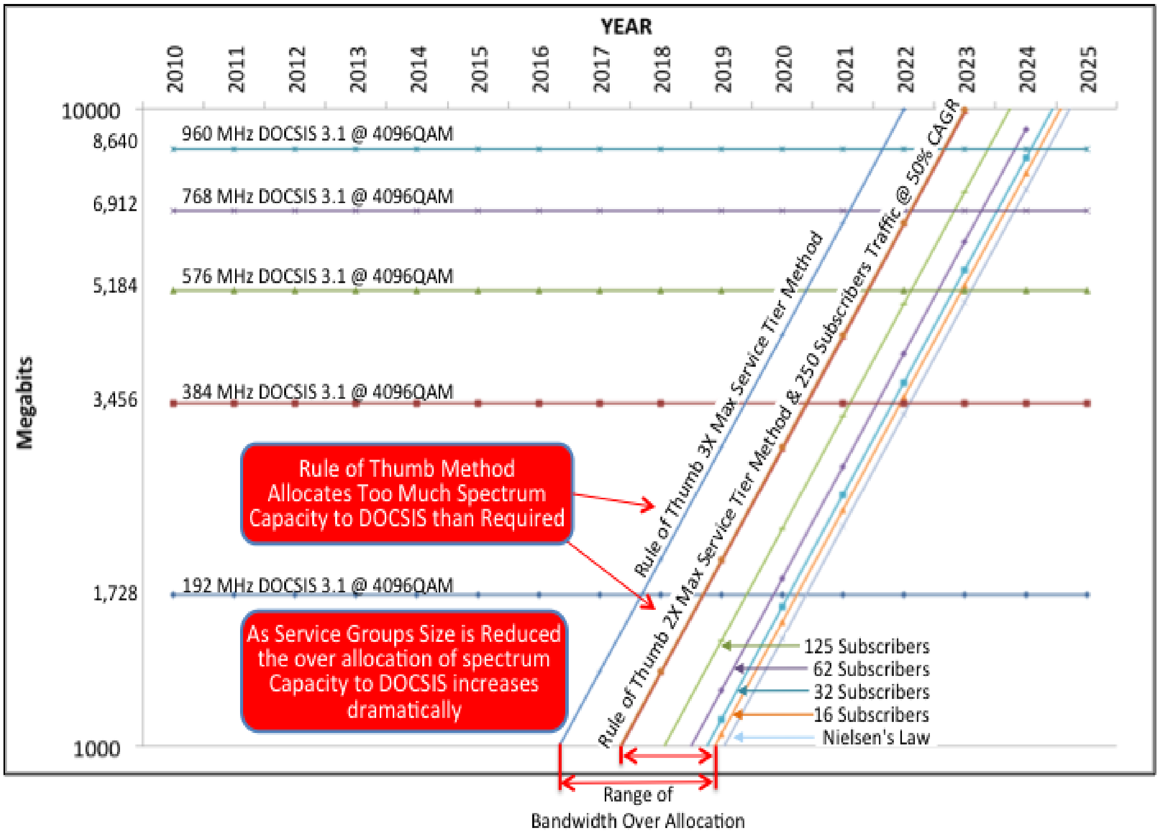


Figure 16: DOCSIS 3.1 Capacity with Rule of Thumb Method and NQoE Formula

Figure 16 uses the NQoE Formula, the combination of Service Tier plus Traffic per Service Group as well as other factors to estimate network capacity needs. The horizontal lines represent spectrum allocation and modulation to determine the network capacity. In figure 16, the red boxes illustrate that the rule of the thumb methods may allocate too much capacity as service group size is reduced. Additionally, since this does not take into account traffic or service group size among other factors, the NQoE Formula may more accurately forecast network capacity requirements.

CONCLUSIONS

Operators will need to track these key drivers and levers that force network change, like Nielsen's Law of Max Services Tier Growth Rate and also Traffic Growth Rates for proper network planning. In terms of targeting a single video service delivery network, this would likely reach the same network capacity, yet other costs need to be considered. The main conclusion is that the migration to a single network for video delivery will likely extend the life of the HFC network and perhaps the service group levels. The figure above illustrates that a typical node and customer take rate could be served if we allocated capacity and spectrum to a single video delivery network. If there were two video delivery networks, MPEG TS and DOCSIS, the allocation of capacity would likely be higher since it may be hard to predict which network a customer would use in a given evening. The use of the Network Quality of Experience (NQoE) Formula will help service providers of all types and technologies size their network more accurately.

Conclusion Summaries:

- Drivers for Traffic Engineering and Capacity Planning:
    - Nielsen's Law will dominate network capacity allocation
    - We must understand the Traffic Composition like OTT video
- Unicast video delivery method may actually extend the life of the network:
    - OTT utilizes unicast video over DOCSIS
    - MSO allocating capacity for unicast and broadcast is over MPEG TS
    - MSOs are really support two video delivery networks
    - The challenge is traffic engineering because which one will the customer use
    - Planning Video Capacity for two networks is costly and a long terms guessing game
    - Which network will the customer use for video (DOCSIS or MPEG) tomorrow, month, year, etc. from now?
    - Eventually migration to full spectrum DOCSIS creates one service delivery network to manage
    - Advanced Video CPEs accommodate consumers TV choices while preserving capacity
- We need a new Traffic Engineering and Capacity Planning formula:
    - NQoE Formula (Network Quality of Experience) Formula for Capacity Planning
    - Achieves Max Service Tier even during busy periods
    - Allocates appropriate amount of network resources

REFERENCES

(1) Nielsen's Law Max Internet Service Tier Offered – Downstream. Source: http://www.nngroup.com/articles/law-of-bandwidth/
(2) Sandvine, Global Internet Phenomena Snapshot: 2H 2013 North America, Fixed Access, Available for Download at www.sandvine.com
(3) Sandvine, Global Internet Phenomena Snapshot: 2H 2013 North America, Fixed Access, Available for Download at www.sandvine.com
(4) Sandvine, Global Internet Phenomena Snapshot: 2H 2013 North America, Fixed Access, Available for Download at www.sandvine.com
(5) Nielsen Holdings, http://www.nielsen.com/us/en/newswire/2010/u-s-homes-add-even-more-tv-sets-in-2010.html
(6) Nielsen Holdings, http://www.nielsen.com/us/en/newswire/2011/factsheet-the-u-s-media-universe.html
(7) United State Census, http://quickfacts.census.gov/qfd/states/00000.html

Not All 4k Is Created Equal: Finding Clarity, Reducing Costs

Theron Trowbridge, Director of Encoding Operations & Advanced Video Technology
Deluxe Digital Distribution

Abstract

4K is the new leading television technology that will capture consumers' imagination. While the industry recognizes the potential, the processing, workflow, and delivery mechanisms must be re-defined for mass scale as subscribers embrace this exciting new viewing experience and 4K UHD devices proliferate.

As 4K unfolds, multichannel video programming distributors (MVPDs) will assume that files delivered by studios and content providers will conform to the optimal viewing experience across the multiscreen device landscape. In reality, all 4K is not created equal.

To prepare for 4K processing, workflows and distribution, MPVDs must sort through multiple issues to ensure that 4K-optimized content measures up to the requirements of a superior customer experience. In this paper, we will begin to frame the conversation to address the many questions that MPVDs will have to answer to distinguish true 4K-optimized quality content from everything else going by that name. We also consider steps operators can take to position themselves as leading providers of very high-quality content and the ways in which use of cloud-based management and storage can greatly reduce the costs of offering a 4K service.

INTRODUCTION

Faster than many experts anticipated, 4K Ultra High Definition (UHD) TVs are moving toward initial commercialization, bringing with it a new set of issues and requirements that will have to be addressed by MPVDs in preparing for and enabling distribution of 4K-optimized content.

These issues revolve around properties specific to two major areas of technology development: the 4K platform itself and the new encoding standard, High Efficiency Video Coding (HEVC) or H.265, which will be required to reduce the bandwidth constrains imposed by distribution of programming in the 4K format. Our goal here is to discuss the broad range of considerations that must be addressed as MPVDs set parameters essential to ensuring their 4K services deliver a quality of experience commensurate with market expectations, at costs commensurate with the pace of 4K UHD TV set penetration and consumer demand.

Fundamentally, it should be understood that there will be variations in the quality of 4K formatted titles from content providers and post-production houses. This is based on nuances having to do with parameters used in the preparation of content from sources that are either 4K native shot and mastered production or in the conversion, restoration, and re-mastering from film or High Definition (HD). Furthermore, to ensure that the playback experience is optimized on target devices, 4K files will need to undergo rigorous testing and quality analysis of frame rates, bitrates, color gamut, bit depth, and decoding algorithms.

Operators, programmers and technology organizations such as CableLabs® will have to work together to determine what the thresholds should be across several parameters to ensure the Quality of Experience (QoE) targets are met. These include:

- Color gamut and depth
- Contrast depth
- Frame rate
- Mode of converting legacy content to 4K
- HEVC profile
- HEVC-encoded distribution bitrate

Beyond these core technical concerns, MPVDs 4K strategies will need to take into account a variety of business and cost-related issues. These include:

- The benefits and technical requirements associated with delivering a pre-4Kservice, where content formatted in 1080 progressive (1080p) HD can be up-scaled by 4K UHD TV sets for high-quality display;
- Determining the right balance between relying on HEVC-capable set-tops to support 4K pay TV delivery vs. IP-based distribution over broadband to HEVC-enabled IP-connected devices, including Smart TVs as well as smaller portable consumer electronic (CE) devices;
- The cost impact of accommodating workflow, ingest, storage and file management requirements of 4K-optimized content and the degree to which CapEx and OpEx might be alleviated through use of cloud-based management and storage.

EVALUATING PARAMETERS ESSENTIAL TO ACCEPTABLE QoE

Color Space

There remains some uncertainty within the industry's production circles over the definition of the ideal range of color gradations for4K, as expressed in three leading color space formats:

- the ITU's Recommendation BT.709 or rec.709, the foundation color depth standard for digital TV since 1992;
- rec. 2020, a new standard for next-generation TV, including 4K UHD;
- Dolby Vision, a proprietary system designed to comport with the requirements of High Dynamic Range (HDR) imaging.

Early iterations of 4K UHD TVs on display at leading events and trade shows in 2013 relied on rec. 709, which prescribes a bitrate of 8-bits per color with 256 gradations per primary color, for a total of approximately 16.8 million possible colors. The industry consensus was that the higher resolution enabled by 4K UHD at 3,840 x 2,160 pixels, twice the pixel width and height density of HD 1080p resolution, was not a strong enough differentiator to drive consumer demand for the new TV sets, given that the color gamut used with those early displays was the same used by HD.

However, with the introduction of rec. 2020 on chipsets and in next-generation displays appearing at events in 2014, the consensus shifted to a thumbs-up on the appeal of 4K compared to HD, whether the display was on a large form factor TV or a small-screen handheld device. With a coding scheme of 10- or 12 –bits per color, rec. 2020 generates about 1.07 billion colors with 10-bit coding and a whopping 68.7 billion colors with 12-bit coding.

At this point, the market is increasingly focused on the 10-bit version of rec. 2020 with color depth of 1024 gradations per primary color. While 10-bit vs. 8-bit might seem to represent a 25 percent increase in encoding bit rate, the actual margin of increase is much lower, owing to compression methodologies and narrower gamut of legacy content. However, 10-bit does represent some increase in bit rate over 8-bit encoding.

Further boiling the waters is the emergence of Dolby Vision as a contender for adoption as the de facto 4K capture and display standard. This is based on what proponents and many objective observers believe to be a significantly better quality picture than seen so far with 4K. Dolby Vision uses the 12-bit version of the rec. 2020 color palette and introduces far greater contrast capabilities in line with HDR imaging.

Some vendors are touting HDR imaging capabilities as well. But Dolby appears to have gone much farther in gaining market penetration with leading Consumer Electronics (CE) manufacturers including Sharp, TCL and Vizio. They are already producing 4K UHD TV models based on Dolby Vision. Several Web video providers, including Amazon, Vudu and Microsoft X-box, have committed to producing HDR-caliber content using Dolby Vision, if and when a consumer market emerges for the format. While there's been much focus on Dolby's demonstration of a prototype display generating eight times the brightness of conventional displays, the real goal of HDR is not so much to increase luminance as it is to generate greater contrast across whatever the range of luminance might be for a given display.

At the present time, the best bet for driving consumer adoption of 4K is 10-bit rec. 2020-based display technology. That said, it's possible the baseline for 4K-optimized content could shift between now and when the threshold of 4K UHD TV set penetration reaches a point where it becomes necessary to introduce 4K services. The prudent thing will be to develop strategies on the assumption that quality parameters will revolve around a rec. 2020 baseline while being nimble enough to adjust should conditions change.

Frame Rate

Another factor that remains an open question is the optimal frame rate for 4K services. A common misconception is that 60 frames per second (fps) is a given requirement with 10-bit rec. 2020. In truth the optimal frame rate is a function of the need to minimize the number of transformations any content file goes through prior to staging it for distribution to consumers. Any time a file goes through a transformation, say, from 23.976 to 60 fps, the process introduces artifacts that far outweigh the benefit of moving to the higher frame rate.

Consequently, it's better to accept files natively captured at the standard film industry rate of 23.976 fps or the TV programming rate of 30 fps rather than insisting they be reconfigured to 60 fps. In any event, where natively stored 30 fps content is concerned, the frame rate can be increased to 60 fps in the distribution transcoding process, since it's a multiple of the native rate. The fact that 120 Hz is now the top frame rate supported by most 4K UHD TV sets greatly simplifies matters, since it is a multiple of all these lower rates. The fact that 120 Hz is a 5x multiple of 24 has contributed to the ability of 4K UHD TV sets to work compatibly with Blu-ray, which is pegged to the 23.976 fps film rate, even to the point of performing upscale processing to enhance the quality of the Blu-ray HD content when displayed on 4K UHD TV sets.

MODES OF CONVERSION TO 4K

Another dimension in a MPVDs' search for baseline requirements is the impact of the various modes of format conversions from HD and film to 4K will have on the quality of content. Early on, most content licensed for network-delivered 4K will likely be content such as older movies that have been converted to 4K.

MPVDs will have to be vigilant in their acquisition of converted content insofar as some processes used in the conversion process won't produce 4K material that measures up to MPVD quality standards. The move to specialized 4K conversion capabilities is better understood by lab-based post-production houses who can address the right balance between quality and economics to build large 4K libraries.

Upscaling

The lowest-cost approach and therefore one that's received early market appeal involves upscaling of the original files. In its most basic form, upscaling does little more than quadruple the number of pixels. However, there are some versions of the upscaling process that attempt to augment the quality through algorithmic processes and add information by "guessing" what a rendering with deeper color depth and contrast would look like.

While upscaling may offer a fast track to library building, operators will have to decide whether any of the upscaling methods will meet the higher quality expectations that subscribers expect and demand.

A special case in addressing this question is how to convert content originally shot in 4K but which has been down-scaled to 2K to accommodate the prevalence of 2K projectors in theaters. Here, the gap to be closed in the scaling process is smaller, but operators will have to decide based on technical analysis of scaled content whether such techniques will be sufficient for processing this type of content.

Scanning

MPVDs will be on much more solid ground to the extent they can rely on suppliers of 4K content who use more advanced conversion methods such as scanning to convert 2K and film to 4K. But there are nuances that will impact quality with these methods as well.

One approach involves scanning each frame individually to create a sequence of uncompressed discreet image files, usually in the Digital Picture Exchange (DPX) format, which in turn are processed with various restoration imaging applications, including applications that bring the 4K color palette into play. An alternative approach involves continuous scanning, basically an enhancement of the old Telecine restoration process, which scans the frames as the film runs through the scanner - slower than real time but faster than the one-second-per-frame speed of the intermittent frame scanning process. Either way, the processing procedure also repairs defects in old films such as scratches and removes dust and other materials, leaving a pristine, very large master file which must then be compressed for storage.

No matter what the scanning methodology might be, it's important to recognize that additional steps must be taken when it comes to encoding the converted content for storage to avoid introducing excessive graininess into the completed file. Older film stock, being grainier than newer films, is especially problematic in this regard.

While film converted and encoded digitally for display on today's HD TV sets has not posed much of a problem, use of HEVC and the higher resolution of 4K UHD TVs will display the picture more accurately than has been possible in the past. HEVC, by virtue of the more advanced processing techniques used with this encoding protocol, captures the graininess with greater clarity than was the case with H.264 or MPEG-2, and 4K makes any graininess far more apparent to viewers. Whether and to what extent this intensification of graininess needs to be remedied through filtering or other additional processing in the conversion process is another issue that content owners will have to address prior to releasing content to MVPDs.

Beyond scanning, there's an even more expensive and comprehensive conversion method where computer generated imaging (CGI) is used to recreate the film digitally in 4K. Experience shows – and is reflected in 4K conversion choices being made by motion picture studios – that a properly administered advanced scanning process can achieve quality levels comparable to native 4K-originated content.

For example, Deluxe Media Services is applying continuous scanning, film grain filtering, artifact removal and related processes to the 4K conversion of growing numbers of movies with outstanding results. As techniques evolve and more usage generates more feedback, Deluxe will continue to refine the parameters to meet MPVDs' requirements.

BITRATES AND HEVC

How MPVDs work through the issues raised in the foregoing discussion will greatly impact their approach to setting transmission bitrates for delivery of HEVC-encoded 4K-optimized content. There has been a lot of discussion in the 4K space that has emphasized how little bandwidth will be needed to deliver a compelling user experience, with some citing 15 mbps as a target rate for its 4K rollout.

Over time, HEVC, as was the case with previous codecs, will become more efficient, allowing bitrates to fall for delivering content at any given level of quality. But at s point, the quality level that MPVDs will likely want to support will require the setting of bitrates for 4Kcontent well above the rates cited by some in the industry.

Early on a key choice to be made is what profile – Main 10 or Main – operators will require for HEVC encoding. Main, the original profile set with the approved standard, was designed to support 8-bit color with a sampling depth of 256 levels – in other words the parameters used with rec. 709. Late last year, Main 10 was added to the standard to enable encoding of content utilizing 10-bit color with 1,024 sampling levels – specifically, the parameters employed with rec. 2020.

At this point both Main and Main 10 limit chroma subsampling to 4:2:0, but extensions are scheduled to be added in 2014 that will raise the levels to 4:2:2 and 4:4:4, along with introducing multi-view video coding for 3D applications. All of this will require additional bandwidth headroom should operators decide to adopt such extensions. Frame rates, too, will be a factor in setting transmission bitrates.

The state of the art in encoding, notwithstanding some aggressive claims, appears to require something on the order of 25 mbps or above for rec. 2020-compatible 4K transmissions from stored content delivered at 60 fps. Live feeds are likely to require higher rates, especially in sports programming. Supporting the chroma subsampling extensions will increase the optimal bitrates as well.

In light of all this, MPVDs will probably want to look at new approaches to QAM allocations, where 4K might use the majority of the 38 mbps available on a 256 QAM while allocating the remainder to other HEVC-encoded or even MPEG-2 or H.264 encoded content. The picture, of course, will change with the coming of 1024 QAM, a feature of the DOCSIS 3.1 spec. By the time 1024 QAM goes into wide use, there will likely have been considerable gains in HEVC efficiency, allowing operators to pack two or three 4K channels per QAM in tandem with the need to introduce more 4K feeds.

PREPARING THE MARKET

Meanwhile, in the run-up to introduce true 4K and HEVC, MPVDs may want to consider adjusting QAM allocations along these lines sooner than later. This is necessary to support a move to 1080p HD as a way of gaining competitive differentiation as providers of extraordinarily high quality content to 4K UHD TV set owners. In other words, by dedicating more QAM bandwidth to transmitting higher bitrate 1080p content, operators would be able to provide content amenable to upscaling on 4K UHD TV sets to quality levels approaching 4K while providing non-4K households a quality of HDTV that offers a 10 to 20 percent or better improvement over customary lower bitrate 1080 interlace (1080i) content.

Upscaling of 1080p by 4K UHD TV sets has become a key selling point, given the improvement upscaling delivers to Blu-ray 1080p content. Operators offering1080p HD to 4K UHD TV set owners would be helping to satisfy those subscribers' hungry for high-quality content, while making them aware that the cable company will be their source for true 4K-optimized content as it becomes available.

The emergence of HEVC opens additional possibilities for cable services, including the opportunity to consume lower bandwidth in unicast deliveries of premium content to IP devices that are equipped to decode HEVC. Many researchers have noted that, given the fast turnover rate and rapid increase in processing power of CE devices like smartphones and tablets, HEVC decoding capabilities will rapidly populate these markets. Multimedia Research Group, for example, says one billion devices capable of decoding HEVC were in the market by year-end 2013.[1] But, the researcher says, penetration of the set-top market will take much longer.

This advantage on the broadband side also applies to 4K UHD TV sets, which are Smart TVs equipped to access IP premium content directly without the need for a set-top box. Opening this conduit for 4K distribution by delivering pay TV service apps to such device would allow operators to expand their 4K service reach faster than they could by relying solely on the introduction of HEVC-capable set-tops.

[1] "HEVC Decoding in Consumer Devices," Multimedia Research Group, March 2013

Cable operators will gain another advantage over competitors as they obtain licenses to first run releases for distribution in4K. The motion picture industry is introducing new security requirements for this extremely high-value content, including forensic watermarking that will associate each unicast of an on-demand file with the viewing household. Operators are likely to be in a far better position to implement advanced security requirements over their managed networks and devices versus those who operate in the open security provisioning environment of the Internet.

LEVERAGING THE CLOUD TO REDUCE COSTS

Whatever approaches MPVDs take to bring 4K-optimized content to market, it's inevitable that most of this content will be delivered from storage in an on-demand mode in the early stages. It will take a considerable amount of time for an end-to-end HEVC/4K ecosystem capable of supporting live 4K programming to emerge. In sports programming, which is seen as the initial driver to live 4K broadcasting, "just at the field" production level will require a new system to be implemented to run in parallel with the legacy system. This is needed to support different degrees of scene panning, approaches to slow-motion replay and other elements affecting the picture and the textual and graphics trappings.

Of course, as the editing systems, distribution pipelines to head-ends and beyond, and equipment in the home are installed to accommodate live 4K programming, the volume of 4K-optimized files positioned for on-demand access will grow, eventually reaching the quantities common to today's HDTV-based video on demand (VOD) systems. The infrastructure and storage burden alone, adding to costs already incurred with expansion of on-demand service into the multiscreen domain, will be immense. This raises the natural question of whether operators can sustain legacy approaches to amassing VOD content.

In contemplating what can be done to alleviate these costs, it's important to recognize what the true costs of sustaining current approaches would be. A thorough analysis of what goes into building an in-house VOD library must include operating expenses, initial capital outlays and ongoing capital investments across a task list that encompasses workflow system development and management, ingest processes, metadata and content processing, quality control, storage, distribution and reporting.

At the same time, further analysis will reveal how much can be saved through a strategic approach that exploits the cost-savings benefits of a sophisticated cloud-based asset storage and management system. For example, in the instance of mounting multiscreen on-demand services, operators are discovering that whether they rely on such a service to process, store and deliver their entire portfolio or use such a service to build on the existing in-house infrastructure, they can cut the costs of operations by hundreds of dollars per title.

Such savings will apply as well to the use of the cloud to support expansion into 4K distribution, especially as that service, too, requires multiple formatting and encodes to reach multiple types of connected devices. To the extent such costs can be shared across multiple users of the cloud, the per-title costs accruing to each operator will be a fraction of what they would be if all these processes had to be performed individually for each pay TV distributor. Moreover, beyond cutting costs, the opportunity to rely on best practices developed by experts in content management, processing and storing 4K-optimized content for access in the cloud will help operators sort through the complex issues that go into establishing benchmarks for delivering a superior 4Kexperience to their customers.

Companies addressing these issues via the cloud include Deluxe Digital Distribution. Through decades of managing master source files for Hollywood studios, they have developed critical know how in meeting the requirements of the new content supply chain including management, ingest, processing, storage and distribution of content. Such best-of-breed cloud solutions eliminate the need for heavy investments in storage and processing infrastructures and workflows while providing operators affordable, easy access to libraries of UHD-optimized titles.

CONCLUSION

While 4K is widely regarded as the inevitable next stage in the evolution of television technology, there are many questions yet to be resolved between now and when 4K content and services enter the cable TV service mix. This starts with issues related to minimum requirements to ensure the movies and TV programming offered on-demand to subscribers meet expectations for a superior viewing experience and delivered at the highest quality and optimal bitrate for bandwidth, transport and storage efficiencies across the multiscreen landscape.

There is and will continue to be great variation in the fundamental characteristics of content rendered and labeled as 4K in postproduction, including differences in color gamut and depth, contrast, frame rate and degrees to which artifacts are removed and film grain levels are matched to the original work of art. It will take time to determine what the minimum parameters should be for cable-caliber service and to identify the types of conversion methods that can be trusted to produce acceptable 4K-optimized content. In tandem with resolution of these issues, service providers will have to determine the amount of bandwidth to be allocated to 4K in the context of setting optimal bitrates for HEVC encoding to address the congestion and capacity issues at the network core and edge.

Clearly, these issues will have to be addressed in the context of a competitive environment. As more subscribers purchase 4K UHD TV sets, operators will be pressured to satisfy demand for higher quality content suited to viewing on large-screen 4K UHD TVs. Consequently, operators may want to begin planning for allocation of QAMs to support delivery of a high-quality HD service in 1080p that can be upscaled on 4K UHD TV sets to provide a superior viewing experience in advance of true 4K-optimized content. Operators will also need to address the question of they want to expand the reach of 4K service by introducing 4Kover IP broadband for direct access on connected TVs and other devices, in parallel with reliance on use of HEVC-capable set-top boxes.

Underlying all these preparations is the cost question. The need to create a support infrastructure, including workflows, ingestion, asset management, storage and other processes for 4K on top of ongoing expansion of resources for traditional VOD and TV Everywhere, imposes cost burdens that will be hard to sustain using old approaches to building VOD libraries. The logical alternative will be reliance on cloud-managed services where the costs of

building and maintaining 4K libraries can be
shared among multiple operators.

Acronym List

Computer Generated imaging (CGI)
Consumer Electronics (CE)
Digital Picture Exchange (DPX)
Frames per Second (fps)
High Definition (HD)
High Dynamic Range (HDR)
High Efficiency Video Coding (HEVC)
Mutichannel Video Programming Distributors **(**MVPDs**)**
Over-the-Top (OTT)
Quality of Experience (QoE)
Ultra High Definition (UHD)
Video on Demand (VOD)
1080 progressive (1080p)
1080 interlace (1080i)

PARTNERSHIP FOR EXTENDED CAPACITY: DOCSIS 3.1 WITH RFOG

Venk Mutalik, Marcel Schemmann, Ayham Al-Bana, Zoran Maricevic
ARRIS Group

Abstract

In an RFoG environment, where many of the physical limitations of the HFC cable plant do not exist, DOCSIS 3.1 could achieve significantly more capacity. A well thoughtout partnership of DOCSIS 3.1 and RFoG can satisfy the anticipated growth in traffic demand while providing MSOs with plant and equipment investment protection well into the future.

As MSOs consider an HFC split change or if moving from HFC to FTTH, it is important for the upgraded plant to have a life of at least a decade or more. To support the bandwidth requirements of a service group of around 100 subscribers, between 10 and 100 Gbps downstream capacity and of several Gbps upstream capacity is likely to be needed.

It is shown in this paper that plain NRZ binary modulation formats cannot meet the above demand in the presence of high fiber dispersion. In this paper, we show a path to a much more efficient use of the RF spectrum, which when used with RFoG technology and features available via the DOCSIS 3.1 standard - and combined with innovative low cost ONU designs using current technology - addresses the anticipated growth rates and bandwidth expansions seamlessly over the next decade and more.

INTRODUCTION

Anticipated growth in data capacity over the next decade may be covered in several ways. These include HFC - possibly with a split change - supporting greater than 10Gbps downstream and several 100 Mbps to Gbps upstream capacity. The various PON variants, chiefly the 10G PON and point to point Ethernet with an optional video overlay could support up to 10 Gbps of data capacity.While RFoG is mentioned, it is often presented as a transitional technology that permits an easy transition to PON for the MSO.

In this paper, a case will be made that technologies such as the 10G PON may not offer enough downstream capacity to justify the cost of a system change sooner than a decade from now. It will be shown that 25-40 Gbps is a more reasonable target for downstream traffic and around 2.5-4 Gbps upstream a reasonable target for upstream traffic capacity.

While 40G PON technology may provide this bandwidth, such a technology requires multiple wavelength receivers or other complex technology in the ONU that is inherently expensive. Here it is shown that an RFoG ONU can be designed cost effectively with current technology that can support these requirements, such that an investment made by MSOs in these ONUs is much more robust. Furthermore such an ONU is compatible with current HE and CPE equipment and further HFC plant extensions, thus extending the life of such investments.

We will first discuss data traffic growth. Next, an RFoG system with an extended downstream band is introduced and SNR and data throughput analysis is performed. An RFoG system with extended upstream bandwidth is also then introduced and discussed.

TRAFFIC GROWTH AND MODELING

Relentless appetite for over the top content and interactivity further fueled by the imminent advent of 4K television standard and higher standards to come prtents a substantial growth in avarge and maximum advertised capacity over the next decade. Traffic is expected to grow to 10-100 Gbps downstream in 2030 for a reasonable service group size (128 HHP), upstream traffic is expected to grow to 2.5-4 Gbps upstream. A target of 40 Gbps downstream capacity and 2.5-4 Gbps upstream capacity is therefore proposed for FTTH deployments in this decade based on average and peak traffic modeling. 10 Gbps downstream capacity may limit service group sizes to 16 or less.

<u>DS Capacity growth as a function of time</u>

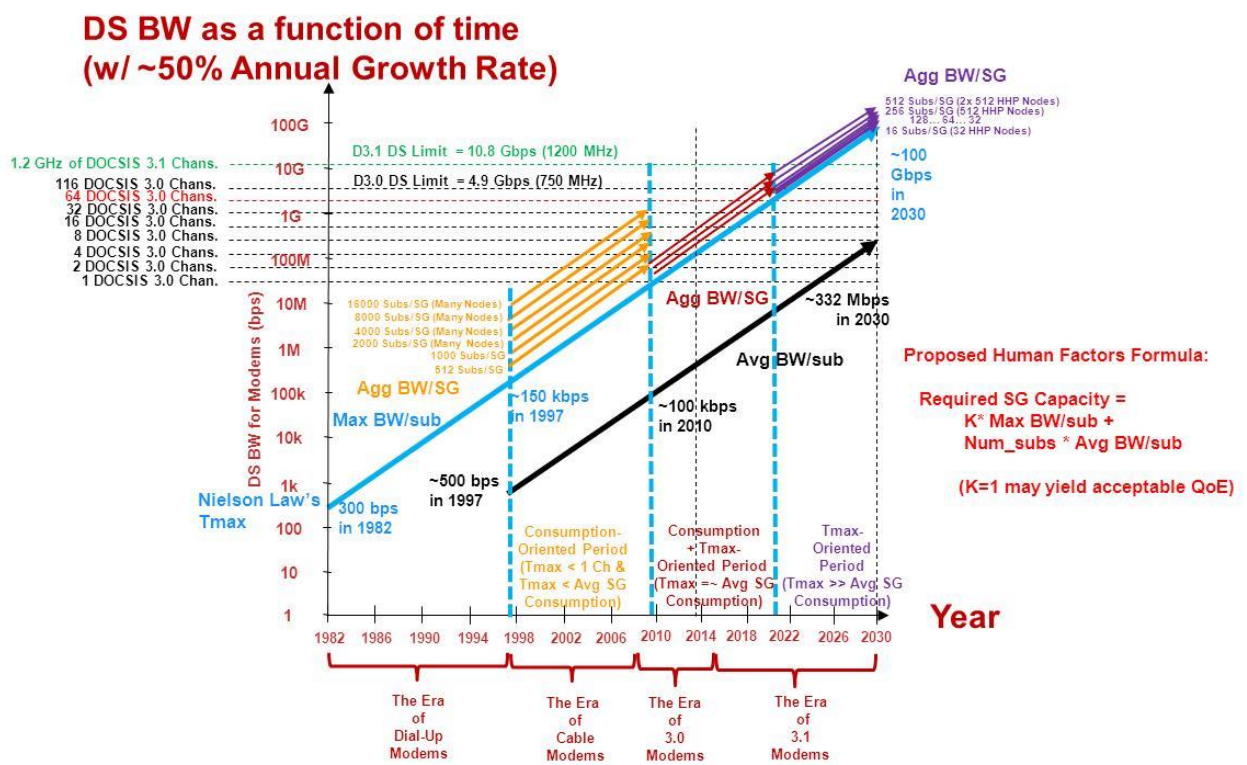
The figure below [1] shows the downstream traffic growth from 1982 to the present and projected to projected to 2030 for various service group sizes and differentiated between maximum bandwidth per user and average bandwidth per user and per service group (SG). Up to 2010 the maximum tier (Tmax) per subscriber is below the average service group consumption for service groups greater than 512 subs.

As service groups get smaller - after 2010 - this changes and the maximum tier (Tmax) starts to drive the required downstream capacity; for service group sizes of between 256 and 512 the average service group consumption and the capacity needed to support maximum tier are approximately the same.

In order to ensure acceptable quality of service an aggregate service group capacity consisting of a sum of weighted (with K=1) max tier (Max_BW_sub) and average bandwidth (Avg_BW_sub) times the number of subscribers (Num_subs) is expected to be sufficient:

Agg_BW_SG = K.Max_BW_sub + Num_subs.Avg_BW_sub

At the end of the decade the service groups are expected to shrink and the extrapolated maximum tier drives the required forward capacity up to around 100 Gbps.



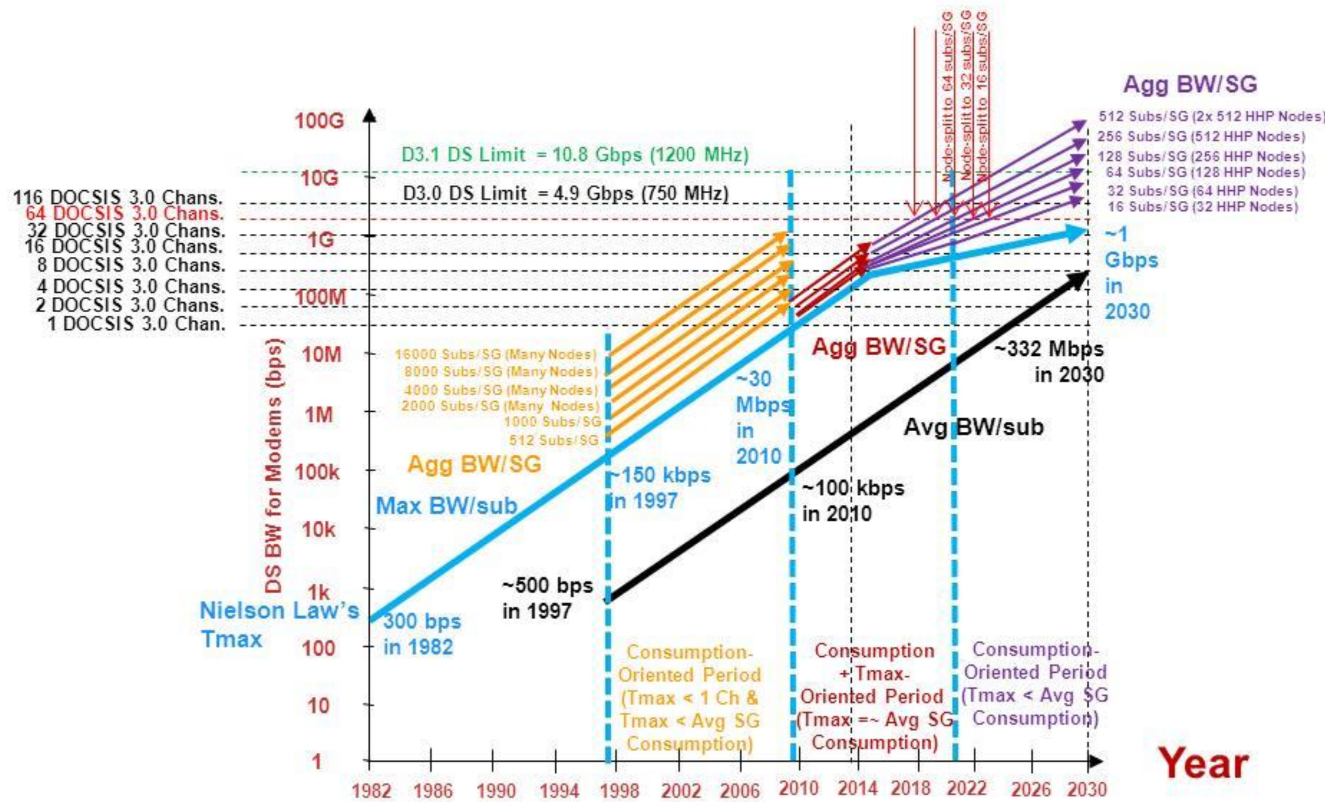
Illustrating the Capacity growth from 1982 to the present and projected to 2030

However, there is a possibility that the projected growth in Tmax does not come to pass (~ 100 Gbps as predicted by Nielsen's law) for maximum downstream service tier and the maximum service tier requirement is limited to a number as low as 1 Gbps to 3 Gbps. This results in a lower aggregate service group capacity that is still dominated by the average service group capacity needs.This is illustrated in the plot below; that shows aggregate downstream throughput requirement for a max tier limited to 1 Gbps and under the assumption that a pipe with enough capacity for the average service group throughput offers sufficient quality of service.

Now the average bandwidth per subscriber can dominate the total bandwidth required so that in practice some additional capacity, up to a factor M, may need to be reserved to avoid contention issues. It is therefore appropriate to rewrite the aggregate capacity requirement formula as:

$$Agg\_BW\_SG = K.Max\_BW\_sub + M.Num\_subs.Avg\_BW\_sub$$

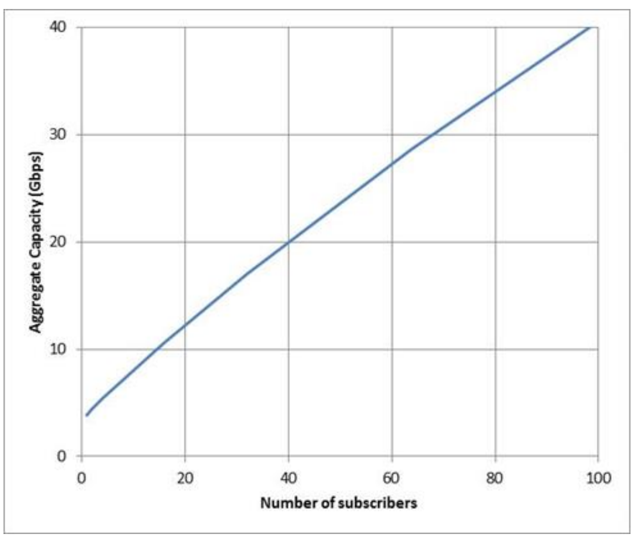
The value of M is generally around 2 for smaller service groups but becomes closer towards 1 as the service group size increases due to the statistical benefits of larger numbers of subscribers. The figure below shows average bandwidth per sub (red), max tier bandwidth and the required aggregate bandwidth (green) in a service group of 128 users as a function of time. The max tier bandwidth (blue) growth post 2013 was reduced to reach 3 Gbps in 2030 as a compromise between Nielsen's law (~100 Gbps) and the rather conservative estimate of 1 Gbps.



Illustrating the cap in Capacity growth from now to 2030

Illustrating capacity needs to avoid contention

The required DS capacity in 2030 can also be plotted as a function of SG size. The figure below shows the required aggregate serive group capacity as a function of the service group size with a factor 2 for the average service group capacity to prevent contention and without an additional factor to prevent contention. The service group size can be up to around 128 for 40 Gbps and is limited to less than 16 for 10 Gbps (for instance 10 GEPON) capacity.



Illustrating capacity per service group

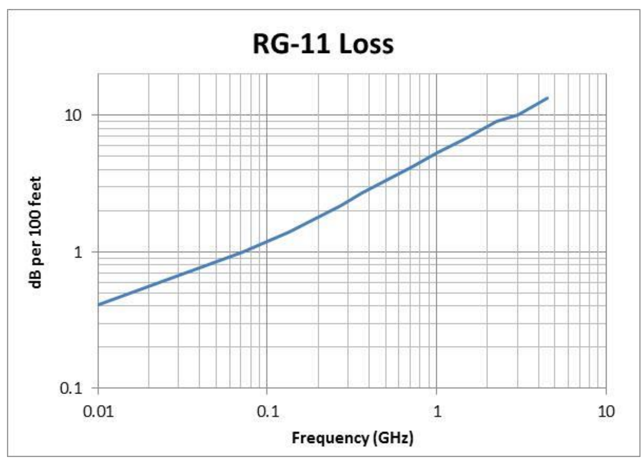
It is concluded that in order to retain a reasonable service group size at the end of the next decade the system should allow 40 Gbps downstream capacity. In the upstream at least 10% of that capacity should be available or 4 Gbps.

RFOG DOWNSTREAM

At 1550 nm RFoG is more suitable for SMF-28 fiber access networks than binary modulation. This surprising result is due to the twin effects of the superior capability of QAM schemes to take full advantage of available bandwidth and convert it to higher throughput relative to the binary formats and the effects of fiber dispersion on links where fundamental limits can cap NRZ binary modulation formats to around 20 Gbps throughput at 25 km and fewer at greater distances. This paper indicates that RFoG when used with the DOCSIS 3.1 standards can support greater than 40 Gbps up to 40 km.

Home wiring

In-home wiring exists to support satellite LNA receivers, bandwidth up to 4.5 GHz is available with RG11 cable. A typical loss curve is shown below [2].



Illustrating loss of a typical RG11 cable

The typical return loss is better than 20 dB over the frequency range such that this cable type is suitable for distributing RF signals up to 4.5 GHz through the home.

SNR of optical link

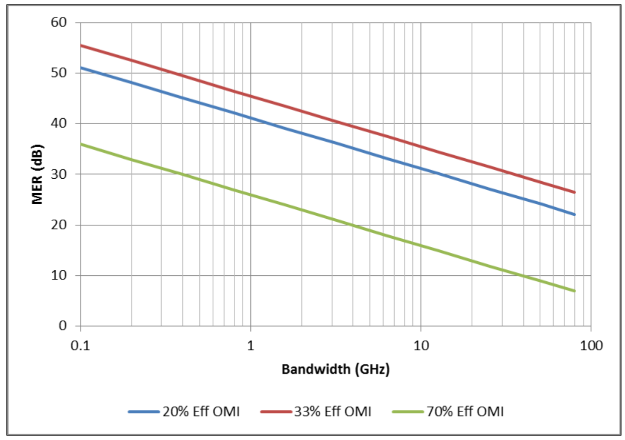
An optical link is characterized by a transmit power, a loss budget resulting in a receive power, receiver characteristics and modulation format and bandwidth characteristics. These will be investigated here for the example case of a 24 dB loss budget, the reference for PON and RFoG systems. The SNR (or MER) of a received signal is calculated as:

$$SNR = \frac{(\mu \cdot Ipd)^2}{B \cdot \left(Ipd^2 \cdot 10^{-0.1 \cdot RIN} + 2 \cdot q \cdot Ipd + In^2\right)}$$

Where Ipd is the photo detector current of the receiver, μ is the effective modulation index, B is the signal bandwidth, RIN is the relative intensity noise of the optical source (can be laser and optionally includes EDFA spontaneous emission beat noise, Rayleigh backscatter noise on the fiber and optical beat interference noise for multiple optical sources), q is the electron charge in the shot noise contribution and In is the receiver equivalent thermal input noise current. Depending on the receiver design, different values of In can be achieved, for higher bandwidth In typically increases:

The detector current follows from the transmit power P_tx, loss budget (24 dB) and detector sensitivity. In the forward direction the effective or composite modulation index is well controlled. For binary signals it is typically 70%, for multi-carrier RF signals it is around 20% and with PAPR (Peak Average Power Reduction) methods that can be improved to around 33%.

With a DFB laser followed by an 18 dBm EDFA the SNR can now be calculated for a wideband TIA (trans impedance amplifier) (5 GHz+ bandwidth, suitable for 10 Gbps link). In the absence of excess phase noise from the demodulator the MER is the same as the SNR



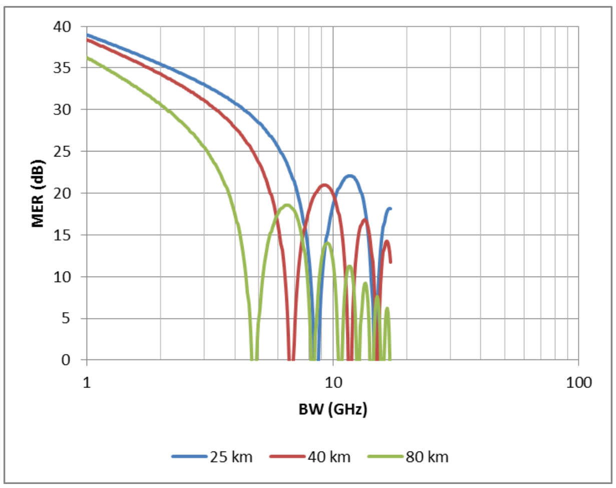
Illustrating MER for various effective OMI with increasing BW for a 24 dB loss budget

For a bandwidth of 1 GHz and a loss budget of 24 dB, the MER is just over 40 dB, as expected for an all-QAM loaded RFoG system with an effective modulation index of 20% (Blue line). With PAPR that could be further improved to around 45 dB or more (Red line). Also shown (Green line) is the expected SNR of a binary signal with merely 3 dBm launch power, for 7 GHz bandwidth (practical bandwidth needed for 10 Gbps) the SNR is just over 15 dB, just over the requirement for 10 Gbps binary communication. While the use of an EDFA to launch 18 dBm is a cost disadvantage for RFoG compared to a 10 Gbps PON this cost is shared over many users and results in more than quadrupling the channel throughput as will be shown later.

Fiber dispersion

In a practical fiber-optic link with dispersion the attainable MER is also limited by second-order distortion. An analog optical system system with a 40 km fiber link can reasonably be expected to deliver a composite second order distortion around 46 dB or better in an application with 1 GHz worth of QAM loading. The composite second order distortion (CSO) is normally understood in the context of AM-VSB signal content but also applies to QAM modulated signals where the CSO acts as a noise source that must be added to the noise of the optical link. CSO amplitude is proportional to fiber length and maximum frequency of operation. CTB amplitude is proportional to the square thereof. In a system with significant fiber dispersion there is also linewidth conversion noise that needs to be accounted for, this is proportional to the square of the product fiber length and frequency such that at higher frequency even for shorter fiber lengths there is a measurable impact.

Furthermore there is fading in an RF modulated fiber-optic system due to dispersion. The periodic fading as a function of frequency causes nulls in the MER as a function of frequency, the MER also drops as a function of bandwidth due to CSO, CTB and SNR and is illustrated below.



Illustrating Dispersion induced fading effect

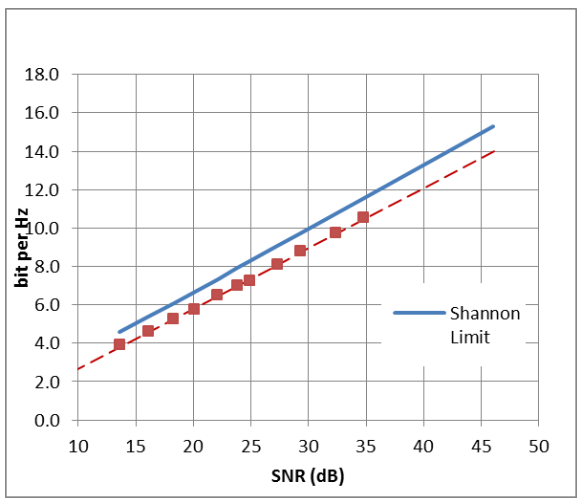
Thus the attainable distance drops quadratically with the signal bandwidth (or bitrate). Error correction can be applied to permit operation for relative pulse spreading up to around 80% of the bit period. The resulting data throughput with plain binary signaling attainable as a function of fiber length is approximately (120 km at 10 Gbps, 30 km at 20 Gbps, 15 km at 40 Gbps).

<u>Optical nonlinearities</u>

We have limited the discussion in its current form to single wavelength operation, where launch power is generally dependent upon the SBS capability of the transmitter. If multiwavelength operation is considered however, additional non-linearites such as 4WM, SRS and XPM xome ito play [3]. 4WM can be eliminated by a good wavelength plan, and SRS is predominantly at lower RF frequencies. The XPM however depends upon fiber dispersion and occurs at higher RF frequencies. In practical higher frequency systems, particular care must be paid to limit XPM by a combination of wavelength spacing and wavelength selection and power launch.

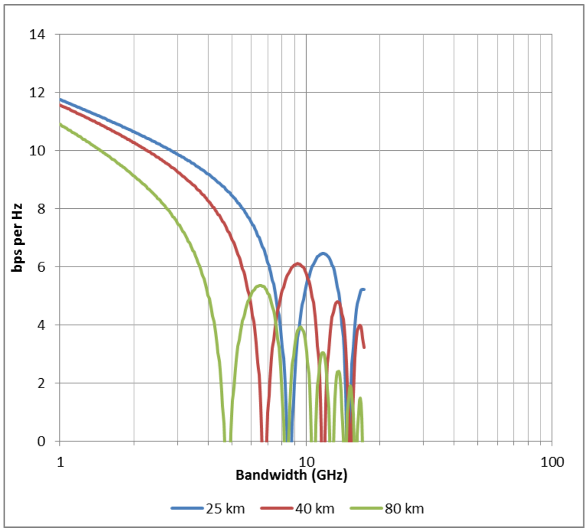
<u>MER performance of OFDM with LDPC</u>

Using OFDM modulation formats with LDPC a throughput in bps/Hz can be estimated from the figure below (from DVB-C2 specification). The data line is fitted through the data points and extrapolated to estimate available throughput in bits per second per Hz as a function of MER. The Shannon limit is shown for comparison [4].



OFDM/LDPC Capacity and the Shanon Limit

It is interesting to note that this result extends down to SNR values as low as 13 dB; almost 4 bps/Hz is still obtained as merely 13 dB of SNR (almost QAM-16 equivalent, realized with QAM64 and 1/3rd of FEC overhead reducing the 6 bps/Hz raw QAM64 throughput to 4 bps/Hz net throughput). Without LDPC QPSK or binary modulation is needed to allow operation at such a low SNR value (each offering only half the throughput; up to 2 bps/Hz at the Nyquist rate). The MER versus bandwidth plot can now be converted to attainable throughput in bit per second per Hz as a function of bandwidth.

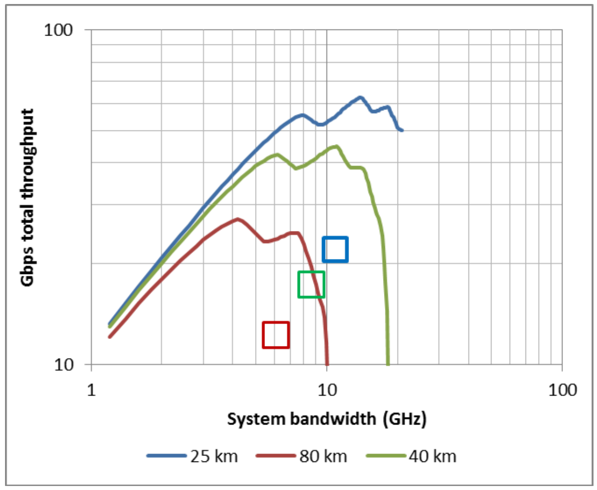


Illustrating Capacity/Hz in RFoG/OFDM/LDPC system

Total Throughput

By integrating bits per second per Hz for a given bandwidth over frequency f the throughput is obtained:

Using, as an example, the TIA front end (11 pA/sqrt(Hz)) but with unlimited bandwidth and the reference receiver input power P_rx (-6 dBm) and laser RIN and modulation index without PAPR (of 20%) the total throughput can be plotted as a function of input signal bandwidth for different fiber lengths (25, 40 and 80 km). The plots show the nulls in the transmission window as areas where bandwidth expansion of the source does not contribute to additional throughput. In practice a system would only be operated up to the first null. 40 Gbps throughput can thus be obtained with around 4 GHz of bandwidth.



Illustrating the Capacity of an RFoG link with OFDM/LDPC vs. Binary Modulation

For a binary system the throughput was simply given as:

$$\mathrm{bin\_bps}(L) := 10 \cdot \sqrt{\frac{120}{L}}$$

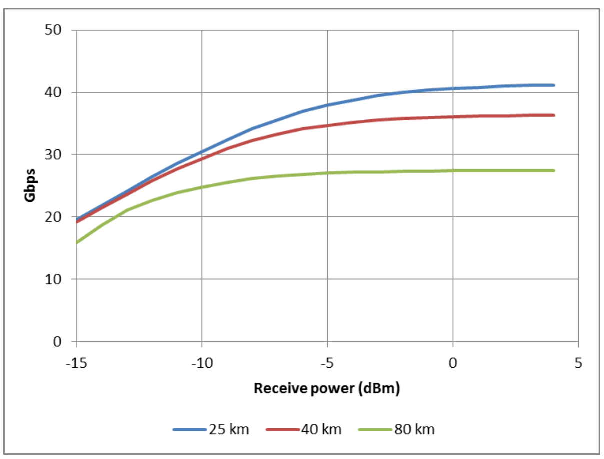
This is plotted as markers for the 3 fiber lengths at their respective Nyquist bandwidths. Whereas the binary modulation permits operation at higher bandwidth (compared to a dual sideband AM modulated signal around a carrier) it does not provide more throughput at any of the distances (note it is assumed that the SNR of the RFoG signal is maintained with EDFAs to arrive at the receiver power of – 6dBm and the EDFA(s) have sufficient input power to prevent significant contribution from the EDFAs to the noise.

At 1550 nm on SMF the binary modulation format provides just over 20 Gbps at 25 km length and 12 Gbps at 80 km length, at each fiber length the binary modulation format performance falls well below the RFoG performance curves for -6 dBm receiver power. The RFoG performance curve can be further enhanced at higher receive power.

It is thus concluded that the addition of an EDFA in the headend, as needed for RFoG, increases the available downstream bandwidth in the system from 10 Gbps (PON) to more than 40 Gbps (RFoG). As shown in the figure even more throughput can be available with wider bandwidth receivers. However this paper is limited to RFoG ONUs that can be realized with currently available low cost components and the discussion is therefore be limited to around 5 GHz receiver bandwidth. It is further concluded that due to dispersion a binary transmission format cannot reach the same throughput as an RFoG system with EDFAs, neither at 25 km reach nor at 40 km reach.

RFoG Receiver performance

Estimated capacity is shown below as a function of receiver power with 4 GHz receive bandwidth at 25, 40 and 80 km. At -6 dBm receive power almost 40 Gbps is achieved, the performance saturates above –3 dBm to 40 Gbps.



Illustrating System Throughput in Gbps as a function of receiver input power

These estimates are for a multi-carrier RF modulated signal (OFDM with LDPC) with a modulation index set to the same value as for current QAM256 operation. With peak-average power reduction (PAPR) the throughput can be increased by around 10% beyond these results.

Subsequent RF amplification to 4 GHz or higher is readily available with low-cost commercial MMICs. Thus the receive side of the ONU can be realized with a subset of the components used in a 10 Gbps PON ONU. The required components can satisfy reduced bandwidth and gain requirements. The transmit side of the ONU that will be proposed here can be based on a regular directly modulated DFB laser with 1 GHz of bandwidth. Given the availability of wideband lasers and amplifiers that is not significanty more complex than RFoG ONUs with less upstream bandwidth.

Discussion on modulation format and throughput in the access

In the access part of the plant the loss budgets are relatively small compared to long haul telecommunication. In RFoG the number of forward wavelengths is also small compared to long haul communication systems. By using optical amplification in the access plant, permitted for one or a few wavelengths on a fiber, the SNR can be increased so much that analog modulation formats, in spite of their shortcomings, readily obtain 4x the throughput of binary modulation formats given practical receiver bandwidths available today using low cost components. Binary modulation formats with optical amplification could permit larger service groups, however, the size of service groups needs to be limited unless much more than 10 Gbps forward capacity can be offered, that is not available today in low cost binary data processing receiver and electronics and its application is limited by fiber dispersion. Therefore the binary modulation formats do not benefit from optical amplification in the access. It can thus be concluded that, unlike in long-haul telecommunication, in the access plant analog modulation formats are superior to binary modulation formats. This should come as no surprise, in most short distance bandwidth limited systems analog modulation formats are used (DSL, CAT-6 10 GbE cables, WiFi and of course DOCSIS).

While it may sound contradictory to classify a fiber system as a bandwidth limited system it is fiber dispersion and the OE conversion (receiver) and associated processing electronics that do pose a practical bandwidth limitation in fiber systems. In fact this bandwidth limitation is acknowledged by 40G PON manufacturers resorting to 4 wavelengths each operating at 10 Gbps rather than attempting to run one wavelength at 40 Gbps. Reality is that aggregate throughput levels are expected to increase to a level that far exceeds the baseband bandwidth of low cost receiver electronics that are used with simple on-off keying on standard SMF fiber. Current low-cost analog technology on the other hand can already provide a transparent pipe with the required throughput capability today.

A WORD ABOUT OBI

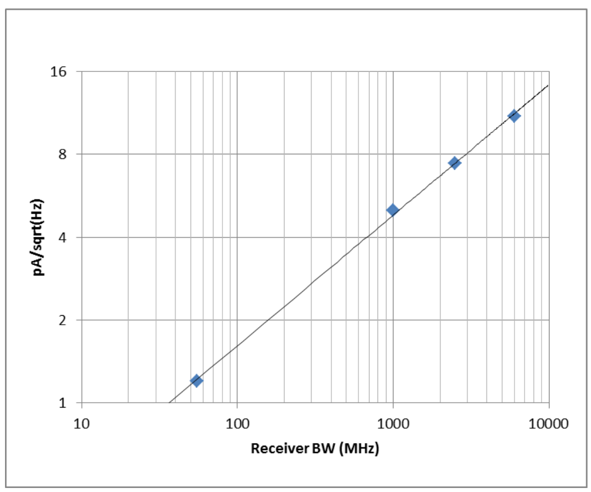
Optical Beat Interference (OBI) is a profound issue in RFoG reverse path and affects the system in debilitating ways. The subject of OBI mitigation is vast and is the subject of intense study even now. While mitigating OBI could be a goal, since it is a statistical and intermittent phenomena, it is not entirely easy to establish observable performance robustness unless these tests are done at a high enough utility contention with representative systems along with the CMTS. Even then, it becomes hard to predict the performance if the SNR/MER requirement change, as would be the case of a move to D3.1 from D3.0. There are however standard ways of eliminating OBI altogether, and these might be used to get predictable and robust performance. In this paper, as OBI is not the subject under discussion, we assume that the issue of OBI is settled and the OBI occurrence is eliminated.

RFoG UPSTREAM

In the upstream SNR is evaluated accounting for laser RIN, laser power and loss budget, the upstream bandwidth and the modulation index that can be allowed per chanel given a total bandwidth, an additional dynamic window to accommodate uncertainty in transmitter OMI setting and the optional use of PAPR (Peak to Average Power Reduction) in the transmission format.

The effective modulation index just under clipping induced BER is 20% for regular RF modulated systems and around 32% for RF modulated systems with PAPR (Peak Average Power Reduction) methods applied.

The return bandwidth can vary, up to 1 GHz should not significantly affect return laser and driver cost, this primarily affects the choice of diplex fiters. However the receiver thermal noise figure is a function of bandwidth as shown below for readily available low cost receivers.



Illustrating the typical receiver noise as a function of the bandwidth

The blue markers show single input receivers that are practically available, 1.2 pA/sqrt(Hz) for an RFoG 55 MHz return receiver, 5 pA/sqrt(Hz) for a CATV 1 GHz forward receiver, 8 pA/sqrt(Hz) for a 2.5 GHz TIA front end and 11 pA/sqrt(Hz) for a 6..8 GHz TIA front end. The fit approximates the results very well and this is used to conservatively model the thermal noise as a function of the reverse bandwidth.

In the return band multiple transmitters can be on at the same time. When these are combined before a receiver then the receiver thermal noise does not change when one or more transmitters are active. The shot noise and RIN noise for all the transmitters needs to be added appropriately.

In RFoG return systems multiple receivers are also combined into a single CMTS port, when that is done the receivers usually have "low squelch", the receiver outputs are squelched (turned off) unless there is an input signal. As a consequence the noise floor to the CMTS varies and the thermal noise contribution of multiple receivers also is a function of the number of transmitters active over a number of receivers. Here we will analyze the worst case condition with all receivers active at the same time although generally few receivers wil be active and the SNR is better.

An RFoG system is not limited by the fixed RF split frequency of the CATV plant. At the ONU the amount of spectrum allotted for the return band can be selected with the diplex filter and a number of different return bandwidths are evaluated

Upstream RFoG MER and throughput

The table below shows the resulting upstream RFoG performance estimates

Note	Plas dBm	Budget dB	BW MHz	DW dB	PAPR Y/N	# Tx on #	Rx ns add	MER dB	Data Gbs
Basic RFoG	3	24	37	12	N	1	N	29.6	0.33
Basic RFoG	3	24	37	12	N	4	N	24.7	0.27
Basic RFoG	3	24	37	12	N	4	Y	23.5	0.26
200 MHz	3	24	200	12	N	1	N	18.8	1.08
200 MHz	3	24	200	12	N	4	N	15.9	0.90
200 MHz	3	24	200	12	Y	4	Y	16.8	0.96
200 MHz	3	24	200	3	N	4	Y	21.8	1.27
200 MHz	3	24	200	3	Y	4	Y	25.8	1.53
200 MHz	10	24	200	3	N	4	Y	32.0	1.91
200 MHz	10	24	200	3	Y	4	Y	36.1	2.17
200 MHz	10	24	200	9	Y	4	Y	30.1	1.79
1 GHz	10	24	1000	3	N	1	N	28.0	8.30
1 GHz	10	24	1000	3	Y	1	N	32.1	9.58
1 GHz	10	24	1000	3	Y	4	N	28.9	8.58
1 GHz	10	24	1000	3	Y	4	Y	25.9	7.65

System throughput for high performance RFoG upstream

For instance a traditional RFoG return system uses a DFB laser, 3 dBm with a 24 dB loss budget resulting in -21 dBm at the headend receiver. Such a receiver has a low noise figure of around 1 pA/sqrt(Hz) and at the effective modulation index of 0.05, that is 12 dB (the Dynamic Window DW) less than the maximum effective modulation index of 20% an MER of 30 dB results. Best case 330 Mbit/sec could be supported with the best OFDM+LDPC scheme utilizing the full 5-42 MHz bandwidth (37 MHz). This assumes no that no PAPR implemented.

When 4 transmitters are on at the same time the MER drops, the situation is worst when these 4 transmitters are each on a different headend receiver ("Rx ns add" column is 'Y' represents receiver thermal noise addition), this could result in MER values under 24 dB. In practical systems the number of channels is limited to just 24 MHz of bandwidth and often the load is less such that this worst case condition does not occur often.

However when bandwidth is expanded to 200 MHz then the available modulation index is spread over a larger bandwidth such that the per channel modulation index drops by about 7 dB and also a typical 200 MHz receiver has a worse noise figure resulting in another 3 dB drop in performance. For a single active transmitter SNR is low but throughput up to 1 Gbps is still possible. However when multiple transmitters are active low SNR could result. A number of measures can be taken to overcome this limitation. Firstly the level could be set up more accurately so that the dynamic window could be reduced to 3 dB, even with multiple transmitters on a reasonable SNR is recovered. Secondly PAPR methods could be applied to permit higher modulation index without clipping. It should also be considered to use higher power return lasers (or equivalently lower noise receivers), such as a 10 dBm devices resulting in more than 30 dB SNR. Combined with PAPR a throughput up to 2 Gbps can be possible. Also shown is a relaxation of the setup accuracy in a 200 MHz return system where acceptable performance is still available.

In case the return traffic grows faster than expected and 4-10 Gbps of return traffic is needed then the return bandwidth could be further enhanced to 1 GHz. Without any of the aforementioned measures the SNR would be too low for operation. However just changing laser power (or using improved receivers) provides enough SNR to enable data rates up to 8 Gbps. Combination with PAPR provides almost 10 Gbps and even with multiple transmitters active a high throughput up to 8 Gbps is still available.

CONCLUSIONS

In this paper, we have shown a path to a much more efficient use of the RF spectrum which when used with RFoG technology and features available via the DOCSIS 3.1 standard. Innovative low cost ONU designs using current technology were presented that can address the anticipated growth rates and bandwidth expansions seamlessly over the next decade and more. Thus a well thoughtout partnership of DOCSIS 3.1 and RFoG can satisfy the anticipated growth in traffic demand while providing MSOs with plant and equipment investment protection well into the future.

REFERENCES

1. Neilsen's Law vs. Nielsen TV Viewership for Network Capacity Planning, Mike Emmendorfer, NCTA 2014
2. Clark CD7511P Plenum 4.5GHz HD/SDI RG11 Coaxial Cable, http://www.clarkwire.com/downloads/CD7511P_Clark_PB24202.pdf
3. When Wavelength Collide Chaos Ensues: Engineering Stable and Robust Full Spectrum Multi-Wavelength Networks, Venk Mutalik, Fernando Villaruel, Marcel Schemmann, Kwang-Yi Wu, SCTE 2011
4. http://www.dvb.org/resources/public/factsheets/DVB-C2_Factsheet.pdf
5. Value of Transparency in the Evolution of the HFC Networks, Venk Mutalik, Marcel Schemmann, Brent Arnold, SCTE 2013

PREDICTIONS ON THE EVOLUTION OF ACCESS NETWORKS TO THE YEAR 2030 & BEYOND

Tom Cloonan (CTO-Network Solutions), Mike Emmendorfer (Sr. Director),
John Ulm (Engineering Fellow), Ayham Al-Banna (Distinguished System Engineer),
Santhana Chari (Sr. Director)

ARRIS

Abstract

As the Cable Industry moves forward into the future, MSOs will experience profound changes in their HFC plants and in their head-end equipment. These changes will be required to support the challenging bandwidth growth that is expected within all of the MSO service types, including DOCSIS HSD, DOCSIS IP Video, SDV, VoD, nDVR, and Digital Broadcast Video (SD, HD, 4K, & 8K resolutions).

These expected changes are forcing many MSOs to wonder about the ultimate evolution of their head-ends and HFC plants. Many questions abound. Will head-end-based CCAPs be able to support the future bandwidth per RF port? Will head-end-based CCAPs be able to support the number of RF ports needed after many rounds of nodes-splits? Will the power and rack-space requirements of future CCAPs fit within head-end budgets? Will MSOs need to use the bandwidth expansion capabilities offered by DOCSIS 3.1? Will MSOs need to use the higher-order modulations offered by DOCSIS 3.1? Will Digital Optics be required to enable the higher-order modulation in DOCSIS 3.1? Will Distributed Access Architectures be required to provide the required scale? If so, which Distributed Access Architectures will likely be utilized? Will it be Remote PHY? Will it be Remote CCAP? Will RFoG offer needed benefits in the future? Will EPON and EPOC offer needed benefits in the future? When do the traffic engineering analyses predict that transitions will take place between these different technologies?

This paper will look into the crystal ball and use basic trend analyses and traffic engineering analyses to create predictions that attempt to answer all of these challenging questions.

BACKGROUND

Traffic Capacity Trends For The Future

By its very nature, traffic engineering is an imperfect science requiring many guesses and assumptions and approximations to be utilized. Sometimes, these guesses can be wrong. However, when implemented properly, traffic engineering predictions can be very useful in developing rough plans for future HFC networks and in the sizing of future HFC networking equipment. It may be valuable to explore some of the current predictions coming from recent traffic engineering work.

Using empirical data collected in the field, trending information on all service types can be analyzed in an attempt to predict the future. However, after doing this for most services, the authors believe that the most problematic service type within the MSO network of the future will undoubtedly be the DOCSIS HSD and DOCSIS IP Video service types. We will therefore focus on this particular service type within this section.

For the past several years, studies have indicated that DOCSIS HSD Downstream traffic has been experiencing a ~50% compound annual growth rate (CAGR). For almost 30 years, this growth rate has shown itself in the Maximum Downstream Sustained Traffic rates (aka the "Billboard Bandwidths") that service providers have offered to their subscribers, and it has also shown itself (with slightly more variation) in the Average Downstream Bandwidth Consumption rates that subscribers have consumed for at least the past 10 years. Upstream Billboard Bandwidths and Average Upstream Bandwidth Consumption Rates display much more variation, with CAGRs at different MSOs showing rates ranging from 10% to 30%.

There are some who believe that the Average Bandwidth Consumption rate of the future may slow as homes begin to receive adequate bandwidth to offer an HD video feed to every pair of eyeballs within the home. There has even been recent slow-downs in the Average Downstream Bandwidth Consumption rates witnessed by some MSOs that seem to validate this belief. Thus, it is very possible that the CAGRs may be reduced below 50% in the future. However, this potential slow-down in Average Bandwidth Consumption growth rate will likely be accelerated again as 4K (3840x2160 resolution) and 8K (7680x4320 resolution) UHD TV begins to gain acceptance in the future and be passed over DOCSIS networks. At a minimum, it is expected that 4K video feeds will likely be viewed on the main television screen in many homes within the next five years. These new video resolutions could lead to increases of up to 16x the required bandwidth of today's 1080p video resolution. In addition, new, yet-to-be-invented machine-to-machine applications of the future may also contribute to increased DOCSIS HSD Bandwidth Consumption demands.

Bandwidth Pressures For The Future

The MSO's competitive landscape has changed rapidly in the last five years-especially from Over The Top (OTT) video providers such as Apple TV, Amazon, Hulu, Netflix and others entering the On-demand video market. The resulting growth in OTT video usage has also led to a requirement for increased investment within the MSO's DOCSIS HSD network due to increased consumer bandwidth usage. Bandwidth growth from these OTT providers switching to 4K or 8K video will likely drive more bandwidth demands in the future.

Verizon FiOS and other FTTH providers that have rolled out networks over the past several years will also remain a threat to the MSO's triple play offering in the future. As an example, it was reported that Verizon will consider an upgrade to their FiOS network to the next generation PON technology that may be capable of offering 10 Gbps downstream and 2.5 Gbps upstream traffic. This would increase the bandwidths available in competing service providers, placing a new stress on the bandwidth demands within DOCSIS networks.

All of these trends will undoubtedly raise the height of the competitive bar that MSOs will have to try to match or beat over time. Most MSOs plan to counter these threats with smooth and sustained bandwidth capacity growth in their HFC network that matches the needs of their subscribers in a just-in-time fashion.

ANALYSIS OF HFC PLANT EVOLUTION

Bandwidth Extrapolations & Node Split Efficacy In Future HFC Plants

The existing HFC cable network infrastructure provides high levels of digital capacity (limited to ~6 Gbps to the home and perhaps ~100 Mbps from the home in pre-DOCSIS 3.1 systems, and limited to 10-15 Gbps to the home and perhaps 1+ Gbps from the home in post-DOCSIS 3.1 systems). The HFC network also offers great flexibility.

If an MSO uses this capacity carefully, it can likely compete well for many years to come. The cable industry is making investments in IP-based video delivery technology and expanding the high-speed Internet IP capacity as well.

The coaxial network is very nimble, and changes to modulation types, Forward Error Correction techniques, node sizes, spectral splits, and upper frequency limits may increase the spectrum allocation beyond the current levels in both directions. The capacity needed in each direction is projected into the future in this section. We will consider the Downstream and Upstream bandwidth demand trends, and then we will couple those trends with node-split changes that MSOs may opt to utilize. We will also consider the impact of new modulation types and Forward Error Correction techniques that are being provided by the DOCSIS 3.1 specification.

Node splits have long been a trusted tool used by many MSOs throughout the years to reduce the bandwidth demands within a Service Group. The basic idea behind the node split is that it divides the subscribers connected to a single Fiber Node into two groups (Group A and Group B), and the subscribers in Group A are re-connected to a single (smaller) Fiber Node and the subscribers in Group B are also re-connected to a different (smaller) Fiber Node. Thus, two separate Fiber Nodes (and the associated feeds for two separate Fiber Nodes) are required to support the bandwidth for the pool of subscribers, so there is a cost associated with the node split.

Node splits offer no change in the Service Group bandwidth requirements for Broadcast services—if the MSO needed 50 QAMs to support Broadcast Video prior to the node split, then the MSO will still require 50 QAMs to support Broadcast Video after the node split.

However, the principle benefit of the node split is associated with Narrowcast services (Switched Digital Video (SDV), VoD, DOCSIS HSD, DOCSIS VoIP, and DOCSIS IPTV). The benefit of the node split is primarily derived from the fact that, in the past, the Narrowcast bandwidth required on the HFC spectrum was roughly halved every time a node split was implemented. Node splitting oftentimes permitted MSOs to “free up” substantial amounts of Narrowcast spectrum whenever they performed the node split. For example, an MSO who had ~30 QAMs of Narrowcast could free up ~15 QAM channels (~600 Mbps) with a single node split.

Many MSOs envision the number of Narrowcast QAMs growing over time because they expect increased QAM counts associated with VoD (as VoD gains popularity and includes more HD video content), SDV (as SDV adds more HD content), and HSD (as users’ Internet bandwidth demands continue to rise). Of these different Narrowcast services, the highest growth rate going forward will undoubtedly be associated with DOCSIS HSD (which may continue to experience a 50% CAGR in its Average Downstream Bandwidth levels as defined in the previous section). Without the use of node splits, an MSO who had 8 DOCSIS Downstream channels consuming some of the HFC spectrum in 2013 would likely experience DOCSIS channel growth (assuming a 50% CAGR) as shown below.

‘13	‘14	‘15	‘16	‘17	‘18	‘19	‘20	‘21	‘22
8	12	18	27	41	61	92	137	205	308

Table 1. Year & Number Of QAM Channels Required For HSD Growth (No Node-Splits)

Obviously, this type of uncontrolled bandwidth growth would fully consume any reasonably-sized HFC spectrum within ten years. As a result of this anticipated problem, many MSOs plan to split nodes to control this Narrowcast QAM growth. The theory is that carefully-timed node splits can halve the QAM count associated with the DOCSIS HSD channels and can keep the total DOCSIS HSD QAM counts to reasonable levels. For example, if an MSO with a 750 MHz (115 QAM) plant wanted to keep the DOCSIS QAM count to levels below 80 channels (assuming the other 35 QAMs were reserved for Digital Video services), then the commonly-held belief is that, for the system shown above, node splits timed to occur in 2019 and 2021 would yield the desirable results shown below.

'13	'14	'15	'16	'17	'18	'19	'20	'21	'22
8	12	18	27	41	61	46 (split)	69	52 (split)	78

Table 2. Year & Number Of QAM Channels Required For HSD Growth (With Node-Splits in 2019 & 2021)

There are, however, two potential flaws in the above logic that may cause MSOs to experience undesirable issues with these planned future node splits. The first issue is that node splitting becomes more and more expensive each time a round of node splits is implemented—primarily because the number of Fiber Nodes that need to be split increases by a factor of two with each round of node splits. This may preclude MSOs from performing the node splits at the rates that they desire.

The second issue is that node splits do not help at all to reduce the QAM counts if the QAM counts are being driven primarily by the Maximum Sustained Traffic Rates (aka the Billboard Bandwidths) that the subscribers are being offered by MSO Marketing teams. These Billboard Bandwidths have also been growing at a 50% CAGR for quite a few years, and if that growth rate continues in the future, it will require that many DOCSIS HSD QAMs will have to be offered on the HFC spectrum within each Service Group *even if* a large number of node splits are implemented. In fact, even in the hypothetical case where numerous node splits are used to reduce the number of subscribers within a Service Group to only one subscriber, a high Billboard Bandwidth of, say, 4 Gbps would require 100 DOCSIS QAMs to be fed into the Service Group- even if the average bandwidth consumed by that single subscriber was only 100 Mbps. This is an interesting phenomenon that requires more study as MSOs begin to plan their node split strategies for the future.

In traffic engineering simulations carried out by the authors, preliminary results indicate that the amount of DOCSIS HSD capacity which might be required to provide an "adequate" Quality of Experience level to a Service Group can be roughly described by simple formulae. One form of the formula is given by:

Required Service Group Bandwidth Capacity = S*Tavg + Tmax (1)

where S is the number of subscribers within the Service Group, Tavg is Per-Subscriber Average Busy-Hour Bandwidth, and Tmax is the Maximum Sustained Traffic Rate (Billboard Bandwidth) offering to the subscribers.

It should be noted that other more complex (and more accurate) formulae are also under study, but this simplified formula will be utilized extensively within this paper to make some key points. To make use of this formula, we will need to make some crystal ball predictions about the future values of S, Tavg, and Tmax. This is challenging, but we will turn to historical data to help make these predictions.

One can look back at the DOCSIS HSD Billboard Bandwidth (the Tmax values) and the DOCSIS HSD Average Downstream

Bandwidth Consumption rates for a single subscriber (the average Tavg values) on a yearly basis. One can then assume that the trends of the past will continue into the future and then extrapolate the resulting curves into the future. This assumption may or may not be true, but we will use it as a baseline, and then we will also explore what might happen if the assumption does not hold true.

One can also predict the per-Service Group Aggregate Average Consumption Bandwidth (Tagg) levels for Service Groups of various sizes, which is simply the per-subscriber Average Downstream Bandwidth Consumption (Tavg) times the number of subscribers (S) within the Service Group. Obviously, Service Groups with different sizes (S) will have different Aggregate Average Consumption Bandwidths (Tagg). In general, Tagg is given by S*Tavg (which is the first term in Formula 1 above).

These predicted, future values are shown in Figure 1 for a traffic model that assumes the 50% CAGR of the past continues for both Tavg and Tmax values moving forward into the future.

There is quite a bit of useful information provided by the plots in Figure 1. The plot has the Downstream Bandwidth displayed in a logarithmic fashion on the y-axis, and it has the years ranging from 1982 to 2030 on the x-axis. These years cover several different eras of modem service, including the Dial-Up Modem era, the Cable Modem era, the 3.0 Modem era, and the 3.1 Modem era.
The first plot to explore is the light blue plot, which shows Nielsen's Curve that identifies the expected Billboard Bandwidth (Tmax) values on a year-by-year basis. This plot illustrates that MSOs will likely have to provide higher and higher Tmax values to their subscriber pool on a yearly basis. If these trends continue, then Nielsen's Curve predicts that the Tmax value for a high-end modem may be on the order of 100 Gbps by 2030. One may wonder what applications could possibly require that kind of bandwidth 16 years from now. The honest answer is that we do not know what those applications will be—nobody does. But it is probably fair to say that the college students of today will help to invent those impossible-to-identify applications. It is also fair to say that we are

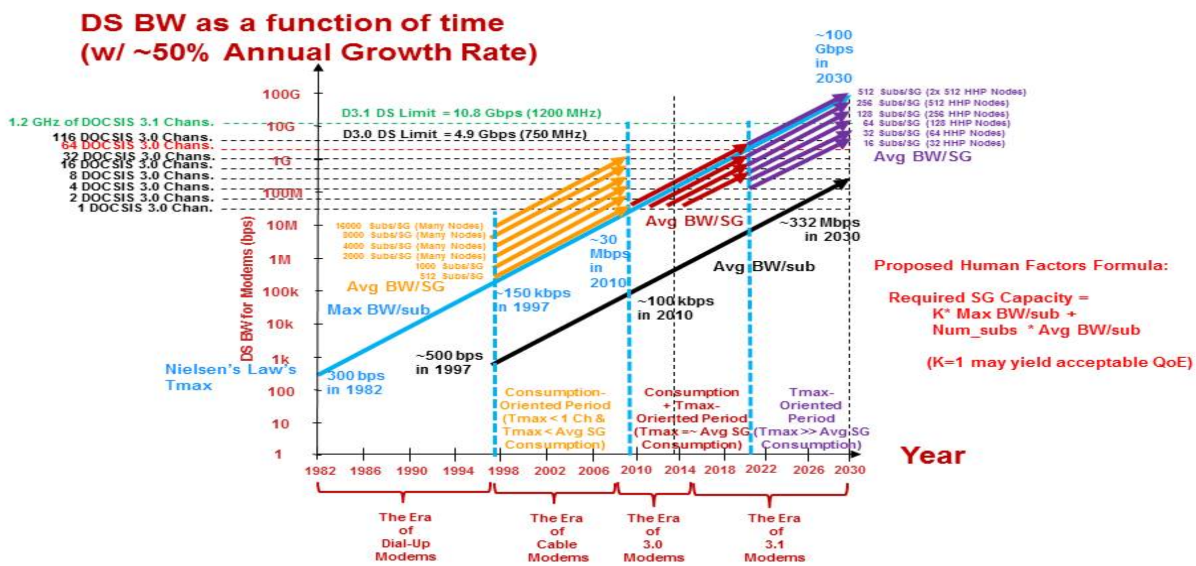


Fig. 1 DOCSIS HSD Downstream Traffic Engineering Predictions with 50% CAGR In Future

probably in the same position as an Internet user of 1998. Back then, they probably had no idea about the types of services and applications that would ultimately be developed that would require high-end users to require 100 Mbps of service by the year 2014 (16 years after 1998). The 100 Mbps service level of today must have appeared astronomical to a 1998 user who was only receiving ~200 kbps of bandwidth capacity at the time. But if trends continue, the Tmax values are predicted to grow quite high within the 2020 decade.

The second set of plots (in Figure 1) that should be explored is the plot set illustrating Average Bandwidth Consumption rates as a function of time. The black plot illustrates the approximate average per-subscriber bandwidth consumed by a single subscriber during the busy-hour period of time (8pm-9pm). This plot calculates an average value using contributions from both active and inactive users, so the average values (Tavg) are much lower than the Tmax values. In fact, the Tavg values for a single subscriber tend to be ~300 times lower than the corresponding Tmax values within a given year! However, the Tavg values in this chart are assumed to grow with a 50% CAGR, so the plotted values start in 1997 with a ~500 bps value and then grow to be ~100 kbps by 2010 and ~500 kbps by 2014 (a typical high-end number for today). The plot shows that if trends continue, then the average bandwidth per subscriber will grow to be ~332 Mbps by 2030.

While a plot for a single subscriber is interesting, it is not very useful for traffic engineering analyses looking at aggregate traffic patterns for many subscribers in a shared pool (like a Service Group). The single subscriber plot (shown in black within Figure 1) can, however, be scaled upwards in a linear fashion to yield the average aggregate bandwidth (Tagg) for a larger pool of subscribers that can share the DOCSIS HSD channels within a Service Group (SG). The orange, maroon, and purple plots contain that scaled up information for Service Groups of varying size that were utilized in past years or that will likely be utilized in future years.

The Service Group sizes shown in the figure include those with 16,000 subscribers per Service Group (in orange), which were utilized in the 2000 time-frame and could have corresponded to a Service Group containing ~24 Fiber Nodes of 2000 HHP each with a 33% take-rate on the DOCSIS HSD service. The Service Group sizes within Figure 1 also include those with as few as 16 subscribers per Service Group (in purple), which may be found in the 2020 decade if MSOs perform heavy node splits and create many 32-HHP Fiber Nodes with a 50% take-rate on the DOCSIS HSD service. It should be noted that two adjacent, parallel lines of color (orange, maroon, or purple) are exactly one node split apart from one another in terms of the bandwidth capacity required to support the Narrowcast services. (Note: The use of 32-HHP Fiber Nodes may be implemented by some MSOs in preparation for an ultimate transition to FTTH architectures in the 2030 time-frame. This transition to FTTH might make use of technologies such as RFOG, PON, or Point-to-Point Ethernet).

The family of curves illustrating the growing Service Group Average Bandwidth levels is divided into three different operating regimes, which in turn define three different periods of time in the life of the HFC plant. Within this paper, the three periods are labeled:

- The Consumption-oriented Period from 1997-2009 (illustrated in orange)
- The (Consumption+Tmax)-oriented Period from 2009-2021 (illustrated in maroon)
- The Tmax-oriented Period from 2021-2030 (illustrated in purple)

The Consumption-oriented Period occurred in the early days of DOCSIS deployment. It was characterized by a period of time when Service Groups were quite large (containing many Fiber Nodes) and each Service Group was serviced by a single DOCSIS Downstream channel. This approach worked quite well in its time, because the Tavg value of each individual subscriber was quite low, and it required the pooling of a lot of subscribers into a large Service Group to generate enough bandwidth to efficiently utilize the bandwidth capacity offered by a single DOCSIS Downstream channel (30-40 Mbps). The Tmax values of the time were also extremely low; they were typically much lower than the Service Group's Aggregate Average Bandwidth Consumption (Tagg) values and much lower than the ~30-40 Mbps bandwidth capacity offered by a single DOCSIS Downstream channel. So Tmax values were practially negligible in the traffic engineering efforts of those days. As a result, the total bandwidth capacity required by a Service Group was dominated by Tagg value (which was equal to the product of the Tavg value and the S value representing the modem count). During the Consumption-oriented Period, the Tmax values were much lower than the Service Group's Aggregate Average Bandwidth Consumption (Tagg) level. Thus, the Tmax values could generally be ignored during the Consumption-oriented Period, and Equation (1) could be reduced to an even simpler formula

Required Service Group Bandwidth Capacity = S*Tavg (2)

As a result, the main task of the traffic engineer in the days of the Consumption-oriented Period was to identify the Service Group's Aggregate Average Bandwidth (Tagg) for the pool of subscribers sharing the single DOCSIS channel and schedule Service Group "de-combining" activities to remove Fiber Nodes from Service Groups whenever the Service Group's Aggregate Average Bandwidth Consumption (Tagg) values began to rise to levels that were close to the ~30-40 Mbps capacities offered by the single DOCSIS Downstream channel. (Note: Oftentimes, this "de-combining" was triggered if the Service Groups Aggregate Average Bandwidth Consumption in the busy-hour grew to be ~70% of the capacity of the single DOCSIS Downstream channel).

Many MSOs transitioned from the Consumption-oriented Period to the (Consumption+Tmax)-oriented Period in the 2009 time-frame. This transition was typically implemented in conjunction with the MSO transition to DOCSIS 3.0 channel-bonding CMTS equipment. The (Consumption+ Tmax)-oriented Mode of operation was characterized by the fact that a single DOCSIS Downstream channel was no longer adequate to support the rising Tmax values required by the subscribers within the Service Group. As a result, MSOs needed to utilize bonded Downstream channels to provide adequate bandwidth capacity to each Service Group. The (Consumption+Tmax)-oriented Period was also characterized by another interesting change—the calculation of the total required bandwidth capacity for a single Service Group could no longer ignore the impact of the Tmax value, because the Tmax values and the Service Group's Aggregate Average Bandwidth Consumption (Tagg) levels tended to be of the same order of magnitude. This seems to be true for Service Group sizes of 1024-HHP per Service Group (512 subs per Service Group), 512-HHP per Service Group (256 subs per Service Group), and 256-HHP per Service Group (128 subs per Service Group). As a result, the calculations for Equation (1) required the traffic engineer to consider both terms within the formula. Since the Tmax value and the Service Group's Aggregate Average Bandwidth Consumption level (Tagg = S*Tavg) tended to be about the same in magnitude, many MSOs modified this formula to produce a much simpler approximation formula given by:

Required Service Group Bandwidth Capacity =2*Tmax (3).

The simplified and approximate formula in Equation (3) is roughly correct as long as the Tmax value is approximately equal to the Tagg (= S*Tavg) value, which is interestingly true for many MSOs right now. By definition, this constraint is roughly satisfied during the (Consumption+Tmax)-oriented Period. The (Consumption+Tmax)-oriented Period will likely exist for many years to come. During these years, MSOs will undoubtedly add more bandwidth capacity (channels) to the DOCSIS service tier to stay ahead of the Tmax growth and the Tagg (=S*Tavg) growth, and they may also perform node splits to periodically reduce the Service Group's Aggregate Average Bandwidth Consumption Tagg (=S*Tavg) level.

However, as a result of repeated node splits that reduce the Aggregate Average Bandwidth Consumption (Tagg) levels in a Service Group and as a result of expected increases in Tmax levels, there will come a point in the future when MSOs will likely transition from the (Consumption+Tmax)-oriented Period to the Tmax-oriented Period. At that point in time, the Tmax value will have grown to be much larger than the Service Group's Aggregate Average Bandwidth Consumption value (Tagg = S*Tavg). This is the period of time when the required Tmax values will begin to dominate the traffic engineering rules, and it is also the period of time when the Service Group's Aggregate Average Bandwidth Consumption level (Tagg = S*Tavg) may become practically negligible relative to the Tmax levels. This implies that the Tmax term in Equation (1) will be much larger than the (S*Tavg) term. Interestingly, node splits only impact the (S*Tavg) term of Equation (1) by reducing the number of subscribers (S) in the Service Group. Node splits do not impact the Tmax term at all. Thus, during the Tmax-oriented Period of the future, node splits will cease to have much impact on the required bandwidth capacity levels for a Service Group. This will have profound effects on the traffic engineering decisions that will be made by different MSOs in the future, with each MSO choosing a potentially different path.

To see why this is the case, it is best to consider Figure 2, which illustrates the Required Downstream Bandwidth Capacity required for different Service Groups of different sizes. (Note: Similar plots with similar dates can also be created for the Required Upstream Bandwidth Capacity, but due to space constraints, we will focus on the Downstream within this paper). This required bandwidth capacity value is calculated using the rule-of-thumb formula given in Equation (1) above. In essence, Figure 2 is the same as Figure 1, but Figure 2 shows the sum of the Tmax value and the Aggregate Average Bandwidth Consumption (Tagg = S*Tavg) value instead of showing each parameter individually. In Figure 2, the same three Periods of time are shown with the same Service Group sizes, but the bandwidth shown is the Required Downstream Bandwidth Capacity (which is equal to the sum of the Tmax and the Aggregate Average Bandwidth Consumption values (Tagg = S*Tavg)). When we compare Figure 2 to Figure 1, it becomes clear that the Service Group's Aggregate Average Bandwidth Consumption (Tagg = S*Tavg) values dominated during the Consumption-oriented Period. Within the figure, it also becomes clear that we are currently in an interesting period of time (defined as the (Consumption+Tmax)-oriented Period and shown in maroon) when the Required Downstream Bandwidth Capacity is driven by both the Tmax value and by the Service Group's Aggregate Average Bandwidth Consumption (Tagg = S*Tavg) value. Finally, it can also be seen that the future Tmax-oriented Period (in purple) may have Required Downstream Bandwidth Capacity values that will be dominated by the Tmax values, and the resulting required Bandwidth Capacity will be very close to the growing Tmax line (shown in light blue).

However, when MSOs move into the future and begin to operate in the Tmax-oriented Period, node splits will not be as effective as they once were. Successive node splits will yield less and less benefit as they produce smaller and smaller percentage

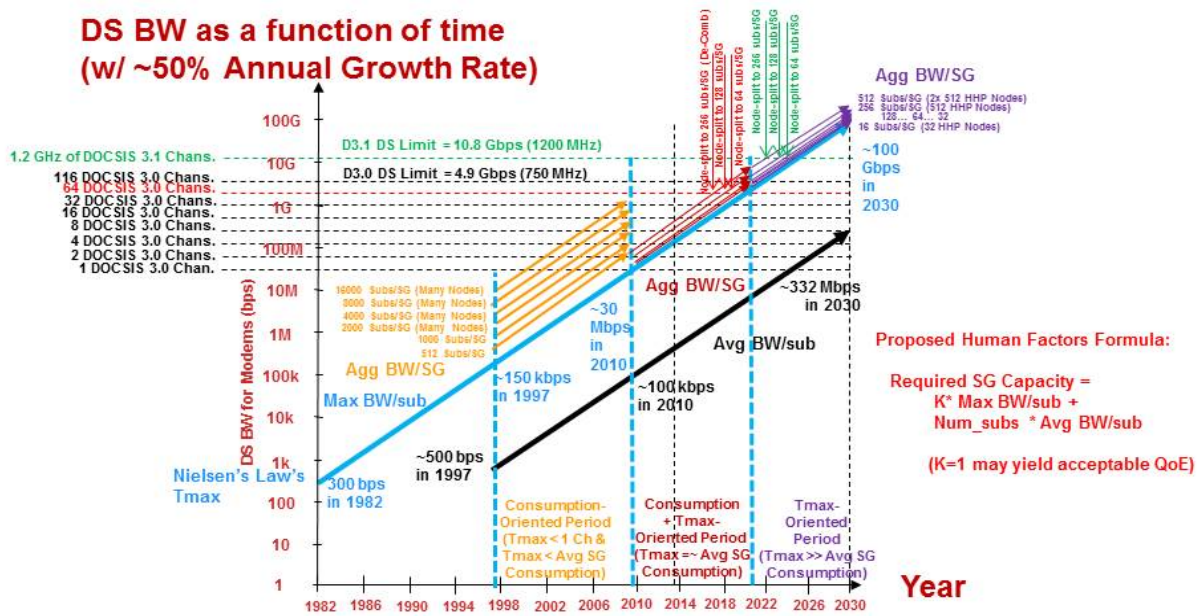


Fig. 2 DOCSIS HSD Required Downstream Bandwidth Capacity with 50% CAGR In Future

Each successive node split in the future will force an MSO's Required Downstream Bandwidth Capacity levels to drop from one curve to the curve directly below it. From the figure, it becomes clear that node splits had a very favorable effect on the Required Downstream Bandwidth Capacity during the Consumption-oriented Period and during the (Consumption+Tmax)-oriented Period, because the node splits tended to produce a ~50% reduction in the Consumption portion of the Required Downstream Bandwidth Capacity within the Service Group.

reductions in the Required Downstream Bandwidth Capacity levels. This is primarily due to the growing dominance of Tmax values in the Required Downstream Bandwidth Capacity levels as Fiber Nodes are made smaller and smaller. Since the node splits only impact the (S*Tavg) value, they will likely not have much of an impact in the distant future assuming bandwidth growth trends and node split trends continue. For the assumptions used in our calculations above, one can easily calculate the percentage reductions for each node split (a single hop between adjacent curves), and these results are shown below in Table 3 below.

Node Split		% Decrease in Req'd DOCSIS Downstream BW Capacity in the Service Group
From (subs/SG)	To (subs/SG)	
16000 to 8000		49%
8000 to 4000		48%
4000 to 2000		47%
2000 to1000		43%
1000 to 512		38%
512 to 256		32%
256 to 128		23%
128 to 64		15%
64 to 32		9%
32 to 16		5%

Table 3: Decrease in Required DS Bandwidth Capacity for Node Splits with Different Sized Fiber Nodes

Another way to look at the problem of diminishing returns on successive node splits is to study the amount of time that a node split "buys" for an MSO before another node split is required. To perform this study, let's assume that the MSO has decided that DOCSIS HSD will be given no more than 64 Annex B channels within the HFC spectrum because the rest of the spectrum is required for Analog, Digital Broadcast, SDV, VoD, and/or nDVR services. Within Figure 2, the bandwidth capacities associated with several different numbers of Annex B DOCSIS channels are shown as dashed horizontal lines, with the number of DOCSIS channels associated with each line designated on the left-hand side of the figure. The bandwidth capacity associated with 64 Annex B channels is shown by the red, dashed, horizontal line. If the red line represents the MSO's ceiling that sets a limit on the number of DOCSIS channels, then every time the Required DOCSIS Downstream Bandwidth Capacity for a Service Group hits that ceiling, then some action (such as a node split) must be taken. Once the node split action is taken, the Required DOCSIS Downstream Bandwidth Capacity is reduced and "buys" the MSO some amount of time before the Required DOCSIS Downstream Bandwidth Capacity levels rise up to the ceiling again (requiring another node split). We can calculate the amount of time that each node split "buys" the MSO before another node split is required, and these results are shown below in Table 4 below.

Node Split		# Months Each Node Split "Buys" the MSO Before Another Node Split Is Required (assuming a 50% CAGR in Tmax & Tavg)
From (subs/SG)	To (subs/SG)	
16000 to 8000		20.0
8000 to 4000		19.5
4000 to 2000		18.5
2000 to1000		16.9
1000 to 512		13.9
512 to 256		11.2
256 to 128		7.7
128 to 64		4.8
64 to 32		2.7
32 to 16		1.5

Table 4: # Months before another Node Split for Node Splits with Different Sized Fiber Nodes

These diminishing returns on node split investments over time are concerning. It implies that at least for DOCSIS HSD services MSOs can expect to get smaller and smaller percentages of their DOCSIS HSD spectrum freed up by their successive node split activities if the conditions described above hold. It is even more concerning when coupled with the fact that successive node splits applied in a ubiquitous fashion across a market will become more and more expensive over time as the number of Fiber Nodes involved in the operation grows by a factor of two with each successive node split.

In addition, the average bandwidth utilization in a Service Group will become very low with more node splits occurring, because most of the bandwidth capacity within the service group will be added to simply provide the bandwidth capacity for the infrequently-used Tmax value. For the cases shown in Figure 2, the expected bandwidth utilizations within the different-sized Service Groups can be calculated, and these results are shown below in Table 5.

Service Group Size	Expected Average Bandwidth Utilization on DOCSIS Channels within a Service Group
16000 subs/SG	98%
8000 subs/SG	96%
4000 subs/SG	93%
2000 subs/SG	87%
1000 subs/SG	77%
512 subs/SG	63%
256 subs/SG	46%
128 subs/SG	30%
64 subs/SG	18%
32 subs/SG	10%
16 subs/SG	5%

Table 5: Expected Channel Bandwidth Utilization for Node Splits with Different Sized Fiber Nodes

Thus, as the Service Group size is decreased, the average channel utilization drops to extremely low levels due to the growing divergence between the Aggregate Average Bandwidth Consumption values and the Tmax values. Some MSO's may view this low channel utilization condition as being wasteful of the investment required to support the Required Bandwidth Capacity. One may wonder how MSOs will deal with this condition. One of the CCAP tools that can help is known as Output QAM Replication. It permits the channels within a CCAP chassis to be split in the digital domain and steered to more than one RF port on the chassis. This permits different service types (*e.g.*, DOCSIS HSD, VoD, SDV) to have Service Groups of different sizes.

It should be noted that VoD and SDV are Narrowcast services that are quite different from DOCSIS HSD. VoD and SDV do not suffer from the complexities caused by the rising Tmax values in DOCSIS. VoD and SDV do not have anything like a Tmax value. With VoD and SDV service types in reasonably-sized nodes, any node split that halves the number of subscribers sharing the Service Group resources will also produce a similar halving of the Required Bandwidth for that Service Group. (Note: For smaller-sized nodes, this rule-of-thumb may break down).

In the Tmax-oriented Period for DOCSIS HSD, MSOs may find it interesting to explore having different Service Group sizes for VoD and SDV and DOCSIS HSD. As an example, the splitting of Fiber Nodes may be of high value for VoD and SDV, but may be of low value for DOCSIS HSD. If this is the case, then the MSO can use Output QAM Replication and permit the DOCSIS HSD streams for a single DOCSIS HSD Service Group to propagate to more than one RF port on the CCAP chassis, whereas the streams for a single VoD/SDV Service Group could propagate to only one RF port on the CCAP chassis. This approach could help the MSO save on the cost of enabling the additional DOCSIS HSD processing on the RF ports. Another approach (which will be described in more detail below) uses both dedicated bandwidth (DOCSIS bandwidth dedicated to a fiber node) and shared bandwidth (DOCSIS bandwidth shared between multiple fiber nodes) to solve the problems of having low bandwidth utilization levels within service groups after performing many node splits.

Due to the four reasons listed above (the reduced impact of the node splits on required bandwidth over time, the reduced impact of the node splits on intra-node split durations over time, the higher cost of the node splits over time, and the low channel utilizations that result from the node splits over time), some MSOs may decide that they will only perform node splits for DOCSIS HSD up to a certain point in the future. After that, they may find that the node splits are too expensive. For example, if the MSO wishes to see at least a 30% reduction in the Required DOCSIS Downstream Bandwidth Capacity for any node split and that they must also see at least an 11 month period of time until the next node split is required, then the last node split that can be performed on a Service Group and accomplish these goal would be the node split that takes the Service Group size from 512 subs/SG to 256 subs/SG. In this case, node splits that reduce the size of the Service Group to values less than 256 subs/SG would be considered to be too expensive or too ineffective to be practical. This implies that many MSOs will likely stop their node splitting activities once their Service Groups fall to a certain size (on the order of 512 subs/SG, 256 subs/SG, or 128 subs/SG).

Due to the diminishing return on investment for node splits to smaller sized Service Groups, many MSOs will likely stop in that size range, placing an apparent limit on the window of time during which the HFC can provide adequate bandwidth capacity for the subscribers on the HFC plant. In other words, the ending of effective node splitting essentially defines a sunset time at which the HFC plant will cease to provide adequate bandwidth for all of the subscribers connected to it.

Many MSOs are trying to determine when this HFC sunset will occur. As can be seen in Figure 2, it is a function of many variables. How much total spectrum is supported by the HFC plant? How much of that spectrum will be dedicated to DOCSIS HSD services? What DOCSIS implementation (3.0 or 3.1) is utilized on the HFC plant? What are the SNR levels on the HFC plant? Which modulation formats are permitted on the HFC plant?

Different MSOs will answer these questions differently. But two examples of the sunset are illustrated in Figure 2 (for the assumed growth rates defined for the figure). The red arrows at the top of the figure locate dates at which an MSO would likely perform its last few node splits before the HFC sunset if they used DOCSIS 3.0 across only 64 QAM channels for its HSD services. These dates occur in the 2018-2021 time-frame and may permit the MSO to operate on the HFC network until ~2022. The green arrows at the top of the figure locate dates at which an MSO would likely perform its last few node splits before the HFC sunset if they used DOCSIS 3.1 across an entire 1.2 GHz plant for its HSD services. These dates occur in the 2022-2025 time-frame and may permit the MSO to operate on the HFC network until ~2026, so four years of extra runway were effectively created by the transition from DOCSIS 3.0 to the more robust DOCSIS 3.1 modulation formats and spectral widths.

One may wonder how these dates can be extended. One obvious way is for the MSO to alter one or more of the terms within Equation (1). As an example, some MSO Marketing teams may have the option of slowing down the growth rates on Tmax values within that formula. This may or may not be possible depending on the Tmax challenges from the competition. But if it is possible, then this can have profound effects on the life-span of the HFC network. This is illustrated in Figure 3, which shows the same types of plots found in Figure 2, but the plots in Figure 3 are for a network where the MSO Marketing team has found it possible to limit the growth rate on Tmax values to a 30% growth rate beginning in 2014. The plots illustrated in Figure 3 also show the growth rate on Tmax values dropping to a 30% growth rate beginning in 2014 as well. (Note: These growth rate reductions can be seen by the flattening of the curves on the right-hand side of the figure).

As illustrated by the red arrows at the top of the figure, this change in the growth rate (from 50% to 30%) of both Tmax and Tavg has clearly extended the sunset of the 64-channel plant to ~2024 or so, and the change also makes node splits to small node sizes valuable again. The green arrows at the top of the figure are even more intriguing, as they indicate that the change in the growth rate of Tmax has greatly extended the sunset of the 1.2 GHz HFC plant deep into the 2030 timeframe. The change in growth rates has also made node splits to smaller node sizes more valuable again, and it also extends the HFC sunset deeper into the future. If MSOs are able to initiate this change in Tmax growth (and a similar change occurs in Tavg values), then a lot of value can be obtained.

But what if MSOs are forced by competitive pressures to keep raising their Tmax values at the current 50% growth rate? And what if the Tavg values continue to grow at their current 50% growth rates? If that occurs, then MSOs would be forced to live on the curves of Figure 2, and they would start to see their HFC plants running out of gas in the 2020's. What can MSOs do to extend the life-span of their HFC plant in this case?

One promising solution based on the use of dedicated and shared bandwidth will be outlined in later sub-sections. Other highly-promising solutions based on DOCSIS 3.1 and RFOG and PON and Point-to-point Ethernet will also be outlined in later sub-sections. We will also explore several variants of another alternative technology known as Digital Access Architectures (DAAs) in later sub-sections. All of these alternative technologies, plus others, will be studied in the next section.

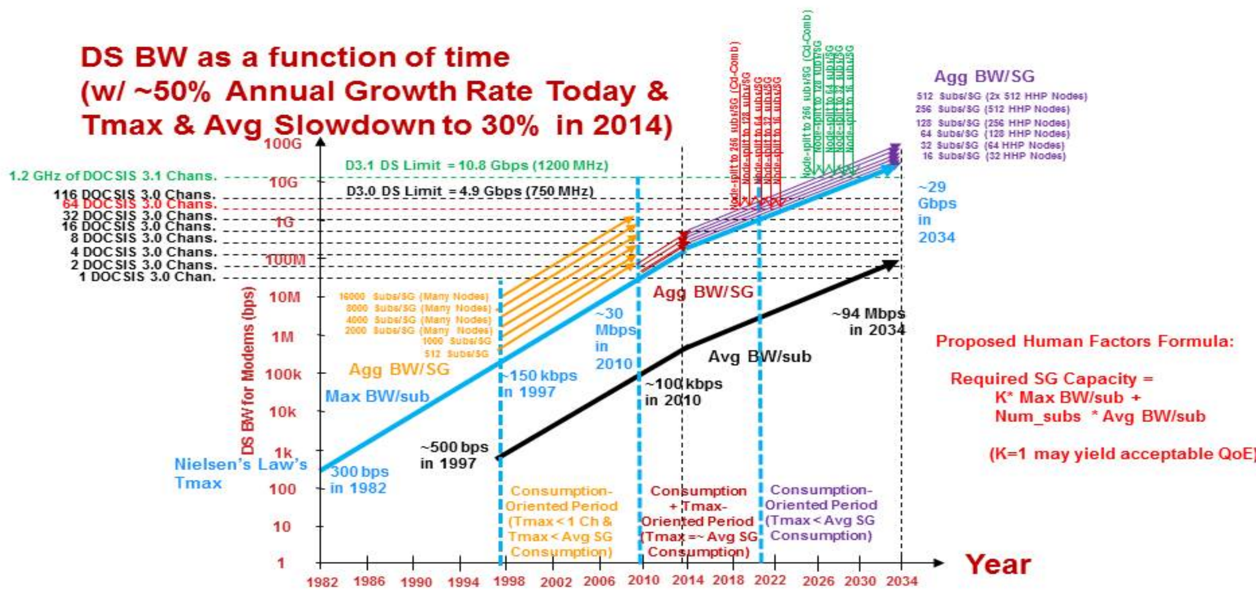


Fig. 3 DOCSIS HSD Required Downstream Bandwidth Capacity with 30% CAGR In Future

USE OF ALTERNATIVE TECHNOLOGIES IN OR AROUND THE HFC PLANT

Overview

There is a plethora of technologies being proposed that offer MSOs evolutionary or revolutionary changes to their existing HFC plant and spectrum to help accommodate the required bandwidth growth rates for all of their different service types in the future. In this section, we will explore several well-known approaches, and we will also define some new proposals. These approaches and proposals include:

- Increasing spectral efficiencies & spectral widths with DOCSIS 3.1
- Increasing spectral widths & Service Group capacities with Traditional Head-end-based CCAP systems
- Improving SNRs with Broadband Compression Forward/ Broadband Compression Return (BCF/BCR) Distributed Access Architectures (DAAs)
- Improving SNRs and Decreasing head-end power/rack space with Remote PHY Distributed Access Architectures (DAAs)
- Improving SNRs and Decreasing head-end power/rack space with Remote CCAP Distributed Access Architectures (DAAs)
- Bridging to the future with RFOG FTTH
- Increasing bandwidth with Extended-Spectrum RFOG FTTH
- Improving bandwidth capacity with Dedicated DOCSIS/Shared Extended-Spectrum RFOG arrangements (based on bonding)
- Capitalizing on various blends of Switched IP Video BCast over DOCSIS/Nailed-Up IP Video BCast over DOCSIS/IP Video VoD over DOCSIS/ BCast Digital Video/ SDV/ VoD Digital Video/ Analog
- Capitalizing on various blends of MPEG2/H.264 /HEVC compression
- Supporting various blends of SD/HD/4K/8K Video resolutions
- Increasing spectral efficiencies & spectral widths with EPOC
- Increasing bandwidth with PON FTTH
- Increasing bandwidth with Point-to-point Ethernet FTTH

From the long length of this list, it should be apparent that the cable industry is entering an interesting period in the history and evolution of the HFC plant. With the existing HFC plant limited in bandwidth capacity, changes are required if subscriber demand for bandwidth continues in the fashion described in the previous sections. Should MSOs augment their HFC plant to accommodate the growing bandwidth demands? Should MSOs deploy new technologies to accommodate the growing bandwidth demands? If so, which technologies should be deployed? All are good questions. And all are challenging questions.

There are a large array of choices that MSOs will be making as they adapt their HFC plants over the next fifteen years. Each MSO will be making these decisions independently, and each MSO will usually have a set of constraints that are unique from all of the other MSOs. As a result, some bifurcation of the HFC market is likely to take place as we move forward into the future and different MSOs select different paths. However, the decisions for any particular MSO can only be made wisely if the MSO is equipped with accurate information on each of the available approaches in the above list. We will attempt to provide a brief, and un-biased synopsis of each approach in the sub-sections below.

Improvements Provided By DOCSIS 3.1

DOCSIS 3.1 had been heavily discussed and publicized since its beginnings at Cablelabs in the mid-2012 timeframe. Due to its importance for the future, the cable industry (consisting of both MSO and vendor communities) has dedicated many resources to the effort to create the new DOCSIS 3.1 specification in record time. As a result, the specification is now available and vendors are rapidly working to deliver DOCSIS 3.1-capable products within the next year or two.

DOCSIS 3.1 is a backwards-compatible augmentation to the DOCSIS 3.0 specification that promises better spectral efficiencies (more bps/Hz) and wider spectral widths for both the Downstream and Upstream paths. The specification provides improved spectral efficiencies via many techniques, including the use of Orthogonal Frequency Division Multiplexing (OFDM) modulation, the use of higher modulation orders (4096QAM and higher), the use of more efficient Low-Density Parity Check (LDPC) Forward Error Correction, and the use of bit-loading to custom-fit the modulation orders to the varying SNRs across the spectrum of the HFC plant. Backwards compatibility is guaranteed by the fact that DOCSIS 3.0 and DOCSIS 3.1 channels can co-exist on the HFC spectrum. In addition, pre-DOCSIS 3.1 CMs will work with DOCSIS 3.1 CMTSs, and pre-DOCSIS 3.1 CMTSs will work with DOCSIS 3.0 CMs.

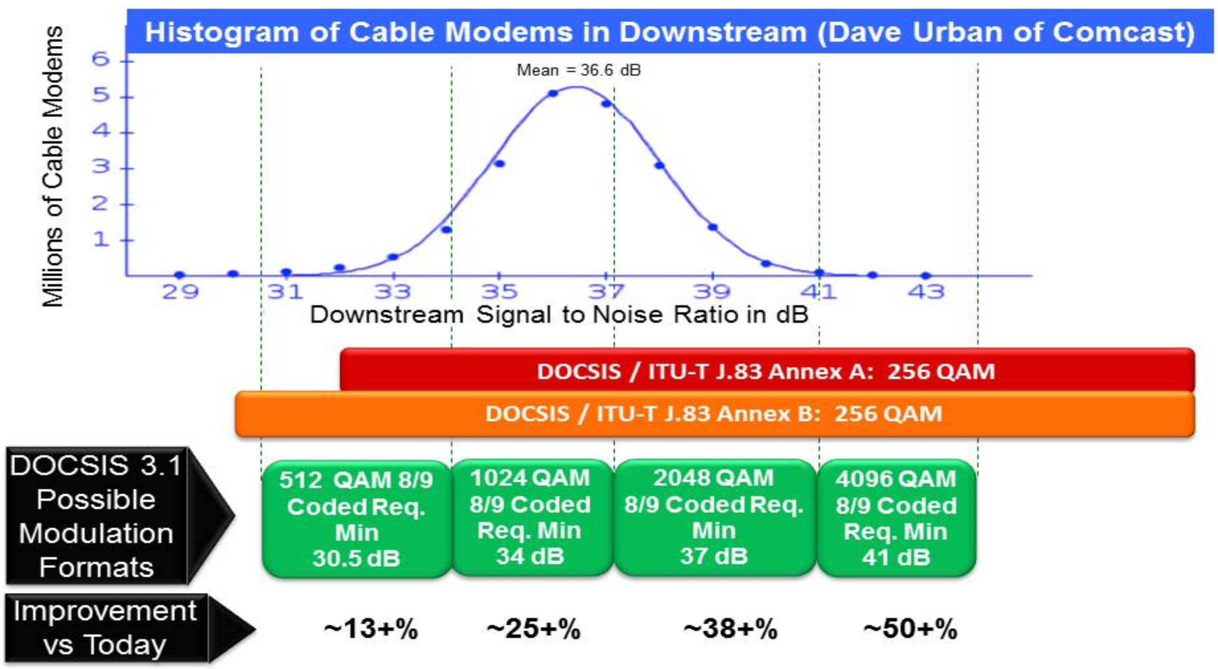


Fig. 4 Histogram of Modem Counts vs SNR and DOCSIS 3.1 Modulation Formats and Spectral Efficiency Improvements vs Today

As a result of its power and flexibility and backwards-compatibility, many MSOs are looking to DOCSIS 3.1 to give them a boost that will extend the life of their HFC plant by (at a minimum) several years. The actual HFC plant life extension that will result from the use of DOCSIS 3.1 depends on many different factors, including the annual subscriber bandwidth growth rates, the number of node splits that are performed, the amount of investment that the MSO is willing to put into their plant to extend its spectral width, and the quality of the HFC plant (i.e.-

SNRs). To better quantify this last point, the bottom of Figure 4 illustrates the likely spectral efficiency gains that may be expected if an MSO transitions from DOCSIS 3.0 to DOCSIS 3.1. These gains are shown to be a function of SNR, and improved SNRs obviously lead to increased spectral efficiency gains. Overlaid on top of the chart is a histogram created by David Urban of Comcast, showing a distribution of sampled field data taken from a large number of modems in HFC plants today [URB1]. Obviously, improving the SNRs in HFC plants would lead to improved spectral efficiencies in the future.

Examples shown in Figures 2 and 3 illustrate that reasonable DOCSIS 3.1 investments (using full-spectrum DOCSIS 3.1 in a 1.2 GHz plant) could provide HFC plant life extensions of 5-6 years. Heavier investments that push the DOCSIS 3.1 spectrum to 1.7 GHz or higher could yield even longer extensions to the HFC plant life.

Improvements Provided By Traditional Head-end-based CCAP Systems

As MSOs move forward, there is no doubt that node splits will be a natural part of the future HFC plant evolution going forward. Due to their importance, the benefits of node splits were discussed in detail within the first sections of this paper. In those paper, it was shown that there may be at least 3-4 rounds of node splits that many MSOs may perform in the next decade (before the value of node splits begins to diminish).

In addition, it is becoming clear that some MSOs will consider increasing the spectral widths on their Upstream spectrum and/or their Downstream spectrum.

CCAP chassis of the future will have to evolve to support these two changes. Well-designed CCAP chassis should be able to accommodate these changes with minimal hradwaree changes (ex: the insertion of new cards into existing chassis). In addition, well-designed CCAP chassis should also be able to support the expected Network-Side Interface and backplane bandwidths that would be required for these node-splits and this spectral expansion. The authors have performed some studies of this topic in a companion paper [ULM1], and it appears that at least a subset of the existing CCAP chassis on the market today should be capable of supporting at least 3 rounds of node splits between now and 2020, while also providing DOCSIS 3.1 support for at least 1.2 GHz of spectrum on each Service Group. This capacity expansion should be achievable without requiring any significant change in chassis power or rack-space requirements. Further expansion may also be possible, but further innovation in RF connector density may be required to support that.

As a result of these facts, many MSOs will be able to rely on their existing Traditional Head-end-based CCAP chassis to support their HFC network needs (DOCSIS and EQAM) for many years to come. And due to the benefits of Moore's Law, most MSOs will likely be able to utilize these Traditional head-end-based CCAP resources without suffering from a need to increase their head-end power and rack-space requirements. Using the plots within Figures 2 and 3, Traditional Head-end-based CCAPs should be able to support MSOs (without head-end power or rack-space increases) until ~2024 (with a 50% CAGR) or until ~2030 (with a 30% CAGR). If MSOs are able and willing to slightly expand the amount of power and rack-space required for CCAP chassis within their head-ends, then these dates can be extended even further into the future.

Improvements Provided By Distributed Access Architectures (DAAs)

As mentioned in the previous sub-section, most MSOs will likely be able to support their video and HSD services using Traditional Head-end-based CCAPs. However, some MSOs may be planning to limit their HSD bandwidth capacity growth going forward (if they can), and as a result of creating that limit, they may foresee benefits to performing more node splits more rapidly than other MSOs. These MSOs may see a need to support more Service Groups than would be easily supported by a Traditional Head-end-based CCAP chassis. In particular, they may run into issues related to the required power and/or rack-space for the large number of CCAPs that they may need to deploy in their head-ends.

There are two other types of MSOs who may also see issues in using Traditional Head-end-based CCAPs. The first type consists of those MSOs who might be planning to use large amounts of Wavelength Division Multiplexing (WDM) on a small number of fibers feeding a large number of fiber nodes (which may be created by active node splitting activities of the future). The second type consists of those MSOs who might be planning to deploy (or have already deployed) long fiber runs to their fiber nodes. Both of these conditions (many wavelengths or long fiber runs) can lead to the introduction of many optical noise side-effects resulting from nonlinear effects within the core of the fiber.

There are numerous sources of these nonlinear effects within optical fibers. Some are due to changes in the refractive index of the medium with optical intensity. Others are due to an inelastic scattering phenomenon. Typical phenomena include:

- Self-Phase Modulation (SPM)
- Cross-Phase Modulation (CPM)
- Four-Wave Mixing (FWM)
- Stimulated Brillouin Scattering (SBS)
- Stimulated Raman Scattering (SRS)

In general, these nonlinear effects generate noise whose magnitude is increased with higher optical powers. As mentioned, these nonlinear effects are also increased as a result of more lambdas on the fiber and/or longer optical fiber runs. The changes that lead to this problem (more lambdas per fiber and longer fiber runs) are definitely expected in the future, and they will exacerbate the noise problems caused by nonlinear fiber effects.

Unfortunately, this implies that noise levels on the fiber will rise (and corresponding SNR levels will fall) at a time when many MSOs may want to deploy the higher QAM modulation orders that are permitted by DOCSIS 3.1. Since the higher QAM modulation orders (like 1024QAM or 4096QAM or even 16384QAM) will require much higher SNR levels to operate with adequate bit error rates over the coaxial portion of the HFC plant, dealing with the aforementioned nonlinear noise issues within the fiber portion of the HFC plant may become a necessity in the future.

One technique for mitigating the effects of the nonlinear optical noise (and increasing the overall end-to-end SNR of the HFC plant) is to use a newer generation of lasers and fibers and receivers. As an example, the use of Externally Modulated Lasers (instead of Direct Modulated Lasers) will greatly help to reduce the dispersive effects of chirp, and that will generally help to reduce noise on the fiber. Studies have indicated that the correct usage of these components will permit operation to the higher modulation orders in most HFC plants.

However, there is another exciting technique for mitigating the effects of the nonlinear optical noise, which also helps to solve the problems of MSOs who have issues with the required power and rack-space within their head-ends. This technique employs Distributed Access Architectures (DAAs).

There are many types of DAAs being proposed for use in the future, and each proposal has its own sets of pros and cons [EMM1]. Due to lack of space, we will give only a brief description of three of them, which are shown in Figure 5.

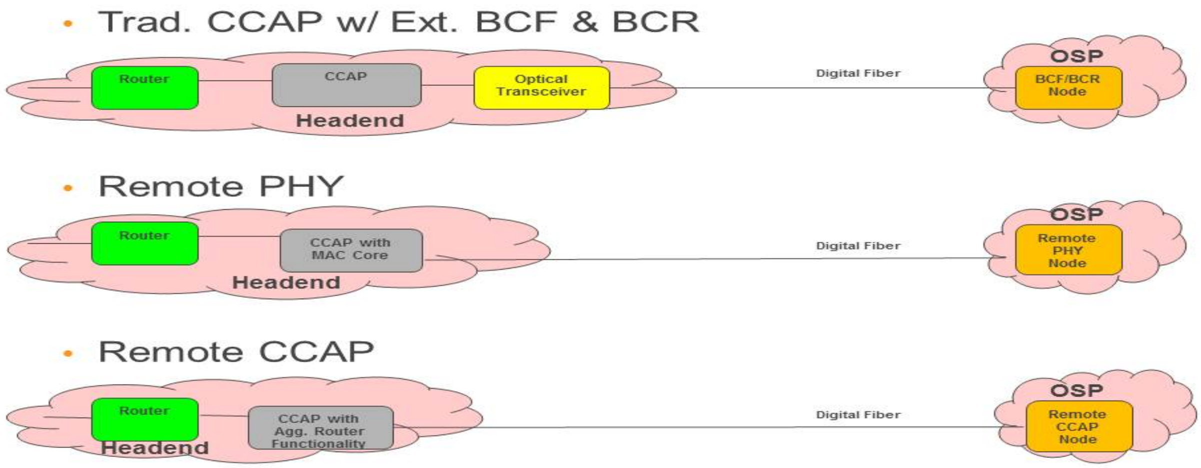


Fig. 5 Example Distributed Access Architectures

Broadband Compression Forward /Broadband Compression Return (BCF/BCR): This approach places new circuitry in the head-end optical transceiver equipment and in the fiber node. It assumes a separate optical shelf receiving RF sources from analog video, Edge QAM, CMTS, CCAP, RF Out-of Band, and RF Test equipment. The Broadband Digital equipment receives multiplexed RF feeds in the head-end and digitizes the spectrum before it is transported via Digital Optics to (or from) the Fiber Node. Key components of this process are the Analog-to-Digital Converters (ADCs) and the Digital-to-Analog Converters (DACs), which convert between analog and digital signal formats. The digital samples of the analog signal can be transmitted in the payload of standard Ethernet packets across the low-cost digital optic fiber system (which now carries digital Ethernet signals that are much more robust to optical noise than their amplitude modulated counterparts). This approach creates the least amount of disruption to the existing head-end equipment, and it is also the only solution that allows for full transparency of the RF feeds over the outside plant. This type of solution is actually in use today for the Upstream direction- it is typically called Broadband Digital Return (BDR). Improvements in DAC and ADC technology speeds now permit this same approach to be utilized in the Downstream direction as well. In its most generic form, the Downstream approach would be called Broadband Digital Forward (BDF). However, if novel compression techniques are used to process the digital packet streams, then the resulting approach can be called Broadband Compression Forward (BCF) or Broadband Compression Return (BCR). BCF/BCR helps with the nonlinear optical noise problem, but it does not help with the head-end power and rack-space problem.

Remote PHY (R-PHY): This approach separates the PHY (Upstream and Downstream) from the head-end and places the full PHY layer (including the FEC, symbol generation, modulation, and DAC/ADC processing) into the Fiber Node. This requires that these functions be removed from the head-end CCAPs, CMTSs, and EQAMs. The DOCSIS MAC processing remains in the MAC Core within the head-end. This approach is slightly disruptive, as it requires many pieces of head-end equipment (ex: CCAPs, CMTSs, and EQAMs) to be modified. There are some similarities between this R-PHY approach and the Modular Headend Architecture (MHA) approach. But there are also many differences, such as the need to support Upstream MAC/PHY separation, the need to support new timing interfaces that work over Ethernet, and the need to add DOCSIS 3.1 support within DEPI and UEPI. However, this approach offers benefits as well. Remote PHY helps with the nonlinear optical noise problem, and it also helps with the head-end power and rack-space problem. Another benefit of the Remote PHY approach is that it permits MSOs to continue to re-use their head-end-based CCAPs as part of the solution. That represents a form of investment protection.

Remote CCAP (R-CCAP): This approach places the entire upper and lower MAC (Upstream and Downstream) and the entire PHY layer functionality (Upstream and Downstream) into the fiber node. In effect, this places all of the CMTS, Edge QAM, and CCAP functions into the Fiber Node and only requires a switch or router to remain in the head-end. As a result, this approach is slightly disruptive. However, Remote CCAP helps with the nonlinear optical noise problem, and it also leads to the maximum amount of power and rack-space savings within the head-end (even more than the Remote PHY approach). It is also possible that existing head-end CCAPs (if appropriately modified) could be used to serve as dense Aggregation Routers (or PON OLTs) feeding the Remote CCAPs as well.

Improvements From RFOG and Extended-Spectrum RFOG

RF Over Glass (RFOG) technology permits MSOs to transmit their standards RF signals (ex; DOCSIS, MPEG-TS Video , Analog) all the way to the subscriber homes over fiber. It requires a special ONU to be placed within each home, and the ONU is responsible for performing an optical-to-electronic conversion function (which is quite similar to the function performed by a typical fiber node). RFOG offers several benefits to MSOs. First, it permits them to begin transitioning their HFC plant into a Fiber-To-The-Home (FTTH) plant (which is likely to be the plant of the future). Second, RFOG eliminates the coaxial portion of the HFC plant, which can lead to improved SNRs and higher modulation orders. Third, RFOG can extend their DOCSIS 3.1 transmission system to spectral widths that exceed the 1.2-1.7 GHz spectral limits of typical coaxial distribution systems within the HFC plant.

The spectral extensions permitted by RFOG can be utilized in both the Downstream and Upstream directions, and it could potentially permit DOCSIS 3.1 systems operating over RFOG to provide much higher bandwidths. As an example, a shared, symmetrical 40+ Gbps DOCSIS 3.1 service to subscribers (as shown in Figure 6) is feasible with this Extended-Spectrum RFOG technology. This type of DOCSIS 3.1 banwidth extension could greatly delay the ultimate HFC sunset deep into the future.

Interestingly, the resulting Optical OFDM DOCSIS 3.1 system could provide important mitigation against fiber dispersion, because OFDM is comprised of many narrow-spectrum subcarriers that are not likely to experience much dispersion.

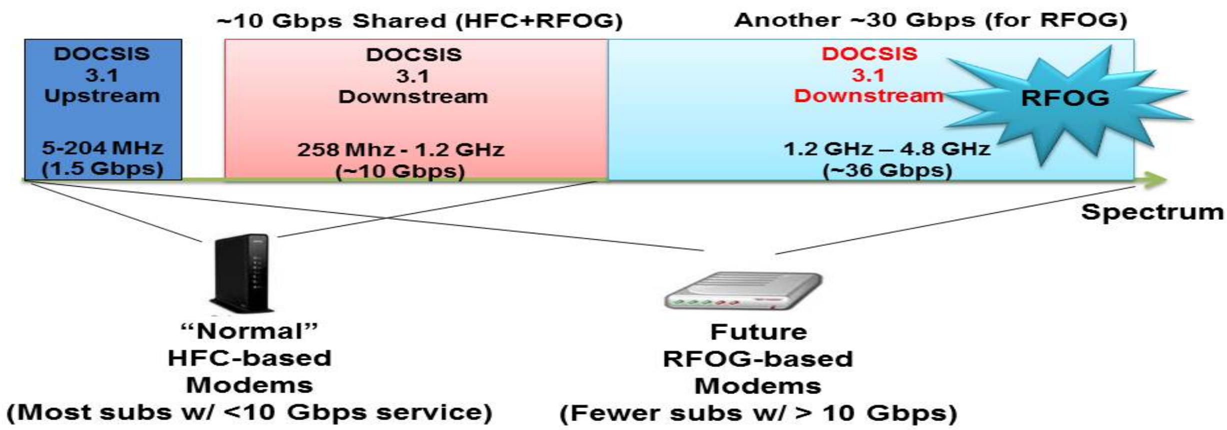


Fig. 6 Example Of An Extended-Spectrum RFOG System

An interesting attribute of Extended-Spectrum RFOG is that it can easily capitalize on the channel bonding capabilities of DOCSIS 3.1. Using this capability, the traditional spectrum below the 1.2-1.7 GHz limit can be channel-bonded with new RFOG-only spectrum that resides above the 1.2-1.7 GHz limit. This approach may provide a nice level of investment protection with the DOCSIS equipment of the past- with the new CPE devices that work above the 1.2-1.7 GHz limit being allowed to also re-use the bandwidth offered below the 1.2-1.7 GHz limit. It can thus permit MSOs to use the RFOG spectrum above the 1.2-1.7 GHz limit as a "boost feed," increasing the normal bandwidth into the subset of subscriber homes who request more Tmax bandwidth than that which is provided by the traditional spectrum below the 1.2-1.7 GHz limit.

Since only a small subset of subscribers may opt for the Extended-Spectrum RFOG technology in any given year, the use of this technology could also provide MSOs with a slow and incremental way to transform their HFC network into a FTTH network- a few subscribers at a time.

<u>Improvements From New Blends Of Video Delivery Techniques</u>

Although High-Speed Data (HSD) is the fastest growing service within the MSO's HFC spectrum, MSO-managed video services still consume the largest percentage of the spectrum today. And those MSO-managed services produce a large amount of revenue, so they will undoubtedly remain a big portion of the spectrum for many years to come. How much spectrum will MSO-managed video consume in (say) 2025? That is an important question, because whatever spectrum is used by MSO-managed video cannot be used by the growing HSD services. (Note: The green horizontal lines in Figures 2 and 3 assume that HSD services are offered 100% of the available HFC spectrum. Obviously, this is not a likely scenario in the near or distant future). As a result, the seemingly un-stoppable growth in HSD bandwidth will eventually hit a limiting ceiling on the capacity that is available, and offering more spectrum to MSO-managed video services will force the HSD bandwidth to hit that ceiling sooner in time. In a sense, MSO-managed video and HSD will become fierce competitors for the precious spectral resources on the HFC plant as time moves forward into the future.

Thus, to accommodate the growing HSD bandwidth, MSOs will undoubtedly look to various technology paths that offer to squeeze the bandwidth of MSO-managed video into a smaller portion of the HFC spectrum. Some of these technology paths will force them to re-consider the technologies used to transport their MSO-managed video. Thus, the future will likely see different MSOs using different mixes of SD BCast Digital Video, HD BCast Digital Video, SDV, VoD Digital Video, and Analog.

Over time, Analog spectrum will be heavily reclaimed (if not entirely reclaimed). DTAs offer a good, low-cost technique for accomplishing that goal, but future Media Gateways with low-cost IP-STBs may also provide similar low-cost alternatives. In the future, as SD BCast Digital Video becomes less popular and HD BCast Digital Video becomes more popular, SD BCast Digital Video will also begin to be turned off, yielding more reclaimed spectrum. SDV is another technique that can help to reclaim spectrum from the BCast Digital Video tier, whereby video streams are only transmitted over a Service Group if a subscriber is viewing that stream. This can save anywhere from 50-66% (or more) of the BCast Digital Video spectrum, permitting MSOs to offer up to 3 times more programs in a given amount of spectrum or permitting MSOs to reduce their spectrum requirements by 66%.

In addition to a transition away from Analog Video towards Digital Video, and in addition to a transition away from BCast Video and towards SDV, many MSOs are also looking to a transition away from MPEG-TS Digital Video delivery to IP Video delivery over DOCSIS. There are several reasons for this trend. First, IP Video delivery over DOCSIS will permit MSOs to capitalize on the higher spectral efficiencies of DOCSIS 3.1- which can yield a 50% improvement over the spectral efficiencies of traditional MPEG-TS delivery. Second, the wider effective channel widths created by DOCSIS channel bonding can help produce statistical multiplexing gains that can provide transport for an additional 20-30% of programs (or can provide commensurate savings in video service bandwidth requirements). Third, the use of IP Video can permit MSOs to capitalize on the bandwidth-saving benefits of Adaptive Bit-Rate (ABR) algorithms for all of their multiplexed unicast video streams. Fourth, more intelligent head-end-based adaptive algorithms that take into account the nature of the video content and the nature of the video receivers can permit even more bandwidth savings [ULM2]. Fifth, IP Video over DOCSIS will permit MSOs to consolidate their head-end servers and use common equipment for both their primary screen video delivery and their multi-screen video delivery. This may result in overall cost savings.

IP Video over DOCSIS may be phased in using many different tricks- the challenge is to find a way to shorten the “Simulcast window,” which is the period of time during which MSO-managed video services need to be delivered in both their legacy MPEG-TS infrastructure as well as in the new DOCSIS infrastructure within a single Service Group. It may be used to transport VoD services (in lieu of the similar MPEG-TS-delivered VoD services). It may also be used initially to transport ethnic video services to a subset of the subscribers who access that ethnic offering. Over time, more and more programming will likely be delivered over IP Video, and less will likely be delivered over MPEG-TS. As more Digital BCast Video programming is moved onto the IP Video delivery infrastructure, MSOs will be able to utilize Switched IP Multicast over DOCSIS as well as Nailed-Up IP Video Multicast over DOCSIS.

Improvements From New Blends Of Video Compression Techniques

As mentioned in the previous sub-section, MSOs will definitely be looking to various technology paths that offer to squeeze the bandwidth of MSO-managed video into a smaller portion of the HFC spectrum. Using these technologies will permit the MSOs to accommodate the rapid growth of HSD bandwidth on their HFC plant for a longer period of time. In addition to changes in the blend of Video Transmission Techniques, MSOs will also be looking very seriously at changes in the blend of Video Compression Techniques.

Many Digital Video Compression technologies are now available for MSOs to consider and utilize. These include the MPEG2, H.264, and HEVC compression techniques. Proprietary statistical-multiplexing technologies that simultaneously compress multiple streams are also a key part of these compression algorithms.

	MPEG2	H.264	HEVC
SD (480i30)	2.25 Mbps	1.25 Mbps	0.75 Mbps
HD (1080i30 or 720p60)	9 Mbps	5 Mbps	3 Mbps
HD (1080p60)	16 Mbps	9 Mbps	5 Mbps
UHD (4Kp60)	65 Mbps	35 Mbps	20 Mbps
UHD (8Kp60)	260 Mbps	140 Mbps	75 Mbps

Table 6: Video Bandwidth Requirements With Different Resolutions And Compression Techniques

The compression efficiencies of these various technologies continue to evolve and improve over time, but the comparative quality of each compression technique varies from technology to technology, from vendor to vendor, from month to month, and from video quality expert to video quality expert. As a result, it is always difficult to compare the efficiencies of these compression technologies using an apples-to-apples comparison that everyone agrees on. Nevertheless, one video quality expert recently gave an assessment of the video compression capabilities of these different technologies as they exist today. That assessment is captured in Table 6.

From the table, it is apparent that a move towards HEVC encoding would provide the best compression ratios. As a result, MSOs will likely want to eventually move in that direction. However, they will need to migrate in that direction from their current MPEG2 equipment, and there are limits placed on them by the fact that the existing equipment deployed in the field may not support the improved compression technologies. As a result, this migration path may be a two-step path, moving from MPEG2 to H.264 first, and then moving from H.264 to HEVC. (Note: The arrival of some 4K video content in the coming years will require that HEVC be instantly utilized for that content).

Challenges From New Blends Of Video Resolution Techniques

While MSOs have a large array of available tools that they can utilize to help compress the amount of bandwidth required for MSO-managed video in the future, there are also some evolutions taking place that will unfortunately push the required bandwidth in the opposite direction- requiring more and more bandwidth capacity as time goes forward. The biggest threat in this area is the arrival of new and higher video resolutions for the future. SD video and HD video have been common-place on the HFC plant for many years. But the next round of video resolution improvements will produce Ultra High-Definition (UHD) video streams that require more bandwidth. 4K UHD content delivery (3840 pixels x 2160 pixels, progressive scan @ 60 frames per second, compressed using HEVC) will consume ~2 times the amount of bandwidth of today's HD content delivery (1920 pixels x 1080 pixels, interlaced scan @ 30 frames per second, compressed using MPEG2). In addition, 8K UHD content delivery (7680 pixels x 4380 pixels, progressive scan @ 60 frames per second, compressed using HEVC) will consume even more bandwidth within the HFC spectrum.

<u>Improvements Provided By EPOC</u>

Ethernet PON Over Coax (EPOC) is a developing standard that will provide a new technology available for MSOs to consider for use within HFC networks. In an EPOC system, Passive Optical Network (PON) technology is used to transmit signals over the digital fiber to the node. As a result, it helps mitigate against nonlinear optical noise issues. Within the fiber node, the PON signals can be converted into QAM-based OFDM signals for final transmission over the coaxial portion of the HFC plant. EPOC will likely utilized LDPC Forward Error Correction as well. As a result, the performance of EPOC is expected to be very similar (or identical) to the performance of DOCSIS 3.1 systems. Since both EPOC and DOCSIS 3.1 operate in an HFC network, the two technologies will likely compete for the interests of the MSOs in the future.

MSOs will likely make their EPOC vs DOCSIS 3.1 decision using the following logic. EPOC systems have the benefit of beginning to utilize a technology (PON) that may be a key technology for the future, but it has the disadvantage that it is not backwards-compatible with existing MPEG-TS or DOCSIS CPE equipment that is already deployed. On the other hand, DOCSIS 3.1 and CCAP offers backwards-compatibility to both existing DOCSIS CPE equipment and existing MPEG-TS CPE equipment. But DOCSIS 3.1 and CCAP also offers a future path to the PON systems of the future (since CCAP specifications incorporate PON functionality- even though no existing CCAPs actually offer it yet).

So the real question seems to boil down to whether MSOs will value a present-day ability to begin deploying equipment that they may be using in the future (ex: PON OLTs) or whether MSOs will value a present-day ability to be backwards-compatible with already-deployed CPE equipment (permitting investment protection and equipment re-use). Those who prefer the former scenario may choose EPOC. Those who prefer the latter scenario may choose DOCSIS 3.1.

<u>Improvements Provided By PON</u>

A Passive Optical Networks (PON) is a technology that provides a direct optical link between the head-end and the subscriber home. The device in the head-end is called an OLT, and the device in the home is called an ONU or an ONT. Many ONUs (or ONTs) can share a single FTTH optical feed from the OLT in the head-end, so the bandwidth capacity provided by a PON is always shared by all of the ONUs (or ONTs) connected to the PON feed.

There are two incompatible PON technologies that have been defined in different standards committees- EPON (driven by the IEEE organization) and GPON (driven by the ITU organization). PON technologies of the future may include bandwidth capacities such as 1 Gbps, 2.5 Gbps, and 10 Gbps. Ultimately, 40+ Gbps bandwidths will also likely be provided. This is an overlay technology to the DOCSIS HFC delivery system, since it does not offer any form of backwards-compatibiltiy to DOCSIS. PON will likely be used in Business Services and MDU environments first, but it will also find great utility in servicing Residential subscribers as well (once Residential subscriber bandwidth demands exceed those that can easily be provided by traditional DOCSIS systems).

PON may find a few competitors in the FTTH space. One FTTH competitor to PON is RFOG/Extended-Spectrum RFOG (which was described in a previous sub-section). If future Extended-Spectrum RFOG systems are created with low-cost DOCSIS 3.1 bandwidths exceeding 40 Gbps, then MSOs will definitely need to consider both approaches when doing comparisons for higher-bandwidth systems. Another FTTH competitor to PON is Point-to-point Ethernet, which will be described in the next sub-section.

Improvements Provided By Point-to-point Ethernet

Whereas PON is a FTTH technology that provides shared bandwidth services, Point-to-point Ethernet is a FTTH technology that provides dedicated bandwidth services. For a particular optical fiber in a Point-to-point Ethernet system, there is one and only one subscriber connected to the fiber. The bandwidth capacities associated with Point-to-point Ethernet will follow the Ethernet bandwidth curves. As a result, 1 Gbps, 10 Gbps, 40 Gbps, and 100 Gbps Ethernet are readily available today. 400 Gbps Etherne tis also becoming available (although it is still quite expensive). There is also work to extend Ethernet to 1 Tbps services in the future.

As a result of these higher bandwidth capacities, Point-to-Point Ethernet may become a popular access technology of the future for Business Services applications and MDU applications. It may also become popular for Residential services if/when the required Tmax bandwidth levels transmitted into each home exceed the bandwidth capabilities of PON or Extended-Spectrum RFOG. (Note: While it is difficult to predict the future, PON and Extended-Spectrum RFOG may peak out at 40 Gbps of banwidth).

PREDICTIONS

Defining The Approach

Predicting the future is always a challenging taks, and predicting the future to the year 2030 and beyond borders on the edge of fool-hardy, because new technologies and new subscriber demands can always develop without warning over a lengthy fifteen year period. As a result, those who attempt to make long-term predictions are likely to be wrong.

The authors expect that to be the case for many of the predictions that will be made within this section of the paper. Nevertheless, the authors are hopeful that there is still some value in laying out predictions about the paths that MSOs may follow into the future, because the prediction exercise requires one to examine the subscriber demands and the upcoming technologies that are likely to be available in the future that can accommodate those demands. Weaknesses and strengths of each technology can therefore be identified, and that information alone can oftentimes be useful. In addition, the creation of the predictions will hopefully stimulate good discussions that lead to better predictions in the future.

The authors will not describe a single path into the future, because it seems certain that the technologies used within HFC plants will likely be quite varied as MSOs move forward into the future. Market bifurcations will undoubtedly occur as different MSOs select different paths based on their different constraints, starting points, and biases.

Instead, we have decided to utilize the information collected in the preceding sections and attempt to describe the higher-probabilities scenarios that are likely to play out at various MSOs between now and 2030. These higher-probability scenarios will be presented in a list. The list will be divided into three sections:

1) scenarios that will impact most MSOs (in general),
2) scenarios that will impact MSOs who choose to extend the life-span of their current HFC plant, and
3) scenarios that will impact MSOs who choose to switch to new technologies early as a means of supplementing the bandwidth on their current HFC plants (without incurring any costs of HFC plant upgrades).

We will utilize the terms "all," "most," "many," "some," and "few" to indicate the probability (from highest to lowest) that MSOs will follow a particular path.

Scenarios Impacting All MSOs

The higher-probability scenarios that will likely play out at most MSOs between now and 2030 include the following:

- All MSOs will begin to face growing demands for more and more bandwidth capacity for both their DOCSIS HSD services (due to standard growth) and their Video Services (due to a gradual shift towards UHD feeds), creating stresses on their HFC spectrum.
- All MSOs will be forced to transport larger and larger numbers of higher-bandwidth 4K UHD video streams within the next several years. Support for some 8K UHD video streams may also be required as we progress into the 2020 decade.
- All MSOs will be faced with an important decision to either extend the life-span of their current HFC plant to support the required bandwidth capacity (upgrading the HFC plant in an effort to delay any transition to new technologies) or to switch to new technologies early as a means of supplementing the bandwidth on their current HFC plants (without incurring any costs of HFC plant upgrades).

Scenarios Impacting MSOs Who Will Extend The Life-Span Of Their Current HFC Plant

The higher-probability scenarios that will likely play out at most MSOs who will choose to extend the life-span of their current HFC plant include the following:

- Most of these MSOs will continue to take advantage of Traditional Head-end-based CCAPs deep into the future.
- Many of these MSOs will converge their video EQAM functionality into their CCAPs in an effort to reduce power and rack-space requirements within the head-end. Some of these MSOs will wait until actual power and rack-space issues develop (likely in the 2020 timeframe) before they begin this transition. Some of these MSOs will plan ahead and begin the convergence earlier.
- Most of these MSOs will perform at least 2-3 rounds of node splits in the next 6-10 years. This "business-as-usual" activity will help them accommodate Narrowcast bandwidth growth over the next decade. As they perform these node splits, the MSOs will capitalize on the increased Service Group counts that will undoubtedly be provided by Traditional Head-end-based CCAP boxes of the future.
- Most of these MSOs will take advantage of DOCSIS 3.1 as a means of increasing the available bandwidth capacity of their existing HFC plants.

- Many of these MSOs will likely come to realize that moving any video streams (whatsoever) from their current MPEG-TS delivery/MPEG2 compression into IP Video over DOCSIS delivery/HEVC compression (with its improved DOCSIS 3.1 spectral efficiencies and improved large-channel stat-mux gains and improved ABR-based compression techniques) will be essential to permit UHD Video feeds to be transmitted over the HFC plant in the future. Any transfer of video content from PMET-TS to DOCSIS will prove beneficial.
- Many of these MSOs will likely move quickly from MPEG2 compression to H.264 compression, and then will move from H.264 compression to HEVC compression. Some will skip the intermediate step of H.264 compression and move directly from MPEG2 compression to HEVC compression.
- Many of these MSOs will begin to slowly move towards IP Video over DOCSIS transport for their video in the next few years. But as those MSOs begin to move video streams to IP Video over DOCSIS deliver systems, they will eventually realize that they should expedite the process to get through the "simulcast window" more quickly, permitting them to eventually reclaim most of the bandwidth associated with MPEG-TS video. They will therefore expedite this process and attempt to retire MPEG-TS video transport as quickly as possible.
- Many of these MSOs will start using Switched IP Multicast for Linear IP Video transmissions to capitalize on the bandwidth efficiencies that can result from the use of Switched IP Multicast.
- Many of these MSOs will increase their Downstream spectrum to 1.2 GHz in an attempt to provide more HFC plant bandwidth capacity.
- Some of these MSOs will increase their Downstream spectrum to 1.7 GHz in an attempt to provide even more HFC plant bandwidth capacity.
- Many of these MSOs will change the split on their Upstream spectrum to be 85 MHz in an attempt to provide more Upstream bandwidth capacity.
- Some of these MSOs will change the split on their Upstream spectrum to be 204 MHz in an attempt to providemore Upstream bandwidth capacity. This may permit Tmax = 1 Gbps Upstream service offerings.
- Some of these MSOs may utilize BCF/BCR to improve SNRs and permit them to use the highest possible modulation orders within DOCSIS 3.1 systems. This may be more likely for MSOs who may be planning to multiplex a lot of wavelengths on a fiber or who may be deploying long fiber runs.
- Some of these MSOs may utilize Remote PHY or Remote CCAP or Remote PON architectures to improve SNRs (increasing modulation orders) and to reduce power and rack-space requirements in the head-end. This may be more likely for MSOs who may be planning to multiplex a lot of wavelengths on a fiber or who may be deploying long fiber runs. This may also be more likely for MSOs who may be performing rapid node splits. This may also be more likely in the 2020 decade, when MSOs may be out-pacing the abilities of CCAP boxes to support further node splits.
- Some of these MSOs may utilize Remote PON (a variant of Remote CCAP) to improve SNRs (increasing modulation orders) and to reduce

power and rack-space requirements in the head-end.

- Some of these MSOs may utilize RFOG as a technique to roll out FTTH technologies to homes without requiring them to change out existing head-end and CPE equipement. This may be more likely for MSOs who are rolling out Greenfield deployments. This may also be more likely for MSOs who are planning to transition their HFC plants into FTTH plants in preparation for the introduction of new technologies (such as PON or Point-to-pint Ethernet) in the future.
- Some of these MSOs may utilize Extended-Spectrum RFOG as a means of increasing the bandwidth capacity within a Service Group as Tmax values rise in the future. This type of bandwidth (in excess of 10 Gbps) and this type of Extended-Spectrum RFOG technology may not be required until the 2020 decade.

Scenarios Impacting MSOs Who Will Switch To New Technologies

The higher-probability scenarios that will likely play out at most MSOs who choose to switch to new technologies early as a means of supplementing the bandwidth on their current HFC plants (without incurring any costs of HFC plant upgrades) include the following:

- Some of these MSOs may begin to use EPOC as an interim technology in preparation for an ultimate transition to PON.
- Some of these MSOs may begin to use PON as a technology to deliver higher bandwidths to their subscribers. They may transition to PON directly from DOCSIS, or they may transition to PON after using EPOC for a while.
- Some of these MSOs may begin to use DOCSIS 3.1-based RFOG as a technology to deliver higher bandwidths to their subscribers.
- Some of these MSOs may begin to use DOCSIS 3.1-based Extended-Spectrum RFOG as a technology to deliver even higher bandwidths to their subscribers.
- Some of these MSOs may begin to use Point-to-point Ethernet as a technology to deliver even higher bandwidths to their subscribers. The use of Point-to-point Ethernet becomes more probable as we move closer to the 2030 time-frame, because the dedicated bandwidths associated with Point-to-point Ethernet may not be required until the 2030 decade.

CONCLUSIONS

In general, the next fifteen years promise to be an extremely interesting time, and it is quite apparent that many changes will be taking place throughout the cable industry as MSOs prepare for the 2030 decade and beyond.

Fortunately, there are plenty of technology options available to MSOs that can help them maneuver their way through the transitions. Different MSOs will select different paths, but all MSOs will undoubtedly be focusing on increasing their bandwidth capacities to deliver high Quality of Experience levels to their subscriber base for all of the applications of the future (whatever they may be).

REFERENCES

[URB1] D. Urban, Comcast study.

[ULM1] J. Ulm and T. Cloonan, "Scaling Traditional CCAP To Meet The Capacity Needs Of The Next Decade," The NCTA Cable Show Spring Technical Forum 2014.

[EMM1] M. Emmendorfer, T. Cloonan, and J. Ulm, "A Side-by-side Comparison of Centralized vs. Distributed Access Architectures," The NCTA Cable Show Spring Technical Forum 2014.

[ULM2] J. Ulm, A. Eshet, and N. Radian, "Unmanaged ABR: How To Control Those Unruly Teenagers," SCTE Cable-Tec Expo 2013.

Premium Cable Content on Any Device, In Any Room Using DLNA CVP-2

Amol Bhagwat, CableLabs

Abstract

The paper first introduces key requirements and challenges for delivering premium cable content on customer owned and managed devices in the home. The paper then provides the details of a multi-industry effort in Digital Living Network Alliance for defining the Commercial Video Profile-2 specifications that enable distribution of premium content over home networks to various CE devices such as TVs, tablets, mobile phones, game consoles, HDMI sticks, and PCs. Benefits of the DLNA CVP-2 specifications to consumers, CE manufactures and service providers are discussed. The paper also presents an overview of the architecture and components of open source implementations of the DLNA CVP-2 Server and Client specifications that are aligned with the Reference Device Kit.

INTRODUCTION

Consumers increasingly expect to consume premium entertainment content on any device of their choice, at any time and at any place (inside or outside of the home). Cable operators are interested in leveraging IP-based home networks to deliver their QAM-sourced content from an in-home video gateway to devices of customers' choice in any room within the home. This serves as a transition strategy for providing content to IP-based devices in the home before migrating to End-To-End IP video architecture. Such a solution would also help U.S. cable operators to fulfill their regulatory obligation of enabling a standards-based IP output from their high-definition set-tops (HD-STBs) to serve retail Consumer Electronics (CE) devices as a part of Federal Communications Commission's (FCC) CableCARD™ order #3 [1].

This paper first discusses cable operator requirements that customer-owned-and-managed (COAM) retail CE devices need to support in order to receive premium cable content over the home network. These requirements include a consistent cable user experience across all devices, support for cable regulatory and contractual services (e.g., closed captions, parental controls, descriptive video services, and Emergency Alert System messages), content protection and quality of service. Consumer Electronics (CE) manufacturers and service providers from all over the world, including north American cable operators, led an effort to define Commercial Video Profile-2 (CVP-2) specifications within Digital Living Network Alliance (DLNA) to enable premium content to retail CE devices [2].

This paper presents a detailed overview of the CVP-2 specifications that include features such as HTML5 Remote User Interface (RUI), Authentication, Diagnostics, Low Power, MPEG-DASH, and DTCP-IP [3]. Benefits offered by CVP-2 to consumers, OEM manufacturers and service providers are also discussed. To support market adoption and implementation of CVP-2, CableLabs has developed an open source implementation of CVP-2 Server and Client reference devices [4] using the same set of base libraries as used in Reference Device Kit (RDK) [5]. The Server and Client reference devices serve as reference platforms for CE device manufacturers and cable operators to test their CVP-2 implementations. An overview of the architecture and various components of CVP-2 Server and Client implementations is also provided.

REQUIREMENTS FOR ENABLING PREMIUM CONTENT ON COAM DEVICES

Traditionally, the model for delivering cable content (or any subscription TV content) has been to supply a cable operator specified set-top box (STB) for each display device in the home. The STB ensures a consistent navigation experience using a cable operator guide and video playback for different video formats and bit rates. Using such a device, cable operators are able to fulfill their contractual agreements with content providers in terms of support for features such as content security, ad insertion, and enhanced TV (ETV) applications including Video on Demand (VoD) and television commerce (T-commerce) transactions. In addition, the STB, in conjunction with a cable operator guide, supports cable operators' regulatory compliance requirements such as providing closed captions, parental controls, Emergency Alert System (EAS) messages, Secondary Audio Programming (SAP), and Descriptive Video Services (DVS). Thus, in order to serve cable operator premium content from a video gateway over the home network to COAM CE devices in any room, without requiring a cable operator STB, COAM devices need to be able to fulfill the aforementioned functional requirements. The subsequent subsection provides detail overview of the key requirements for COAM devices.

Application Framework

COAM devices need to support an application framework that allows cable operators to serve their guide to COAM devices, either from a video gateway in the home or directly from the cloud. This application framework needs to enable playback of video inside the user interface. Support for cable operator contractual and regulatory services (e.g., closed captions, parental control, EAS, SAP, and ad insertion) needs to be supported by this application framework. Information about these services for cable content is carried in-band as elementary streams of the MPEG-2 transport streams (TS). So, the application framework needs to support mapping of these elementary streams to the application layer. In order to enable rapid application development cycle, the application framework needs to support a "write once and run anywhere" model.

Content Protection

As premium cable content is streamed from a video gateway over the home network to COAM devices, the content needs to be protected to prevent unauthorized copying. Thus, a video gateway and COAM devices need to support a content protection solution that is acceptable to the content owners.

Media Formats

In order to support a full set of cable services such as live/linear content, Video On Demand (VoD), Digital Video Recorder (DVR), and Pay-Per-View (PPV), COAM devices need to support an appropriate set of audio and video codecs with specific resolution, bit rate, and frame rate. QAM-sourced cable content predominantly uses MPEG-2 video encapsulated in MPEG-2 TS, and H.264/AVC in MPEG-2 TS to a lesser degree. In addition, support for adaptive bit rate streaming needs to be considered as cable operators may have a need to stream video over Wi-Fi networks to portable devices.

Network Quality of Service

Cable operators, as well as content providers, want to ensure that their services are offered with the highest quality when the content is streamed over the home network from video gateway to COAM devices. Thus, it is necessary to avoid congestion or interference of home network traffic that could degrade the quality of user experience.

Therefore it is necessary to consider that video gateway and COAM devices support a home network technology with throughput in excess of 100 Mbps (enough to support 3 MPEG-2 video HD streams). In addition, support for either priority-based or parameterized quality of service (QoS) needs to be considered.

Home Network Diagnostics & Management

As premium cable content is streamed over the home network from a video gateway to COAM devices, cable operators need a mechanism to diagnose and troubleshoot home network related issues remotely. Such a mechanism needs to support the ability to test the home network's connectivity between a video gateway and COAM devices, provide network topology, and information about network throughput. In addition, the ability to query information about COAM devices such as device model, manufacture, and, firmware version needs to be enabled by this mechanism.

Energy Save Operation

In order to meet consumer expectations and cable operator requirements for energy consumption, cable operator STBs and gateways implement energy saving operations, including various types of sleep modes. To avoid a consumer having to explicitly wake up the video gateway when the consumer wants to watch cable content on a COAM device, it is necessary that the COAM device is able to wake up the video gateway from sleep mode.

Device Authentication

When a COAM device requests a service from a cable operator video gateway, the cable operator wants to ensure that the COAM device meets aforementioned requirements so that cable operator guide, video and related services are presented correctly to consumers. Thus, it is necessary that the COAM device supports a secure mechanism for device authentication based on digital credentials.

SOLUTION: DLNA COMMERCIAL VIDEO PROFILE (CVP-2)

To enable secure distribution of premium content from an in-home video gateway to COAM CE devices, major cable operators in the U.S. (Comcast, Cox, and TWC) and CableLabs led an effort, in partnership with CE manufacturers and other service providers all over the world, to define Commercial Video Profile-2 (CVP-2) specifications within Digital Living Network Alliance (DLNA) [2].

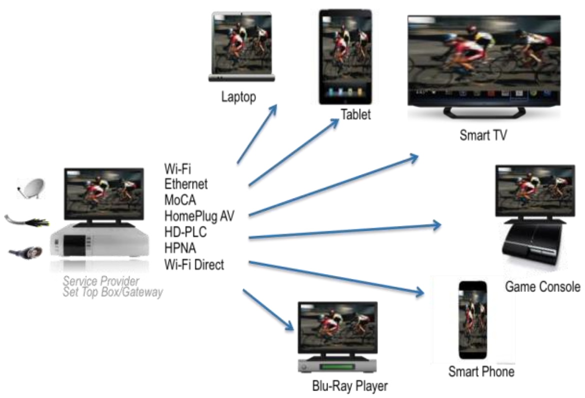


Figure-1: DLNA CVP-2 Overview

Using CVP-2 specifications, cable operators can stream various cable services (e.g., live/linear VoD, DVR, PPV) from a video gateway to COAM CE devices, such as TVs, game consoles, tablets, mobile phones, and laptops, with a consistent cable operator user interface across different devices without the need of a dedicated cable operator supplied STB per device.

DLNA CVP-2 specifications guidelines were published in March 2014 and the certification program is scheduled to be launched in October 2014 [2].

The DLNA CVP-2 Specifications define the following set of features for CVP-2 Server and Client [3]:

- HTML5 Remote User Interface (RUI)
- MPEG-2 and AVC media formats
- DTCP-IP Link Protection
- Diagnostics
- Low Power
- Authentication
- 3D Media formats; conditionally mandatory
- HTTP Adaptive Delivery; mandatory for Client, optional for Server
- Priority-based QoS
- Digital Media Server (DMS); mandatory for Server only
- Digital Media Player & Digital Media Renderer; mandatory for Client only

The following sub-sections describe various features defined in the CVP-2 specifications.

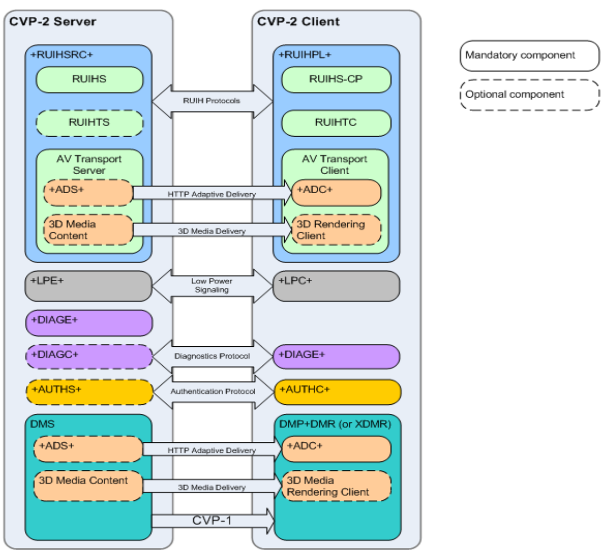


Figure-2: DLNA CVP-2 Architecture

<u>HTML5 Remote User Interface (RUI)</u>

In order to support consistent cable operator user interface to different form factors of COAM CE devices (e.g., TVs, tablets, mobile phones, and, game consoles) and requirements identified in the Application Framework subsection, DLNA CVP-2 specifications specify support for an HTML5-based Remote User Interface. DLNA HTML5 Remote RUI specification defines a profile of W3C's HTML5 specification [6] and other related specifications such as Cascading Style Sheets (CSS), Web Sockets, XMLHTTPRequest (Ajax), and FullScreen.

HTML5 is a widely adopted industry standard supported by a broad range of browsers on a wide variety of devices. Thus, it enables cable operators to develop their guide once and offer it on a wide range of platforms resulting in reduced development costs and faster time to market for new services/applications. It also enables cable operators to offer their guides directly from the cloud, thereby enabling them to rapidly evolve their services and applications to consumers.

A cable operator video gateway advertises that the Uniform Resource Locator (URL) of the cable operator HTML5 guide and CVP-2 devices discover the URL using the UPnP RUI Discovery mechanism [7]. Cable operator's HTML5 guide can be served either from the in-home video gateway or from the cloud. Using the <video> tag defined in the HTML5 specification, cable operators are able to display video within their guide user interface pages. DLNA HTML5 RUI Specification defines DLNA specific extensions to support playback Digital Transmission Copy Protection (DTCP) over IP link protected video content using <video> tag. In addition, the DLNA HTML5 RUI specification defines extensions to HTML5 <video> tag to support time-based seek and playspeed trick modes so that a consumer is able to pause, rewind and forward the video from the HTML5 guide page.

CableLabs developed a specification [8] that defines a standardized mechanism for exposing information about cable operator regulatory and contractual services, such as closed captions, content advisories, SAP, DVS, and ad insertion carried in the MPEG-2 TS video stream as HTML5 audio, video and text tracks, so that cable operator HTML5 applications can provide these services to consumers. DLNA HTML5 RUI requires implementation of this specification, so that cable operators can fulfill their regulatory and contractual obligations while offering cable services to CVP-2 devices. DLNA HTML5 RUI Specification also requires support for W3C's Server Sent Events (SSE) specification [9]. Using SSE, cable operators are able to provide EAS messages to cable operator HTML5 RUI applications running on CVP-2 devices. Figure-3 shows various HTML5 RUI entities and their functions.

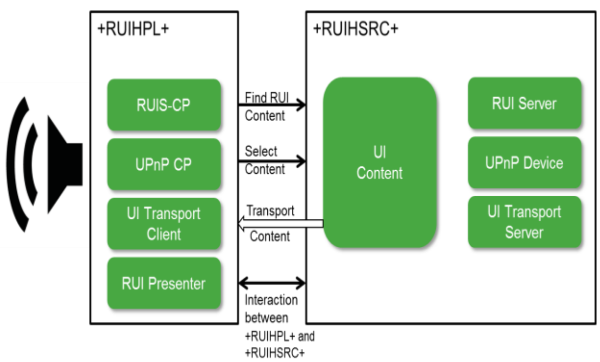


Figure-3: CVP-2 HTML5 RUI Usage Model

HTML5 RUI (RUI-H) Source capability (+RUIHSRC+) has the role of exposing and sourcing RUI-H content and includes RUI-H Server (RUIHS), RUI-H Transport Server, and an optional DLNA Media Transport Server (for serving media content):

- RUIHS provides UPnP RUI Server device functionality, which enables CVP-2 Servers to offer one or more remote UIs based on HTML5, and to handle UPnP RUI Server service actions.

- RUI-H Transport Server and RUI-H Transport Client are the device functions for transport of the RUI-H content between a client and server.

- RUI-H Pull Controller (+RUIHPL+) has the role of finding and loading RUI-H content that is exposed by a +RUIHSRC+ capability, rendering the UI content, and interacting with it. RUI-H Pull Controller includes RUI-H Server Control Point (RUIHS-CP), RUI-H Transport Client, RUI-H User Agent and an optional DLNA Media Transport Client.

- RUIHS-CP is a controller for browsing and selecting an HTML5 remote UI offered by a RUI-H Server.

- RUI-H User Agent functionality on a RUI-H Client is responsible for retrieving, decoding, presenting and interacting with the RUI-H content received from the RUI-H Server.

<u>MPEG-2/AVC Media Formats:</u>

In order to ensure baseline interoperability between the CVP-2 Server and the CVP-2 Client, the DLNA CVP-2 specifications define a required set of Media Format profiles for both CVP-2 Server and Client for a particular geographic region (e.g., North America, Europe). This set of media format profiles is representative of premium content sourced by service providers in that particular region.

MPEG-2, as well as AVC/H.264 video encapsulated in MPEG-2 TS with resolutions up to 1080p is required. Support for audio codecs such as AC-3, E-AC-3, AAC, MP3, and MPEG Layer-1 & 2 is required as a part of this media format profile set. Additionally, AVC video encapsulated in MP4 containers needs to be supported to enable interoperability with portable devices. CVP-2 Server and Client devices are also required to support DLNA specified trick modes (byte seek, time seek and playspeed) and DTCP-IP link protection for this set of media format profiles. Due to this mandatory set of media format profiles, as long as cable operators offer their content using one of the media format profiles from the CVP-2 server implemented in the video gateway, a CVP-2

Client device will be able to play back the content over the home network.

DTCP-IP Link Protection:

In order to meet content provider expectations and requirements, DLNA CVP-2 specifications leverage Digital Transmission Content Protection over Internet Protocol (DTCP-IP) Link Layer protection technology to secure content from unauthorized copying and misuse within the home as it is streamed from a cable operator video gateway to a CVP-2 client device. DTCP-IP is a link protection specification published by Digital Transmission License Administrator [10].

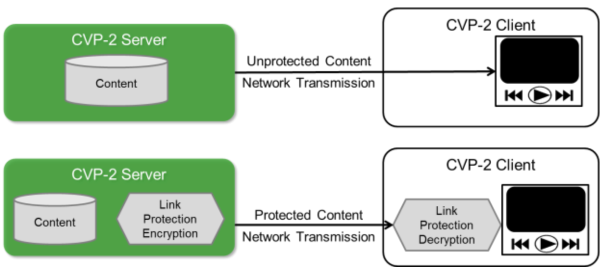


Figure-4: Secure content transmission using DTCP-IP

This is a critical enabler for multi-device viewing experiences involving premium subscription TV content. DTCP-IP is automatically negotiated between devices and has been designed to protect content as it moves across the local home network. In accordance with the CVP-2 specifications, digital content can be shared securely between products in a user's home, but not with third parties outside the home network.

Diagnostics:

The DLNA CVP-2 Diagnostics feature focuses on the collection of data about the home network conditions and devices through a set of actions and queries, so that a cable operator or a user can take appropriate steps to troubleshoot and diagnose service-related issues. The CVP-2 diagnostics feature relies on UPnP Device Management [11] as a required functionality, and IEEE 1905.1 [12] as an optional functionality. UPnP Device Management provides the ability to collect layer-3 & layer-4 diagnostics information such as IP-connectivity, network bandwidth, device information, and device status. IEEE P1905.1 provides layer-2 diagnostics information such as layer-2 link information, status, and layer-2 topology information.

Figure-5 shows various DLNA Diagnostics logical entities and their functions.

- A Diagnostics Endpoint (+DIAGE+) capability has the role of offering diagnostics services and responding to diagnostics action requests by implementing UPnP Basic Management Service v2 [13] as a required service and UPnP Configuration Management Service v2 [14] as an optional service. DLNA CVP-2 Specifications requires certain actions to be implemented, such as Ping, Trace Route, and NSLookup. Both the CVP-2 Servers and Clients are required to support diagnostic Endpoint capability.

- Diagnostics Controller (+DIAGC+) has the role of providing a diagnostics application and a control point for issuing action requests to a +DIAGE+. However, a Diagnostics Controller is optional for CVP-2 device profiles, although it is expected that a Diagnostics Controller may be included on a CVP-2 server to allow the service provider's support staff to diagnose issues within the consumer's home. The diagnostics application drives the Diagnostics Controller to access diagnostics data and capabilities. Cable operators remotely access the diagnostics application running on the CVP-2 server using a TR-069 or SNMP management interface. Alternatively, a cable operator technician or end-user may access the diagnostics application through a browser or screen interface as shown in Figure-5.

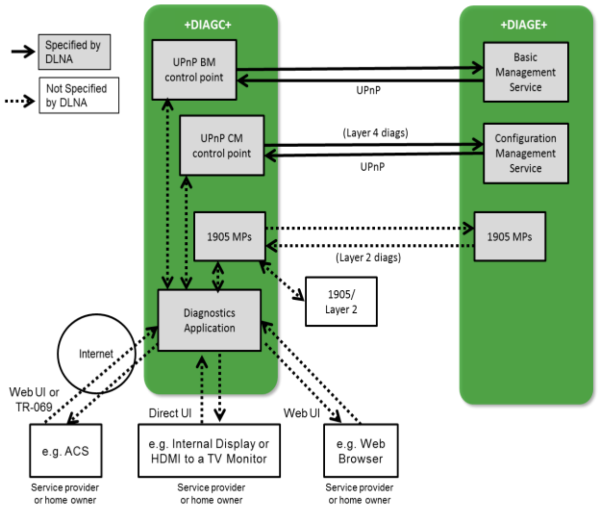


Figure-5: CVP-2 Diagnostics Architecture

<ins>Low Power:</ins>

To account for service provider STB/video gateway devices implementing energy saving operations, e.g., different levels of sleep modes, the DLNA CVP-2 specifications provide wake-up or reservation mechanisms to CVP-2 client devices. The specifications enable DLNA devices to convey energy management and sleep-mode capabilities for each of its network interfaces. This facilitates the awareness of the availability of DLNA functionality, even in the presence of power-saving mode operations. The CVP-2 Low Power feature is based on the UPnP Energy Management Service [15].

Power savings is modular within a physical device. In the context of DLNA networked devices, as shown in Figure-6, each physical network interface can have various power modes. Some of these power modes can allow layer-2 or layer-3 connectivity to still be present even when many other device components are powered down. Other physical components, such as screens, hard drives and similar resources, can also support different power modes.

The CVP-2 Low Power feature consists of the following entities:

- Low Power Endpoint (+LPE+) capability implements UPnP Energy Management Service and has the role of responding to action requests, including requests to provide information on network interface mode, and requests to access services based on subscriptions.

- Low Power Controller (+LPC+) capability implements a control point for the UPnP Energy Management Service and has the role of issuing action requests to a Low Power Endpoint or a Low Power Proxy.

The CVP-2 Server is required to implement Low Power Endpoint (+LPE+) capability, and the CVP-2 Client is required to implement Low Power Controller (+LPC+) capability. This enables CVP-2 Clients to query information about power save mode operations of a service provider's CVP-2 Server and invoke appropriate actions to wake-up the CVP-2 Server when its services are needed for the consumer. Waking up a CVP-2 Server from the low-power mode can introduce some latency and longer response time, so it is expected that a CVP-2 Client provides appropriate messages to the user to provide a good user experience.

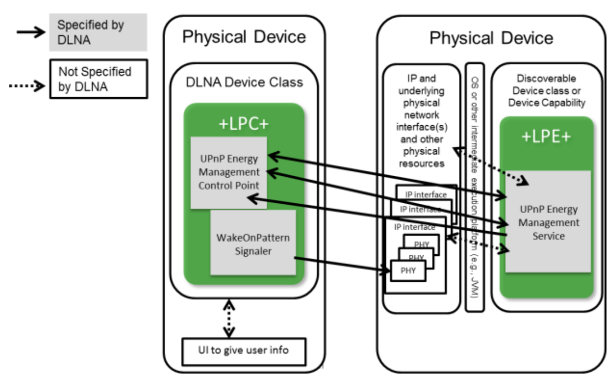


Figure-6: DLNA Low Power Architecture

<ins>HTTP Adaptive Delivery:</ins>

The HTTP Adaptive Delivery feature of CVP-2 enables service providers to describe content as adaptive content; i.e., in timed segments at various bit rates and in various media formats. In the event of network congestion, which is likely to happen over Wi-Fi, a client rendering devices can maintain smooth streaming of content for display by switching between streams at different bitrates. A Media Presentation Description (MPD) file provided by a server includes segment information such as timing, URL, and, media characteristics (e.g., video resolution and bit rates). This feature leverages Moving Picture Expert Group Dynamic Adaptive Streaming (MPEG-DASH), over HTTP (ISO/IEC 23009-1) standard [16]. Additionally, DLNA CVP-2 specifications mandate support for ISO-based media file format (ISOBMFF) Live, ISOBMFF On-Demand, and MPEG-TS Simple profiles defined in the MPEG-DASH specification.

Different logical entities of the HTTP Adaptive Delivery feature are shown in Figure-7.

CVP-2 Clients are required to support HTTP Adaptive Delivery device option and aforementioned HTTP Adaptive media format profile. Support for HTTP Adaptive delivery is optional for a CVP-2 Servers, but if it is supported, then the CVP-2 Server is required to support at least one of the HTTP Adaptive media format profiles.

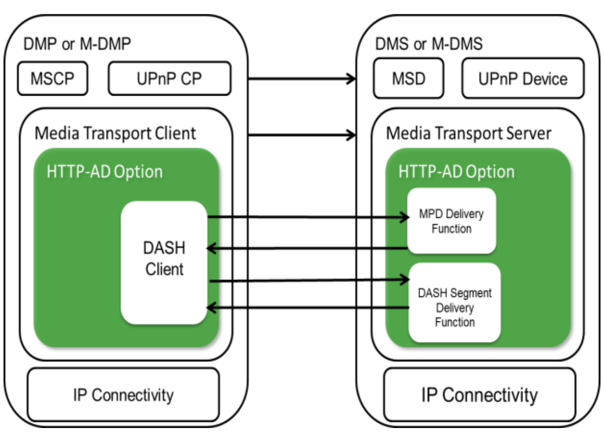


Figure-7: HTTP-Adaptive Delivery Entities

On the CVP-2 Server, the HTTP Adaptive Delivery device option has the role of exposing and sourcing content using the HTTP Adaptive Delivery mode. This includes exposing and sourcing both the MPD and the media itself (segments for different representations). This functionality maps to the MPD delivery function and segment delivery function in MPEG-DASH. On the CVP-2 client side, the HTTP Adaptive Delivery device option has the role of requesting appropriate content MPD and media representation (segments), and assembling and rendering the media while adapting to changing network conditions.

<u>Authentication:</u>

By utilizing the CVP-2 Authentication feature, service providers can verify that the CVP-2 client has been certified to the DLNA CVP-2 specifications. This provides confidence to service providers that a CVP-2 Client is able to display their HTML5 RUI guide, meet regulatory requirements, and deliver content services appropriately to meet consumer expectations.

The CVP-2 authentication feature also supports authentication of a CVP-2 Server by a CVP-2 Client. A CVP-2 Client can optionally authenticate a CVP-2 Server to ensure that the Client is talking to a legitimate CVP-2 Server to protect consumers from rogue servers.

Upon DLNA certification of a CVP-2 device (Client or Server), a device manufacturer obtains a DTLA CVP-2 Certificate, which has the same format as the legacy DTLA DTCP certificate used for DTCP-IP link protection, except that it has a special field that indicates the device is DLNA CVP-2 certified. The same certificate is used by the device for CVP-2 device authentication as well as for DTCP-IP link protection. This avoids including additional certificates in the device and saves cost for the device manufacturer. If a service provider authentication server is located in the cloud, then it obtains a CVP-2 X.509 certificate from DTLA.

DLNA CVP-2 Authentication uses Transport Layer Security Supplemental Data (TLS-SD) extensions, defined in RFC 4680 [17], to carry CVP-2 client's DTLA CVP-2 certificate over Hypertext Transfer Protocol over Transport Layer Security (HTTPS). Standard Transport Layer Security [18] protocol only supports transport of X.509 certificates. A TLS-SD extension allows transport of arbitrary pieces of information over the TLS protocol.

The HTML5 RUI browser implemented by the CVP-2 Client is responsible for performing authentication using HTTPS with cable operator Authentication Server. Cable operator Authentication Server verifies that the device requesting service is a DLNA Certified CVP-2 device based on the DTCP CVP-2 certificate supplied using the DLNA CVP-2 authentication protocol.

Figure-8 shows various CVP-2 authentication logical entities:

- Client Authentication is a device option that supports client credentials and the protocols to allow a client to be authenticated by an Authentication Server.

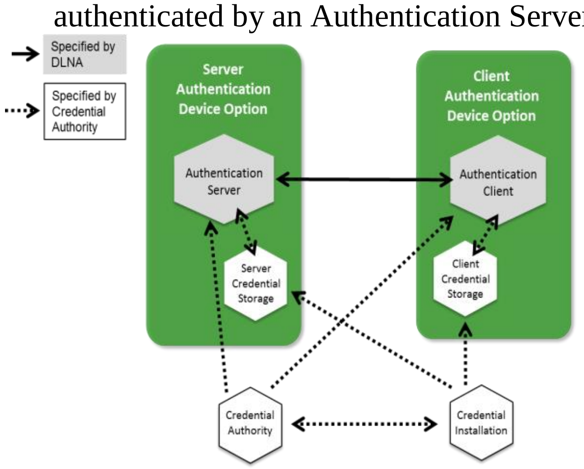


Figure-8: CVP-2 Authentication Entities

- Server Authentication is a device option that supports server credentials and the protocols to allow a server to be authenticated by an Authentication Client.

The DLNA CVP-2 Authentication supports two different scenarios for the Client/Server Authentication:

1. In the first scenario, shown in Figure-9, the Authentication Server is in the cloud and authentication must be accomplished with a cloud-based server. In this scenario, the server uses trusted X.509 CVP-2 certificate and client uses DTLA CVP-2 certificate.

2. The second scenario is shown in Figure-10, where the Authentication Server is located in the home (in a video gateway/STB) and all authentication protocol exchanges are performed within the home network. In this scenario, the server uses trusted or self-signed X.509 certificate signed with DTLA CVP-2 certificate, and client uses DTLA CVP-2 certificate.

<u>Other CVP-2 Features:</u>

- **Digital Media Server (DMS)**: CVP-2 Server is required to support DLNA DMS device class. This provides essential functions of device discovery, content streaming with support for trick modes (pause, rewind, forward).

- **Digital Media Player (DMP)/Digital Media Renderer (DMR):** CVP-2 Client is required to implement DLNA DMP and DMR device classes. These provide essential functionality for content streaming with support for trick modes. DMR provides device discovery and "Play To" scenario where a phone or tablet can establish and control content streaming between a DMS and DMR.

- **Priority-based Quality of Service (QoS):** DLNA CVP-2 requires prioritized QoS solution where video streams are given a higher priority over data/background traffic over the home network. The majority, if not all, of home networking technologies (e.g., Ethernet, Wi-Fi, MoCA, HomePNA, and HomePlug) support traffic prioritization when packets are marked with layer-2 802.1 p/q tags. The CVP-2 Server is required to mark video packets with diffserv codepoints (DSCP), as well as with layer-2 802.1 p/q tags, so that video traffic receives appropriate priority when streamed over the home network.

- **3D Media Formats:** DLNA CVP-2 specifications conditionally mandate support for 3D media formats for CVP-2 Clients and Servers. DLNA has defined a set of frame-compatible stereoscopic-3D media formats (Side-by-Side and Top-and-Bottom), which are representative of content supplied by service providers. If the CVP-2 client supports rendering of 3D video, then it is required to implement support for these DLNA defined 3D media formats.

CVP-2 DEPLOYMENT SCENARIOS

The DLNA CVP-2 Specifications support two deployment scenarios: Hybrid In-Home + Cloud scenario, and In-home only scenario.

Hybrid In-Home + Cloud Scenario:

In the hybrid In-home+Cloud Scenario, the cable operator's HTML5 RUI server and authentication server reside in the cloud, but all other functions of CVP-2 server reside on an in-home video gateway or STB. A CVP-2 Client discovers URL of the cable operator's cloud guide from an in-home CVP-2 gateway/STB. The CVP-2 Client is authenticated with a cloud Authentication Server, which may be co-located with the cloud RUI server (server uses trusted X.509 CVP-2 certificate). Upon authentication, the CVP-2 Client downloads cable operator HTML5 guide from the cloud. The HTML5 guide has links to video content that point to the in-home gateway/STB. Thus, actual video content is served from in-home gateway/STB to the CVP-2 Client.

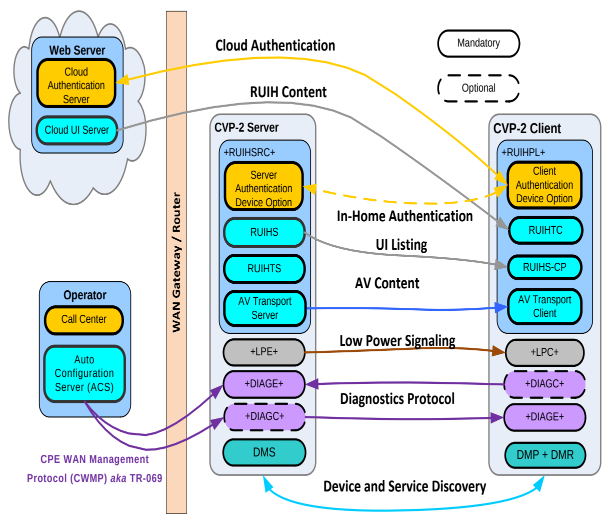


Figure-9: Hybrid In-home + Cloud Deployment

In-Home Only Scenario:

In the In-home only deployment scenario, the cable operator's HTML5 RUI server and Authentication Server reside in the in-home gateway/STB along with all other CVP-2 Server functions. A CVP-2 Client discovers URL of the cable operator's guide from an in-home CVP-2 gateway/STB, which is served from within the home from the same gateway/STB. The same gateway/STB also hosts the Authentication Server.

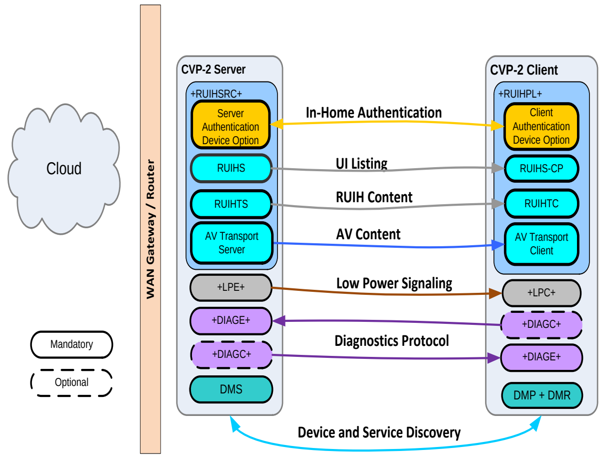


Figure-10: In-home only Deployment

The CVP-2 Client is authenticated with the in-home Authentication Server (the server uses self- signed or trusted X.509 certificate signed with CVP-2 certificate). Upon authentication, the CVP-2 Client downloads cable operator HTML5 guide to access content services from the in-home gateway/STB CVP-2 Server.

OPEN SOURCE IMPLEMENTATIONS

CableLabs, in partnership with industry participants such as Intel and ARM, has developed open source implementations of CVP-2 Server and Client [4]. These implementations are aligned with libraries used by Reference Device Kit (RDK), an integrated software platform initiative for cable operator customer premise equipment (CPE) led by major cable operators in the U.S. and Europe [5]. The main objectives for the CVP-2 open source implementation efforts are as follows:

- Provide reference devices to DLNA to help launch CVP-2 certification program
- Provide reference devices to the industry for testing and development of CVP-2 products
- Foster CVP-2 adoption and speed time to market

The following sub-sections provide an overview of open source implementations of CVP-2 server and client.

CVP-2 Server:

CableLabs' open source CVP-2 Server implementation is Linux-based and built on the same libraries as RDK (e.g., gUPnP, OpenSSL, etc.), so that it can potentially be integrated in the RDK. Figure-11 shows various components used by the CVP-2 server. Rygel, an open source DLNA DMS implementation provided as a part of the Gnome project [19], is the foundation for the CVP-2 Server. CableLabs developed CVP-2 specific extensions to Rygel such as UPnP RUI Service, Diagnostics, and Low Power. As a part of this project, CableLabs also provided upstream contributions to the OpenSSL library to include support for TLS-SD to enable carriage of DTLA CVP-2 Certificate over HTTPS for the purposes of CVP-2 authentication. The CVP-2 Server implementation also supports an Apache-based web server for delivery of HTTP-Adaptive content and serving of HTML5 RUI content.

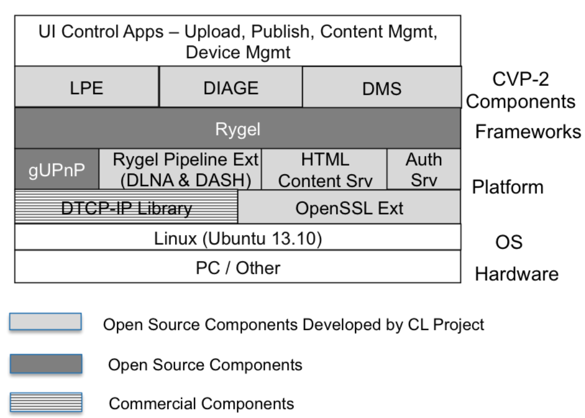


Figure-11: CVP-2 Server Implementation Components

As shown in Figure-11, CVP-2 Server implementation uses a commercial DTCP-IP library. Due to DTCP-IP requirements and content protection/encryption needs, open source implementation of DTCP-IP is not available. Thus, DTCP-IP stack is not part of the open source distribution of CVP-2 Server. However, the open source implementation of CVP-2 Server provides appropriate APIs so that vendors can include their own DTCP-IP library to have a complete CVP-2 Server solution that includes DTCP-IP support. CableLabs' CVP-2 server implementation is available at [4].

CVP-2 Client:

CableLabs' open source CVP-2 Client is Linux-based and built on the same libraries as RDK (such as gUPnP, Gstreamer, and WebKit). The CVP-2 Client also uses Rygel to provide basic DLNA functionality of Digital Media Renderer. It uses dLyena [20], which is an open source project by Intel that provides UPnP control point functionality for implementation of DLNA Digital Media Controller (DMC) and DMP. CableLabs, in partnership with Intel, developed extensions to dLeyna to support CVP-2 specified features such as Low Power Controller and Diagnostics Controller. Figure-12 shows the CVP-2 Client different software components.

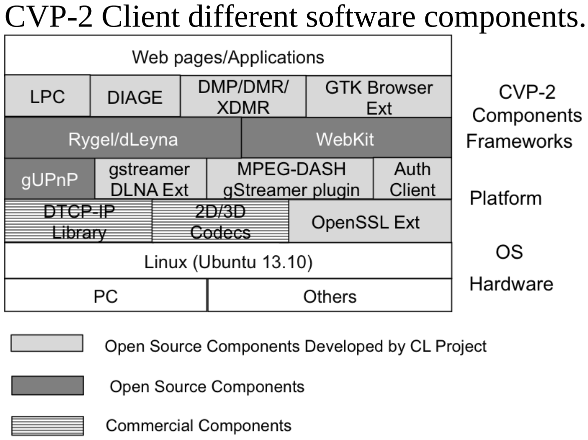


Figure-12: CVP-2 Client Implementation Components

HTML5 RUI browser implementation that meets DLNA CVP-2 requirements is the key component of the CVP-2 Client implementation. Figure-13 shows the details of the HTML5 RUI browser implementation that is part of CableLabs' CVP-2 Client. The HTML5 RUI browser uses WebKit [21] and GTK [22] at its core. CableLabs contributed extensions to WebKit to support DLNA HTML5 RUI requirements such as mapping of MPEG-2 TS elementary streams carrying TV services such as closed captions, SAP, ETV, and ad insertion.

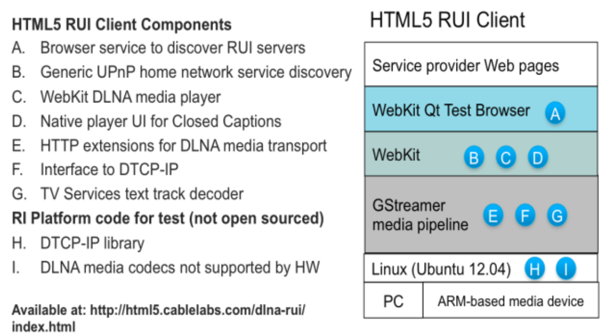


Figure-13: HTML5 Browser Implementation Components

The CVP-2 Client (both HTML5 RUI browser as well as DMP/DMR components) use Gstreamer [23] for media pipeline. CableLabs developed extensions to Gstreamer 1.0 to include support for DLNA streaming with trick modes and DTCP-IP support. CableLabs also developed a Gstreamer plug-in that supports playback for HTTP-Adaptive (MPEG-DASH) content. The CVP-2 client implementation also includes Authentication client, based on OpenSSL [24], that supports client authentication using DTLA CVP-2 certificates. As in the case for CVP-2 Server, commercially available DTCP-IP library is part of the CVP-2 Client implementation and is not part of the open source distribution. The CVP-2 client supports APIs, however, for vendors to plug in their own DTCP-IP library. Similarly, the CVP-2 Client uses commercially available 2D & 3D codecs to include support for required DLNA media formats. CableLabs' CVP-2 Client implementation is available at [4].

CVP-2 BENEFITS

The CVP-2 specifications offer benefits to consumers, service providers, and CE manufacturers alike. Consumers will be able to consume premium subscription TV content such as live/linear HD programs, VoD, and DVR content on devices of their choice anywhere in the home without having to download different platform-specific applications or obtain additional equipment from service providers. A consistent user experience is deliverd across all devices. Using CVP-2, service providers are able to offer all their services to a wide variety of COAM CE devices by maintaining their own user experience on all the devices. Using CVP-2 HTML5 RUI, service providers are able to evolve their services more rapidly and reduce time-to-market for new services and products. Auto service discovery feature supported by CVP-2 facilitates easy installation and setup, which is a benefit to both consumers and service providers. The Diagnostics feature allows service providers to remotely diagnose and troubleshoot any service related issues.

CVP-2 authentication provides assurance to service providers and content providers that only certified CVP-2 devices access their services and provides assurance for their user experience on CE devices. CVP-2 offers a single, interoperable solution to CE manufacturers to enable premium subscription TV services from different service providers. This avoids having to develop one-off solutions and platform specific development, which is likely to speed time-to-market for their products.

CONCLUSIONS

The industry has been investigating solutions for enabling premium cable content to retail CE devices in a manner that meets cable operators' contractual, regulatory and business requirements for a long period of time. The DLNA CVP-2 specifications, with its key constituent features such as HTML5 RUI, Device Authentication, a well-defined set of media formats and diagnostics, offers a solution that meets these requirements. It allows cable operators and other subscription TV providers the ability to offer all their subscription TV services with consistent user experience to retail CE devices with different form factors and platforms, without the need for providing additional STBs.

CVP-2 offers consumers more choices in consuming their subscription TV content. With CVP-2, CE manufacturers need to only implement a single unified platform to receive subscription TV content from different service providers. Open source implementations of CVP-2 Server and Client are available to the industry for development and testing of CVP-2 products.

ACKNOWLEDMENTS

The author thanks Ken Barringer, Joan Branham, Jud Cary, Christie Poland, Darshak Thakore, and Eric Winkelman for their review and contributions to the paper.

REFERENCES

[1] Federal Communication Commission's Third CableCARD Report & Order, October 14, 2010 http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-181A1.pdf

[2] DLNA CVP-2 Press Release, March 18, 2014, http://www.dlna.org/docs/default-source/press-releases/the-digital-living-network-alliance-releases-cvp-2-guidelines-for-viewing-subscription-tv-content-on-multiple-home-devices.pdf?sfvrsn=4

[3] DLNA Guidelines, http://www.dlna.org/dlna-for-industry/technical-overview/guidelines

[4] Open Source Implementations of CVP-2 Server and Client, CableLabs, http://html5.cablelabs.com/dlna-cvp-2/index.html

[5] Reference Device Kit, http://rdkcentral.com

[6] HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Candidate Recommendation 04 February 2014, http://www.w3.org/TR/2014/CR-html5-20140204/

[7] RemoteUIServer:1 Service Template Version 1.01, For UPnP™ Version 1.0, September, 2, 2004, http://upnp.org/specs/rui/UPnP-rui-RemoteUIServer-v1-Service.pdf

[8] Mapping from MPEG-2 Transport to HTML5, I03, CL-SP-HTML5-MAP-I03-140207, Cable Television Laboratories, Inc. Specifications, Web Technology, February, 7, 2014

[9] Server Sent Events, W3C Candidate Recommendation, 11 December 2012, http://www.w3.org/TR/eventsource/

[10] DTCP Volume 1 Supplement E, Mapping DTCP to IP, Revision 1.4 ED3, June 5, 2013, Digital Transmission License Administrator, http://www.dtcp.com/documents/dtcp/info-20130605-dtcp-v1se-ip-rev-1-4-ed3.pdf

[11] UPnP Device Management: 2, http://upnp.org/specs/dm/dm2/

[12] IEEE 1905.1, IEEE Standard for a Convergent Digital Home Network for Heterogeneous Technologies, 2013, http://standards.ieee.org/findstds/standard/1905.1-2013.html

[13] BasicManagement:2, Service Template Version 1.01, For UPnP™ Version 1.0, February 16th, 2012, http://upnp.org/specs/dm/UPnP-dm-BasicManagement-v2-Service.pdf

[14] ConfigurationManagement:2, Service Template Version 1.01, For UPnP™ Version 1.0, March 4th, 2013,http://upnp.org/specs/dm/UPnP-dm-ConfigurationManagement-v2-Service.pdf

[15] EnergeyManagement:1, Service Template Version 1.01, For UPnP™ Version 1.0, August 30, 2013, http://upnp.org/specs/lp/UPnP-lp-EnergyManagement-v1-Service.pdf

[16] ISO/IEC 23009-1:2012, Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57623

[17] S. Santesson, TLS Handshake Message for Supplemental Data, IETF RFC 4680, September 2006, http://tools.ietf.org/html/rfc4680

[18] T. Dierks, et al, The Transport Layer Security (TLS) Protocol, Version 1.2, IETF RFC 5246, August 2008, http://tools.ietf.org/html/rfc5246

[19] Gnome's Rygel Project, https://wiki.gnome.org/action/show/Projects/Rygel?action=show&redirect=Rygel

[20] dLeyna Project, Intel Open Source Technology Center, https://01.org/dleyna

[21] The WebKit Open Source Project, http://www.webkit.org

[22] The GTK+ Project, http://www.gtk.org

[23] GStreamer, Open Source Multimedia Framework, http://gstreamer.freedesktop.org

[24] The Open SSL Project, https://www.openssl.org

QoE MONITORING OF IP VIDEO SERVICE

Eshet, Amit Lutsker, Mark Cohen, Uzi
ARRIS

Abstract

The IP Video service offered by MSOs is about to enter its 2nd phase, extending from the 2nd screens to the big screen. With this transition the expectations and requirements for QoE will be going up to be at par with the legacy video QAM service.

IP Video introduces new challenges when it comes to QoE monitoring. One of the most important paradigm shifts is an extremely wide range of screens and consumption habits: with screens that are few inches in size to ultra HD TVs with 80" screen; with a laid back to fully engaged experience. To make things more interesting, in some cases the format of the content itself would dramatically vary (sub-VGA to Ultra HD) while in others the exact same content, say HD, may be viewed both on a tablet and on a big-screen TV. Other critical game changers when it comes to IP video QoE monitoring are WiFi, OTT delivery of video over best effort networks which in some cases are not even owned by the operator (e.g., OTT, off-net), and of course the fact that the decoding device itself may be CE with, at best, a limited ability of the operator to control and guarantee QoE. And of course, on top of all this, operators are rightfully looking for a single QoE Monitoring solution applicable to all screens and all use cases.

In this paper we will start by discussing the differences between QoE and QoS and between QoE and video quality. We will then compare different methodologies for video quality and QoE monitoring, including full-reference vs. reduced-reference, vs. no-reference; compressed vs. pixel domain; statistical vs. exhaustive. We will conclude with a review of alternatives for embedding QoE probes in the end-to-end IP Video architecture and their ability to collect true and effective QoE information.

QoS, QoE, and Video Quality

Quality of service (**QoS**) is the overall objective performance of a network, particularly the performance seen by the users of the network. To quantitatively measure quality of service several related aspects of the network service are often considered, such as error rates, bandwidth, throughput, transmission delay, availability, jitter, etc.

Quality of Experience (**QoE**) is a subjective measure of a customer's experiences with a service. QoE systems will try to measure metrics that customer will directly perceive as a quality parameter (e.g., channel change time). In short, QoE provides an assessment of human expectations, feelings, perceptions, cognition, and satisfaction with respect to a particular product, service or application.

QoE is related to but differs from QoS, which attempts to objectively measure the service delivered by the vendor, with QoS measurement is most of the time not related to customer, but to media (customers will never tell you : the jitter is too high). It is tied closely to the black and white of a contract and measures how well the vendor lives up to its end of the bargain.

A vendor may be living up to the terms of a contract's language, thus rating high in QoS, but, the users may be very unhappy, thus causing a low QoE. Conversely, the users may be very happy with a product or a vendor, resulting in an artificially high QoE if the vendor is not, in fact, doing what he was paid to do, thus rating low in QoS.

Finally, **subjective video quality** is a subjective characteristic of video quality. It is concerned with how video is perceived by a viewer and designates his or her opinion on a particular video sequence.

As such, although video quality is definitely part of the broad definition of Video QoE, it is definitely only a sub-set. Channel change time, number of black frames in the transition between content and an ad, the contribution of the device, light conditions, and distance of viewing are just a subset of the attributes of QoE that are not related to the video quality itself.

MEASURING QoE AND VIDEO QUALITY

Full, Reduced, and No Reference

There are three basic schemes for measurement of video quality, Full Reference, Reduced Reference, and No Reference. When looking at a function or a sub-system of the network that introduces degradation to the video signal, video quality at the output of the subsystem can be measured as follows:

- Full reference involves comparing the video signal at the output of the subsystem to the uncompressed digital source.
- Reduced Reference involves comparing the video signal at the output of the subsystem to the video signal at the input to the subsystem.
- No Reference involves evaluation of the video at the output of the subsystem without using any reference.

Figure 1, Video Quality Measurements, depicts Linear IP Video delivery architecture and overlays it with video quality measurement. In this example, to use a Full Reference scheme one would need to get access to the uncompressed source video available to the content providers before it is even encoded and sent to the MSO. Assuming it was available, Full Reference can be used, as an example, for video quality measurement at the output of the IRD or of the ABR transcoder. For Reduced Reference video quality measurement the IRD or Transcoder output can be used as the reference for the video arriving to the home gateway or end device. Finally, the video quality of the video arriving to the home gateway or the end device can also be measured, on its own leveraging a No Reference scheme.

The advantage of a Full Reference scheme, such as SSIM, is that it provides the ability to separate the artifacts inherent to the original video signal from the artifacts introduced by the delivery network, including compression artifacts. A Reduced Reference scheme enables a good measurement of the degradation of the video quality of the signal passing through the subsystem. The only scheme that actually attempts to truly measure video quality and not degradation is the Non Reference scheme. As such, to truly measure video quality the above schemes need to be revised:

- No Reference applied to the uncompressed digital source plus Full Reference comparing the video signal at the output of the subsystem to the uncompressed digital source.
- No Reference for the input to the subsystem plus Reduced Reference involves comparing the video signal at the output of the subsystem to the video signal at the input to the subsystem.
- No Reference applied to the video at the output of the subsystem

One should ask, if No Reference is mandatory for video quality evaluation why not rely solely on the No Reference scheme. The challenge there is that there is no widely acceptable and standardized No Reference video quality measurement scheme. Moreover, in the MSO space, access to the original uncompressed digital source is not available either. As such, at the end of the day we are left with measuring degradation rather than absolute quality using a Reduced Reference scheme. This is "translated" to video quality under the assumption that the video signal ingested by the MSO is of perfect quality. For "true" video quality measurement one can consider using proprietary No Reference protocols.

When it comes to Reduced Reference video quality degradation measurement, MSE, and PSNR are the most common. That said, none of them was proven to achieve high correlation to human perception across a wide variety of content, a wide variety of artifacts, and a wide range of the severity of the artifacts.

.

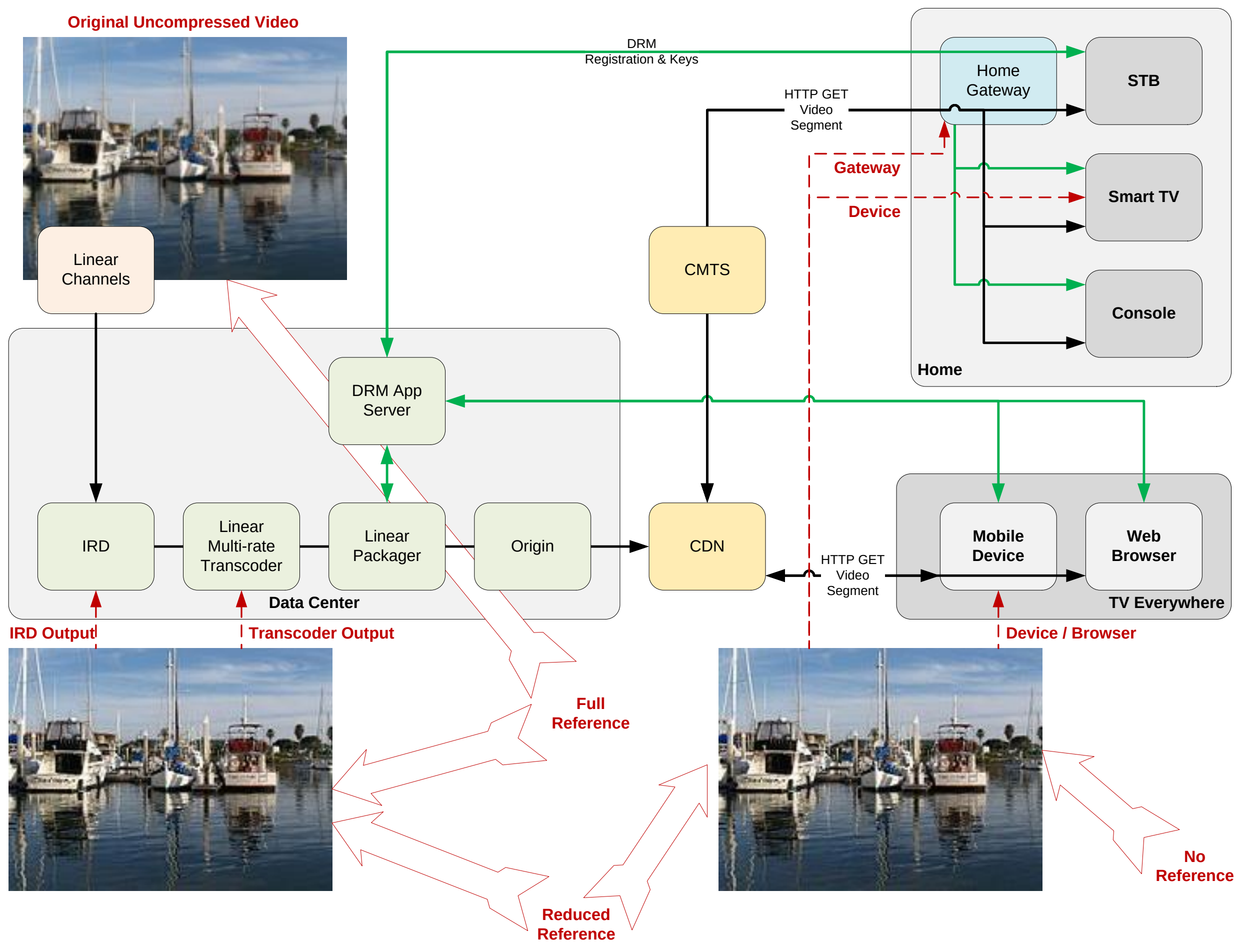


Figure 1 - Video Quality Measurement

Pixel and Compressed Domain

Video going through the MSO network is compressed, and in many cases even encrypted. No Video Quality measurement scheme exists for encrypted video, however, in most cases, no QoE degradation is expected while the video is encrypted as lossy video processing are not applied once the video is encrypted (although QoS degradation may take place, e.g., packet drop).

Video Quality estimation in the compressed domain is very attractive as it doesn't mandate extra decoding, which can pose a problem, especially with high scale. The most common compressed domain parameter used for video quality (degradation) estimation is the quantization parameters. The higher the parameters the higher the quantization noise is. Compressed domain schemes can offer a good tradeoff of performance and accuracy.

In the pixel domain multiple schemes exist. MSE, PSNR, and SSIM were mentioned before. On top of these, many proprietary schemes are leveraging techniques to identify blockiness, blurring, and noise enhancement. Moreover, spatial and temporal tools can be used to identify objects across a video stream and use that to identify artifacts. Once artifacts are identified the video quality degradation can be evaluated based on parameters such as the number of artifacts, their position, and severity.

Statistical vs. Exhaustive

Exhaustive video quality measurement implies that all the data, both spatial and temporal, is being used. Statistical approaches would analyze just a portion of the video frames, just a subset of the pixels / spatial data, or just a subset of the Chroma components (e.g., luma only). Very good results can be achieved while performing spatial decimation or by relying only on the luma component.

OVERLAYING QoE AND VIDEO QUALITY PROBES OVER THE VIDEO DELIVERY SUBSYSTEM

Figure 1 suggests key locations where video quality and QoE measurement probes can be inserted into the IP video delivery network.

The first place where MSOs impact the video quality is the IRD. Since the IRD is controlled by the content provider this is an excellent place to take reference measurement for the video quality. Any further degradation is under the responsibility of the MSO. In the case of VOD the equivalent would be to measure the video quality of the original assets.

A key place for a second probe is the output of the ABR transcoder. Reduced Reference is a very effective tool to compare the video quality at the transcoder output to that of the IRD output as both are likely to be co-located and serial.

The packager, Origin server, and CDN are not expected to generate any video quality artifacts. However, the packager may still create QoE degradation if the segmentation process is not done properly. The first time to check that would be the qualification of the packager. A real-time option involves taking the packager output, decrypting it (the packager is also used to apply DRM) and using a probe that simulates the behavior of an ABR client to check for QoE degradation (e.g., lost data at the seam, degradation at the transition between segments of different profiles).

For an on-net service involving transcoding in the home, the home gateway (or the home transcoder) can be used to measure video quality degradation. Reduced Reference can be used in case the platform is capable of measuring both input and output. If not, one option is to use No Reference scheme for the output. Another option is to extract key parameters from the output and make them available to a QoE Estimator located in the network, which may have access to the network transcoder output thus leveraging a Reduced Reference scheme.

Finally, the MSO application running on the device can be used to collect critical QoE information. This information may include the device type, the ABR profile of each segment, decryption problems, decoder buffer underrun or overrun, decoder resets, as well as other decoding problems. With this information made available to a QoE Estimator located in the network, a complete QoE picture can be made available for the MSO. To create this complete QoE picture, the QoE Estimator would cross the QoE data coming from the devices with video quality information associated with the video content coming from the network transcoder as well as potentially from the IRD. Figure 2 provides a simplified network diagram showing the network QoE Estimator and its interfaces to the various probes.

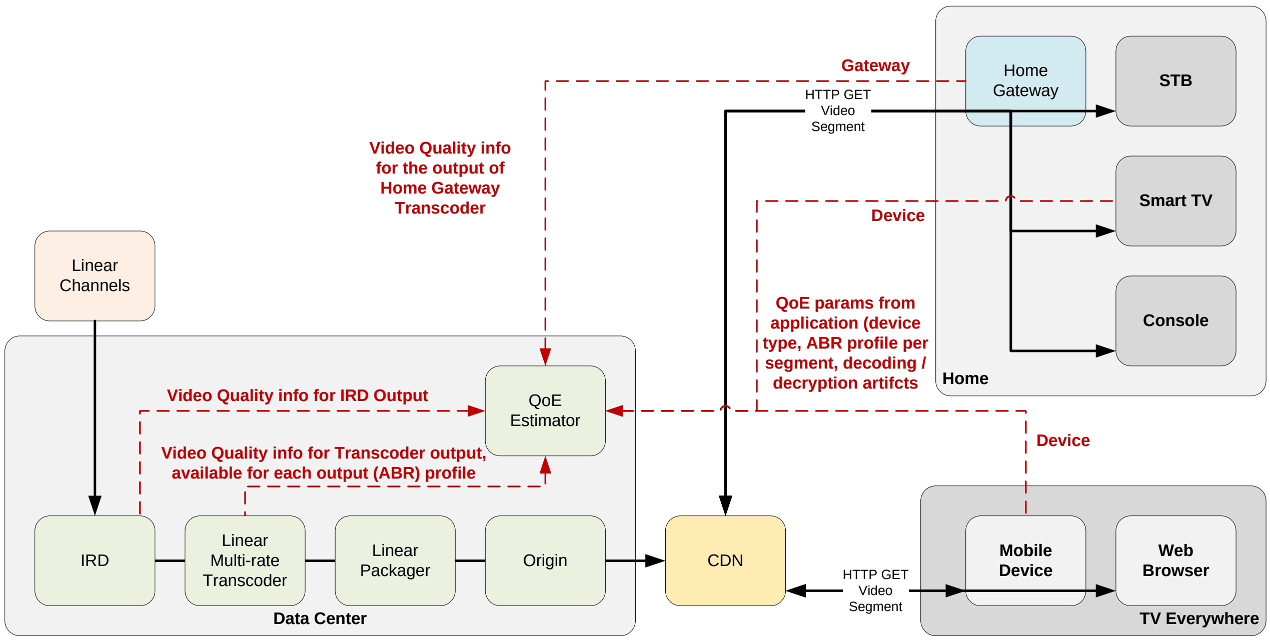


Figure 2 - QoE Estimator

SUMMARY

IP Video provides new challenges and opportunities when it comes to monitoring QoE. The uncast nature of the IP Video delivery calls for the ability to monitor each viewer independently, as QoE will vary from viewer to viewer. At the same time the IP Video delivery architecture allows for a highly effective QoE measurement solution leveraging a network QoE Estimator, taking advantage of the centralized nature of the video processing and the ability to retrieve critical QoE information from the ABR clients at minimal effort and complexity.

Remote PHY for Converged DOCSIS, Video and OOB

John T. Chapman, CTO Cable Access BU & Cisco Fellow
jchapman@cisco.com
Cisco

Abstract

Remote PHY refers to the technique of moving the PHY circuit out of a device such as a CCAP and putting the PHY circuit at the end of a network. Remote PHY builds upon the work started with Modular CMTS (M-CMTS) and Modular Headend Architecture (MHA) at CableLabs.

Remote PHY is an evolving set of specifications and products. This white paper will focus on the expanded definition and the updates to the transport and timing for Remote PHY and how they apply to DOCSIS, MPEG-TS video, and Out-of-Band signaling for STB.

TABLE OF CONTENTS

TABLE OF CONTENTS 1

INTRODUCTION 2
- Disclosure 2
- What is Remote PHY? 2
- Remote PHY Lineage 2
- Why is Digital HFC Interesting? 3
- Review of Comparable Technologies 4
  - *BDR/BDF* 4
  - *Remote PHY* 4
  - *Remote MAC* 5
  - *Remote CMTS* 5

CURRENT MHA TECHNOLOGY 5
- Overview 5
- DEPI 5
  - *D-MPT Pseudowire* 6
  - *PSP Pseudowire* 7

REMOTE PHY TECHNOLOGY 8
- Overview 8
- Specifications and Applications 8
  - *Remote PHY Pseudowire* 8
  - *UEPI* 9
  - *R-DTI* 9
  - *R-OOB* 9

NEW PROTOCOL EXTENSIONS 9
- Packet Length 9
- MPLS 9
- MCM: Multi-Channel MPT 10
- BFD 10
- DOCSIS 3.1 11
  - *DS OFDM Channel Pseudowire* 11
  - *DS PLC Channel Pseudowire* 11
  - *D3.1 Upstream Pseudowires* 11
- Video 12
- OOB 12

REMOTE PHY TIMING 13
- R-DTI – One-Way Reverse 13
- R-DTI – Two-Way Reverse 14
- R-DTI – Two-Way Forward 14
- Remote Scheduler 14
- Details 15
  - *Packet Formats* 15
  - *Support for DTP* 15
  - *Distances* 15

NEXT STEPS 15

SUMMARY 16

REFERENCES 16
  - *Related Papers by the Same Author* 16
  - *Related CableLabs Specifications* 17
  - *Related IETF Specifications* 17

INTRODUCTION

Disclosure

The ideas described in this paper are being considered to become part of a formal set of public Remote PHY specifications.

The Remote PHY specifications are still under development. The following represents the author's current thoughts on the requirements, form, and functionality of the Remote PHY protocol. Since actual bit positions within the protocol are still being finalized, they are not explicitly specified in this white paper.

However, this white paper should be a fairly accurate guide to what the Remote PHY protocol can accomplish and how it operates.

What is Remote PHY?

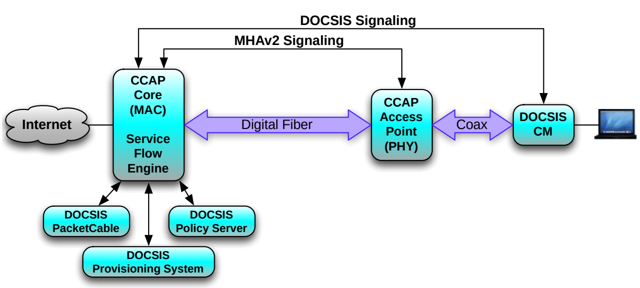


Figure 1 - Remote PHY CCAP Network

Remote PHY is an architectural strategy that removes the PHY element from a product and places that PHY element in a separate access point that is interconnected with an IP network (even simple Metro Ethernet networks or just EPONs qualify as they use IP packets). This is shown in Figure 1 and explained in the white paper [8].

Restated, Remote PHY allows you to put your main chassis at one end of a network and your PHY chip at the other end of the network. This is a useful technique when the PHY chip needs to be close to an access network, but the desire is to put the intelligence and complexity in a central location that has more room and is more serviceable.

Remote PHY infers centralized software. The least amount of complexity is placed remotely; the most amount of complexity is retained centrally.

Remote PHY-like strategies (similar in concept but different in implementation) have been used in adjacent markets such as:

- WiFi access points
- LTE access points
- Ethernet over Coax (EoC)
- EPON over Coax (EPoC)
- xDSL

Remote PHY in this context was applied initially to a DOCSIS CMTS. Remote PHY is also now being applied to traditional MPEG-TS video and to out-of-band (OOB signaling).

Remote PHY Lineage

The first instance of Remote PHY technology was the Modular CMTS (M-CMTS) specifications from CableLabs in 2005. Architecture and tutorial discussions of M-CMTS can be found in white papers [11], [12] and [13].

The specifications included:

- DEPI – DOCSIS External PHY Interface
- DTI – DOCSIS Timing Interface

- ERMI – Edge Resource Management Interface
- M-OSSI – M-CMTS Operations Support System Interface Specification

These specifications where targeted at combining CMTS and Edge QAMs (EQAM) together into one system. While the original intent was to lower the system cost, the real benefit ended up being the ability to greater customize the CMTS for cost and performance.

At the time, another specification was written but was not published, as there was no market application for it. That spec was:

- UEPI – Upstream External PHY Interface

UEPI builds upon the DEPI pseudowire concepts and applied them to the upstream CMTS MAC-PHY interface. In 2006, UEPI was used as a MAC to PHY interface for vendor silicon for I-CMTS applications.

In 2008, a second round of specifications that focused on the EQAM were published. Since the set of specifications now referred to more than just the CMTS, they were renamed as the Modular Headend Architecture (MHA) specifications. These specifications were:

- EQAM-PMI – Edge QAM Provisioning and Management Interface
- EQAM-VSI – Edge QAM Video Stream Interface
- MHA Technical Report

In 2012, the Chinese national regulatory body for the cable industry known as SARFT (State Administration of Radio, Film, and Television) adopted DEPI and UEPI as C-DOCSIS (China DOCSIS) Type III.

Also in 2012, the Remote PHY technology won a China CRTA Scientific and Technological Innovation Award, shown in Figure 2.

Figure 2 - Remote PHY Innovation Award

It is anticipated that the work described in this white paper will evolve into industry specifications as well.

Why is Digital HFC Interesting?

One of the product goals of Remote PHY is to put Remote PHY technology into an Optical Node in an HFC (Hybrid Fiber Coax) Plant. This would allow the fiber portion of the HFC plant to become digital. To date, the HFC forward path has used analog optics and the reverse path has used both analog and digital optics.

The conversion of the HFC plant from a linear optics plant to a digital optics plant would be applicable anytime the plant is to be segmented or upgraded. The most compelling case is deep fiber. Today's plant is typically N+5, which refers to an optical node plus 5 amplifiers deep. A deep fiber plant design would be N+0 or N+1, meaning an optical node plus no or one amplifier.

In a deep fiber plant, there are many more optical nodes, head end optical lasers and receivers, as well as CMTS ports to be purchased. A conversion to digital fiber may yield a better investment decision.

Digital forward and reverse paths are interesting for technical and strategic reasons. The technical reasons are:

- Longer distances (80+ km vs. 40 km)
- More wavelengths (80 vs. 16)
- Lower cost optics (based upon 10G Ethernet)
- Higher throughput (more bits per Hertz) if the DOCSIS 3.1 PHY is located after the coax segment.
- Lower maintenance costs
- Higher Reliability

If the digital fiber is also an IP network (or Metro Ethernet or EPON/GPON network), then there are also additional strategic benefits:

- Compatibility with digital access networks which are now appearing internationally
- Good scaling for deep fiber
- The same IP-based access network can be used for residential and commercial use.
- The same IP-based access network can be used to support DOCSIS (with Remote PHY) and fiber to the home (FTTH).

Review of Comparable Technologies

Table 1 - Technology Comparison

Criteria	BDR, BDF	Remote PHY	Remote MAC	Remote CMTS
IP Network	No	Yes	Yes	Yes
I-CMTS impact	Low	Low	High	Low
Remote SW	0%	5%	50%	100%
Remote HW	5%	10%	40%	100%

There are several ways to address the market need for digital HFC. These methods are shown in Table 1 and they are ranked to several basic criteria that define them. [9]

BDR/BDF

The most basic approach would be to do a baseband digital forward (BDF) in a similar manner to how baseband digital reverse (BDR) was done. This has been too costly to date to do, but may become feasible in the future. The main advantage of this approach is that it is transparent to any modulation or service that goes across it. The main disadvantage is that it maintains a single hop proprietary fiber interface. Even if BDF where to adopt a network packet format, the network traffic would be very high bandwidth and continuously on 100% of the time. It also does not allow the CMTS to Remote PHY path to traverse a generic IP network.

There are three variations on the DOCSIS CMTS theme, all of which support IP networking.

Remote PHY

Remote PHY, the subject of this white paper, puts the bare minimum hardware and software into a remote entity. Remote PHY keeps all the complexity centralized where it can be more easily scaled and maintained. The advantage of Remote PHY is its simplicity. The disadvantage is that the PHY definition has to be committed to and not changed. (Note that an overlay network could deal with changes, but the long term goal is that an overaly network should not be required).

Remote MAC

Remote MAC puts the DOCSIS PHY, the MAC, and a good amount of packet processing and software into the remote entity. Some examples are C-DOCSIS Type II and the PASI interface that was under development in CCAP but was later cancelled.

The advantage of the Remote MAC approach is that the CMTS-Core could be replaced with alternative products such as generic BRAS (Broadband Remote Access Server). The disadvantages of this approach are that there is no compatibility with an I-CMTS and the remote phy entity is more complex.

Further, to prevent truck rolls and upgrades, the remote node has to implement full spectrum DOCSIS MAC processing years before full spectrum DOCSIS is needed. It also requires the millions of lines of DOCSIS specific CMTS code to be decomposed and ported to the BRAS and MAC PHY node. This is a huge code stability concern, and an unnecessary risk.

Remote CMTS

Remote CMTS puts the entire CMTS into the node. The advantage of this approach is that the software model for the CMTS is not split in half, as is the case with Remote MAC. The disadvantages of Remote CMTS are that there are now 1000’s of CMTS to configure and the node is not a secure location that was always an assumption for BPI (Baseline Privacy).

Of these choices, this white paper would like to suggest that Remote PHY provides the optimum balance of supporting IP networking and still striving for maximum simplicity and compatibility with existing CCAP devices.

CURRENT MHA TECHNOLOGY

Overview

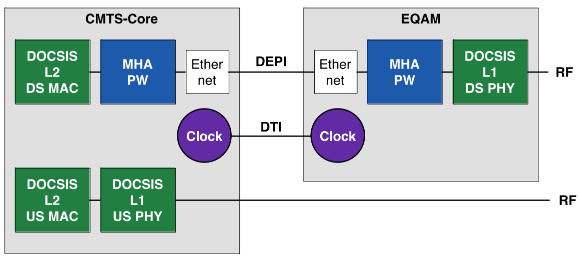


Figure 3 - MHA System Diagram

In a current MHA system, the downstream DOCSIS PHY is located externally in an EQAM while the upstream DOCSIS PHY remains in the CMTS-Core. The interface for the downstream PHY is DEPI.

DEPI

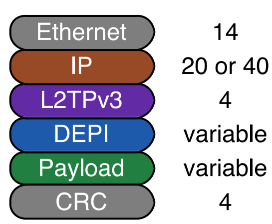


Figure 4 - DEPI Pseudowire Format

DEPI is the Downstream External PHY Interface. In an Integrated CMTS (I-CMTS) system, it can be the interface between the MAC chip and the PHY chip. In a Remote PHY system, it is the protocol interface between the MAC interface in the CCAP-Core and the PHY chip in the Remote PHY entity.

The general pseudowire format for DEPI along with a byte count is shown in Figure 4. The sub-elements of the DEPI packet are as follows:

- Ethernet header
- IPv4 or IPv6 header
- L2TPv3 header
- DEPI header
- DEPI Payload
- CRC

The first three headers are well defined by IEEE and IETF specifications. The published version of DEPI does include an optional UDP header ahead of the L2TPv3 header. This option was not used in practice and will be eliminated from the specification.

This also simplifies the L2TPv3 header. The L2TPv3 header contains a single 32-bit session ID. If the session ID is all zeros, the packet is a control plane packet. If the session ID is non-zero, then it is a data packet. The control plane will associate a session ID with a pseudowire type and sub-type. The base L2TPv3 protocol is in [20].

The format of the DEPI header is specific to DEPI. The DEPI header also specifies the format of the DEPI payload. There are two DEPI pseudowire types.

- D-MPT which is the DOCSIS MPEG Transport pseudowire
- PSP which is the Packet Streaming Protocol pseudowire

Each pseudowire can have a sub-type. Examples of sub-type are “DOCSIS 3.0 Downstream” and “DOCSIS 3.1 Downstream”. The significance is that the pseudowire type is part of the normal L2TPv3 control plane while the pseudowire sub-type is advertised in the DEPI specific extensions.

D-MPT Pseudowire

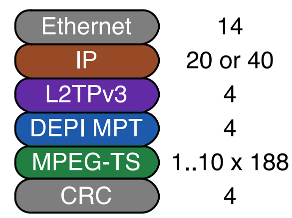


Figure 5 - D-MPT Pseudowire Type

The DOCSIS MPEG-TS pseudowire is the only pseudowire that was deployed in DOCSIS 3.0 based M-CMTS systems. As the name implies, there is a DEPI header followed by a number of 188 byte MPEG-TS packets. For M-CMTS, the DEPI packet size limited the number of MPEG-TS packets to seven.

The D-MPT header is 32 bits in size and contains the basic DEPI header that is used in all the pseudowire types. It contains:

- A V bit to permit L2TPv3 payload multiplexing within an L2TPv3 session.
- A S bit to indicate if the sequence field is valid
- Two H bits that allow DEPI payload multiplexing within a DEPI session. This is used for DLM, the DEPI Latency Measurement packet.

- Three-bit flow ID. Used for indicating the QoS of the payload.
- Segment count for PSP.
- Sequence number for the pseudowire session.

The V bit, S bit, and the sequence number are recommended fields from L2TPv3. All the other fields are DEPI specific.

PSP Pseudowire

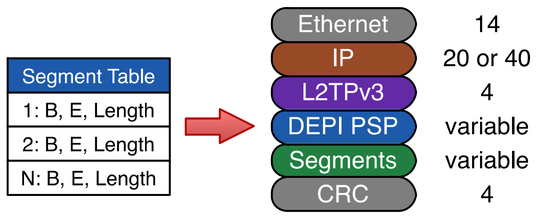


Figure 6 - PSP Pseudowire Type

The PSP Pseudowire has a payload that is divided into one or more segments. Packets are mapped to segments. A segment can contain the beginning, middle, or end of a packet. Payload packets can be split across multiple DEPI packets. Unlike CCF (Continuous Concatenation and Fragmentation) from DOCSIS 3.0, a segment cannot contain the end of one packet and the beginning of the next.

The basic PSP pseudowire sub-type only places packets in the segments. In other sub-types, segments can be used to carry pre-pended or post-pended information that is per packet.

The PSP pseudowire header uses the same 32-bit header from D-MPT, plus it adds a segment table. There is one entry in the segment table for each segment present in the payload. The entries contain:

- Begin bit
- End bit
- Segment length

The “begin” and “end” bits are used for packet reassembly. The segment length is used to find the next segment start byte.

REMOTE PHY TECHNOLOGY

This white paper will focus on the Remote PHY transport. In later white papers, the control protocols and management/OSS protocols will be covered in more detail.

Remote PHY technology builds upon the original DEPI specification.

Overview

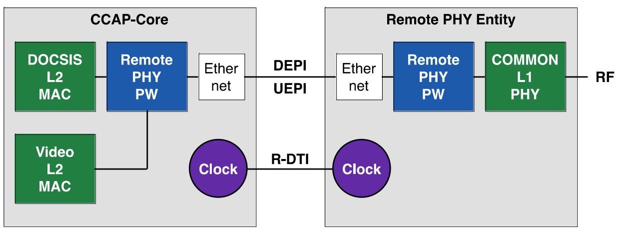


Figure 7 - Remote PHY CCAP System

A CCAP (Converged Cable Access Platform) is a combination of a DOCSIS (Data Over Cable Service Interface Specification) CMTS (Cable Modem Termination System) and an EQAM (Edge QAM).

An I-CCAP (Integrated CCAP) has the CMTS and EQAM in one chassis. In a Remote PHY system, the PHY circuitry from the I-CMTS is removed and put into the Remote PHY Entity (this will get a proper CCAP name later). The remaining parts of the CCAP are called the CCAP Core. This is shown in Figure 7.

Note that hybrid approaches are equally valid. A CCAP system could have both integrated MAC-PHY line cards and MAC-only line card that support Remote PHY.

The Remote PHY protocols use the concept of a pseudowire (PW). A pseudowire is just a cooler and updated name for an IP tunnel between two points in an IP network. Pseudowires can take a specific packet from one point on an IP network and move it to another point. In the case of Remote PHY, pseudowires are used to move MPEG-TS packets and DOCSIS frames between the MAC and the PHY.

One of the significant properties of the Remote PHY Protocols is that it is an IP packet that can traverse any kind of network. It can traverse a layer 2 forwarding network, a layer 3 routed network, an MPLS network, or even a lambda over fiber network.

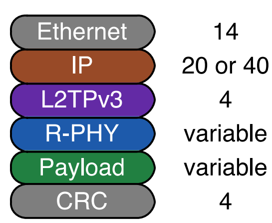
Specifications and Applications

Table 2 - Applications and Transports

Transport	Application		
	DOCSIS	Video	OOB
DEPI	▯	▯	▯
UEPI	▯	▯	▯
R-DTI	▯	▯	▯

Table 2 shows the basic transports and timing used in Remote PHY and how three different applications or higher layer protocols are mapped to these transports. Each of the transports and applications will be described next.

Remote PHY Pseudowire



The generic Remote PHY Pseudowire is the same as the MHA DEPI pseudowire with the exception that there are many more sub-types. In MHA DEPI, there was just the D-

MPT and PSP pseudowire types. There are now a collection of sub-types that add extensions to D-MPT and PSP.

UEPI

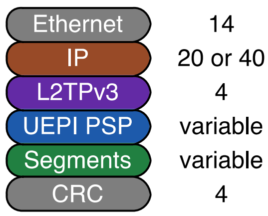


Figure 8 - UEPI Pseudowire

UEPI is the Upstream External PHY Interface. UEPI is new for Remote PHY and was not part of the original M-CMTS and MHA specifications.

UEPI is an extension to L2TPv3. UEPI uses the PSP pseudowire type from DEPI and uses the same control plane as DEPI. UEPI has unique pseudowire types for the upstream direction that are explained in a later section. UEPI for DOCSIS 3.0 is described in [1].

R-DTI

R-DTI is the Remote DOCSIS Timing Interface. The local version, just known as DTI, is part of the M-CMTS specifications. The original DTI protocol defines how a timing server can run timing to a CMTS-Core and an EQAM and how both devices can have the same timestamp. The DTI protocol is described in [10].

R-OOB

R-OOB is Remote Out-of-Band. OOB is the signaling channel in the downstream and upstream that is used to control Set-Top Boxes (STB).

R-OOB use DEPI and UEPI so that the pseudowires can be setup and torn down with the same control plane protocols as rest of the Remote PHY Pseudowires. R-OOB is carried as a separate specification since it is a distinct and potentially separate application.

NEW PROTOCOL EXTENSIONS

Packet Length

DEPI was originally defined with a 1500 byte Ethernet packet. In D-MPT mode, this allows for up to 7 MPEG-TS packets per DEPI packet. In the new Remote PHY specifications, that length will be increased to approximately 2000 bytes. This will allow up to 10 MPEG-TS packets per DEPI packet.

MPLS

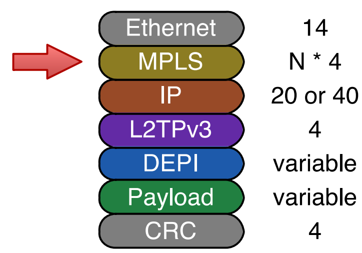


Figure 9 - DEPI over MPLS

MPLS is Multiple Protocol Label Switching, and is a method of moving packets across a network using labels which can be popped on and off at each network forward point. MPLS routes are setup using LDP, the Label Distribution Protocol. MPLS routes are calculated using protocols such as MPLS-TE that is MPLS Traffic Engineering.

Technically, there are two underlying types of pseudowires used in network – MPLS and L2TPv3. The MPLS pseudowire is used for native MPLS networks and the L2TPv3 pseudowire is used for native IP networks. Each protocol has its own control plane. If DEPI had been originally designed as an MPLS pseudowire, it would have added protocol extensions to MPLS-TE.

Since DEPI is just an IP packet, it can be sent over any network, including an MPLS network. Since the DEPI control plane is already well defined and working, the current direction is to not rebuild DEPI with MPLS-TE. Instead, the DEPI forwarding plane and control plane will remain as L2TPv3 and will (optionally) run on top of MPLS. This is shown in Figure 9.

Thus, DEPI does not use an MPLS pseudowire. DEPI uses an L2TPv3 based pseudowire which can be run over an MPLS network.

MCM: Multi-Channel MPT

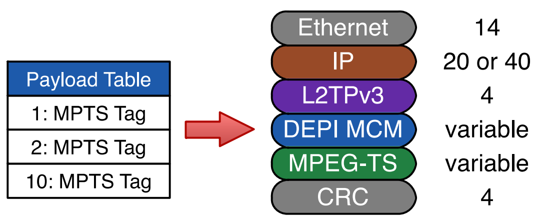


Figure 10 - DEPI MPT/MCM Pseudowire

In the current usage of the MPT pseudowire, there is one pseudowire for each QAM channel in the Remote PHY. The advantage of MPT is that it provides a simple point-to-point connection from a MAC channel to a PHY channel. The disadvantage is that additional latency is incurred in building up MPEG-TS packets into a DEPI packet. The latency of ten MPEG-TS packets at 1 Gbps is 1.6 usec.

While this is not much of an additional delay, MCM also offers an improvement by allowing multiple QAM channels to share the same pseudowire. This also helps with scaling. For example, instead of having 160 pseudowires for 160 QAM channels, technically, there could be one really busy DEPI pseudowire. Typically, there would be a smaller number of DEPI pseudowires with a number of pseudowires.

MCM is a sub-type of the main MPT pseudowire. MCM uses a table in the DEPI header that lists which MPEG-TS packet in the DEPI payload goes to which QAM Channel. A table format in the header is used rather then tagging each MPEG-TS packet so that in implementation, the table can be read and then executed against the rest of the packet. This is shown in Figure 10.

Note that while any number of QAM channels can be supported within a pseudowire, each DEPI packet can only contain up to 10 MPEG-TS packets in one packet at one time.

BFD

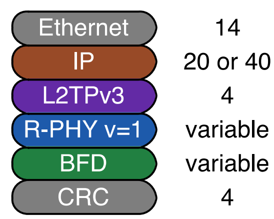


Figure 11 – DEPI or UEPI with BFD

BFD refers to Bi-Directional Forwarding Detection. BFD is essentially a loopback packet that can be used for testing data path integrity. By integrating BFD directly into the packet, the entire transmission path from the DOCSIS MAC channel to Remote PHY QAM channel can be tested.

Note that BFD can also be sent in a stand-alone UDP packet. BFD is an industry standard extension to L2TPv3 and applies to all DEPI and UEPI pseudowires types and sub-types. BFD is described in [17], [18] and [19].

DOCSIS 3.1

DOCSIS 3.1 is a new version of DOCSIS that has recently been released from CableLabs. DOCSIS 3.1 uses a new physical layer based upon OFDM (Orthogonal Frequency Division Multiplexing) and an error correction scheme called LDPC (Low Density Parity Check). The data path consists of one or more OFDM channels. Each OFDM channel has a PLC (PHY link Channel) for initializing CMs (Cable Modems). DOCSIS 3.1 is discussed in more detail in [3], [4] and [5].

In DOCSIS 3.1, the downstream OFDM channel is assigned a list of profiles. The profile lists the modulation to be used for each sub-carrier in an OFDM channel. The profile can be different for each LDPC block and is dependent upon which CM is receiving the OFDM block. This is done to allow the optimization of the transmission path to those CMs that can tolerate a higher modulation. The management of DOCSIS 3.1 profiles is described in [2].

The PLC Channel is composed of message blocks (MB). The following message blocks are defined in DOCSIS 3.1:

- Timestamp MB
- DLS (DOSIS Light Sleep Mode) MB
- Trigger MB
- Message Channel
- Null MB

The timestamp MB is generated locally at the Remote PHY. The other MBs are transparently passed from the CCAP-Core to the Remote PHY Entity.

In DOCSIS 3.1, the downstream and upstream frequency ranges also can be altered to provide more throughput. DOCSIS 3.1 can support 1 to 2.5 Gbps in the upstream, and 5 to 10 Gbps in the downstream. [7]

DS OFDM Channel Pseudowire

Each OFDM Channel is assigned to a PSP pseudowire with a subtype of DOCSIS 3.1 Data Path. The PSP header is extended so that its table contains the OFDM profile number for each packet in the PSP stream.

DS PLC Channel Pseudowire

Each PLC is assigned to a PSP pseudowire with a subtype of PLC. For the DLS MB, Trigger MB and the Null MB, the MBs are placed directly into a PSP segment and referenced by the PSP table. A 32-bit timestamp can optionally be prepended to a segment. This is useful for the trigger MB. The packets in the payload of the message channel are mapped to PSP segments.

D3.1 Upstream Pseudowires

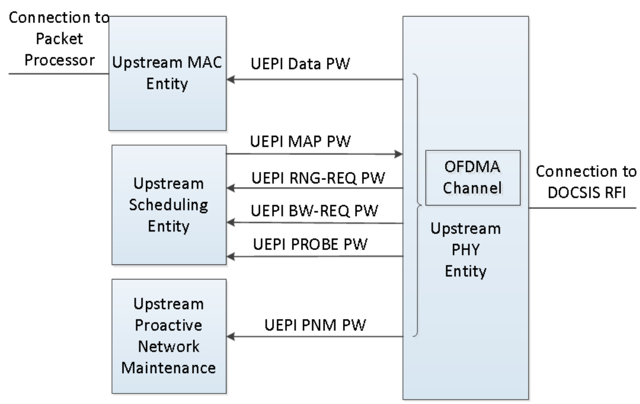


Figure 12 - DOCSIS 3.1 UEPI Pseudowires

There are a number of pseudowires associated with the DOCSIS 3.1 Upstream. All pseudowires have a pseudowire type of PSP and each have a unique sub-type. The DOCSIS 3.1 UEPI pseudowires are similar to the DOCSIS 3.0 UEPI pseudowires, but have differences in formats. The Probe PW is unique to DOCSIS 3.1 Remote PHY.

Video

For video, the same Remote PHY philosophy that was used for a DOCSIS CMTS is applied to a video EQAM. This works extremely well now that both the CMTS and EQAM functionalities are contained in the same CCAP device and use the same PHY chip.

The video in the form of SPTS or MPTS is received by the CCAP-Core from the network. The CCAP core performs all the same EQAM functions that it normally performs. These include a jitter buffer with PCR re-stamping, SPTS assembly into MPTS, Conditional Access, PID remapping, etc.

The resulting fully formatted MPTS stream is then sent over a DEPI pseudowire to the Remote PHY entity. Due to some jitter that will be introduced by the network between the CCAP-Core and the Remote PHY entity, the Remote PHY needs to contain a second smaller video jitter buffer. This jitter buffer will be smaller than the one in the CCAP-Core as it does not need to filter server jitter.

Since video is already in MPEG-TS format, video can be sent using a variant of the MPT/MPT and MPT/MCM pseudowires. The variant is that an MPEG-TS frame counter is added in addition to the L2TPv3 sequence counter. The additional MPEG-TS packet counter helps in error recovery.

OOB

OOB refers to the Out-of-Band protocols that control STB operation. There are two main OOB systems in use and they are defined in SCTE specifications.

- SCTE 55-1: This is the Motorola/Arris system that uses MPEG-TS packets in the downstream and an Aloha polling system in the upstream.
- SCTE 55-2: This is the SA/Cisco system that uses ATM over a modified T1 frame.

Analysis showed that the 55-1 system could be implemented using the pure Remote PHY strategy. However, the 55-2 system has a 3 millisecond turn around time from upstream to downstream that would of made it extremely sensitive to network latency. As a result, the decision was made to include the OOB framer along with the OOB PHY in the Remote PHY Entity. All the signaling for OOB and the carousel generation would remain centralized.

This also allows the OOB system to be driven directly by provisioning systems and OOB controllers independent of the CCAP-Core if so desired.

OOB would use a dedicated DEPI and UEPI pseudowire so that the links between the CCAP-Core and the R-PHY Entity can be reconfigured in conjunction with the other links for DOCSIS and video.

REMOTE PHY TIMING

Timing in a Remote PHY Network is managed with the Remote DTI (R-DTI) protocol.

The goal of R-DTI is also to get approximately the same timestamp value in the CCAP-Core and the Remote PHY. However, the technology for R-DTI is completely different than the technology used for DTI. Only the name is similar. Whereas DTI was a local protocol that ran over a maximum distance of 200 meters from DTI Server to DTI Client, R-DTI is required to work over long distances.

The R-DTI specification specifies multiple modes of operation that provide different levels of performance. There is a simple low cost mode at one option and a complete IEEE-1588 mode as another option.

R-DTI – One-Way Reverse

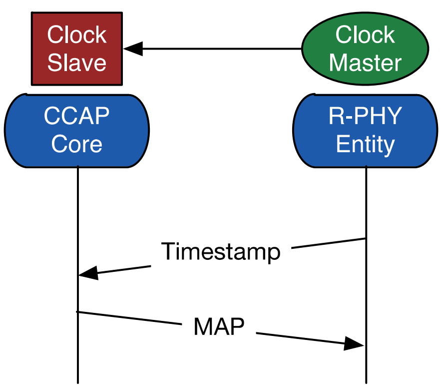


Figure 13 - R-DTI One-Way Reverse

The reverse direction refers to the R-PHY Entity being a clock master and the CCAP-Core being a clock slave.

The one-way refers to the Remote-PHY entity sending a timestamp to the CCAP-Core. The CCAP-Core filters the timestamp traffic, adds a MAP advance time, and uses the result to generate MAPs.

The advantage of the reverse method is that the physical clock in the R-PHY Entity does not need to be adjusted. The act of aligning the clock in the R-PHY entity to the CMTS-Core may cause enough disturbances in the DOCSIS baud rate to cause the R-PHY entity to fail the DRFI specifications. The disadvantage of the reverse method is the scaling required at the CCAP-Core. The CCAP-Core may now have to track hundreds of clocks. This, however, is not unlike what EQAMs have to do for video dejittering.

In this method it is not known what the delay is from the R-PHY entity to the CCAP-Core. And maybe it does not really matter. That delay, whatever it may be, becomes engineering margin. The longer the network distance, and hence the delay, the more margin is required at the CMTS-Core. The more jitter the network introduces, the more margin that is required at the CMTS-Core. So, it is somewhat of a self-compensating mechanism.

R-DTI – Two-Way Reverse

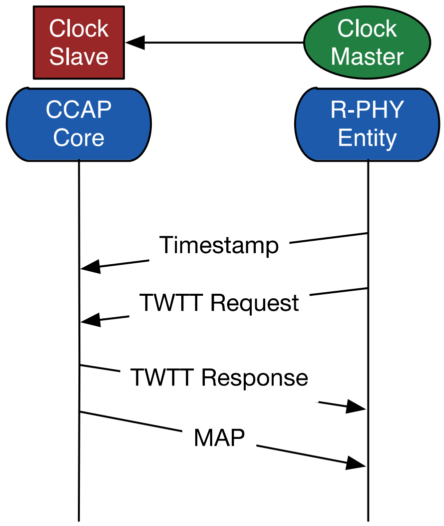


Figure 14 - R-DTI Two-Way Reverse

When the CMTS-Core receives a timing message from the R-PHY entity, it returns a timing message, creating what is known as a TWTT or Two-Way Time Transfer. TWTT collects outbound and inbound timestamps in each direction. Using these four timestamps, the one-way network delay can be calculated. That network delay can then managed more precisely with the MAP advance.

There are variations on the style of signaling as to whether or not the timestamp from the R-PHY Entity is in the same message as the request/response messages that actually get stamped on ingress and egress.

This method is superior to the one-way reverse method only if the time stamping of the timing packets happen in a timely and accurate manner. Such mechanisms usually are hardware based.

R-DTI – Two-Way Forward

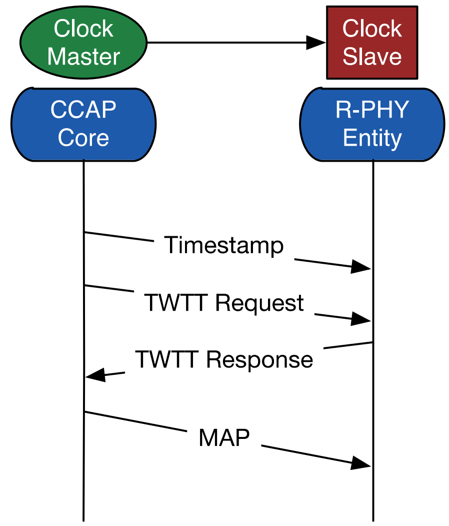


Figure 15 - R-DTI Two-Way Forward

In the two-way forward mode, the CCAP-Core is the clock master and the R-PHY entity is the clock slave. A TWTT protocol is run between the CCAP-Core and the R-PHY entity.

This is the classic clocking network, and is included to allow a full-blown level of accuracy. This is the classic IEEE-1588 implementation. This method would be applicable when the network is 1588 compatible.

Remote Scheduler

The Remote PHY specification permits the relocating of the DOCSIS scheduler from the CMTS-Core into the Remote PHY Entity. With the scheduler remotely located, there is no longer any timing requirement between the CMTS Core and the Remote PHY entity.

Although this sounds attractive at first, there are several drawbacks to remotely locating the scheduler. The first drawback is that the complexity of the Remote Node increases. The Remote Node entity now become subscriber aware and requires state information, QoS information, policy information and needs to track historical information for implementing of rate shaping. This is definitely an increased level of complexity.

The second reason that a remote scheduler is not attractive is interoperability. If the CCAP-Core is from one vendor, the scheduler and node are from another vendor, and something goes wrong, who is at fault? How is troubleshooting done without a centralized scheduler as a reference point?

Our current analysis shows that a centralized scheduler will work under currently known network conditions and that the timing design is quite manageable. (This will be the subject of a future white paper). The option for a remote scheduler is a future option should it ever be needed.

Details

Packet Formats

The two-way forward method is intended to be an IEEE-1588 implementation. It may also include a Synchronous Ethernet (SyncE) implementation for frequency alignment.

The one-way and two-way reverse methods could be implemented with a light version of IEEE-1588 or a Remote PHY specific control packet.

Support for DTP

DTP is the DOCSIS Timing Protocol that was developed by this author and is being introduced with DOCSIS 3.1. DTP allows a CM to provide timing services such as IEEE-1588 and Synchronous Ethernet derived from DOCSIS timing. This is explained in a white paper study [6] and in the DOCSIS 3.1 specification [14].

Distances

DOCSIS define a CMTS-to-CM, PHY-to-PHY distance of 100 miles (160 km) for DOCSIS 3.0 and 50 miles (80 km) for DOCSIS 3.1. Remote PHY maintains these distances.

Because Remote PHY separates the DOCSIS MAC and PHY, there is an additional distance specification. That is the MAC-to-MAC, CMTS-Core to CM distance. This distance is relevant when the CMTS-Core is removed from the hub and placed in the head end or regional data center.

In theory, there is no distance limitation to R-DTI. In practice, longer distances add jitter and latency that may impact services such as scheduling. While the maximum operating distance has not been chosen at the time this paper was written, research is being conducted to see if R-DTI and the scheduling applications it supports could run over a distance up to 2000 km. This number comes from the Indian market. The Norway market requires about 1500 km if everything was driven out of Oslo.

That is more than 10x the original DOCSIS distance of 160 km!

This additional distance is intended to allow the CCAP-Core to be located at the head end or regional data center instead of at the hub site.

NEXT STEPS

This white paper defined the operation of the Remote PHY transport. The next steps will include:

- agreeing on the bit definition of the various protocol headers,
- extending the DEPI control plane to cover UEPI, and
- defining a configuration and operational model for the Remote PHY entity.

SUMMARY

Remote PHY is an approach that literally takes the PHY chip out of a box and puts it at the end of an IP network. One of the philosophies of Remote PHY is to put the least amount of hardware and software at the end point and keep the complexity centralized.

Remote PHY infers centralized DOCSIS software. This allows the same software model to be used for I-CCAP and Remote PHY CCAP. Remote PHY, I-CMTS and M-CMTS can all co-exist in the same chassis and use the same software base and configuration systems. This is a very powerful concept for feature velocity and backwards compatibility.

Remote PHY works and works well. The design of remote PHY is built on top of open standards such as Ethernet, IP, L2TPv3, and CableLabs MHA.

Remote PHY will allow CCAP devices to be deployed in more creative manners such as using digital fiber in the HFC plant. For the Cable Operators, this will allow their network to have higher performance with lower OPEX, lower CAPEX, and an evolutionary path the FTTH.

REFERENCES

Related Papers by the Same Author

[1] John T. Chapman, "DOCSIS Remote PHY," SCTE Cable-Tec Expo, October 21, 2013

[2] John T. Chapman, "The Power of DOCSIS 3.1 Downstream Profiles", NCTA Spring Technical Forum, June 13, 2013

[3] John T. Chapman, "DOCSIS 3.1 – Redefining the State of the Art", Cable Congress, March 6, 2013

[4] John T. Chapman, "What the Future Holds for DOCSIS", Keynote speech, Light Reading Conference, Denver, May 18, 2012

[5] John T. Chapman, Robert Howald, Mike Emmendorfer, "Mission is Possible: an Evolutionary Approach to Gigabit-Class DOCSIS", NCTA Spring Technical Forum, May 12, 2012

[6] John T. Chapman, Rakesh Chopra, Laurent Montini, "The DOCSIS Timing Protocol – Generating Precision Timing Services from a DOCSIS System", NCTA Spring Technical Forum, June 14, 2011

[7] John T. Chapman, "Taking the DOCSIS Upstream to a Gigabit per Second", NCTA Spring Technical Forum, May 12, 2010

[8] John T. Chapman, Jeff Finkelstein, "Maximizing Legacy and DOCSIS 3.0 Bandwidth in the Forward Path using M-CMTS and EdgeQAM Technologies", SCTE Cable-Tec Expo, October 30, 2009 http://www.johntchapman.com/Legacy-and-DOCSIS3-with-M-CMTS-WP-%282009-10-30%29.pdf

[9] John T. Chapman, "Next Generation CMTS – An Architectural Discussion",

SCTE Emerging Technologies, January 15, 2008
http://www.johntchapman.com/NG-CMTS-WP-071125.pdf

[10] John T. Chapman, "DOCSIS Timing Interface – Timing for DOCSIS Networks", National Institute of Standards and Technology (NIST) Keynote Address, March 10, 2007
http://www.johntchapman.com/NIST-jchapman-070310.pdf

[11] John T. Chapman, "Modular CMTS Tutorial", CableLabs Summit, December 2006
http://www.johntchapman.com/M-CMTS-Tutorial-Public-070220a.pdf

[12] John T. Chapman, Andrew Page, Carmelo Iaria, Glenn McGilvray, "Transitioning to M-CMTS", NCTA National Show, April 2006

[13] John T. Chapman, "Modular CMTS Architecture", NCTA National Show, April 3, 2005
http://www.johntchapman.com/ncta-2005-modular-cmts-wp.pdf

Related CableLabs Specifications

[14] "DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification", Revision I02, CableLabs, March 20, 2014.
http://www.cablelabs.com/specification/mac-and-upper-layer-protocols-interface-specification/

[15] "DOCSIS Timing Interface (DTI) Specification", Revision I05, CableLabs, December 9, 2008
http://www.cablelabs.com/specification/docsis-timing-interface-specification/

[16] "Downstream External PHY Interface Specification", Revision I08, CableLabs, June 11, 2010
http://www.cablelabs.com/specification/downstream-external-phy-interface-specification/

Related IETF Specifications

[17] David Katz, David Ward, "Bidirectional Forwarding Detection (BFD)", RFC-5880, Internet Engineering Task Force (IETF), June 2010
http://tools.ietf.org/html/rfc5880

[18] David Katz, David Ward, "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6", RFC-5881, Internet Engineering Task Force (IETF), June 2010
http://tools.ietf.org/html/rfc5881

[19] David Katz, David Ward, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification", RFC-5885, Internet Engineering Task Force (IETF), June 2010
http://tools.ietf.org/html/rfc5885

[20] J. Lau, Mark Townsley, I. Goyret, "Layer Two Tunneling Protocol – Version 3 (L2TPv3)", RFC-3931, March 2005
http://tools.ietf.org/html/rfc3931

Remote PHY: Enabling Immediate Access to Extra Bandwidth Capacity in Existing Networks

Oleh Sniezko, Oleh-Lightcom Consulting, LLC
Rei Brockett, Dave Baran, Michael Field, Steve Hopkins
Aurora Networks – A Pace Company

Abstract

During the last several years, the increasing demand for network capacity created significant uptick in the activity directed towards analysis of the optimal methods of expanding the capacity of wireline telecommunications networks, including broadband HFC networks. Multiple papers described the reason for the increased demand, the ways to expand the network bandwidth and the ways to improve the bandwidth efficiency through deployment of the next generation PHY.

This paper contributes to this discussion by describing the least disruptive method of increasing capacity of the existing networks. Specifically, network design analysis and modeling proves that the existing HFC network can support immediate increase in capacity without resorting to bandwidth expansion and without waiting for the next-generation PHY of DOCSIS® 3.1[1] *and EPoC. Distributed Broadband Access Architecture, a.k.a. remote PHY*[2] *improves the existing network performance from a network capable of supporting 256-QAM Reed-Solomon (R-S) signals, with or without Trellis Coded Modulation (TCM), to a network capable of supporting 500-1000 MHz of 1024-QAM signals of the same J.83 format. The paper presents design guidelines for achieving this capacity increase without having to re-construct either the "F" or the "C" portions of the HFC network.*

The paper takes a snapshot of the status quo of deployed CPE devices to determine the feasibility of taking advantage of this improved network performance with the existing equipment. The paper also presents a partial inventory of silicon chips that can support the transmission of 1024-QAM R-S signals with and without TCM and lists examples of CPE devices capable of the same.

The paper presents test data of 1024-QAM signal delivery to validate the network analysis and modeling and to detail the design guidelines for capacity expansion.

In the summary, the paper presents a set of bandwidth and capacity expansion methods for HFC networks and their ranking based on the outcomes and required effort and investment. This summary encompasses the analysis presented in the previous paper[3]*, including a relatively straightforward expansion of the network capacity by expansion of bandwidth from 750 MHz to 1002 MHz with remote PHY.*

INTRODUCTION

The demand for increased capacity in broadband telecommunications networks comes in waves. Cycles of innovation create new services generated both from within the telecommunication and entertainment delivery industries, and externally by newcomers seeking new markets. Heightened levels of customer expectations force all providers to compete and upgrade. Capacity demand is also fueled by the consumer electronics (CE) industry continuously searching for new sources of revenue in CE technology replacement cycles. We are witnessing insatiable capacity demand in wireless communications networks fueled by the same drivers and, after a period of temporary lull in wireline networks, we see the same build up of capacity demand driven by over-the-top service providers who, despite the wishful thinking of the network operators, do not disappear but are actually growing exponentially in their service delivery offerings; and by the consumer electronics industry that presents to the public 4K HD and 8K HD video sources and video displays of increasing size, often with increased frame repetition rate and pixel depth and in 3D format. The first trend will result in duplication of the same programming, delivered to the same group of households from different providers, which would not and could not be addressed or remedied by multicasting. The second trend, despite the progress in compression algorithms, will increase requirements for the capacity per stream. These trends, combined with increased numbers of display devices per household, often watched simultaneously by the same subscriber and often requiring different resolutions even for the same programming and video streams delivered at the same time, leads to the realization that the increase in capacity to 10 Gbps and beyond for wireline networks will be required within the next 10 years[45].

So, what can network operators do to counterbalance these trends and actually cash in on them in a manageable way that takes into account the cost of the capacity expansion and provides for a reasonable ROI? The place to start is with the outlay part of the ROI equation. The authors will leave the other parts of the equation – revenue and profits – to the wizards of marketing and – operating; and variable costs – to the practitioners of network operations, where cost is partially defined by the outlays, and purchasing magicians of the network operators.

PERFORMANCE MARGINS

Traditional HFC with Analog Forward Links

There are two essential methods of increasing the capacity of broadband networks: increasing their bandwidth and improving the bandwidth efficiency. The method chosen often depends on the all-inclusive[a] cost of achieving the capacity expansion. Sometimes it is less costly to increase the network bandwidth, especially when many network elements are capable of supporting the bandwidth expansion and additional network elements (terminal equipment) facilitate taking advantage of this bandwidth expansion capability. Sometimes it is less costly to increase bandwidth efficiency. In many cases, the tools for achieving both are the same and both could be achieved at the same time thus maximizing network capacity expansion effectiveness[b].

Let us analyze the network performance margins and the weak links that can be improved upon, in search of bandwidth efficiency improvement opportunities.

Headend

The major contributor to the QAM signal impairments in the headend is the signal source itself. Unfortunately, short of replacement of these signal sources, there is little that can be done to improve their performance. Even the little improvement that can be achieved by elimination of the RF combining network in the headend requires replacement of the traditional signal sources with CCAP equipment that performs signal

[a] Cost of the upgrades to distribution network, including in-house wiring, and terminal devices, both in the headend and on customer premises.

[b] Effectiveness of the upgrade for capacity expansion by extending BW and improving BW efficiency at the same time leads often to lower cost per unit capacity gain. The capital outlay management will decide whether to perform both.

combining in the digital baseband domain by signal multiplexing. It is important to note that the major degradation in the RF combining network is not the thermal noise generated by its active components but the signal crosstalk between QAM signals carrying disparate information. This degradation[c] does not practically exist for analog signals which carry the same signals to different outputs of the RF combining network.

The alternative to the replacement of the signal sources is their relocation and creation of a Distributed Broadband Access Architecture with remote PHY. The benefits of this approach will be apparent pending further analysis.

Analog Forward Optical Links

The analog optical links can be incrementally improved but the improvements are bounded by theoretical limits. The typical upper boundaries of analog link performance are listed in Table 1 as "0 km fiber distance" cases. These boundaries assume unrealistic fiber conditions and do not take into account any degradation caused by linear and nonlinear fiber effects (no fiber in the link). For comparison, the typical performances of single-wavelength systems and 20-wavelength systems over 20 to 60 km of fiber are also listed in Table 1. The optical analog links are generating signal impairments at the level comparable to or higher than impairments generated by headend components.

The approaches to improving the optical link performance materially include:

- Split-band dual wavelength approach resulting in increased optical modulation index (OMI) spectral density at the cost of total fiber capacity. This approach requires two fibers per service segment if MWVL system degradation is to be avoided.
- Priority load approach where some part of the bandwidth is allocated higher OMI spectral density. This approach is used in analog optical links supporting NTSC analog signals and QAM signals with higher OMI spectral density allocated to analog signals.
- Significant limits to the number of wavelengths per fiber and careful wavelength allocation to limit nonlinear MWVL fiber impairments (at the extreme, using single wavelength per fiber).
- Deployment of different analog optical link technologies for different distances to optimize technology for the required launch powers and linear and nonlinear fiber impairments.
- Fiber replacement for fibers with better performance for the selected wavelength range and distances.

None of the solutions listed above is low cost and some introduce additional operational burdens (e.g., different alignment processes and design guidelines for different technologies and fibers). Even with unlimited existing fiber count, the opportunity cost eventually adds to the cost of the solutions listed above on top of the analog link component costs.

RF Distribution Network

The RF distribution network with bandwidth up to 1 GHz contributes relatively low signal impairments if designed properly and with short cascades. The major limitation comes when the multiple active cascades need to be expanded in bandwidth beyond 1 GHz. These limitations are diminished in passive coaxial networks (Fiber Deep architecture) in which bandwidth can be expanded significantly up

[c] Multipath degradation of analog signals occurs at much higher crosstalk levels than those introduced by headend downstream RF combining/splitting network, especially after considering multipath differential delays.

to 1.8 and even 2 GHz with simultaneous improvements in bandwidth efficiency. Table 2 presents coaxial network performance in the designed bandwidth based on the designs analyzed in the paper presented at NCTA Spring Technical Forum in 2013[6]. The table does not list the RF cascade CNR performance for the Fiber Deep architecture since the performance of the RF section is defined by the node performance alone and no RF active cascade degradation occurs.

Table 1: Analog Optical Link Performance

Optical Link Description	RF Load	Fiber Distance	Optical Receiver Input	CNR Performance
1310 nm Single Wavelength	74 NTSC analog and 75 QAM channels (6 dB lower)	0 km	0 dBm	53.5 dB analog/46.5 dB QAM
		20 km	0 dBm	52.2 dB analog/45.2 dB QAM
		40 km	0 dBm	51.7 dB analog/44.7 dB QAM
1550 nm Directly Modulated Single Wavelength		0 km	0 dBm	52.2 dB analog/45.2 dB QAM
		20 km	0 dBm	50.2 dB analog/43.2 dB QAM
		40 km	0 dBm	48.2 dB analog/41.2 dB QAM
1550 nm Externally Modulated Single Wavelength		0 km	0 dBm	50.7 dB analog/43.7 dB QAM
		48 km	0 dBm	50.7 dB analog/43.7 dB QAM
1310 nm Single Wavelength	149 QAM channels	0 km	-3 dBm	48.5 dB
		20 km	-3 dBm	48.1 dB
		40 km	-3 dBm	46.4 dB
1550 nm Directly Modulated Single Wavelength		0 km	-3 dBm	47.7 dB
		20 km	-3 dBm	46.4 dB
		40 km	-3 dBm	45.0 dB
1550 nm Externally Modulated Single Wavelength		0 km	-3 dBm	45.0 dB
		60 km	-3 dBm	45.0 dB
20-Wavelength 1550 nm Directly Modulated	149 QAM channels	20 km	-1 dBm	40.0 to 44.1 dB (different frequencies, worst wavelength)
		40 km	-6 dBm	38.9 to 41.9 dB (different frequencies, worst wavelength)

Table 2: RF Active Cascade Performance

Design Scenario	Worst/Best Case CNR for Longest Cascade [dB]
750 MHz N+5 "Moscow" system	49.8/52.2
860 MHz N+5 "Stalingrad" System	56.0/56.2

In-House Wiring and CPE

The signal levels at the outlet of the in-house wiring in combination with the terminal equipment performance define the signal impairments contribution by this network element. The analysis of multiple system designs is presented in Table 3. The results show that the median levels are comfortably higher than the minimum design levels. This data is confirmed by statistical data collected by the industry for millions of terminal equipment modules (cable modems and set top boxes). Nevertheless, some terminal devices will fall below the design levels.

Table 3: Outlet Signal Levels and Performance

Original design Freq & Type	Input Levels @ Highest Freq. [dBmV]	CPE CNR @ Highest Freq. [dB]
750 MHz N+5	-3.02	44.96
860 MHz N+5	-2.22	45.76
860 MHz FD	-6.39	41.59
1002 MHz FD	-2.70	45.28

Table 4: Typical EOL Signal Performance

Original design Freq & Type	Headend performance [dB]	Optical Link performance [dB]	RF Cascade performance [dB]	Median CPE CNR @ Highest Freq. [dB]	Design CPE CNR @ Highest Freq.	EOL CNR @ Highest Freq./Median Level	EOL CNR @ Highest Freq./Design Level [dB]
QAM Signal Performance							
750 MHz N+5	42.2	39.0	49.8	45.0	38.0	36.4	34.5
	42.5	40.0	52.2	45.0	38.0	37.1	34.9
	42.5	45.0	52.2	45.0	38.0	39.0	36.0
860 MHz N+5	42.2	39.0	56.0	45.8	38.0	36.7	34.6
	42.5	40.0	56.2	45.8	38.0	37.3	35.0
	42.5	45.0	56.2	45.8	38.0	39.3	36.0
860 MHz FD	42.2	39.0	NA	41.6	38.0	35.9	34.6
	42.5	40.0	NA	41.6	38.0	36.5	35.0
	42.5	45.0	NA	41.6	38.0	38.0	36.1
1002 MHz FD	42.2	39.0	NA	45.3	38.0	36.7	34.6
	42.5	40.0	NA	45.3	38.0	37.3	35.0
	42.5	45.0	NA	45.3	38.0	39.3	36.1
Analog Signal Performance							
750 MHz N+5	60.0	51.0	57.0	52.2	49.2	47.7	46.4
860 MHz N+5	60.0	51.0	63.2	52.2	49.2	48.1	46.7
860 MHz FD	60.0	48.0	NA	52.2	49.2	46.4	45.4
1002 MHz FD	60.0	48.0	NA	52.2	49.2	46.4	45.4

EOL Signal Performance

Table 4 presents typical EOL QAM signal performance for median and design levels at the terminal devices. For comparison, the table also lists performance of analog NTSC signals at the levels into the terminal devices at 0 and 3 dBmV. One may notice a significant difference that is only partially justified by the difference in terminal device input levels.

The brief analysis of the table clearly points out that the additional major factors contributing to this difference are:

1. headend impairments differences; and
2. optical link impairment differences.

The headend performance for analog video signals is significantly better than the headend performance for QAM signals. There are many reasons for that: the analog signals demand much higher EOL and the limitation of the RF coaxial network and analog optical links demanded almost pristine signal performance at the source of the analog signals at the input to the HFC network. Even more critical was the cost management of the network. Without the high signal quality in the headend, the RF coaxial network cost would increase dramatically. If the headend high signal quality could not have been achieved, then the terminal device input levels would have to be raised during the design process, again drastically increasing the cost of the RF network (and dashing the hopes for bandwidth expansion up to 1000 MHz).

The typical headend performance for analog video signals was historically defined at 60 dB CNR. This is drastically better than 43 dB MER required of QAM signals sources, which are further degraded by 50-52 dB crosstalk levels between narrowcast QAM signals. The crosstalk will be all but eliminated in future replacement of edge QAM modules and headend RF combining networks with CCAP equipment.

Another source of difference comes from the traditionally lower OMI spectral density allocated to QAM signals in analog optical links, again justified by significantly higher analog signal performance requirements.

As already mentioned above, the third contributor to disproportionately lower EOL performance of QAM signals comes from the common practice for designing much lower levels into QAM terminal devices. This is again justified by preserving high levels for analog NTSC signals and managing the cost of the network. The success of digital TV and HSD (carried on QAM signals) resulted in proliferation of QAM terminal devices and hence inflation of the number of outlets per household. Providing high input levels to all of them, short of installing house amplifiers that are not preferred for many reasons[d], would inflate the network cost.

All of these sources of significant EOL performance differences may be completely eliminated or materially remedied with DBAA.

RECAP OF DISTRIBUTED BROADBAND ACCESS DIGITAL HFC ARCHITECTURE

Distributed Architecture Synopsis

For detailed descriptions of several approaches to DBAA, the authors refer to the paper[7] presented at NCTA Spring Technical Forum 2013. Here we just remind the readers that in DBAA (see Figure 1 for an example), signal sources are relocated from headend to the nodes thereby eliminating headend RF combining network and analog optical link contributions to signal impairments. This move also simplifies and granulates future upgrades[e] for better signal source performance that may be required for NG PHY. This relocation will also enable some RF coaxial level management to increase the levels into terminal devices after the move to all digital (QAM-only) signal load on HFC networks. The results of the testing to support this claim follow.

The referenced paper presents detailed analysis, showing what can be achieved in HFC with DBAA in terms of significant bandwidth expansion under different outlay costs and of significant improvements in bandwidth efficiency (with or without bandwidth expansion).

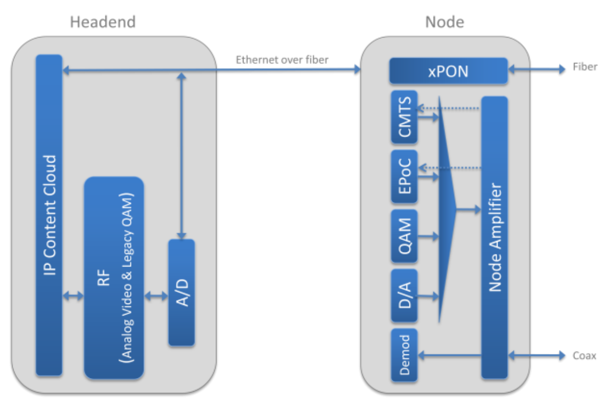


Figure 1: All-Digital Broadband Access Architecture

Immediate Opportunity for HFC Capacity Expansion with DBAA

Implementation of DBAA delivers the immediate opportunity of improving EOL

---

[d] House amplifiers hinder future BW expansion and upstream level management to mention a couple issues.

[e] The network upgrades can be performed on a node-by-node or segment-by-segment basis depending on demand for increase network capacity, allowing more efficient management of capital outlays.

signal performance and hence bandwidth efficiency for HFC networks. Table 5 summarizes EOL performance improvement for the QAM portion of the network under the assumption that there is no performance degradation from analog NTSC signals and other QAM signals carried on the analog optical links, if they are still used[f]. Table 5 shows that an improvement of more than 2 dB in EOL (from the lowest of 34.5 dB to the lowest of 36.6 dB and higher for other values) can be achieved without additional level increase to terminal devices while more than 4 dB improvement (from the lowest of 34.5 dB to the lowest of 38.7 dB) can be achieved after taking advantage of QAM-only RF active load and increasing levels to terminal devices without reworking in-house wiring (see further test results and Table 6). As the results of the testing show, these EOL performance levels are more than adequate to support 1024-QAM Adv PHY signals with R-S coding, with or without TCM.

The only remaining question is whether we can take immediate advantage of the EOL signal performance improvement to increase HFC network capacity without having to fall back to the more fiscally and operationally costly method of capacity increase through bandwidth expansion.

<u>RECOVERED MARGIN MONETIZATION: CAPACITY</u>

<u>1024-QAM with R-S and TCM: Performance Requirements</u>

To confirm many published numbers and define the performance requirements for 1024-QAM signals, testing of the BER versus SNR was performed in the test setup presented in Figure 2. The testing was performed for 256-QAM J.83 signals Annex C and Annex B with three different data patterns and for 1024-QAM J.83 signals Annex C. Due to the lack of reliable test equipment, the testing for 1024-QAM signals Annex B was not completed. However, the coding gains between Annex C and Annex B 256-QAM signals can be extrapolated to 1024-QAM signals, especially after accounting for the very close correlation of test results and theoretical values. The test results are presented in Figure 3.

The test results confirm 27-28 dB SNR (MER dominated by noise) requirement for post-FEC QAM 256 J.83 Annex B signal and 29-30 dB SNR requirement for post-FEC 256-QAM J.83 Annex C signal. This difference confirms approximately 2 dB TCM gain after R-S coding[g] is applied. The results also confirm the 6 dB required improvement in SNR to support a post-FEC error-free environment for 1024-QAM J.83 Annex C signals (35-36 dB SNR). By extrapolation, 33-34 dB SNR performance is required to support error-free 1024-QAM J.83 Annex B signal transmission.

The comparison of the numbers in Table 4 with the test results and theoretical numbers indicates that the HFC networks are designed and operate at 6 dB and higher operating margins, and at 10 dB operating margins for median EOL performance. It also confirms that this margin would drop to zero for the design performance and many CPE devices would not be able to operate in 1024-QAM J.83 Annex B environment. Table 5 shows that simple implementation of DBAA enables transmission and reception of 1024-QAM J.83 Annex B signals, albeit at lower operational margins of approximately 3.5 dB for the design performance levels and 6-7 dB for median plant performance.

[f] To accomplish this, relatively straightforward bandwidth allocation management guidelines and signal filtering in the node can be implemented.

[g] Different levels of R-S coding for Annexes A and C and Annex B.

Table 5: EOL Signal Performance after DBAA Implementation

QAM Signal Performance							
Original design Freq & Type	Modulator Performance [dB]	RF Cascade performance [dB]	Median CPE CNR @ Highest [dB]	Design CPE CNR @ Highest Freq.	EOL CNR @ Highest Freq./Median Level	EOL CNR @ Highest Freq./Design Level [dB]
750 MHz N+5	43.0	49.8	45.0	38.0	40.3	36.6
	43.0	52.2	45.0	38.0	40.6	36.7
860 MHz N+5	43.0	56.0	45.8	38.0	41.0	36.7
	43.0	56.2	45.8	38.0	41.0	36.7
860 MHz FD	43.0	NA	41.6	38.0	39.2	36.8
1002 MHz FD	43.0	NA	45.3	38.0	41.0	36.8

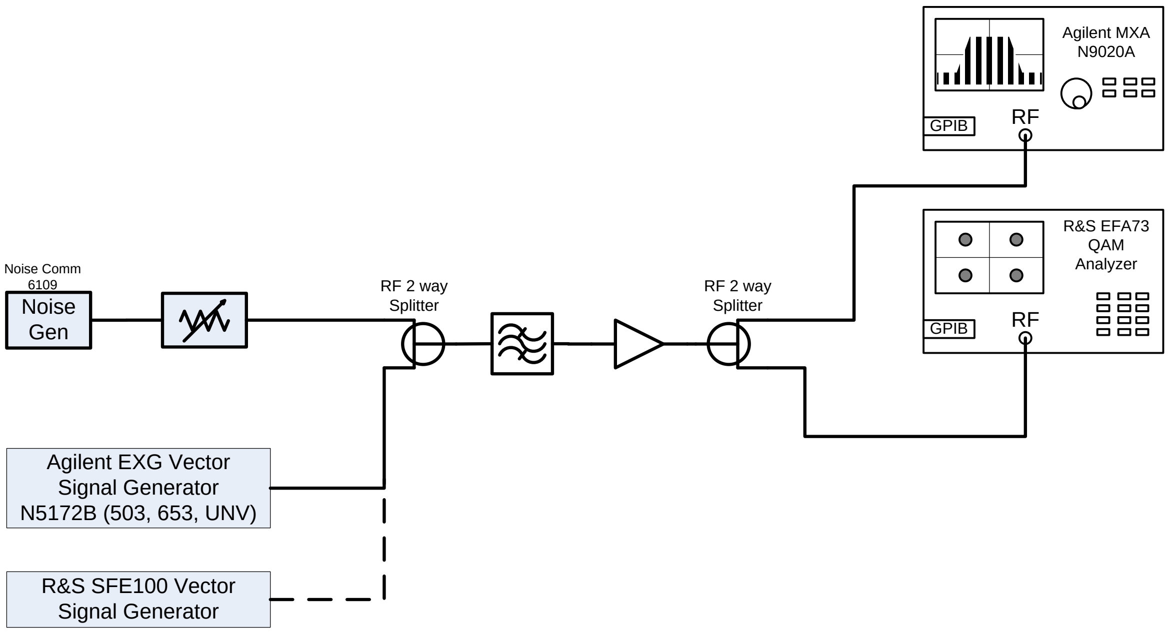


Figure 2: Test Setup – QAM Signals Sensitivity to Gaussian Noise

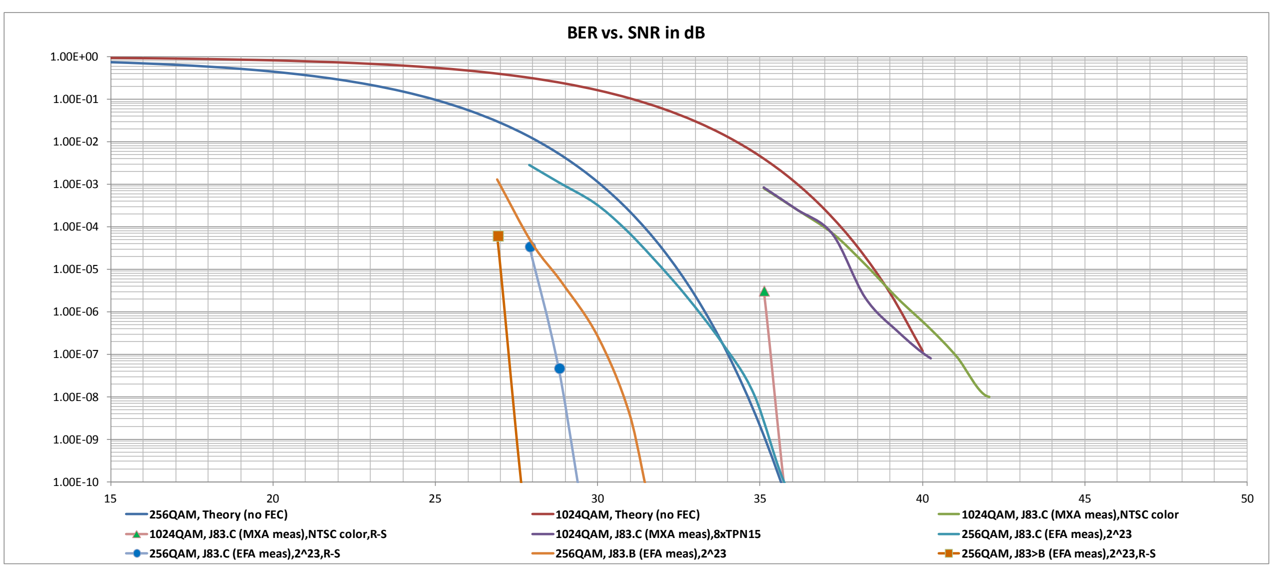


Figure 3: Test Results: QAM Signal Sensitivity to Gaussian Noise

Higher operating margins are critical for:

- achieving and maintaining low operating cost at high level of service satisfaction;
- supporting the operational outliers among terminal devices[h].

The test results also show that TCM gain is approximately 4 dB before R-S coding is applied. Total gains are lower after R-S coding due to the additional FEC gains realized with R-S coding in Annexes A and C.

Analog Forward Optical Links

The performance of the optical links is already pushed to the clipping limits with minimal margin left. The linearization technologies evolved over the years and took advantage of any theoretical margins. Lowering these margins further or eliminating them altogether would bring limited performance improvement in the order of 1 dB but would push the operation of optical links to the brink of disaster even if anti-clipping and AGC circuitries are implemented. To achieve any sizeable improvement approaching that secured by DBAA deployment is practically unachievable even with all the remedies listed previously (even if cost is not a concern).

For example, deploying two single-wavelength split-band systems would require 2 wavelengths per node segment and would yield approximately 48 dB SNR in 1 GHz QAM-only links. This performance, in combination with 42.2 dB headend performance, would yield approximately 41 dB SNR, a far cry from the guaranteed DBAA performance, and would still be a major contributor to the EOL performance (a dominant contributor for the median performance terminal devices).

[h] The statistical data collected by the industry shows that approximately 5% of terminal devices operate at performance lower than the design goals and some are barely above the required performance levels.

RF Distribution Network and CPEs

As is apparent from Table 5, even with DBAA, 1024-QAM signal deployment would lower the operating margins by 2-3 dB. To address this degradation in operating margins, the output levels of RF actives must be increased accordingly to avoid major upgrades to in-house wiring. To verify whether there is sufficient margin in the RF network actives to achieve this, a series of test were conducted at the upper range of the RF actives operating output levels. The test setup was similar to the test setup used to determine whether analog optical links have any OMI margin available before clipping occurs. This is presented in Figure 4.

The test results (presented in Figure 5) show that the BER for channels at 999 MHz approaches 10E (-4) for 1024-QAM J.83 Annex C signals at the RF active levels higher by 1.5 dB at 1 GHz than the levels of QAM signals at 1 GHz in nodes loaded with 75 CW carriers and 75 QAM signal channels (set 6 dB lower). Results also show that for these BER numbers, there are no errors after R-S. The plots in Figure 5 also show results for 256-QAM J.83 Annex B and C signals. These test results show more than 1 dB gain in output level for the same pre-FEC BER between Annex C and Annex B. Similar gain is expected for 1024-QAM J.83 Annex B signal.

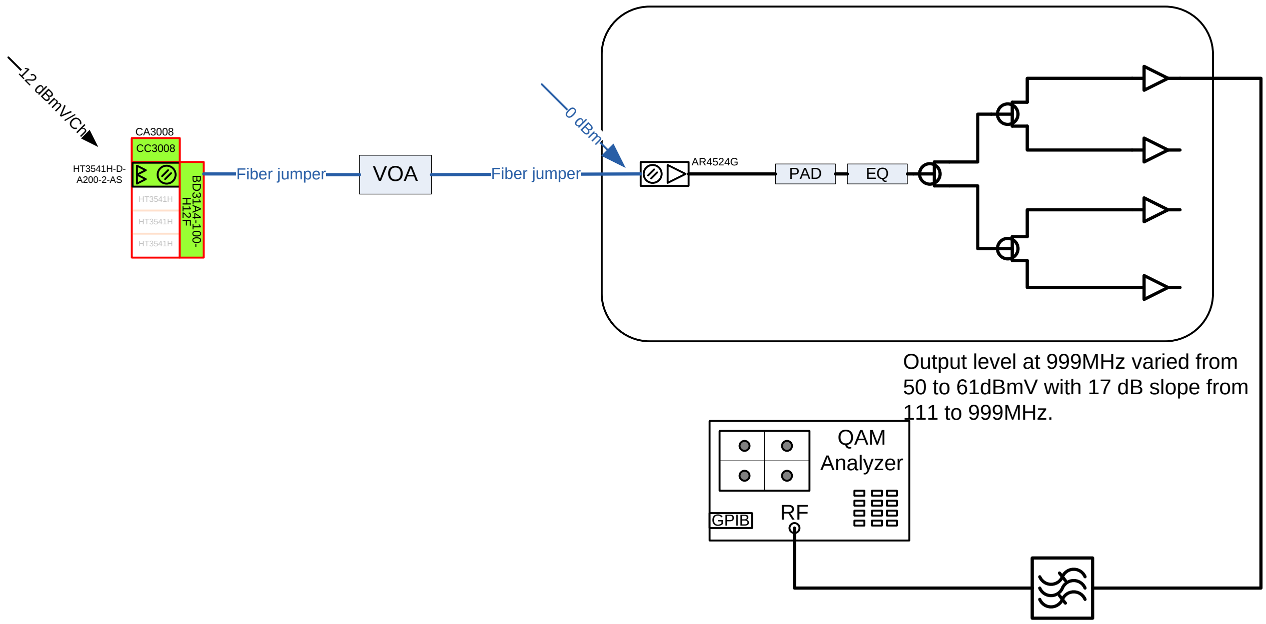
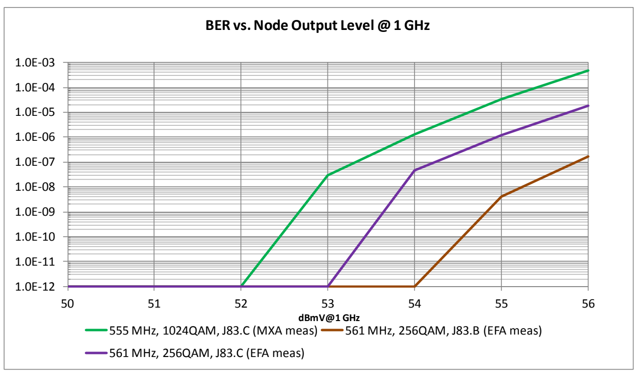
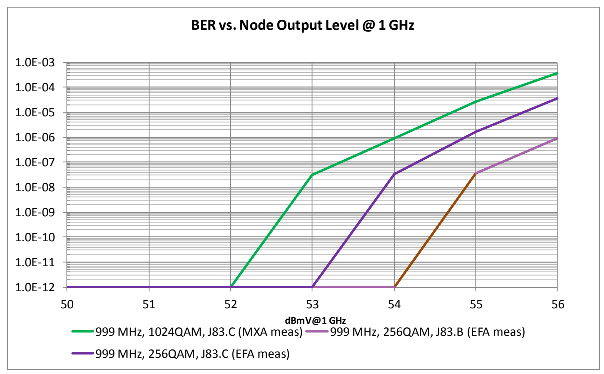


Figure 4: Test Setup: Performance versus RF Active Output Levels





a) Channel at 555 or 561 MHz **b) Channel at 999 MHz**

Figure 5: BER Performance versus RF Active (Fiber Deep Node) Output Levels

The tests were conducted for the highest levels in nodes used in a Fiber Deep architecture with no additional RF actives. In a traditional HFC architecture with RF actives, the RF active output levels are de-rated accordingly and are much lower than the clipping levels. Hence increasing their levels by up to 3 dB will result in a minimal degradation of BER, if at all, and in CIN noise cascaded at much lower levels[i].

These test results lead to an assertion that the levels in all RF actives loaded exclusively with QAM channels, especially with Annex B QAM channels, can be increased by 2 to 3 dB.

If the 2-3 dB increase in the RF active output level is difficult to maintain and provides too low an operating margin, then the split band settings with priority load that does provide sufficient operating margin can be implemented instead

Figure 6 presents the BER results as a function of the node output level with 74 channels of 6 dB higher priority load and 75

---

[i] Cascading of CIN is on a 10*log basis while the CIN levels drop at a higher rate than 1-to-1 with the drop in RF active levels, although the rate of change varies depending on the frequency and on the order of nonlinearity causing the CIN.

channels of lower priority load. For comparison, it also shows BER results for 1024-QAM J.83 Annex C signal for even level load. The results indicate that quite a significant bandwidth of priority load can be supported with 3 dB higher levels (even up to the full 894 MHz load) to 9 dB higher levels (for 444 MHz of priority load) than the levels of QAM signals with 75/75 analog/QAM load. For Annex C, level increase for acceptable performance will be 2 dB lower but more than 500 MHz of priority load at levels 3-7 dB higher can be supported.

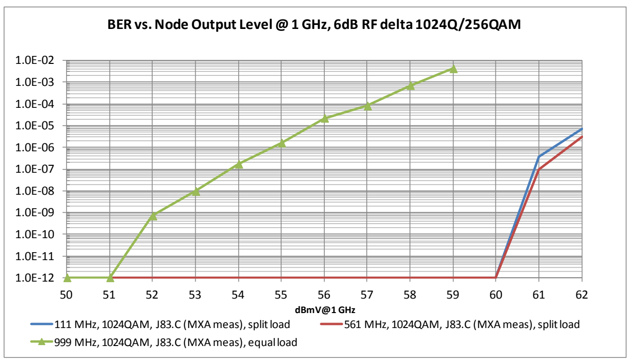


Figure 6: BER Performance versus RF Active (Fiber Deep Node) Output Levels for Split Band Load (with Priority Signals)

Table 6 presents the EOL performance for the network with DBAA and with levels increased by 3 dB (whether for the entire bandwidth or for the priority load bandwidth). Note that the results presented in the table are for the highest design frequency. Terminal devices will see higher levels at lower frequencies unless significant forward slope reaches the in-house wiring.

The results presented in Table 6 show that we achieved similar operating margins for 1024-QAM J.83 Annex B signals as in the original network design (see Table 4) with analog optical links. These limited network upgrades thus allow delivery of 1024-QAM J.83 Annex B signals to terminal devices within the entire design bandwidth or at least in a sizeable portion. Hence, we can increase network capacity by up to 25%. This translates to the following capacity gains:

- 1 Gbps in 256-QAM only load 750 MHz network (from 4 Gbps to 5 Gbps);
- 1.2 Gbps in 870 MHz network (from 4.8 Gbps to 6 Gbps);
- 1.4 Gbps in 1002 MHz network (from 5.6 Gbps to 7 Gbps).

Table 6: EOL Signal Performance after DBAA Implementation and 3 dB RF Active Level Increase

QAM Signal Performance						
Original design Freq & Type	Modulator Performance [dB]	RF Cascade performance [dB]	Median CPE CNR @ Highest [dB]	Design CPE CNR @ Highest Freq.	EOL CNR @ Highest Freq./Median Level	EOL CNR @ Highest Freq./Design Level [dB]
750 MHz N+5	43.0	52.8	48.0	41.0	41.5	38.7
	43.0	55.2	48.0	41.0	41.6	38.8
860 MHz N+5	43.0	59.0	48.8	41.0	41.9	38.8
	43.0	59.2	48.8	41.0	41.9	38.8
860 MHz FD	43.0	NA	44.6	41.0	40.7	38.9
1002 MHz FD	43.0	NA	48.3	41.0	41.9	38.9

TERMINAL DEVICES

The previous sections showed that with the deployment of DBAA and after realignment of RF actives in the HFC network, we can improve the network performance to enable transport of 1024-QAM J.83 signals for up to 25% capacity expansion without bandwidth upgrades while maintaining similar operational margins. The question is whether network operators can take immediate advantage of this development. The authors analyzed the availability of terminal devices capable of transmitting and receiving 1024-

QAM J.83 signals. Implementation of this level of modulation would allow extra margin of time for the deployment of DOCSIS 3.1 PHY with all its increased complexity and cost related to the technology replacement cycle for only 3 dB performance gain at the cost of higher overhead[j] and hence lower capacity gain. A slightly longer equalization tap chain[8] would allow up to 32 channel pooling for up to 1.5 Gbps communication trunk size with 1024-QAM J.83 Annex B signals. Some silicon vendors did implement higher modulation levels in their silicon despite the fact that they were not specified by the standard. This silicon has been used in many CPEs deployed by network operators and is commercially available. As for the signal sources, some CMTS are capable of generating 1024-QAM signals. But most importantly, DBAA modules that are FPGA based are designed to support 1024-QAM transmission and hence can deliver video and high-speed data at 25% higher speeds to the CPE devices.

<u>CPE Devices for 1024-QAM Signal Reception: Are They Capable and How To Make Them Ready</u>

When assessing a cable modem or set top box for its ability to support 1024-QAM demodulation, three things need to be taken into consideration: the performance of the tuner; the capabilities of the demodulator; and the ability of the software to accommodate the higher data rates and the additional modulation rate.

1024-QAM demodulator support in cable modem and set top box silicon has been present in certain chipsets for over a decade. See Appendix A for a partial list. However, the ability to reliably manufacture affordable tuners with sufficient SNR lagged by several years, due to the difficulty in mitigating traditional impairments such as frequency error, microphonics, linearity, thermal noise, and phase noise. With only a few exceptions, the tuners of adequate performance to support 1024-QAM reception and demodulation have all been silicon-based implementations.

Silicon support for 1024-QAM is not sufficient for end-to-end 1024-QAM transport support in cable modems and set top boxes. CPE firmware must also be enabled to direct the device to lock onto and demodulate 1024-QAM signals, as well as accommodate the higher data processing rates required. Until the recent introduction of miniCMTS products, there were no commercial deployments of 1024-QAM, so CPE vendors disabled or removed firmware support for 1024-QAM, in order to improve channel ranging and acquisition speeds and reduce memory requirements.

The process of re-enabling 1024-QAM support in silicon-capable devices differs among different device models and implementations, depending on the current version of firmware. In some, it can be as simple as using a CLI parameter to enable the functionality. In others, a compile-time option includes or excludes 1024-QAM support. On the other end of the difficulty spectrum, multiple libraries and routines might have to be modified. Note that "silicon-capable" refers not just to the tuners and demodulators with support for 1024-QAM, but also to the amount of RAM required to support the larger firmware images.

<u>CPE Device Snapshot: 1024-QAM J.83 Capability</u>

Some current-generation cable modems and set top boxes can support 1024-QAM modulation (see Appendix B for just a few examples). Some older, deployed devices

[j] NG PHY at 3 dB higher performance than that required for 256-QAM J.83 Annex B signal allows 1024-QAM modulation at <17% capacity gain while 512-QAM J.83 if available would deliver 12.5% capacity gain at 3 dB better performance

may also be able to support 1024-QAM, either with existing firmware or with new firmware. Because the logistics of supplying 1024-QAM-capable equipment is much simpler for a node-wide population (rather than plant-wide), increasing spectral efficiency is a viable method of adding plant bandwidth (both data *AND* video) without having to expand the bandwidth or replace terminal devices for ones capable of DOCSIS 3.1 PHY or having to split the nodes and incur the extra space/power/cooling burden of the replicated headend equipment.

One benefit of the approach analyzed in this paper, which cannot be provided by DOCSIS 3.1 PHY, is that downstream data delivery over DOCSIS 3.0 and its older versions and MPEG-2 video delivery can both benefit from 1024-QAM signals in J.83 format while DOCSIS 3.1 improves only IP traffic options. Basically, DOCSIS 3.1 can improve only IP traffic bandwidth efficiency (data, IP video and other IP services) while 1024-QAM J.83 signals can improve bandwidth efficiency by 25% without bandwidth expansion and CPE equipment replacement (pending the system-by-system and operator-by-operator assessments of the CPE equipment capability) with deployment of DBAA modules and CPE upgrades (often with remote upgrade implementation). Considering the deployed volume of CPE equipment with J.83 capability, this additional benefit of increasing the digital video bandwidth efficiency is not something to disregard lightheartedly.

CAPACITY EXPANSION OPTIONS

The DBAA implementation provides several options for upgrading capacity. The opportunities for the capacity expansion differ in possible timing and cost of implementation. Table 7 lists several opportunities ranked starting from the ones that are available sooner and at a lower burden to network operators to ones that could be implemented at higher cost and at a later time based on terminal equipment technology availability. The following subsections describe these options in detail and list some approaches and guidelines for their implementations.

Performance Upgrade within Existing Bandwidth

Deployment of DBAA modules with 1024-QAM J.83 capability makes this option the most appealing from cost and time-to-market points of view.

In traditional HFC networks, supporting 1024-QAM modulation plant-wide is a tricky proposition due to the variability of end-of-line network performance. However, in a distributed architecture model, the use of 1024-QAM is more easily phased in by virtue of the fact that QAM devices are addressable (uniquely identifiable), and 1024-QAM-capable plant conditions are created node-by-node (or service group-by-service group). The limited population makes it easier to upgrade or exchange CPE devices, and enables service upgrades to be done as local needs dictate. For example, if an area serviced by a single node was in need of increased data bandwidth, then by upgrading that node to Hybrid DBAA and using node-based modulators only for the DOCSIS traffic, the frequency allocated to that service could support a 25% increase without switching to new technologies. The only changes, beyond installing the node-based modulators, would be to upgrade the M-CMTS and cable modems for that node to support 1024-QAM. Narrowcast video would require an upgrade to the VoD/SDV resource manager, to accommodate the higher throughput.

After deployment of DBAA, the RF coaxial part of the network becomes the only contributor to the signal impairments but by itself is not materially contributing to the impairments of even 1024-QAM signals as long as it is reasonably maintained. It also can be re-aligned for higher levels in the entire or in a significant part of the bandwidth to increase CPE device input level and support 1024-QAM with operational margins, similar to the margins for 256-QAM signals, to which the network is designed today. It is actually expected that the elimination of the analog optical links with all the intricacies of fiber design will improve operational margins and lower the maintenance costs.

The most important benefit of this option is the potential for immediate implementation as a selective approach or as a wholesale upgrade of an area or system with CPE device upgrades (firmware) and re-allocation/replacement if some CPE devices are not capable of supporting 1024-QAM J.83 signals. It is also of the lowest cost and poses the least, if any, service disruption.

Table 7: HFC Network Capacity Expansion Options

Option Description	Optical Link Activity	Coaxial Plant Activity	Terminal Equipment Activity	Upper Boundary of Capacity Expansion	Earliest Implementation Timing
Performance upgrade within the existing bandwidth	Replacement (or addition to) analog optical link with DBAA node modules	Realignment of RF actives for higher output levels	Activation of 1024-QAM J.83 mode in CPEs and CMTSs if possible.	1 to 1.4 Gbps depending on upper network frequency (to 5 or 7 Gbps)	Immediate
Bandwidth and performance upgrade to 1 GHz (in the existing 1 GHz networks, refer to the option above)	Replacement (or addition to) analog optical link with DBAA node modules	Upgrading RF actives to the new 1 GHz (or 860 MHz) BW if needed with higher output levels	Activation of 1024-QAM J.83 mode in CPEs and CMTSs if possible. If needed, upgrading CPEs to 1 GHz.	From the existing capacity to 7 Gbps in 1 GHz upgraded network (from 4, 4.8 and 5.6 Gbps)	Immediate at cost of RF active upgrades if needed. No re-spacing required.
Bandwidth and performance upgrade to 1.2 GHz	Replacement (or addition to) analog optical link with DBAA node modules	Upgrading RF actives to the new 1.2 GHz BW with higher output levels. Possibly selective upgrades of passives.	Activation of 1024-QAM J.83 mode in CPEs and CMTSs if possible. If needed, upgrading CPEs to 1 GHz. Implementing NG PHY technology above 1 GHz or in the entire BW	From the existing capacity to 8.5 Gbps with NG PHY 2048-QAM to 10 Gbps with NG PHY 4K/8K QAM in the entire upgraded BW	Immediate to 1 GHz with keeping 256-QAM operation and/or activating 1024-QAM J.83 and adding NG PHY technology when available
Bandwidth and performance upgrade to 1.5 and 1.8 GHz	Replacement (or addition to) analog optical link with DBAA node modules	Upgrading coaxial plant to FD or N+0 passive coaxial network	Any combination of the above technology up to complete replacement of all with NG PHY technology	From the existing capacity to 12 Gbps with NG PHY 2048-QAM up to 1.5 GHz to 15 Gbps with NG PHY 4K/8K QAM in 1.8 GHz BW	Immediate to 1 GHz with keeping 256-QAM operation and/or activating 1024-QAM J.83 and adding NG PHY technology when available

Bandwidth and Performance Upgrade in Sub-1 GHz Networks

This approach shares many characteristics of the options described previously. A detailed analysis of the approach and the following options is presented in *Distributed Digital HFC Architecture Expands Bi-directional Capacity* from NCTA Spring Technical Forum 2013. The RF actives do not require re-spacing, the technology to expand the network to 1 GHz exists today, the passives deployed during the last 15 years are 1 GHz capable and the 1024-QAM J.83 terminal equipment availability is the same as in the option described above. The only additional requirement is to upgrade RF actives (nodes only in FD, N+0 architecture) to 1 GHz if they are not already 1 GHz capable. Their re-alignment requirements for level and BW mirror those stated above.

The upgrade can be performed selectively per node/segment or per area/system and can be implemented immediately and the capacity can be also expanded immediately. An additional benefit of this option is that expanded capacity can be realized even without upgrading terminal equipment to 1024-QAM J.83 capability while adding close to 1.5 Gbps capacity with 42 256-QAM channels between 750 MHz and 1002 MHz. The time-to-market may be even shorter than in the previous option with this approach and the upgrade of terminal devices to 1024-QAM J.83 capability can be implemented in the next step.

Bandwidth and Performance Upgrade to 1.2 GHz

This option is a reasonable step to implementation of NG PHY as an addition to and replacement of the existing digital service delivery technologies. The upper limit of this upgrade is dictated by passive performance and network upgrade limits with DBAA[9]. Most 860 MHz networks can be readily expanded to 1.2 GHz in bandwidth without RF active re-spacing. Indeed, the upper limit can be expanded to 1.3 GHz (without re-spacing) based on the performance of passive equipment. Fiber Deep (N+0) networks can be upgraded to higher bandwidth (1.5 GHz to 1.8 GHz) as dictated by the limitation of the passives and operator readiness to upgrade the passives. The RF plant upgrade in HFC networks is predicated on availability of higher bandwidth RF actives and is easier to accomplish in FD networks where only the RF section of the node needs to be upgraded. In FD networks, the RF section upgrade can actually be avoided by deploying DBAA (remote PHY) modules with BW capability above 1 GHz and combining the existing RF module outputs with couplers or filters (guard bands required) as output level capabilities for both (RF node module and DBAA module) allow.

However, the timing to take advantage of the BW upgrades is defined by availability of NG PHY terminal devices (remote PHY modules and CPE devices).

Bandwidth and Performance Upgrade to 1.5 and 1.8 GHz

This option requires implementation of FD or N+0 (passive coax) architecture in DBAA format (with remote PHY). However, in the existing FD networks, it is no different than the previous option where DBAA modules may deliver signal above 1 GHz without upgrading node RF modules unless operational and power consumption benefits promote an RF module upgrade with new silicon amplification technologies for higher reliability and lower power consumption.

In traditional HFC networks with cascades of RF actives, expanding bandwidth beyond 1.3 GHz will require drastic cascade shortening and the benefits of a passive coaxial network would usually promote FD (N+0) solutions if such an upgrade is to be undertaken by the operator to significantly increase the capacity.

The passive coaxial network can support in that case >10 Gbps downstream capacity with >2 Gbps upstream capacity (total network capacity >16 Gbps) as well as TDD technologies with all its cost and traffic management advantages.

SUMMARY: DBAA ENABLER OF IMMEDIATE AND FUTURE CAPACITY EXPANSION

It is agreed that the need for increased network capacity will continue. While there are a multitude of options available to network operators, the challenge for them today is to find the optimum solution to meet that increased capacity need.

Building upon analysis presented in earlier papers, this paper shows that the Distributed Broadband Access Architecture, a.k.a. remote PHY improves the existing network performance from a network capable of supporting 256-QAM Reed-Solomon (R-S) signals, with or without Trellis Coded Modulation (TCM), to a network capable of supporting 500-1000 MHz of 1024-QAM J.83 signals. This results in a 25% increase in network capacity within the frequency range allocated to these signals. With the presented availability of CPE devices to support this higher level modulation, this is a very efficient first step to increasing capacity.

Moreover, combining the DBAA (remote PHY architecture) with higher bandwidth options – up to 1.8 GHz, and eliminating the RF actives in the network, the potential for network capacity is significantly increased, to more than 15 Gbps, a net increase of more than 10 Gbps.

With installed coax drops into the home and the recent advances in architectures and technologies, cable operators are well-positioned to continue to lead the charge for the subscriber!

ACKNOWLEDGEMENTS

The authors express their deepest appreciation to Dawn Emms who for many years carried the burden of taking care of editing our papers and keeping our language clear and to the point. Without her help, the value of our contributions would drastically diminish.

We also want to thank the entire Aurora team, and specifically Doug Combs (who co-authored many papers) and Steve Hopkins (who co-authors this one) for always providing significant and critical contribution to network analysis and testing of the many possible permutations of choices available and hypothesis posted.

Finally, for many papers, the contribution of our vendors has been critical. Listing them all is hardly possible but they know who we mean.

Thank you all.

ABBREVIATIONS AND ACRONYMS

AGC	Automatic Gain Control
BER	Bit Error Rate
BW	Bandwidth
CCAP	Converged Cable Access Platform
CE	Consumer Electronics
CIN	Composite Intermodulation Noise
CLI	Command-line Interface
CMTS	Cable Modem Termination System
CNR	Carrier-to-Noise Ratio
CPE	Customer Premises Equipment
CW	Carrier wave
dB	decibel
DBAA	Distributed Broadband Access Architecture
DOCSIS®	Data over Cable Service Interface Specification
EOL	End-of-Line
EPoC	EPON Protocol over Coax
FD	Fiber Deep
FEC	Forward Error Correction
FPGA	Field-Programmable Gate Array
Gbps	Gigabits per second
HD	High-definition
HFC	Hybrid Fiber Coaxial
HSD	High-speed Data
IP	Internet Protocol
Mbps	Megabits per second
MER	Modulation Error Ratio
MWVL	Multiwavelength
NTSC	National Television System Committee
OMI	Optical Modulation Index
QAM	Quadrature Amplitude Modulation
R-S	Reed-Solomon
RF	Radio Frequency
ROI	Return on Investment
SDV	Switched Digital Video
SNR	Signal-to-Noise Ratio
TCM	Trellis Coded Modulation
TDD	Time Division Duplex
VoD	Video on Demand

APPENDIX A

CPE Components with 1024QAM support

Tuners and Demodulators

Chipset	**Features**	**Year**
BCM3252	Dual Front-end DOCSIS 2.0+ STB with channel bonding	2007
BCM3255	Triple Front-end STB with DOCSIS 2.0	2007
BCM3348	DOCSIS 2.0 cable modem	2003
BCM3349	DOCSIS 2.0 cable modem	2005
BCM3360	DOCSIS 1.1 gateway cable modem	2002
BCM3361	DOCSIS 2.0 gateway cable modem	2004
BCM3367	DOCSIS 2.0 MTA	2006
BCM3368	DOCSIS 2.0 cable modem with dual VOIP	2006
BCM3371	DOCSIS 2.0 cable modem gateway	2011
BCM3378	DOCSIS 2.0 cable modem with VOIP and integrated tuner	2008
BCM3379	DOCSIS 2.0 cable modem with VOIP	2009
BCM3380	DOCSIS 3.0 (8x4) cable modem gateway	2009
BCM3381	DOCSIS 2.0+ (3-channel) cable modem	2006
BCM3382	DOCSIS 3.0 (8x4) cable modem gateway	2013
BCM3383	DOCSIS 3.0 (16x4)cable gateway	2012
BCM3384	DOCSIS 3.0 (24x8) cable modem	2013
BCM33843	DOCSIS 3.0 (16x4) cable modem	2013
BCM3385	DOCSIS 3.0 (32x8) cable gateway	2013
BCM3409	STB Low Power Direct conversion cable tuner	2007
BCM3419	DOCSIS 2.0 Direct conversion cable tuner	2006
BCM3420	DOCSIS 2.0 Low power Direct-conversion cable tuner	2006
BCM3421	DOCSIS 2.0 Direct conversion cable tuner	2006
BCM3422	DOCSIS 3.0 Direct-conversion cable tuner to 1 GHz	2007
BCM3520	ATSC/NTSC/QAM Cable ready TV receiver	2006
BCM3545	QAM digital receiver	2008
BCM3560	Analog and DTV STB	2006
BCM7002	DTA	2011
BCM7003	SD Cable Interactive Receiver with USB DVR	2009
BCM7004	Basic STB SD Receiver	2009
BCM7013	Basic STB SD Interactive Receiver with Ethernet and USB DVR	2009
BCM7014	Basic STB SD	2009
BCM7110	STB with PVR and DOCSIS 2.0 cable modem	2003
BCM7115	STB with PVR	2003
BCM7118	HD STB with DOCSIS 2.0 modem	2007
BCM7583	Fullband Capture CATV Tuner + DVB-C HD STB	2014
BCM7584	Fullband Capture CATV Tuner + DVB-C HD DVR STB	2014
Temic 4937 - 3x7702	Dual conversion cable modem tuner	2002?

APPENDIX B

Examples of 1024QAM-Capable CPE Devices

Castlenet	CBC33843D	DOCSIS 3.0 (16x4)cable modem
Castlenet	CBC3383D	DOCSIS 3.0 (8x4) cable modem gateway
Kathrein	DCV8400	DOCSIS 3.0 (8x4) cable modem gateway

[1] J. Chapman, M. Emmendorfer, R. Howald, S. Shulman. *Mission is Possible: An Evolutionary Approach to Gigabit-Class DOCSIS.* NCTA Spring Technical Forum 2012

[2] O. Sniezko, D. Combs, R. Brockett. *Distributed Digital HFC Architecture Expands Bi-directional Capacity.* NCTA Spring Technical Forum 2013

[3] O. Sniezko, D. Combs, R. Brockett. *Distributed Digital HFC Architecture Expands Bi-directional Capacity.* NCTA Spring Technical Forum 2013

[4] R. Howald, S. McCarthy. *Accounting for Techies: Taking it to the Ultra.* NCTA Spring Technical Forum 2013.

[5] Robert Howald, Sean McCarthy. *Bits, Big Screens, and Biology*. NCTA Spring Technical Forum 2012.

[6] O. Sniezko, D. Combs, R. Brockett. *Distributed Digital HFC Architecture Expands Bi-directional Capacity.* NCTA Spring Technical Forum 2013

[7] O. Sniezko, D. Combs, R. Brockett. *Distributed Digital HFC Architecture Expands Bi-directional Capacity.* NCTA Spring Technical Forum 2013

[8] J. Chapman, M. Emmendorfer, R. Howald, S Shulman. *Mission is Possible: An Evolutionary Approach to Gigabit-Class DOCSIS.* NCTA Spring Technical Forum 2012

[9] O. Sniezko, D. Combs, R. Brockett. *Distributed Digital HFC Architecture Expands Bi-directional Capacity.* NCTA Spring Technical Forum 2013

Remote PHY: Why and How

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

Abstract

The success and growth in narrowcast services continues to require MSOs to expand the capacity dedicated to those services. Most MSOs have deployed 4 and 8 DOCSIS™ QAMs and considering expanding beyond that. Deployment of VOD and SDV QAMS continue to expand with HD and network DVR. And, even more QAMs are needed as service groups are segmented further in the network.

As the narrowcast service growth continues, MSOs are putting emphasis in the evolution of the access technologies to make them more efficient. DOCSIS 3.1 was created to enable more efficient use of spectrum, especially by using higher modulation orders and by taking advantage of newer technologies such as better LDPC codes and OFDM modulation.

MSOs and suppliers alike are considering one more aspect of the evolution: moving the RF modulation downstream into the network. By moving the RF modulation from the headend to the node, known as Remote PHY, it is possible to achieve important gains, such as:

<u>Performance increase:</u> by eliminating the analog laser and reducing cascades as segmentation naturally progresses it is possible to support significantly higher order modulations as SNR performance increases both in the DS and US.

<u>Cost reduction:</u> it seems quite clear that replacing the analog forward link and the analog or digital return link for an Ethernet optical link will be less expensive both from a capital and operational perspective.

<u>Operational improvements:</u> undoubtedly, the Ethernet optical link will be easier to set-up and maintain than the current HFC links, and should carry a lot more capacity at longer distances.

To that end, this paper will cover the following areas:

1. Overview of the rational for Remote PHY

2. Elaborate on the options available to implement Remote PHY, and

3. Explain ways in which it could be implemented.

Typical HFC Networks Today

Most MSO's hybrid fiber-coax (HFC) networks have been designed to either 750 or 860 MHz of spectrum capacity. If not fully utilized, it is expected that use of their capacity will be increased to the point of exhaustion as the use of DOCSIS® increases for the higher high-speed data (HSD) service tiers, additional high-definition (HD) programs for both broadcast (BC) and especially narrowcast (NC) services such as video on demand (VOD) and switched digital video (SDV) are deployed, or new services such as internet protocol (IP) video and cloud-based digital video recorder (cDVR) are added.

Proportionally few HFC networks have been deployed to operate up to 1 GHz, although all equipment available today can support the use of spectrum up to 1 GHz and even 3 GHz for some components.

In recent years the growth in, and demand for, HD programming has resulted in the need for allocation of large numbers of EIA channels 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 EIA channels than previously required. For example, a typical digital multiplex including 10 to 15 programs would require an additional 3 to 5 EIA 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

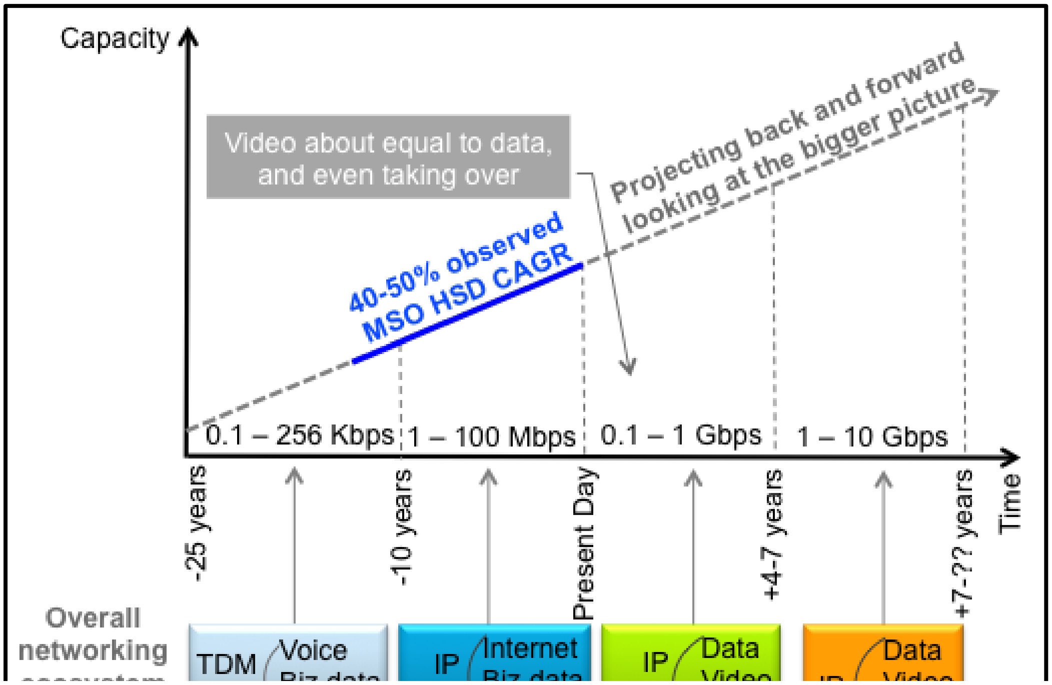


Figure 1: Example of HSD capacity utilization over time

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 distribution of its content. The difference is that the HD and SD versions of the program are not distributed unless a subscriber is requesting 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 a full-array of HD video-on-demand services. 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.

Finally, the growth in HSD services continues. Network operators have observed an increased use of HSD service capacity for well over a decade now, as shown in **Error! Reference source not found.**, which amounts to a year-over-year compounded growth of 40% to 60%. The applications have changed throughout this time, but the demand has continued to increase at the same relentless rate.

How does this compare to other operator’s data services and a longer period? As shown in **Error! Reference source not found.**, projecting the MSO’s HSD service growth back in time to when Internet services started as shown in the diagram, 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-80s, to 56 Kbps/V.42, into ISDN services.
This demonstrates that the growth seen in MSO’s HSD services is typical over a much longer period of time, rather than an exception observed by MSOs in recent years.

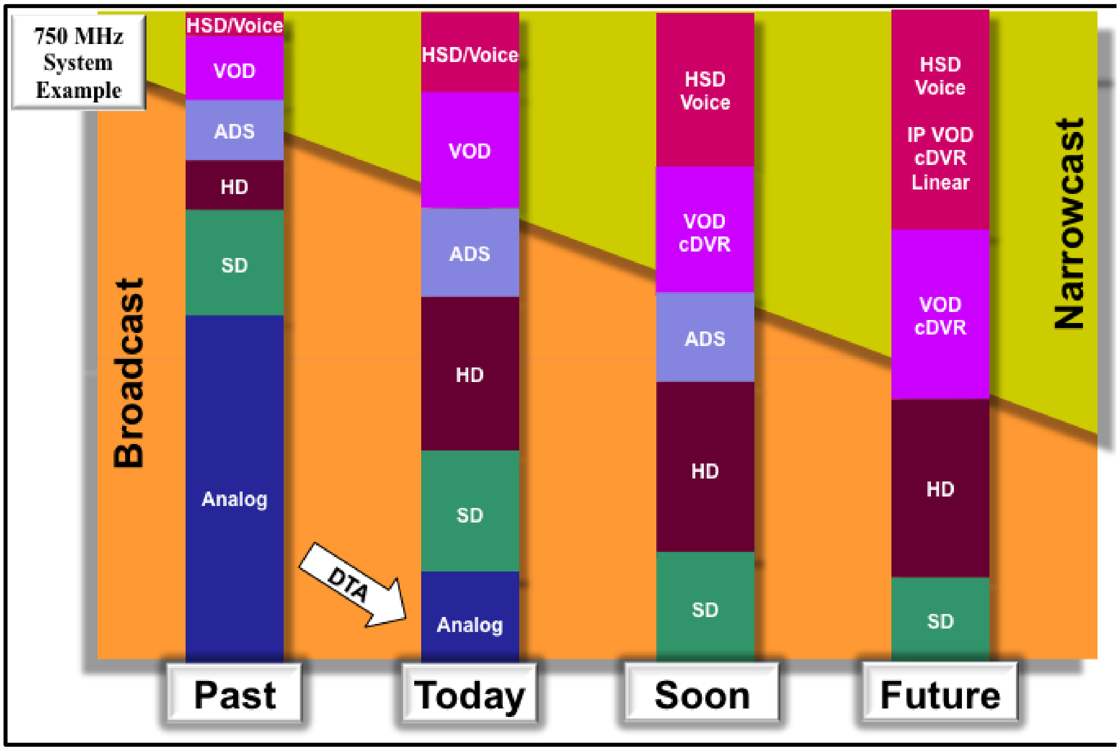


Figure 2: Example of narrowcast service growth over time

Growth Projections

From all of the above, it then follows that, should the usage growth pattern continue at the same rate as in the past, networks will be required to provide HSD services in the range approximating 1 Gbps within the next few years. This growth, coupled with the surge in HD video formats, and more personalized narrowcast services, will result in a significant growth in narrowcast capacity, as shown in **Error! Reference source not found.** below.

To support this growth, MSOs have deployed, or are considering deployment of, bandwidth reclamation tools such as SDV for digital broadcast, digital terminal adapters (DTAs) for analog services, or a combination of both. These tools have been extremely valuable to MSOs, and their operational complexity and cost well justified.

In the case of SDV, early predictions several years back from industry analysts projected that the efficiency of SDV would reach 40% (e.g., programs requiring 10 EIA channels could be carried in 6). This has proven to be understated, since it was based on the use of SDV for reduction in bandwidth required for existing services. As SDV's role in the network grew, the efficiencies have been even greater, especially as SDV has been used to introduce niche services that have low viewership and would have otherwise been difficult to deploy.

The benefit of DTAs has been just as, or perhaps even more, striking. MSOs deploying DTA devices are able to eliminate the need to distribute the analog channels in the network. Even if DTAs are distributed to top analog tier customers, such as only to subscribers of the traditional expanded basic subscribers, such deployment would reduce a channel line up from perhaps 50 EIA channels dedicated to 50 analog programs to perhaps as little as 4 EIA channels dedicated to transport the 50 programs in their equivalent digital 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 MSOs will use one or both tools, in general these tools won't be used by every MSO for all applications.

Finally, there are also significant potential gains to be achieved from the use of advanced video CODECs (AVCs) and variable bit-rate (VBR). In the case of AVCs, coding efficiencies of approximately 50%, depending on implementation and content type, can be obtain with H.264[1] and/or MPEG-4 Part 10[2]. Furthermore, with the recent release of the H.265[3] standard in April of 2013, it is possible to achieve a 50% improvement over H.264. And the use of VBR promises to result in a capacity efficiency gain of as much as 70% versus CBR[4]. The combined gains from using the above approaches could be very significant.

However, these are difficult tools to take advantage on the network since proportionally relatively few legacy set-tops still support AVCs and VBR, especially the latter. These tools will likely enjoy significant support in newer, IP-video based services equipment moving forward.

But, this approach will require additional capacity on the network. This is especially true when considering that the deployment of these advanced video services will result in an additional simulcast of video programs, at least initially, which is expected since its deployment will not at least initially replace the currently deployed services.

Furthermore, ubiquitous support for such devices would require considerable spectrum if the legacy services are maintained for an extended period, as it is expected since legacy devices are and will continue to be deployed. Moreover, this increase in simultaneous use of advanced, IP video services while maintaining legacy services will be especially impacting over time as its penetration increases.

All of the above, coupled with the success experienced by MSOs in recent with business services, will likely require the deployment of IP capacity beyond what can be supported today, requiring the development of tools for increased efficiency in the use of spectrum and/or unlashing of additional spectrum in the HFC network. The following sections of this paper will enumerate ways in which this can be achieved.

The Advent of DOCSIS 3.1

As it has been pretty well advertised in the media, DOCSIS 3.1 is under development. NOTE: For further details on the DOCSIS 3.1 technology and its implementation, see the DOCSIS 3.1 Symposium planned for the SCTE Expo 2013 show.

The key motivation for the new version of the DOCSIS specification is, in a nutshell, to scale DOCSIS more efficiently, both from the cost and operations perspectives.
While for the first 10 years or more it was possible to offer Internet services and support its growth with just 1 DOCSIS channel, services today require many more channels. This is because 1 DOCSIS channel provides almost 40 Mbps, which was well above the data rate of the services offered in the past. However, the year-over-year growth drove service speeds well above the initial levels, to 20, 50 and even higher Mbps tiers today, which can't be supported by the single channel. MSOs then went to multiple DOCSIS channels, now reaching 4 and even 8 channels, and soon requiring well beyond that.

To that end, the 3 key goals and features of DOCSIS 3.1 are:

1. Much more efficient use of spectrum, with up to 50% improvement in bandwidth efficiency (or bps/Hz, resulting from:

[1] ITU-T Recommendation H.264: 2005, Advanced Video Coding for generic audio-visual services

[2] ISO/IEC 14496-10: 2005, Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding

[3] ITU-T Recommendation H.265: 2013, High efficiency video coding

[4] Capacity, Admission Control, and Variability of VBR Flows, CableLabs Winter Conference, February, 2009

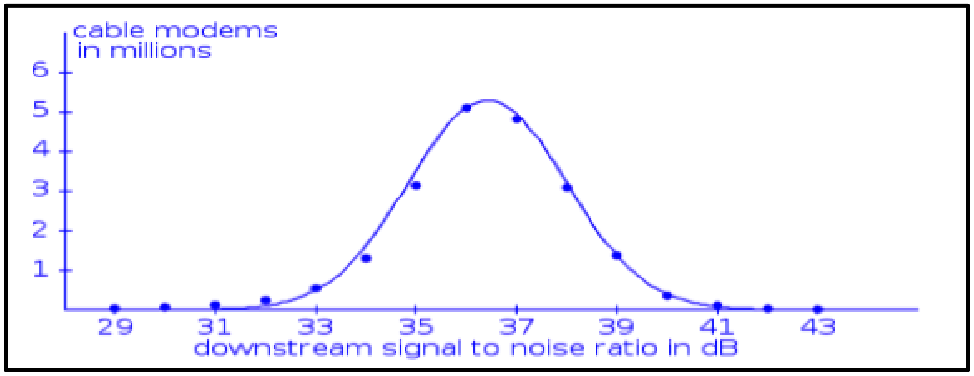


Figure 3: Example of downstream SNR for a large population of cable modems

a. The use of more efficient forward error correction (i.e., replacing the older and less efficient Reed-Solomon approach for the more modern and far more efficient Low Density Parity Check, and

b. Addition of the higher-order modulations 1024 and 4096 QAM downstream and 256 and 1024 QAM upstream.

These new modulation schemes provide 2 and 4 bits/Hertz/second improvement in both upstream and downstream, while the use of the new forward error correction approach provides approximately 5 dB better RF performance. The end result is that MSOs will be able to transport 1 Gbps of DOCSIS capacity in about 120 MHz of spectrum while doing the same with the current DOCSIS approach using single-carrier QAM requires about 180 MHz of spectrum.

2. Cost reduction, mainly by leveraging technologies commonly used in other transmission media, such as the inclusion of Orthogonal Frequency Division Multiplexing, which is used extensively in wireless and wireline transmission media. Specifically, the addition of OFDM for the downstream and OFDMA for the upstream should enable MSOs to reduce costs while “packing” more bits in the HFC network more efficiently since these technologies likely result in a larger supplier ecosystem, increasing innovation and fueling competition.

3. Enable a simple and orderly transition strategy, both with respect to compatibility with previous generation CMTS and CM equipment while supporting an expanded spectrum capacity in the HFC network.

Specifically, DOCSIS 3.1 cable modems will operate with DOCSIS 2.0 and 3.0 CMTS/CCAP equipment, enabling deployment of DOCSIS 3.1 CPE as soon as available. Similarly, DOCSIS 3.1 CCAPs will support DOCSIS 2.0 and 3.0 CPE allowing MSOs to upgrade headend equipment without having to change any of the existing CPE. And, both DOCSIS 3.1 CM and CMTS equipment will support the currently required upstream and downstream spectrum, plus an expansion of the upstream to 85 MHz and beyond, and of the downstream up to 1.2 GHz.

Error! Reference source not found. depicts the downstream signal-to-noise ratio (SNR) as reported by a very large population of cable modems [5]. This data verifies that many cable modems will be able to support the high-order modulation profiles included in DOCSIS 3.1. However, others will not without an increase in SNR.

Assuming an 8/9 coding ratio, **Error! Reference source not found.** shows the required SNR for the modulation rates included in 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[6]	39 dB
16384 QAM	42 dB

Table 1: SNR required for DOCSIS 3.1

Applying the SNR requirements from **Error! Reference source not found.** to the population

[5] Data collected by Comcast and reported to the DOCSIS 3.1 working group

[6] 8196 QAM and 16384 QAM are included for future consideration in the DOCSIS 3.1 specifications

of modems shown in **Error! Reference source not found.**, we can easily 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.

The Analog Modulated Forward Link in HFC Networks

As their name indicates, hybrid fiber-coax networks use a fiber transport between the headend and the coaxial cascade. This fiber link, intended to reduce the size of cascades, mainly driven to improve performance, was originally developed with analog modulated lasers and receivers in both directions, upstream and downstream.

Over time, the performance of the upstream link was improved by replacing the analog modulation with a digital transport. This change 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 vendors' 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 necessary 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 to date. And, while headend equipment is currently capable of launching signals with >47 dB MER performance, which would be sufficient to generate and transport 16,384 QAM signals, analog lasers are limited to about 35-38 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 the best of the cases.

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 QAM edge 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. Migration to digital video, either partially for now or already completely.

With this technology evolution, it is conceivable to remove the RF combiner network, and instead implement it digitally in the edge device, such as the CCAP.

Figure 4 conceptually depicts the output of a CCAP device.

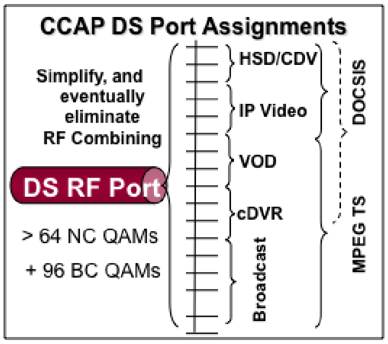


Figure 4: CCAP Downstream Port Functions

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 5, as follows:

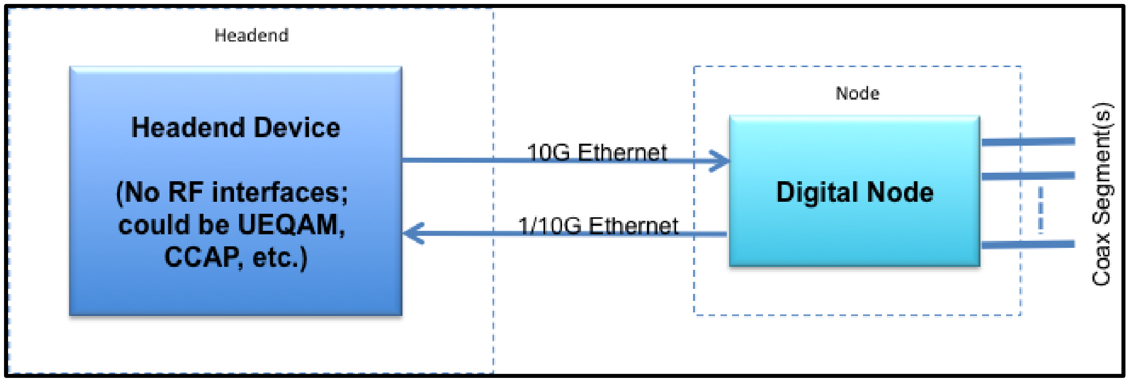


Figure 5: Digital Forward – High Level Architecture

- The headend device, such as a CCAP, which would be a high-density edge QAM comprising QAM modulation for the entire spectrum,
- The node would contain components normally implemented in the edge QAM or CCAP which generate the RF signals,
- The link between the headend device and the node would be comprised of a digital interface, such as an Ethernet link.

There are then various approaches for how a digital forward link can be implemented to replace the currently used analog link. These various approaches for distributing the various components can be categorized into 4 groups, plus 1 option that would still leave an RF generation at the headend device, as outlined in Table 2:

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 is removed from the headend device and placed in the node 5.b IP frames are transported from the headend to the node.

Table 2: Categories of options for implementing a digital forward link.

Comparison of Options for Digital Forward Link

There are pros and cons for each of the options. The following sections outline these trade-offs.

Option 1: RF remains in the headend

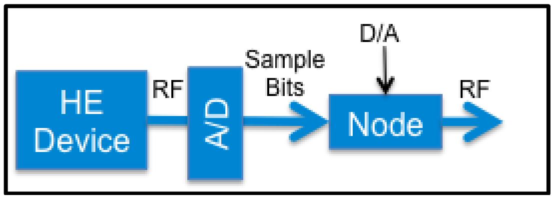


Figure 6: Block diagram for Option 1

- Equivalent to digital return, the RF output from the headend device is digitized, transported digitally, and converted back to RF in the node.
- Maintains HFC transparency
- This option results in the highest bitrate over fiber; the capacity for multiple nodes would not fit into the available capacity of one 10G fiber
- There is a loss of MER in the double conversion, so this option provides the least performance improvement
- Results in the least intelligence placed in the node, but an additional conversion stage is added in the headend

Option 2: Digital-to-analog conversion is moved to the node

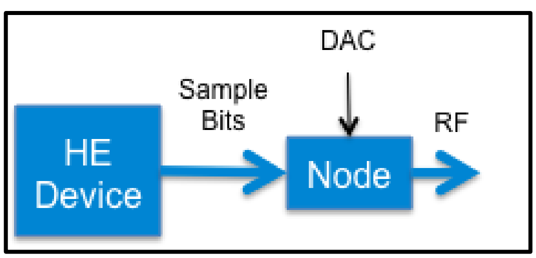


Figure 7: Block diagram for Option 2

- Requires separation of the digital-to-analog conversion from the modulator
- Together with Option 1, results in the least intelligence in node
- Similar high bitrate over fiber as Option 1; capacity for multiple nodes would not fit into the available capacity of one 10G fiber

Option 3: Lower PHY is moved to the node

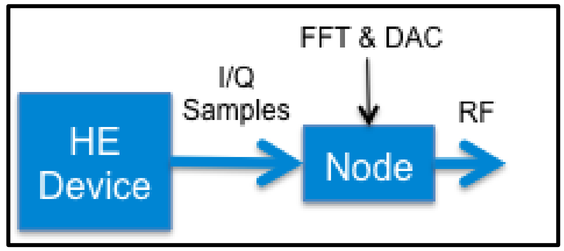


Figure 8: Block diagram for Option 3

- The PHY layer needs to be split into two components: upper and lower PHY
- More intelligence than in either of the previous options is placed in the node
- Although lower than the previous options, this option also results in a very high bitrate over fiber
- This option would require and industry proprietary point-to-point link between the headend port and the node to transport the I and Q samples
- Implementation of 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

Option 4: Entire PHY is moved to the node

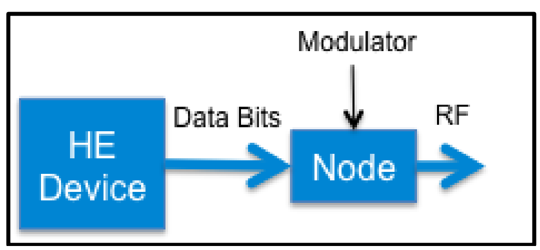


Figure 9: Block diagram for Option 4

- More intelligence is placed in the node than with all previous options
- This option results in the lowest bitrate over fiber; multiple nodes fit into the capacity of a 10G fiber

- Enables a packet-based link between the headend and node, which results in significant benefits outlined later in this paper

- Could use existing/planned silicon devices, and thus may be the easiest and quickest to implement

- Offers the best MER performance improvement over analog

Option 5: Move PHY and MAC to the node

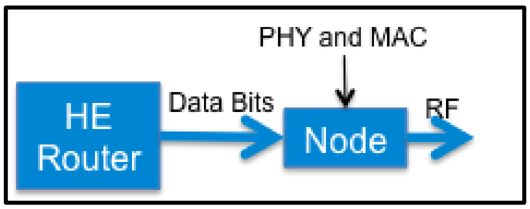


Figure 10: Block diagram of Option 5

- 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

- Same packet-based network benefits as Option 4

- Same highest MER performance as Option 4

Proposed Tenets for Digital Forward Link

In considering the various approaches for implementing the digital forward link, and the 5 options in which these approaches could be categorized, it might make sense to consider tenets for its implementation.

The following list outlines proposed basic, underlining tenets for digital forward link:

1. Headend and node devices for digital forward link should be interoperable

2. Limit the specifications to the areas that are absolutely needed for interoperability

3. Minimize the electronics that is housed in the node to the extent possible

4. Minimize the software that is placed to run in the node

5. Minimize the amount of capacity needed in the optical link

6. Keep as much of the higher layers as possible in the headend

7. Make the timing requirements for the node as simple as possible

8. Keep the independence between the DS and US as much as possible

9. Maintain the digital forward link independent from the DOCSIS version

Additional objectives could be established that would further limit the options to be considered for the digital forward link. What follows are additional proposed objectives:

A. Develop an architecture that enables scalability as capacity is needed over time

B. Minimize the need for replacing the node components as additional capacity is needed

C. Leave system components that scale with capacity in the headend

D. Use technologies used in other communications protocols when possible

E. Minimize space and power requirements in the headend

F. Minimize power requirements in the node, targeting the power consumption of a line extender as the maximum power requirement

G. Enable the use of the digital forward link for other networking functions

		BDF/BDR	Remote DAC	Remote Lower PHY	Remote PHY	Remote CCAP
Basic Tenets	1	✓	Should be	Should be	✓	Should be
	2	✓	✓	✓	✓	Should be
	3	✓	✓	✓	✓	✗
	4	✓	✓	✓	✓	✗
	5	✗	✗	Not likely	✓	✓
	6	✓	✓	✓	✓	✗
	7	✓	Should be	Should be	Should be	Should be
	8	✓	✓	✓	✓	Not likely
	9	✓	Should be	Should be	Should be	Not likely
nal Objectives	A	✓	✓	✓	✓	✗
	B	✓	✓	✓	✓	Not likely
	C	✓	✓	✓	✓	✗
	D	✗	✗	✗	✓	✓
	E	✗	Not likely	✓	✓	✓

Table 3: Comparison of tenets and additional objectives for digital forward link options

Comparison of Digital Forward Link Options

is an analysis of the pros and cons for each of the 5 options considering the tenets and additional objectives outlined above:

Given the comparison of the applicability of each of the proposed tenets and proposed additional objectives outlined above, it appears clear that Option 4, Remote PHY, is the best target for the proposed tenets and objectives.

High-Level Overview of Remote PHY

At the highest level, the Remote PHY separates the PHY from the CCAP device, and places it in the node. As shown in Figure 4, the CCAP device continues to provide all MAC and higher-layer functions, while the PHY modulation and demodulation is moved to the Remote PHY node (RPN).

The interface between the CCAP and the RPN could be any digital link. However, Ethernet links are very appropriate for this application given their preponderance in the market, ability to scale as capacity growth demands it, and low cost resulting from very wide market use.

The downstream link capacity would have to support at least 5 Gbps, which is required for an all-digital. Therefore, a 10 Gbps Ethernet interface is appropriate for such link.

However, the upstream capacity could be significantly lower. Given today's requirements, and considering an expected growth of the upstream capacity, it seems that an upstream link

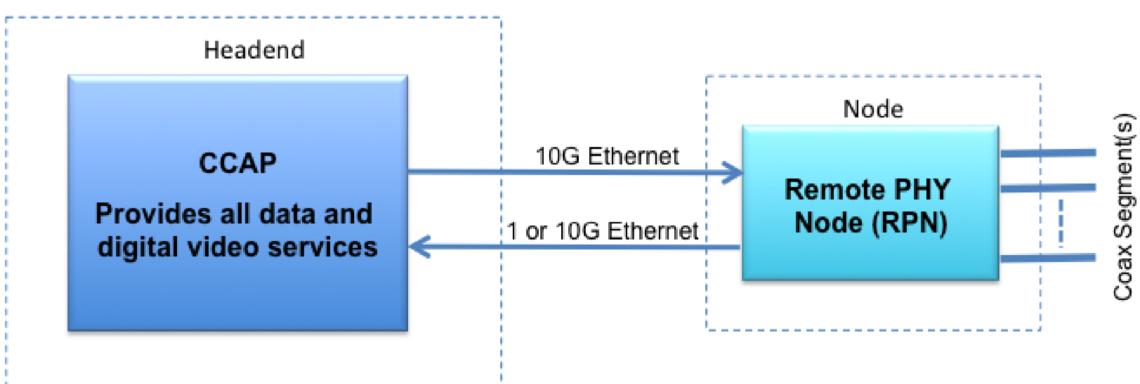


Figure 11: High-level overview of Remote PHY

with a capacity of 1 Gbps might be sufficient. This may reduce cost and simplify implementation of the Remote PHY node. However, for simplicity it may just be easier to use a 10 Gbps symmetrical link.

Benefits of Remote PHY

As described above, one key benefit of Remote PHY 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 at 5 dB of improved signal-to-noise ratio at the end of the line. This gain will result in higher capacity/Hz as it will be possible to run higher order modulations for more of the cable modems in the network.

In addition, Remote PHY will offer the benefit of enabling 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 RPN would enable MSOs 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 MSOs, especially as segmentation of the network continues towards smaller service groups, for which additional CCAP equipment needs to be deployed.

A third benefit from Remote PHY is improved reliability of the optical link. It is well known that analog links require period 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 MSOs.

Increased Headend Equipment Density

The implementation of Remote PHY 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 Remote PHY line card would be 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 8 or perhaps 12 RF ports because of the printed circuit board space required 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=]p0-o9w2q chassis.

Finally, it is possible to consider "daisy chaining" RPNs off of a single CCAP Ethernet port. This is because, on the one hand the capacity of an 10 Gbps Ethernet link would support the capacity needed for a single RPN, plus in addition it is possible to generate an RPN "channel line-up" by transmitting the broadcast content once to multiple RPNs. As depicted in Figure 12, 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 line-up content to recreate the individual RPN channel line-up. In this way each service group served by the CCAP port would contain the same broadcast line-up but its individually different narrowcast line-up.

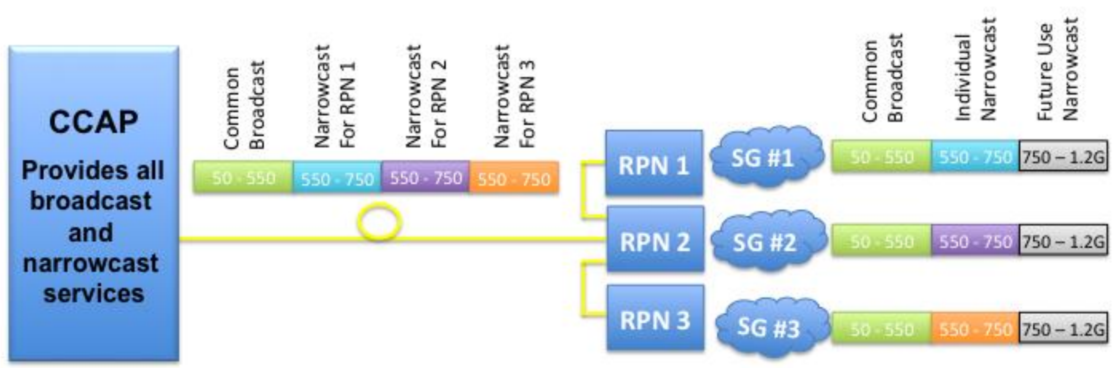


Figure 12: Reuse of broadcast capacity across multiple RPNs

Then, as the narrowcast line-up capacity grows over time, CCAP ports would be segmented to support less RPNs, akin to the way service groups are split today to support more narrowcast capacity as it is required.

As summarized in Table **4** below, the combined effect of the 3 factors described above is very significant, ranging from 8x to 18x of headend capacity gain. From a space and power perspective, this is hugely impacting savings.

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

Table 4: Remote PHY headend density gain

But, just how meaningful is this headend density gain?

Considering that a migration from an HFC architecture with an average of N+5 (i.e., a node plus 5 amplifiers in average) to N+0 would require about 10x the number of nodes, the headend density benefits resulting from Remote PHY would neutralize the potential increase in CCAP equipment.

It is then quite clear that from a space and power savings, Remote PHY takes the benefit of CCAP to a whole new level.

Integration of HFC and Fiber Services

One of the largest areas of growth for MSOs is business services. MSOs have deployed business services via both cable modems and fiber-based infrastructure. The fiber-based services are either point-to-point, using Ethernet and wave-division multiplexing (WDM), or point-to-multipoint, using PON technologies (either EPON or GPON).

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 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 subscriber, which is used for commercial services.

Given the use of a digital fiber in both the forward and the return for Remote PHY, and especially because this digital fiber is based on Ethernet technology, it is possible to collapse both of these networks into a single infrastructure.

Therefore, the implementation of Remote PHY with an Ethernet interface between the CCAP and the RPN would make it be possible to implement a PON interface at the RPN.

The benefits from this integration include:

- Reduce the optical link for PON to the distance between the node to the customer premise
    - Since the typical distance from a node to a customer premise in an N+0 architecture would be 1-2 kilometers. This would virtually eliminate any distance limitations for PON, making it possible to implement the largest possible densities as network capacity enable.
    - 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 especially 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.

Migration Strategy

Clearly, one of the more concerning issues to MSOs is the migration strategy.

Any migration that requires synchronized cut-overs, 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 Remote PHY allow for unsynchronized changes.

Furthermore, ideally the migration to Remote PHY 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 an MDU to increase capacity.

As it turns our, Remote PHY 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 Remote PHY.

Starting with the components of the network on both sides of the Remote PHY, 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 Remote PHY, and any additional MIBs for management and/or commands for configuration as needed can be added well before the first Remote PHY CCAP line card or node are deployed. With respect to customer premise devices, these would not be affected in any way in order to deploy Remote PHY, and any enhancements that are made possible through the introduction of Remote PHY would be implemented in CPE equipment that can be introduced before or after the migration to Remote PHY.

The critical portion of the network were changes need to be made are in the headend and the plant.

To begin with, the changes required in the headend are principally 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 Remote PHY line cards in the same chassis. While some MSOs may chose to deploy a separate CCAP platform for Remote PHY, it is certainly possible to support both types of line cards in the same chassis. Of course, these Remote PHY 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 Remote PHY line cards or nodes.

Turning our attention to the plant, it is similarly possible to migrate regular nodes to Remote PHY 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 Remote PHY.

Figure **13** depicts a single HFC node connected to a CCAP device.

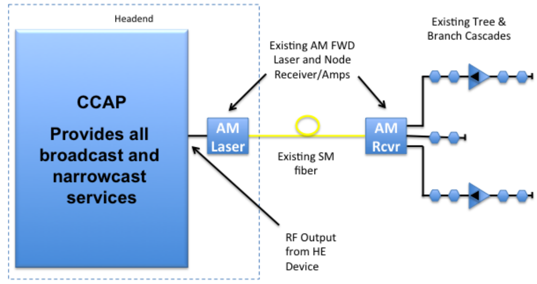


Figure 13: Single traditional HFC node

Figure **14** shows how the HFC node would be converted to Remote PHY while the rest of the HFC network remains unchanged. The Remote PHY line card in the CCAP would have been deployed in the headend a priori, and even the Remote PHY node could have been deployed before the day of the cut-over. Then, the day of the change the fiber cable could be swung in the headend from one AM laser to the CCAP Remote PHY line card, and in the field from the HFC node to the Remote PHY node. Of course it is not necessary to perform the migration in such a fashion, but it would be possible if desired.

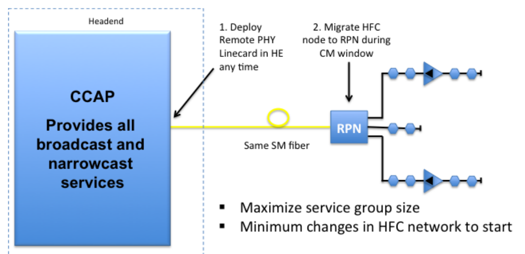


Figure 14: Remote PHY deployment step 1

Figure **15** depicts a possible step 2 in the process, whereby additional Remote PHY nodes are installed to segment the original service group further. These additional Remote PHY nodes could be daisy chained from the original Remote PHY node 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 Remote PHY nodes in an N+0 HFC architecture with some turn-around of passive components.

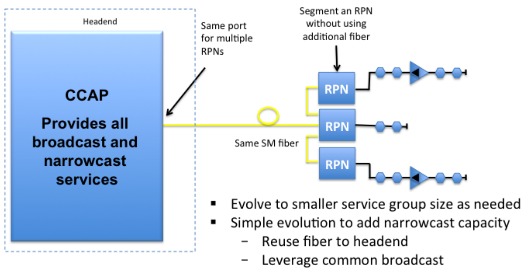


Figure 15: Remote PHY deployment step 2

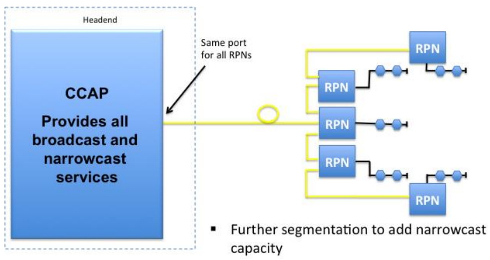


Figure **16** shows how further segmentation could take place by replacing the remaining amplifiers in the network with Remote PHY nodes.

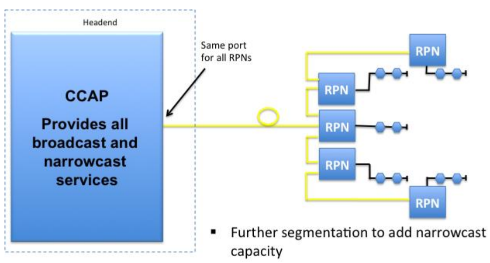


Figure 16: Remote PHY deployment step 3

Figure 17 shows the Remote PHY service group depicted above is segmented as additional narrowcast capacity is required. In this example, 2 of the RPNs from the Remote PHY service group shown in Figure 16 are split into separate service groups using separate CCAP ports.

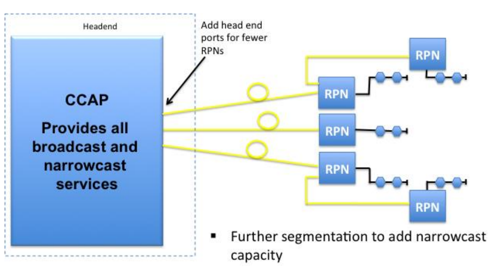


Figure 17: Remote PHY deployment step 4

Eventually each of the RPNs could be connected to an individual CCAP Remote PHY port. 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.

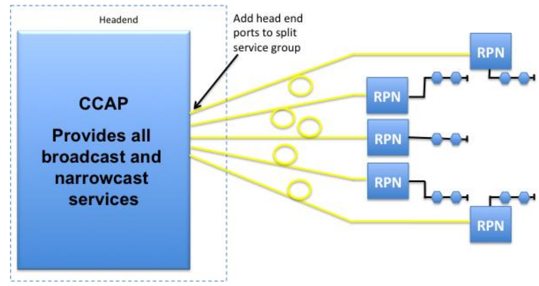


Figure 18: Remote PHY deployment step 5

Similarly, the Remote PHY 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.

Implementation and Interfaces

Having determined that a digital forward link is beneficial, and that Remote PHY is the best approach for doing so, the next step is to determine how best to move forward.

One key consideration for implementing Remote PHY is to enable interoperability between CCAP and RPN implementations. This would enable supplier with expertise in CCAP to focus their efforts on implementing the denser line cards that would interface with the RPN, and suppliers that have developed an expertise on nodes to implement the RPN device.

While one of the tenets outlined above is to minimize specifications to what is absolutely needed (see tenet #2 above), clearly describing an open interface for this purpose will be valuable to the market. Examples abound for how standards such as the one referenced here have improved choice and reduced costs for MSOs, and made the supplier ecosystem stronger. With such interface it would not only be possible to maintain and even increase the system supplier base, but also to develop a strong silicon supplier base.

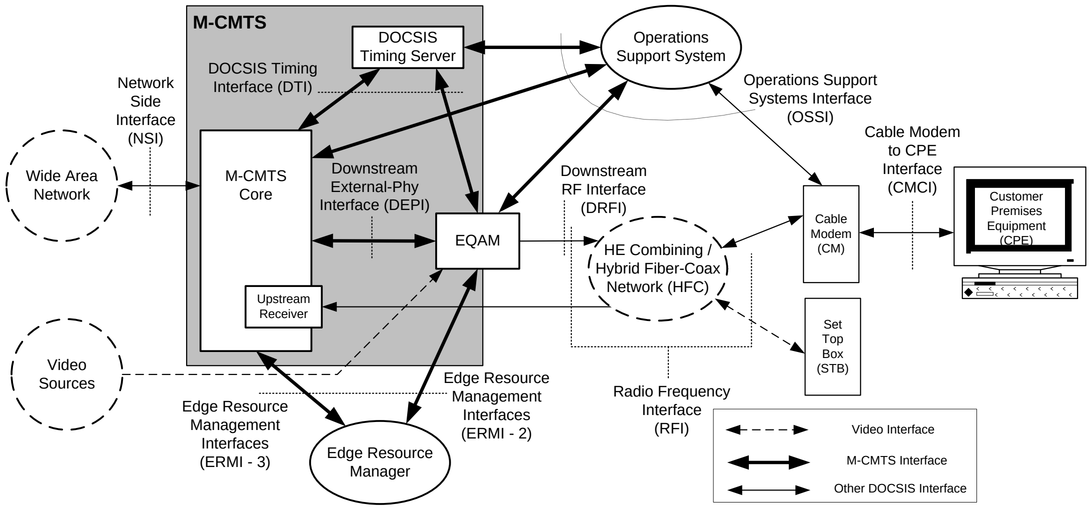


Figure 19: CableLabs Modular Headend Reference Architecture

To that end, the logical path is to expand on specifications that already exist in the market. A set of specifications developed by CableLabs a few years ago, collectively known as Modular Headend Architecture[7] (MHA) already describe the separation of the physical layer in the forward direction.

As shown in Figure 19, MHA defines several interfaces, including the Downstream External PHY Interface (DEPI) and the DOCSIS Timing Interface (DTI), amongst others. These interfaces make it possible to separate the RF downstream modulation from the DOCSIS CMTS into an external Edge QAM.

At the time a separation of the upstream RF demodulation was considered, but not described in the form of a specification. However, such interface, known as Upstream External PHY Interface (UEPI), was actually developed privately and implemented between the commercially available upstream burst receivers and the upstream media access control (MAC) layer.

It then makes sense to extend these existing specifications to support Remote PHY. The following would be required in order to do so:

A. Expand the DEPI pseudo-wire specification to support video, since the current version of DEPI only supports DOCSIS. In addition, it makes sense to update DEPI to support the latest version of the DOCSIS specifications, DOCSIS 3.1.

B. Similarly, expand the UEPI pseudo-wire specification to support DOCSIS 3.1, and publish the specification.

C. Develop a new specification to support the physical separation of the resulting CCAP Core from the RPN. Unlike in the MHA architecture, where the Edge QAM is located within a very short distance from the CMTS Core (i.e., within the same building, a few racks apart), in the Remote PHY architecture the distance from the CCAP Core to the RPN will span 10s and perhaps even 100s of kilometers. This will be a new specification, called R-DTI, which will leverage advances in timing sourcing and recovery developed by the Institute of Electrical and Electronics Engineers (IEEE), known as IEEE-1588 and included in DOCSIS 3.1.

The above 3 specifications, DEPI, UEPI and R-DTI, are being developed and should be published shortly.

However, by themselves the above interface specifications do not define the operational requirements for the RPN, or how the Remote PHY CCAP line card should be implemented to support operator’s requirements such as to handle video out-of-band (OOB). Such requirements will be included in a set of product requirements.

[7] DOCSIS Modular Headend Architecture, CM-TR-MHA-V02-081209

Remote PHY Product Requirements

Following the completion of the DEPI, UEPI and R-DTI specifications, Cable MSOs will develop a set of specifications to define the operational requirements for Remote PHY. Similarly to other efforts where suppliers require guidance from operators to develop equipment, the Remote PHY Product Requirements will define the following:

I. RPN Hardware and Functions. These requirements will include information on topics such as: environmental requirements for the RPN enclosure, number and capacity of interfaces, RF requirements, scaling factors, handling of OOB, etc.

II. Remote PHY CCAP Line Card. These objectives and requirements will include requirements such as interfaces, scaling, redundancy, etc., and will be included in the CCAP Hardware and Functions Specification.

III. Remote PHY Configuration and Management. The requirements contained in this specification will allow an RPN device developed by any supplier to be configured by any Remote PHY-capable CCAP developed by any other supplier. In addition, the requirements contained within this specification will enable an MSO to directly monitor and manage a RPN, much in the same way as it is done with DOCSIS Cable Modems today.

Conclusions

Capacity for narrowcast services in HFC networks continues to increase. MSOs continue using well-known techniques for increasing capacity through service group segmentation.

With the advent of DOCSIS 3.1 MSOs will have the opportunity to use much higher modulation orders, which will result in more efficient use of RF spectrum. Higher end-of-line SNR than currently implemented will be required to enable such higher modulation orders for the majority of cable modems.

One area where improvements can be made is in the optical link from the headend to the node. The performance of the currently used analog-modulated link can be improved by converting it to digital.

While there may be several approaches to convert the link to digital, the one that seems most appropriate is that of moving the RF modulation and demodulation to the node. In this approach the physical layer is moved to the node, called Remote PHY node, or RPN, while all other functions remain in the CCAP. The link between the CCAP and the RPN is then implemented using Ethernet interfaces.

Remote PHY offers many benefits, including: improved performance, enable longer distances, and improved reliability.

Remote PHY makes it possible to increase headend equipment density. This results from several factors, such as: combined upstream/downstream line cards, increased port density, and “daisy chaining” RPNs. The combined effect of these density factors results in a density gain of 8x to 18x.

In addition, the use of an Ethernet link between the CCAP and the RPN makes it possible to integrate fiber-based services into a single consolidated network. For example, it is possible to implement a PON OLT interface from the RPN, which being within 1-2 kilometers from the customer premise would enable higher splits and/or the use of lower cost optics.

The migration to Remote PHY would be a very smooth one, requiring no synchronization between network and customer premise changes. Migration could begin with the deployment of CCAP Remote PHY line cards, followed by migration to Remote PHY nodes on a case-by-case basis.

Finally, implementation of Remote PHY will be described in both interface and product specifications. The interface specifications will be derived from the existing CableLabs Modular Headend Architecture, and will include: an updated version of DEPI, an updated and published version of UEPI, and a new R-DTI timing interface. The product specifications will include: RPN Hardware and Functions Specification, Remote PHY CCAP Line Card (to be included in the CCAP Hardware and Functions Specification), and the Remote PHY Configuration and Management Specification.

Scaling Traditional CCAP To Meet The Capacity Needs Of The Next Decade

John Ulm, Tom Cloonan
ARRIS, Inc.

Abstract

Traditional CCAP systems have come under siege from new distributed access architectures such as Remote PHY and Remote CCAP. This paper takes a very detailed look at how traditional CCAP systems might evolve over the next decade to meet the expected growth demands.

The paper shows how traditional CCAP + AM Optic systems could achieve a 10X increase in Service Groups per head end within today's existing footprint. In addition to this, a 10X increase in capacity can be expected thanks to DOCSIS® 3.1. These capacities can even be achieved at extended distances (e.g. 40-80km) and with dozens of wavelengths.

INTRODUCTION

Much recent industry discussion has focused on new distributed architectures such as Remote PHY and Remote CCAP. Their claims have been that traditional head end based CCAP systems can't scale to meet the space and power requirements for fiber deep architectures needing many node splits; and they can't support modulations rates to take full advantage of DOCSIS® 3.1. Is this really true? The purpose of this paper is take a very detailed look at traditional CCAP systems and see how they may adapt to head ends with significant growth in Service Groups (SG).

We start by reviewing the current state of head ends and the impact of installing today's first generation CCAP platform. Topics discussed will include space, power, RF combining and Ethernet interconnection issues. From here, we investigate anticipated improvements we can expect to see in CCAP technology over the coming years. This analysis includes insights from Koomey's Law and Dennard's Scaling Law, lesser known cousins to Moore's Law.

We also take a look at AM Optic technology and the implications of pushing fiber deeper. As SGs get smaller with each successive node split, there is a double whammy of reduced capacity gains and increased costs. How deep should operators push fiber before reaching the point of diminishing returns?

As we pull together all of this information, the results may surprise some operators as to the longevity of traditional CCAP solutions in today's head end. Don't be fooled so quickly by shiny new objects.

MSO CONCERNS WITH TRADITIONAL CCAP

Head End Space & Power

With the continued 50% growth rates in capacities as shown in [CLOONAN1], some operators are concerned that they may need to continue to split nodes until they reach N+0 systems and need a dozen or more times the number of Service Groups (SG) than they have today. There is a fear among some that traditional CCAP boxes will not keep pace with this growth in SG. This could result in operators running out of both space and power in their existing head end facilities. We will show in this paper that this fear may not be well-founded.

Capacity Limitations due to AM Optics

Today's classic HFC network sends broadband signals down the fiber portion using analog based AM optics. Nonlinear Optical noise distorts QAM signals as they propagate over this fiber portion of the HFC plant. The Nonlinear Optical noise increases with longer fiber runs and more WDM lambdas per fiber. But longer distances and more wavelengths are two trends that are likely to occur with more node-splits and head end consolidations in the future.

Nonlinear Optical noise can significantly decrease SNRs and limit supported QAM modulation rates. This becomes more important with the introduction of DOCSIS® 3.1 that requires downstream support up to 4096-QAM, with optional support for 16384-QAM modulation.

While one approach to solve the Nonlinear Optical noise issue is using Distributed Access Architectures (DAA), our paper will discuss some improvements that are occurring in traditional AM optics to address these trends as well.

Even with these improvements in AM optics, there may be some use cases where a digital fiber link is desired (e.g. extreme distances &/or wavelengths). As discussed in [EMMEN], operators have a choice to either go with DAA, or they can add the digital fiber capabilities to their traditional CCAP head end systems. This approach is called Broadband Compression Forward, or BCF for short.

BCF gives the operator all of the digital optic benefits as any DAA approach such as Remote PHY. These benefits include:

- Longer Fiber reaches
- More Lambdas
- Higher SNRs, higher order QAM
- Smaller components, lower power
- “Set it & Forget it” operation

BENEFITS FROM TODAY’S CCAP, A CCAP CASE STUDY

A detailed analysis of the space and power benefits for using CCAP today was given in [ULM]. It showed that CCAP delivers on the promise of many benefits, including:

- Frees Rack Space
- Reduces head end power
- Less Network + RF Interconnections
- Fewer Boxes to Manage

The case study looked at a range of head ends from different operators: from moderate sized suburban hubs to massive urban master head ends; and from integrated CMTS to modular CMTS systems. In addition to these sites, another Urban Hub site that was “bursting at the seams” was also selected.

Chassis & Power Reductions

The case study shows that there is a significant reduction in the number of unique chassis in the system. This benefit is seen across all types of head ends and ranges from 80% to 95% reduction in the total number of devices in the head end. This provides operational savings as well.

The power savings from the reduced chassis are also dramatic with the larger head ends savings 50% to 63% of their CMTS + EQAM power. In addition to total power, the power per DS channel is also reduced by a factor of ten while supporting four times the narrowcast capacity.

Rack Space Savings

For most of the head ends in the case study, the CMTS equipment accounted for the bulk of the equipment rack space. For one site older EQAMs were more significant in an M-CMTS site. For four of the five head ends in the case study, equipment space savings ranged from 60% to 68%.

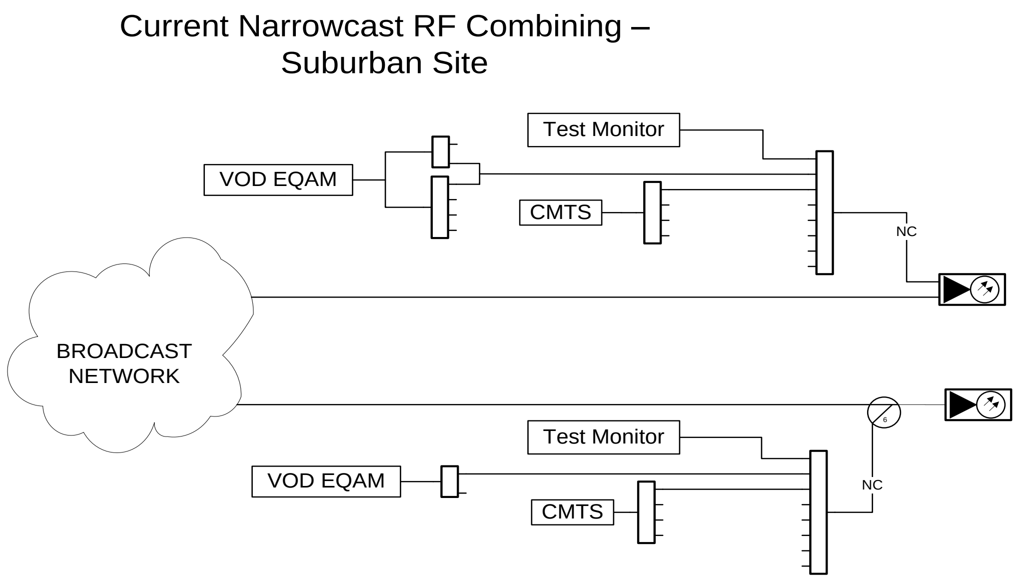


Figure 1a – RF Combining Example: Existing

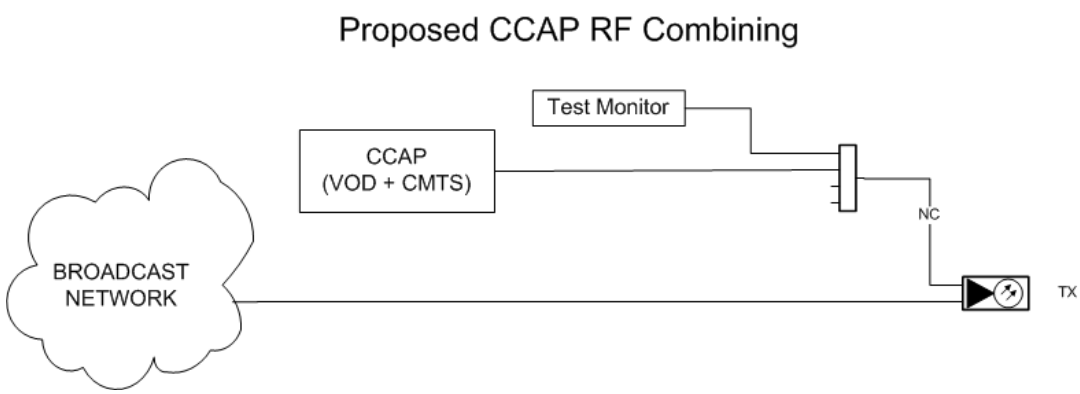


Figure 1b – RF Combining Example: After CCAP Migration

Many of these space savings were then matched with space savings from simplified RF combining. For the case study, a head end design team performed a detailed analysis for collapsing the RF combining with CCAP.

Figure 1a shows the existing RF combining design for one of the suburban hubs, followed by the CCAP design in Figure 1b. Notice that the CCAP design is still a fairly conservative design as a four way combiner was left in the CCAP path to allow for test monitoring with two spare inputs. This means that the case study numbers could be improved even further if needed.

Interestingly, one urban site saw most of its space gains from equipment reduction while the other head ends saw a more equal savings from equipment and RF combining. So in general, the RF combining savings is an equally important point to the CCAP migration. The total space savings seen at both urban master head ends, results in a dozen racks being recovered.

Interconnection Savings

There is more to the CCAP space savings story than just the savings from reduced equipment chassis. There is also a significant savings from the simplified RF Combining that comes with a “Wire Once” strategy.

The “Wire Once” strategy provides significant rack savings from simplified RF combining and makes SG splits to be operationally simpler. The case study shows that RF interconnections are reduced by about 50%.

The study shows even larger gains in the Ethernet port with reductions on the order of 80%. All of this leads to simpler operations and maintenance of the head end.

CCAP Case Study Conclusion

The CCAP case study created some before and after rack elevations to visualize the space savings from installing CCAP. This is shown in Figure 2.

The case study showed that the SG “multiplier” factor ranged from 3.7X to 5.1X. This indicates roughly the number of SG that could be fit within the existing footprint using today’s CCAP technology. The study found that the net effect of the combined equipment and RF Combing space savings is that operators can now roughly quadruple their SG count within their existing footprint.

In addition to quadrupling the SG count, 1st generation CCAP devices will also quadruple the narrowcast channel capacity for every SG. This means that today’s CCAP can enable a 16-fold increase in narrowcast capacity within existing head end footprints.

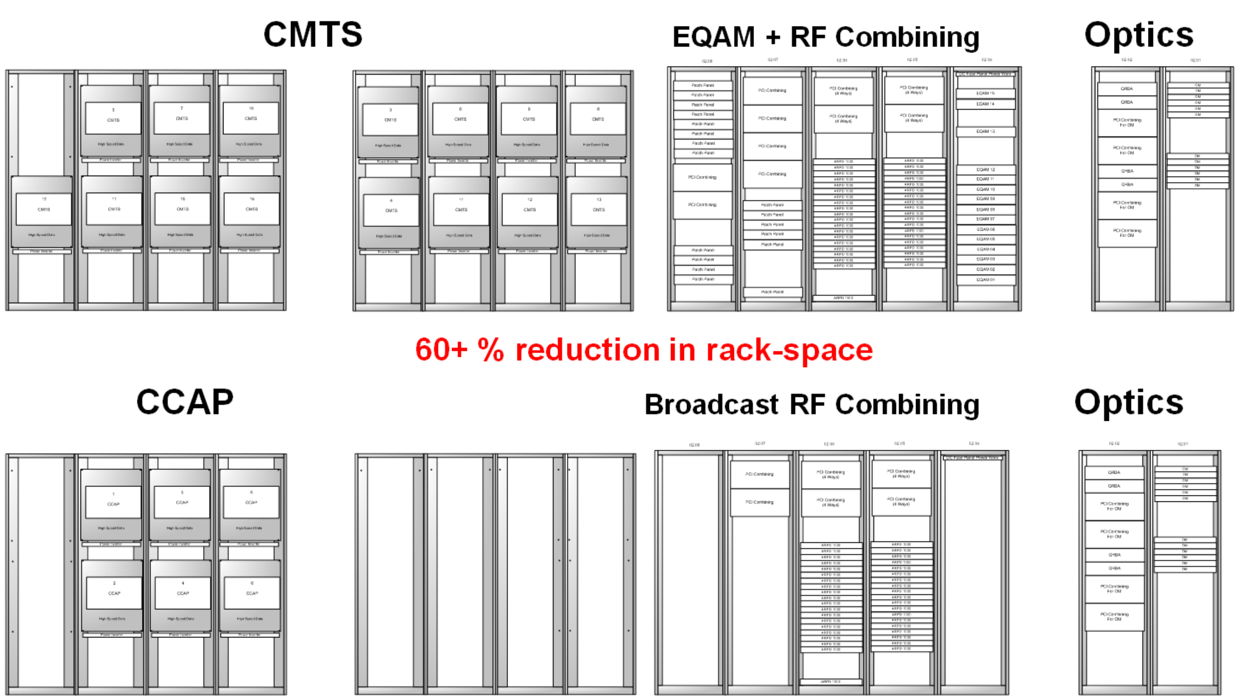


Figure 2 – CCAP Space Savings Example

Configuration	Space Needed For ~200 SG	SG per 1 Rack	Relative Scale
2012 Head End – CMTS, EQAM, RF Combining, Optics	~10 Racks	~20 SG	1X
2013 Traditional CCAP (56 SG) + Optics Shelf (60 SG per 12RU)	~3 Racks	~70 SG	3.5X

Table 1 – CCAP Space Savings Example, Today's CCAP

TODAY'S CCAP + OPTICS SAVINGS

The case study had certain limitations. First, it only considered the integration of narrowcast channels into the CCAP system. The broadcast channels were left external to the CCAP box and provide potential for further space savings in the future.

Also, the case study focused on just CMTS, EQAM, and RF combining, but did not delve into the optics components. For this paper we used the case study as a starting point and evaluated the entire head end footprint. While there are significant variations between head ends, we focused on a conservative "normalized" head end footprint that is represented in Table 1. Our model head end requires 10 racks of space today to support about 200 Service Groups (SG). That's an average of 20 SG per rack. This is shown in row 1 of the Table and is our baseline for our analysis.

Our next step is the migration to today's CCAP platforms and existing optic shelf technology. This could squeeze 200 SG into 3 total racks. This uses 56 SG per CCAP based on 2013 availability and optic rack density of 60 SG per 12RU. That results in an average of 70 SG per rack which results in a 3.5X improvement over our baseline configuration. This means that we might fit 700 SG into the existing 10 rack footprint. This is shown in row 2 of the Table.

While this is a great start, will this be sufficient as we go into the next decade? What if an operator needs to scale SG by a factor of a dozen? We'll now take a deeper look into how CCAP and Optics in the head end might scale over time to see what might be achieved by the year 2020.

SCALING CCAP TO CY2020

MSOs will undoubtedly experience profound bandwidth growth in most of their service types, including DOCSIS® HSD, IP Video, SDV, VoD, nDVR, and Digital Broadcast Video (SD, HD, 4K, and 8K resolutions). But with intelligent planning and carefully-phased deployments of equipment designed to support this growth, MSOs should be able to ride on their HFC infrastructure deep into the 2020 decade and potentially into the 2030 decade. If MSOs select certain paths for their equipment evolution; CCAPs may also provide a very smooth and cost-effective transition from the HFC infrastructure that exists today to the FTTH infrastructure that will likely be used by many MSOs by the 2040-2050 timeframe.

We will now look into some forward-looking ideas on how to accommodate the expected growth rates of the future and present some potential ideas on how CCAPs may evolve in the future. Since these ideas are forward-looking in nature, they should be viewed as proposals that may be altered down the road.

Our analysis will take a detailed look at many aspects that go into the SG density of a CCAP chassis. Scaling issues considered include silicon, backplane, RF connector and Ethernet connector technologies. And just as important, a close look is done at the projected power consumption over time.

Silicon Scaling

Moore's Law is an interesting empirical observation that has held true for 40 years, implying that the number of transistors in chips will double every 2 years (i.e. grow by a "change factor" of 1.42 every year). Over a 7 year time span from 2013 to 2020, this would result in a growth of:

Moore's Law: $(1.42x)^7 = 11.6x$

However, Moore's Law does NOT consider transistor speeds or power consumption, so it really does not give any clues as to the trends in processing capacity or power consumption for chips.

Koomey's Law is another (less famous, but more useful) empirical observation implying that for computing hardware in silicon, the number of computations per Joule of energy will double every 1.57 years (i.e. grow by a "change factor" of 1.56 every year)

If we assume that CCAPs will have fixed power per chassis between 2013 and 2020, then application of Koomey's Law to CCAP systems over the next 7 years predicts that the CCAP silicon processing capacity should grow from its current processing capacity level to:

Koomey's Law: $(1.56x)^7 = 22.5x$

Interesting, but this was for processors, so let's check those results by looking at it another way, by using Dennard's Scaling Law.

Dennard's Scaling Law is more useful in predicting trends in chip processing capacity and power consumption… and it can also be used to predict more complex behaviors like the trends in chip processing capacity when constrained by specific power consumption limits.

Application of Dennard's Scaling Law can be used to show that for computing hardware in typical silicon systems, the number of computations per Joule of energy will grow by a "change factor" of 1.37x every year.

Assuming that CCAPs will have fixed power per chassis between 2013 and 2020, then application of Dennard's Scaling Law to CCAP systems over the next 7 years predicts that the CCAP silicon processing capacity should grow from its current processing capacity level to:

Dennard's Law: $(1.37x)^7 = 9.1x$

This is a bit more conservative than Koomey's Law and gives ourselves a large amount of head-room.

Does this Processing Growth match reality? Consider the evolution for an example commercially available Multi-Core Processor family:

- 2009: 16 cores
- 2011: 32 cores
- 2013: 48-72 cores
- 2014: 128 cores
- 2015: 144 cores

Note that they had 9x increase in 6 years or an 1.44x Annual Change Factor (which is quite similar to the 1.37x number from Dennard's Law). This would lead to a 7 year growth of:

Multi-core NPU example: $(1.44x)^7 = 13.0x$

Based on all this, it looks reasonable to assume that the silicon processing on a CCAP card will grow by at least 8-10x by the year 2020:

CCAP Processor Scaling = 8-10x

Digital Backplane Scaling

In 2013, typical CCAP backplanes support KR interfaces operating at 10-20 Gbps between each Switch Fabric Card and each Client Card (2.5 Gbps per lane).

By 2015, typical CCAP backplanes will likely support KR4 interfaces operating at 40-80 Gbps between each Switch Fabric Card and each Client Card (10 Gbps per lane).

Current R&D being carried out has 33 Gbps per lane being transmitted across typical backplanes…

So, by 2020, this should easily permit CCAP backplanes to support interfaces operating at 100 Gbps+ between each Switch Fabric Card and each Client Card (25 Gbps per lane). This is 10x the backplane speeds of today:

CCAP Backplane Scaling = ~10x

RF Connector Scaling

24 MCX connectors are easily positioned on the CCAP faceplates today for upstream cards. It is believed that the use of MMCX connectors could increase the density of MCX connectors by a factor of ~1.4 and could increase the density of F-connectors by a factor of ~4. As a result, 32 connectors per face plate will likely be possible.

When combined with other techniques such as frequency stacking, it may be possible to create the equivalent of 32-64 connectors per card in the future. Since typical Downstream CCAP cards have ~8 connectors per card today. This results in a 4x to 8x increase in Downstream card connectors. Since typical Upstream CCAP cards have ~24 connectors per card today. This results in only a 1.3x to 2.6x increase in Upstream card connectors. However, the upstream spectrum is significantly less than the downstream spectrum, so frequency stacking can be used more aggressively on upstream connectors to keep pace if needed.

CCAP RF Connector Scaling = 4-8x

Ethernet Connector/Interface Scaling

For our Ethernet interconnections analysis, an example CCAP chassis is used that is constructed from a "typical" ATCA chassis. An ATCA chassis has a card faceplate with dimensions of ~13.5" x 1.2". Twelve SFP+ connectors (i.e. 12x10 Gbps) are easily positioned on these faceplates today. Eight CFP4 connectors (i.e. 8x100 Gbps) will be easily positioned on these faceplates by 2020.

The use of both front and back faceplates can double that density to be 16 CFP4 connectors per card (1.6 Tbps). Thus, the total increase results in a 13x increase in Ethernet interface bandwidth:

CCAP Ethernet NSI Scaling = 13x

Potential Head End Power and Space Issues

From the analyses in the previous sections, it can be seen that great improvements will be seen in silicon performance levels, backplane performance levels, Ethernet performance levels, and connector performance levels between now and 2020. Taken as a whole, the most limiting factor will likely be found in the RF connector densities, which may provide only a 4-8x improvement in RF connector density between now and 2020.

As a result, MSOs can likely expect to see at least four times the number of RF ports on their CCAP chassis in 2020 time-frame than they do today (without experiencing any significant increases in chassis sizes or chassis power levels). This should permit typical CCAP chassis of the 2020 time-frame to support 200 Service Groups or more per CCAP chassis.

However, the CCAP is only part of the head end solution. We also need to look closely at the Optical Shelf and see how that will scale over time.

OPTICAL SHELF RACK DENSITIES

The CCAP case study did not factor into its analysis the impact of optic shelf rack densities. As CCAP densities increase and the RF Combining is eliminated, then the optic shelves start to become the limiting factor for head end space requirements.

Figure 3 takes a closer look at some example rack densities from multiple vendors. This shows the rack space required to support 80 SG with a 1:1 ratio for upstream and downstream optics. This represents several generations of optics and it is apparent that there could be almost a three to one difference in optical rack density depending on vendor.

The optic solution on the left side of the figure represents the latest state of the art for the year 2014. It squeezes 80 SG of optics into 12RU. Currently, the upstream optical receivers are twice as dense as the downstream optical transmitters.

Down the road, once the optical TX catches up to the optical RX, we might see optical rack densities of 80 SG in 8RU. This is the number that we'll use for our year 2020 analysis. We feel this is very conservative and feasible in this time frame. Longer term we may actually see a trend towards integrating multiple WDM wavelengths into optical components which could yield another 2X or 4X in optical shelf rack densities.

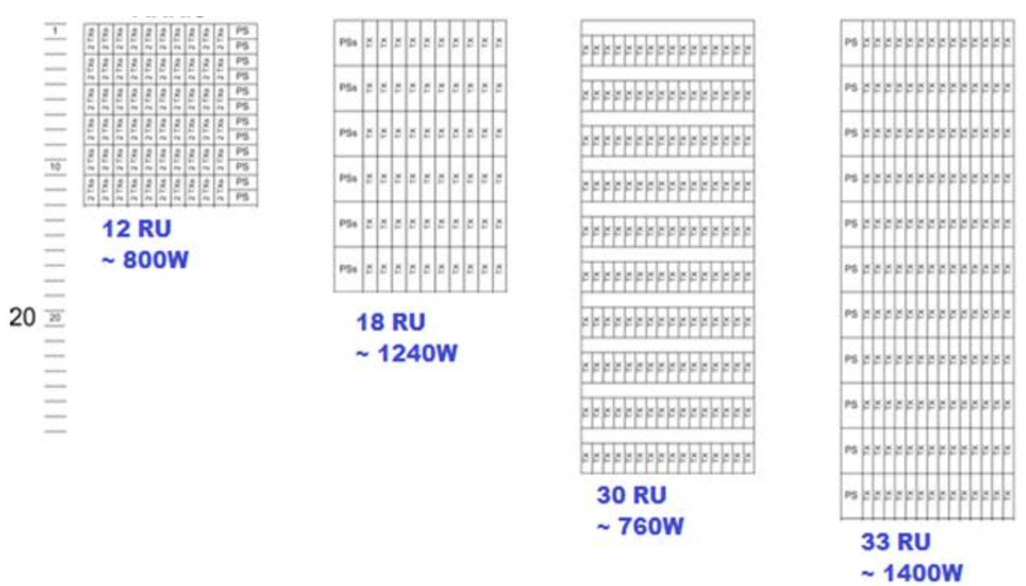


Figure 3 – Example Optic Shelves Rack Densities for 80 SG

An alternate path for future optics might be integration of pluggable optics directly into the CCAP chassis. This may be reasonable in some scenarios, but there are a myriad of trade-offs that must be considered. Today's external Optical Shelves offer the full range of TX & RX optics necessary to cover the many different HFC plant conditions encountered today. It is also expected that external Optical Shelves can provide advantages as operators look to integrate WDM capabilities as well.

SCALING HEAD ENDS TO CY2020

Now let's put together all this information that we've gathered. As the next step in the CCAP evolution, we'll assume that the 2nd generation CCAP devices can achieve at least 25% increase in SG density. This should be quite reasonable given our previous analysis and pushes the CCAP up to ~70 SG per chassis.

At the same time, we've seen optic shelf rack density increase in the last year from 60 SG per 12RU up to 80 SG per 12RU.

Using these two inputs, the next step in the CCAP evolution should get us down to 2 racks to support 200 SG. That's an average of 100 SG per rack for a 5X increase over our baseline of today's CMTS/EQAM based head ends. This is shown in the third row of Table 2 below.

As we push towards the end of this decade, our previous analysis shows that traditional CCAP systems can just about quadruple densities of today's CCAP by CY2020. That would put us around 200 SG per CCAP chassis. This then combines with the expected continued advances in optical shelf rack densities to 80 SG per 8RU. The result is that we have a clear line of sight to achieving 200 SG in a single rack within this decade. That provides a 10X increase in SG growth within today's existing head end footprint. See the last row in Table 2.

In addition to this SG growth, DOCSIS® 3.1 will also give a giant boost to the SG capacity. Starting from today's CCAP system that provides about 1 Gbps (i.e. 32 DOCSIS® channels), DOCSIS® 3.1 can provide more than 10Gbps per SG.

Configuration	Space Needed For ~200 SG	SG per 1 Rack	Relative Scale
2012 Head End – CMTS, EQAM, RF Combining, Optics	~10 Racks	~20 SG	1X
2013 Traditional CCAP (56 SG) + Optics Shelf (60 SG per 12RU)	~3 Racks	~70 SG	3.5X
2nd Gen CCAP (~70 SG) + 2014 Optics Shelf (80 SG per 12RU)	~2 Racks	~100 SG	5X
Future 2020 CCAP (~200 SG) + Optics Shelf (120 SG per 12RU)	~1 Rack	~200 SG	10X

Table 2 – CCAP Space Savings Example, Future CCAP

How Many More Node Splits will there be?

Eventually operators will reach a point where it no longer makes sense to split a node further. As discussed further in [CLOONAN2], a fundamental transition is occurring in cable system traffic engineering. Previously with extremely large SG, average traffic load dominated the analysis and splitting nodes would cut the average traffic roughly in half. However with today's Downstream Service Group sizes around 500 subs, the max burst traffic rate, Tmax, plays a significant role. Some people now use a rule of thumb of taking 2x or 3x the Tmax rate of the highest service tier to determine the amount of DOCSIS® capacity needed.

As operators continue to split nodes past this point, the SG size becomes so small that average traffic load is in the noise and the Tmax burst rate dominates the traffic engineering. This means that further node splits will provide diminishing returns. It will become more important to increase burst capacity through DOCSIS® 3.1 introduction with plant upgrades to 1.2GHz.

Thus, for most MSOs who will perform no more than 3-4 rounds of node-splits within the next decade or two, it is clear that the traditional head-end-based CCAP chassis of the future will be able to accommodate their needs. With a 10x SG increase, future CCAP might take today's 500 sub per SG down to an average of 50 subs per SG. The CCAP chassis of the future will provide increases in RF port (Service Group) densities that can keep up with the demand created by the node splits.

AM OPTIC CONSIDERATIONS

As discussed at the beginning of the paper, the other major concern making operators think about migrating down the DAA path is the impact Nonlinear Optical noise on the potential capacity gains from DOCSIS® 3.1.

Nonlinear Optical noise can significantly decrease SNRs and limit supported QAM modulation rates. This becomes more important with the introduction of DOCSIS® 3.1 that requires downstream support up to 4096-QAM, with optional support for 16384-QAM modulation.

AM Optics – Distances and Wavelengths

In recent years, there continues to be significant improvements that are occurring in traditional AM optics to address the issue of Nonlinear Optical noise. This means that the latest generation of optics can support longer reach capability as shown in Figure 4, where the red represents recent optic improvements.

Newer AM optics also supports more lambdas, such as 44 wavelengths. That's almost a factor of three improvements over previous generation optics. Figure 5 shows an example of the trade-off between # of wavelengths and distances. Note that 44 wavelengths can still be achieved at 40km distances. Only longer distances results in a reduction in wavelengths supported.

It is also important to note that these devices are full 1.2GHz spectrum products that can take full advantage of DOCSIS® 3.1. There is a reduction in power too which is important for scaling the head end SG capacity.

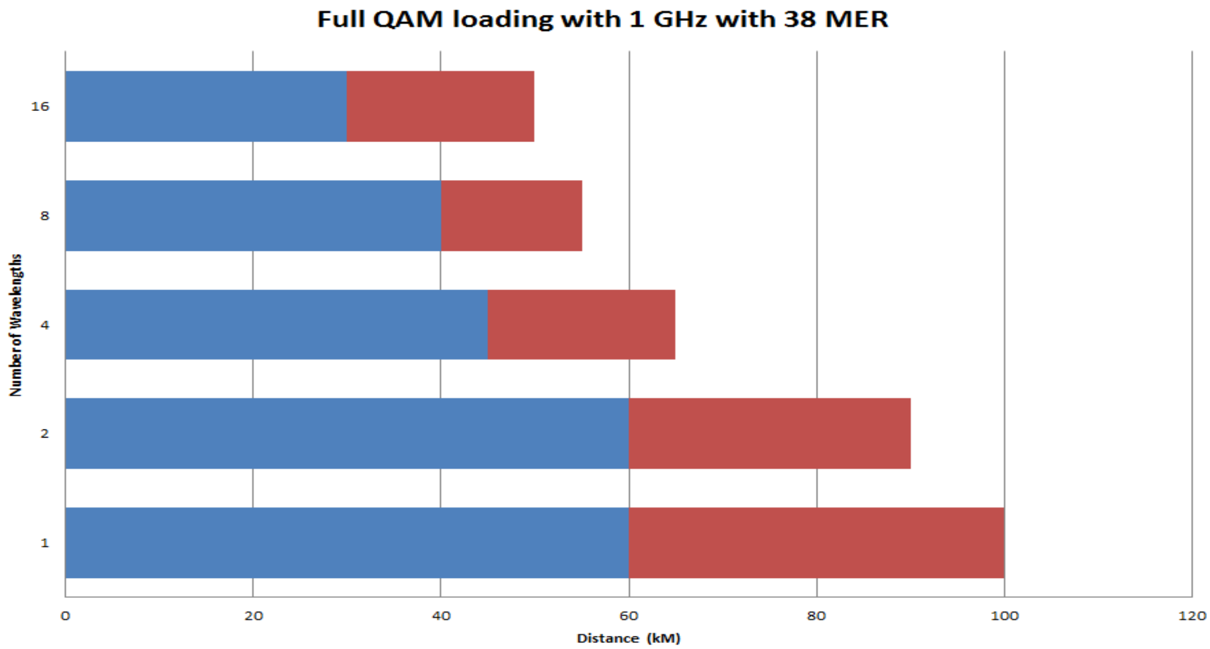


Figure 4 – CY2014 AM Optic Distance Improvements

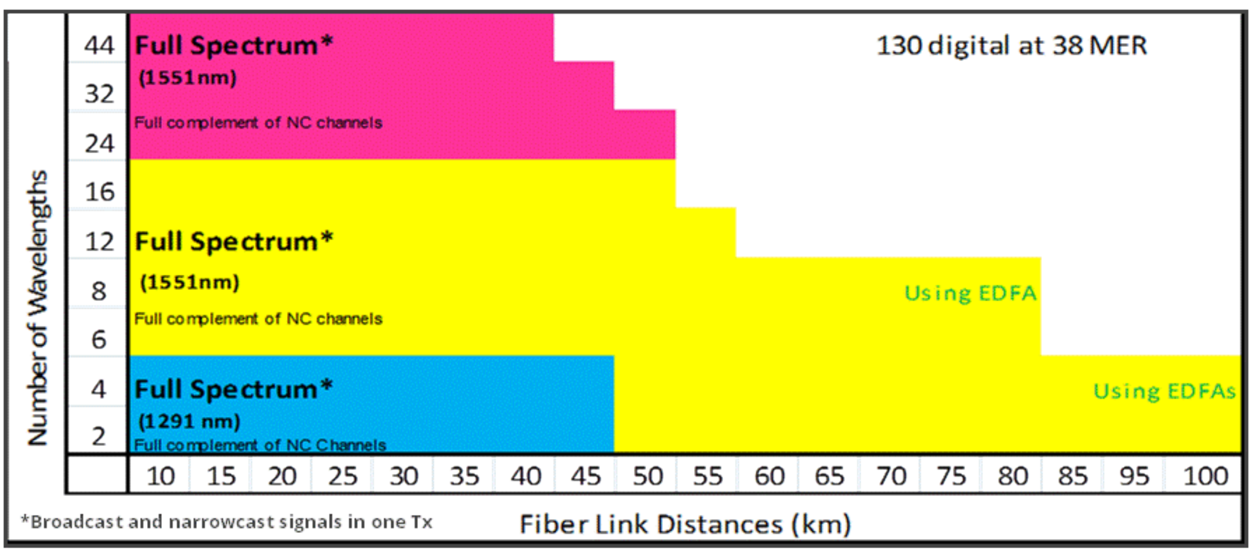


Figure 5 – CY2014 AM Optic Capabilities: Wavelength vs. Distance

<u>AM Optics and DOCSIS® 3.1 Capacity</u>

So, a major question for AM optics becomes what kind of capacity can be achieved with DOCSIS® 3.1? This topic is investigated in detail in [EMMEN]. A very informative chart from that paper is shown in Figure 6. The chart provides the PHY capacities for various optic configurations, both AM optics and digital optics (e.g. Ethernet).

The first four bars in the chart are various AM optic configurations. The first four bars represents Full Spectrum AM optics at distances of 80km, 40km, 25km and 10km respectively. The last two bars on the chart represent different types of digital fiber systems. Note that “Remote Gadgets” refers to either Remote PHY or Remote CCAP.

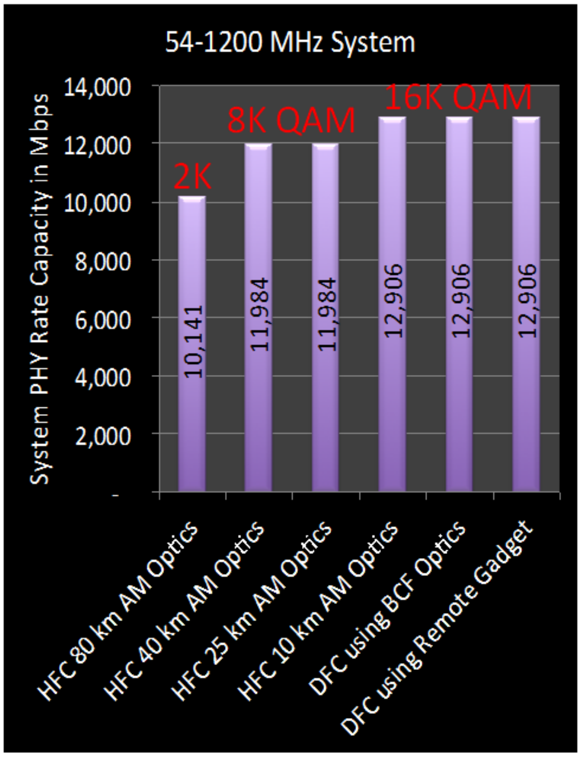


Figure 6 – PHY Capacity for Various Optic Configurations

As can be seen in Figure 6, the digital fiber systems provide maximum DOCSIS® 3.1 capacity. The value of 12,906Mbps represents full spectrum OFDM channels operating at 16384-QAM modulation. Note that AM optics at 10km also achieves the theoretical maximum as well.

The AM optic capacities start to drop off as the distance increases. The capacity at 25km and 40km corresponds to 8192-QAM modulation. It should be noted that 4096-QAM is the highest mandatory modulation for DOCSIS® 3.1 and is what initial products will support. 8192-QAM and 16384-QAM are future options and it is not clear when or if they may be deployed.

At 80km distances, the AM optics capacity drops to 2048-QAM. An interesting note is that the total PHY capacity at this distance is still just over 10Gbps. This means that it can still match cable plant capacity with a Remote Gadget being fed (and limited by) a 10G Ethernet link.

We now take a look at these results in a tabular form in Table 3. For the relative gain, we use 40km AM optics as the baseline performance for comparison purposes.

As the table shows, 40km AM optics supports more than the full DOCSIS® 3.1 mandatory requirements of 4096-QAM. Even once the optional modulations are introduced, 1638-QAM will only provide operators with a best case gain of 7.5% over the 40km AM Optic baseline. These modulations will also have increased SNR requirements which might require more robust FEC that could eat into that gain. Similarly, operating AM optics at 80km only results in a 15% hit compared to the baseline, but only an 8% hit to total PHY capacity compared to the D3.1 mandatory maximum modulation of 4096-QAM.

Technology	D3.1 QAM Modulation	Relative Gain
Digital Optics	16,384-QAM	+7.5%
AM Optics, 10km	16,384-QAM	+7.5%
AM Optics, 25km	8192-QAM	0
AM Optics, 40km	8192-QAM	0
AM Optics, 80km	2048-QAM	-15%

Table 3 – PHY Capacity for Various Optic Configurations

CONCLUSION

Traditional head-end-based CCAP and AM Optic systems will capitalize on many improvements in silicon, packaging, interconnection, and cooling technologies over the next 10 to 15 years. Our estimated gains based on well known Moore's Law and lesser known cousins Koomey's Law and Dennard's Law, show that these improvements will permit extensive increases in bandwidth per RF port (permitting >10 Gbps per RF port) and will also permit extensive increases in the number of RF ports (e.g. allowing more than 200 Downstream RF ports per CCAP chassis) by the end of this decade.

Over the next 4-6 years, DOCSIS® 3.1 will also enable a 10x increase in capacity per SG from today's ~1 Gbps HSD (e.g. 24-32 DOCSIS® 3.0 channels) to 10+ Gbps (e.g. 5 or 6 192MHz OFDM channels). The paper shows that AM optic technology advances will allow operators to still take advantage of this, supporting 40km distances with 44 wavelengths at DOCSIS® 3.1 4096-QAM modulation rates.

So, the bottom line is that traditional head end systems can leverage CCAP + optic advances to get both a 10X increase in SG counts in conjunction with 10X increase in capacity per SG before the end of this decade. Those increases should permit traditional head-end-based CCAPs to provide more than enough bandwidth capacity and RF port capacity than most MSOs will require as they perform expected node splits in the coming 10 to 20 years.

Acknowledgements

The authors would like to thank Mike Emmendorfer for his assistance in helping to pull together material for this paper and for his review of the paper.

REFERENCES

[CLOONAN1] – "HFC Spectrum Planning in the Era of the HSD Bandwidth Explosion"; T. Cloonan Ph.D., S. Chernak, SCTE Cable-Tec Expo 2012

[ULM] – "CCAP Case Study: Enabling Converged Video + Data thru Space & Power Savings"; J. Ulm, J. Finkelstein, S. Rahman, J. Salinger, The Cable Show Spring Technical Forum 2013

[CLOONAN2] – "Predictions on the Evolution of Access Networks to the Year 2020 & Beyond"; T. Cloonan Ph.D., M. Emmendorfer, J. Ulm, A. Al-Banna Ph.D., The Cable Show Spring Technical Forum 2014

[EMMEN] – "A Side-by-side Comparison of Centralized vs. Distributed Access Architectures"; M. Emmendorfer, T. Cloonan Ph.D., J. Ulm, The Cable Show Spring Technical Forum 2014

Software Defined Networking, DOCSIS Provisioning, and MSO Commercial Services

Kevin A. Noll
Time Warner Cable

Wesley George
Time Warner Cable

Richard Loveland
Alcatel-Lucent

Abstract

In recent years, cable operators have realized significant growth outside the traditional residential market. The North American cable industry is set to top $8.5 billion in revenue from offering network services to commercial and carrier customers. In a $130 billion telecom services market, cable operators have plenty of opportunities to continue growing this segment of their business. However, to better compete and realize revenue faster, cable operators need to offer new and innovative services, move them to market faster, simplify operations and significantly increase their rate of customer turn-up while decreasing the time to delivery.

This paper will examine how Software Defined Networking concepts and the Cablelabs DOCSIS, L2VPN, and DPoE specifications can be integrated to enable cable operators to realize fully automated end-to-end provisioning of commercial services and dynamic network reconfiguration through a single and simplified provisioning interface.

INTRODUCTION

In the mid to late 1990's Internet access was slow or expensive (usually both) and dominated by the telephone companies and dial-up connections. The cable industry was investing heavily in two-way plant infrastructure for interaction between customers and services. Much of this investment was dominated by Hybrid Fiber Coax (HFC) builds. This investment was the stepping-stone to allow the cable operators to push into the always-on high-speed Internet access service. By utilizing the two-way capabilities of HFC, the cable operators were able to economically begin to offer residential high-speed Internet access at affordable prices.

In order to take advantage of the HFC investment for high-speed Internet access, the cable operators defined a protocol called Data Over Cable Services Interface Specification (DOCSIS) of which the first version was released in 1997. This specification allowed vendors to create cable modem systems that would allow users to access the Internet and its growing collection of content in a more user friendly and faster method.

The introduction of cable modem systems and the tremendous growth of users demanded a set of provisioning and operational tools to support large scale transactions for automated provisioning and troubleshooting. These tools developed over time into an operational support system that is used by cable operators around the world to manage cable modem services with thousands of transactions a day.

As Cable service providers expand into serving enterprise as well as residential customers with fiber services, that same scale and transactional operation model is needed. Debate is ongoing within the MSO community whether a new OSS system should be developed for these new network models, but the prevailing direction has been to augment fiber access systems so they can be provisioned in a nearly identical manner allowing the reuse of massive investments in DOCSIS OSS systems that have served the industry well.

As a result in 2009 the operators, along with Cable Labs, began to define a set of specifications called DOCSIS Provisioning of EPON (DPoE) to specify how an EPON fiber access system should be provisioned. These specifications leverage the investment in cable modem operational support systems emulating the provisioning of a cable modem (CM) and Cable Modem Termination System (CMTS).

As part of this emulation, the Optical Line Terminal (OLT) Emulates a CMTS and creates a software instance of a cable modem called a virtual Cable Modem (vCM). Network Function Virtualization (NFV), which the vCM is a form of, has become one of the latest driving trends in the telecommunications industry. The goal of NFV is to reduce the cost of devices by relocating much of the intelligence for network devices, especially in the home, up into the network. This keeps the physical hardware as simple as possible with a standardized applications programming interface (API) to set up the hardware device – in this case an Optical Network Unit (ONU). DPoE provides an API to the ONU called DPoE OAM. In addition the vCM is provisioned through the standard CM configuration file allowing the provisioning of the vCM to be an abstracted view of what of the service being offered to the user. Reuse of the CM provisioning and operating model provides a standard method of defining the service while it reduces the CAPEX needed for OSS development and allows automated provisioning of these fiber systems. This helps reduce the time-to-market of the volume rollout of EPON. Of course EPON could be used on a much smaller scale with vendor provided EMS applications but those systems do not make a large scale deployment possible in the timeframe desired.

From the network perspective, DPoE enables an endpoint (vCM) to be provisioned to utilize network protocols such as MPLS and BGP signaling to automatically set up connections across the network to the far end of the MPLS tunnel creating a layer-2 virtual private network (L2VPN). This ability follows the desire of another trend in the network called Software Defined Networking (SDN). This ability to set up these network wide services can dramatically improve service delivery velocity increasing the number of transactions per day and greatly reduce errors in setting up complex network connections. The combination of DPoE and SDN concepts creates a powerful tool for the cable operators.

This paper is intended to examine the commonalities and benefits of combining DOCSIS-based provisioning with SDN and NFV to show how the cable operators can benefit from implementing such an architecture for its customers, It will discuss how these concepts create L2VPN services in a highly automated method.

SOFTWARE DEFINED NETWORKS

Software Defined Networking (SDN) is an increasingly common buzzword, especially when discussing the interaction between applications, networking, and the infrastructure they share to provide services “in the cloud.” Unfortunately, SDN is an overloaded and broad term that if one were to ask three different “SDN Experts” what SDN is, it would likely result in five different answers.

Typically, a discussion of SDN will start with separation of the network's control plane from the forwarding plane, decoupling forwarding and policy decisions (what traffic goes where) from network transport and topology (how the devices are interconnected). This is the focus of projects like the Open Networking Foundation's OpenFlow. This project is defining a set of specifications allowing software (the SDN controller) running on a general purpose computing platform to perform the control-plane computations and functions required to dictate the flow of data through the network. The SDN controller uses OpenFlow to directly manipulate the forwarding tables of the switches and routers in the network. It is important to stress that SDN and Openflow are not interchangeable terms, and a network can be Software Defined without the use of Openflow.

Many service providers operate their network with little or weak central control over its configuration and management. This means that the network configuration and state is effectively stored in a giant distributed database. This is not inherently a bad state of affairs, but network operators aren't always good at getting the information in that giant database into a form that is usable for making business decisions that optimize the use of the network and the services that run over it.

Separation of control and forwarding enables programmatic orchestration and provisioning of network resources and optimal, secure flow between application services. It should be clear that consolidated control should lead to provisioning activities that are consolidated through a single programmatic interface located at the controller. It should also follow that centralizing the flow of network metrics would reduce the complexity of dealing with that giant distributed database.

With consolidated provisioning flow it becomes necessary to create a well-defined and standardized API through which the business can interact with the controller. Ideally, this API would create a layer of abstraction between the business systems and the network. Through abstraction and standardization, the network operator is able to reduce the time to develop and deploy new services. The operator is able to reduce (if not eliminate) human-induced network configuration errors.

Especially in the context of complex Commercial Business offerings, the operator is able to reduce the service provisioning process to a transaction rather than a complex series of steps involving multiple systems and humans.

In short, SDN is a method to enable a high degree of automation in provisioning and managing network services. While this can be done by directly manipulating forwarding tables as with OpenFlow, it can also be done by leveraging existing control plane and configuration methods such that some amount of traditional, distributed network control plane remains in place. Automation in this context is primarily focused on improving efficiency, speed, and accuracy. While there are certainly places where automation can eliminate time-consuming and repetitive manual processes, automation does not have to be hugely complex logic that expects the network to make most decisions via autonomic intelligence without human involvement. The goal isn't to build Skynet (of the Terminator franchise), because the tools and expertise necessary to build that level of independent, intelligent network are fairly limited. Instead, it can be a set of pre-defined scenarios that a human selects and executes, such that it serves as a force multiplier to enable the smart people running the network to do more, do it faster, and do it more accurately.

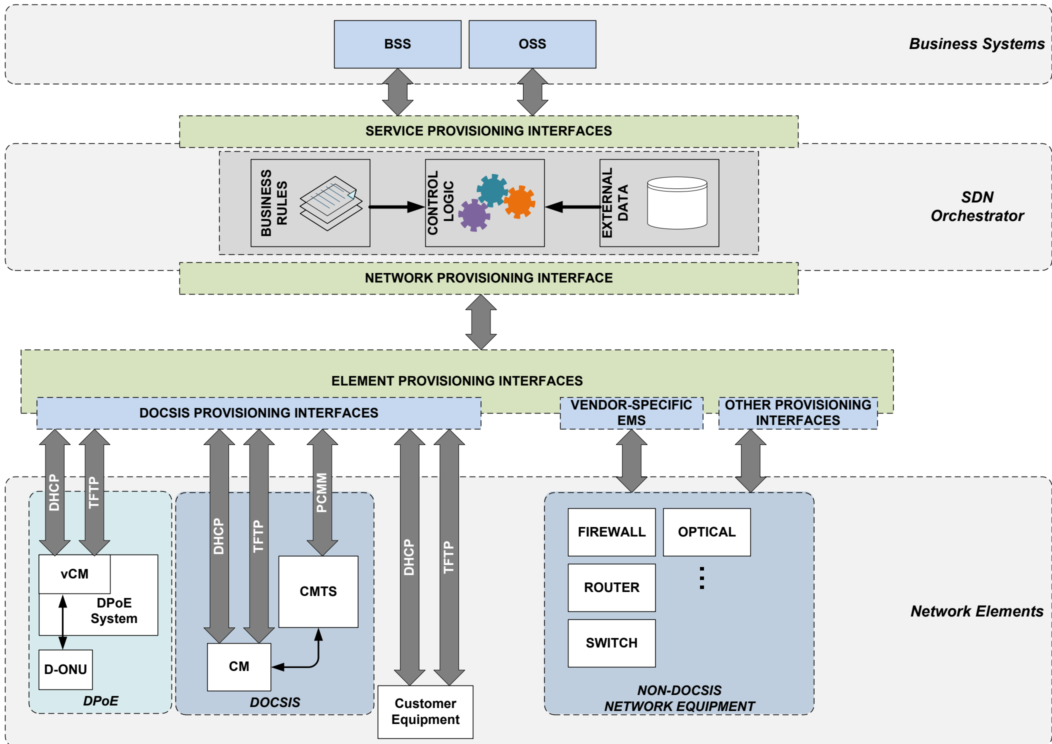


Figure 1 - SDN Architecture with DOCIS-based Provisioning

Automation is something that should enable flexibility and speed when it comes to defining services in the network, whether configuring and deploying existing services, chaining multiple services together to make a new integrated service, or rapidly building new and innovative services, as well as troubleshooting them when they break.

The remainder of this paper explores this aspect of SDN and how the combination of SDN and DOCSIS-based provisioning can help realize these goals.

<u>An SDN Architecture using DOCSIS and DPoE</u>

A provisioning architecture that incorporates SDN with DOCSIS and DPoE would position a cable operator to leverage the advantages of SDN without requiring a forklift upgrade of its existing access network equipment and front-office provisioning tools. Such an architecture can easily be envisioned and one is depicted in Figure 1.

It is important that the key requirements of such an architecture are understood.

- The SDN+DOCSIS architecture must be capable of provisioning the network services that are offered by the cable operator today. The network services typically being offered today are high-speed Internet access and Layer-2 VPN (ELINE, ELAN, ETREE) services.
- The architecture must be easily extended to support new network services (for example, L3VPN) in the future.
- The architecture must provide a rich business-facing API that is capable of describing the offered services without intimate knowledge of the underlying network. In fact, the API should not need to know the particular network technology that is in use.
- The architecture must provide the ability to chain network services to higher-layer services such as hosted

firewalls, application load balancing, cloud-computing resources, etc.

The Business Systems layer depicted in the architecture of Figure 1 is that set of applications and systems that are responsible for customer relationship management, billing, reporting, and workflow management. For most cable operators this will be their BSS (for example, Icoms and CSG). Given the weaknesses of traditional cable BSS to support Commercial Services some other system may interact with the architecture at this level.

The Service Provisioning Interface (SPI) is the API that business systems use to interact with the SDN Orchestrator. The Service Provisioning interface exposes a programmatic method by which the business systems can provide a "business-style" definition of the service(s) to be provisioned along with directives describing the action to be taken (move, add, change, delete, etc.) with the definition being provided to the SDN Orchestrator.

The SDN Orchestrator is responsible for translating the received service provisioning requests into a working network configuration. The SDN Orchestrator is expected to supplement the SPI input with business rules and data collected from the network and other sources.

Using all these inputs the SDN Orchestrator computes the most optimal network configuration that will meet the service requirements and pushes that configuration through the Network Provisioning Interface.

The Network Provisioning Interface (NPI) is responsible for the presentation of service-enabling parameters to the Element Provisioning Interface (EPI) that speaks directly to the network elements and their respective management systems.

The Element Provisioning Interface (EPI) implements the protocols that the individual network elements use to configure themselves for communication on the network and, if necessary, download a configuration from an element-provisioning server. The EPI will contain modules that are customized to the particular network element and manufacturer.

Ideally there would be only one module in the EPI. This would reduce the integration work required in the SDN systems over the long-term, but currently there is no single standard or protocol that can achieve complete coverage. As a result, most models are built with the expectation that multiple modules will be necessary, with the EPI serving as a method to abstract the element-specific configuration method(s) from the upper layers. Candidates for the EPI include currently developing standards like OpenDaylight as well as common off-the-shelf modular provisioning systems that leverage partnerships with network equipment vendors to build and maintain vendor-specific modules, or that use standard extensible configuration protocols such as NETCONF [RFC 6241] and YANG [RFC 6020].

The authors claim here and provide evidence in later sections that DOCSIS-style provisioning implemented in the EPI can carry a significant portion of the service-level provisioning functionality in the cable operator's network, and the presence of an EPI as an abstraction layer can allow support for other non-DOCSIS provisioning methods in concert to provide a complete end-to-end service provisioning solution.

Requirements of a Service Provisioning API

It is important that the API provided by the SPI create an abstraction between the business systems and the provisioning systems such that network-specific knowledge is not required in the business systems. When provisioning network transport services

(Internet access, Layer-2 VPN, etc.), when viewed from a business-layer perspective, it is sufficient to provide only the details required to describe (not completely specify) an attachment circuit and the backbone transport associated with each subscribed service in order to completely specify the end-to-end service.

The authors propose here that a network-layer service can be fully specified by the business systems by providing the following elements:

1. Unique service identifier that is common among the end points attached to a service
2. Specification of the service type (for example: IP, L2VPN, L3VPN)
3. Specification of the role that the end point plays in the service (for example: node, root, leaf)
4. Specifications of supplements to be instantiated on the service (for example: [Y1731], or [1588v2])
5. Unique Identifier of the Access Equipment (described below)
6. Port Identifier on which the service is expected to appear
7. Description of user traffic to accept and/or not accept into the service
8. Bandwidth profile for both upstream and downstream flows

Note that these attributes do not contain any data or imply any *a priori* knowledge of the network infrastructure *other than* an assumption that (1) the business systems have chosen the most appropriate access technology to be used (which could be done in the SDN Orchestrator) and (2) the network can be provisioned purely through communication with the end points (as in [DOCSIS]) or an entity acting on the end point's behalf (as in [DPOE]).

This model is not limited to [DOCSIS] and [DPOE] access network technologies. [EXPO2012] proposed a provisioning model for point-to-point Ethernet based on [DOCSIS] and an implementation of that proposal has been demonstrated [OLIVERDEMO] with common off-the-shelf routers and switches.

Of course, the list of elements above is not sufficient to describe higher-layer services (hosted firewalls, cloud computing resources, etc.). Any Service Provisioning API would need to provide a robust and extensible language to describe those services as well as methods to chain network transport services to those higher-layer services. This present paper will not attempt to address the requirements for those types of services because the authors recognize the insufficiency of DOCSIS-style provisioning to fully describe those services in an effective manner.

Further requirements of the API should specify the actions that can be performed through the API. For example, the API needs to allow the business systems to specify actions of Move, Add, Change or Delete (MACD). The API also needs to be queryable – in other words the business systems must be able to query the status of the requested service instance and the general network through the API.

The API must also allow flexibility in the actual implementation of the requested services. For example, the network operator may choose a network architecture that permits multiple services to be implemented on the same physical port. In this use case, the API must not make assumptions that a single physical port can have only one service. In fact, the SDN Orchestrator would use the business rules as input to determine whether multiple services are allowed on a single port or not.

This is a high-level and incomplete list of requirements for the Service Provisioning

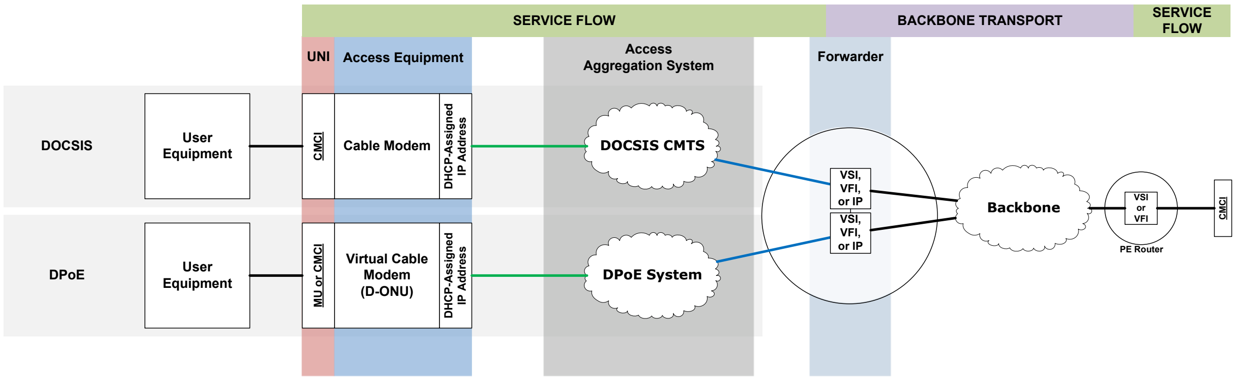


Figure 2 - Network Model for DOCSIS/DPoE Provisioning

API. They do, though, provide a good starting point for discussion of a more detailed set of requirements.

DOCSIS-STYLE PROVISIONING

The DOCSIS specification defines a high-speed data access network operating on a coaxial cable plant. Embedded in [D3-MULPI] is an element and network-provisioning framework. The following sections are a high-level tutorial of the DOCSIS provisioning model and are designed to support the claim that DOCSIS-style provisioning is capable of provisioning a rich and extensible set of services and fits well with the SDN model.

Of particular interest in this study is [D3-MULPI], [L2VPN], and [DPOE]. Specified within these documents is a network model, and protocol for initial provisioning of a cable modem (virtual cable modem in [DPOE]) and the service-description language to be used when configuring access services on a DOCSIS-style network.

Network Model

Rather than choosing to define a common network reference model to be applied to all the specifications, Cablelabs has chosen to weave a network model into each specification, [DOCSIS], [DPOE] and [L2VPN]. Luckily each model is very similar to the others and all are based largely on the [DOCSIS] model. Here we summarize the major elements in each specification.

The network model, shown in Figure 2, consists of:

- *User Equipment* – For example, a PC or router.
- *Network Access Equipment* – Equipment that is specific to the network type and associated with a subset of subscribers on the access network. A Cable Modem in a DOCSIS network, or a D-ONU in a DPoE network.
- *Access Aggregation System* – Equipment specific to the network type and aggregating users from one or more Network Access Equipment and typically located in an MSO's headend or hub. This would be a CMTS or DPoE System.
- *Forwarder* – A logical entity that is typically located in the Access Aggregation System. The Forwarder is responsible for moving customer data from the access network ports to a Network Side Interface (NSI). Forwarders could take the form of an IPv4 or IPv6 interface, an IEEE 802.1 bridge, an MPLS virtual interface, etc.

Within this model, user data is accepted on the CMCI (or MU in DPoE) and must be transported across the access network to the access aggregation system. Inside the access aggregation system, the user's traffic must be matched to a Forwarder instance.

The Forwarder then acts on the user's traffic according to the forwarder's configuration. An IPv4 forwarder will inspect the user's packets and route them according to the rules of IPv4 and the Forwarder's configuration. Similarly, an L2VPN forwarder will encapsulate the user's frames and bridge them according to the configuration assigned to that user and the forwarder.

A *Service Flow* is the construct that directs a user's data (or subset thereof) from the CMCI to the proper Forwarder instance. User data is classified as it enters the CMCI. Based on the classification, a service flow is matched and the frames are encapsulated and forwarded according to that service flow's configuration.

The preceding text describes traffic flowing upstream, meaning from the user toward the Network Services Interface (NSI). Treatment of traffic flowing from the NSI toward the user (downstream) follows a similar path. Traffic entering the NSI must be received by a Forwarder, classified into a service flow, forwarded through the service flow to the CMCI and delivered to the user. The remainder of this paper will describe traffic flow in the upstream reference, but the discussion applies equally to downstream flows.

DOCSIS Provisioning Framework

Any provisioning framework must provide two essential functions. The first is a language with which to describe the configuration to be applied to a network element. The second element is a method to deliver the configuration to the element being provisioned.

The original architects of [DOCSIS] chose a provisioning model in which the cable modem is the primary consumer of the service configuration (referred to in this document as *end-point provisioning*). This choice and the list of available protocols at the time sent the designers down a path of using DHCP and TFTP as the delivery mechanism. DHCP is required because the cable modem must have an IP address to be manageable on the network and TFTP because it is a simple protocol to implement for file transfers.

The configuration language chosen in [DOCSIS] is a Type-Length-Value (TLV) system. The Type indicates a parameter that is to be specified (for example, Bandwidth). The Value indicates the value being assigned to that instance of the Type. Length simply indicates how long the data is for the Type being specified – aiding the receiver in decoding the configuration file.

(In reality, the DOCSIS architects chose the system of TLV configuration files delivered via DHCP/TFTP because early manufacturers of data-over-cable equipment had developed and cable operators had already deployed systems based on these methods.)

The same system is, of course, adopted by [DPOE] for EPON networks.

The provisioning model used by [DOCSIS] and [DPOE] makes several key assumptions about the network and the elements that compose it.

First, [DOCSIS] is only intended to configure a network element at the service-specific level. In other words, there is no method in the [DOCSIS] model to provision the operational aspects of a CMTS or, in [DPOE], a DPoE System. For example, the

DOCSIS configuration file is not able to carry information to create or configure a cable interface on a CMTS. This function is best served by other methods like Command-Line Interface or systems based on NETCONF [RFC 6241] and YANG [RFC 6020].

Second, the end-point based model assumes that the cable modem (CM) or virtual cable modem (vCM) will be the primary consumer of the configuration file. In other words, the CM or vCM is the element that initiates the configuration file download and will convey the required information to other network elements (this is not true in the strictest sense in [DOCSIS] – the CMTS may actually acquire the configuration file contents by capturing the TFTP packets as they are sent to the cable modem).

Finally, the end-point based model assumes that any end-point may cause a forwarder to be instantiated by providing the essential parameters for the forwarder via the end-point's configuration file.

The model described here accomplishes initial provisioning of a user's service. If a change needs to be made to the service instance, it is necessary to reboot the access equipment. A weakness that may be perceived in the [DOCSIS] provisioning model is the lack of a dynamic method of updating a service.

PacketCable Multimedia (PCMM) is the first attempt to address this in the CableLabs specifications. The use of PCMM has not, yet, been defined in [DPOE]. PCMM only has a data model to describe changes in bandwidth or to setup and destroy service flows. PCMM does not currently have a data model to describe Forwarder instances. The latter point is significant in the Business Services space.

It is typical for an IP forwarder instance for Internet access to be pre-configured on the CMTS. Therefore PCMM simply needs to be able to configure service flows that will attach to that forwarder. However, in the Business Services space, it may be necessary to create a new forwarder, modify an existing forwarder, or attach a service flow to a specific forwarder. PCMM cannot perform this function today.

[DPOE] has defined a dynamic update method based on the initial provisioning mechanism. In this method, an external system sends an SNMP SET message indicating to the vCM that it must re-download its configuration file. Upon receipt of the SNMP SET, the vCM executes its initial provisioning sequence (DHCP and TFTP) *but does not reset*. The D-ONU and DPoE System continue forwarding traffic according to the previous configuration. Upon receipt and validation of the new configuration, the D-ONU and DPoE System are reconfigured and begin forwarding according to the new configuration.

<u>DOCSIS Toolbox</u>

The DOCSIS provisioning language, including extensions in [DPOE] and [L2VPN], currently is capable of describing a user's attachment to natively routed IP services and layer-2 VPN services. It is easy to conceive that the TLV dictionary could be extended to represent layer-3 VPN or other service types.

The following sections provide a *sampling* of the key elements contained in the current DOCSIS TLV dictionary. This section is not intended to provide an exhaustive list and description of the TLV dictionary. Interested readers should review [DOCSIS], [DPOE] and [L2VPN].

<u>Equipment Identification</u>

During provisioning it is necessary to identify the consumer of the provisioning data and the equipment to be provisioned. In

[DOCSIS] the consumer of the provisioning data and the equipment to be provisioned are the same – they are both the cable modem. In [DPOE] the vCM is the consumer and the D-ONU is the equipment, but they appear as the same entity to the DOCSIS-based provisioning system.

MAC Address

[DOCSIS] uses the 48-bit IEEE 802.3 MAC Address of the cable modem's RF interface for equipment identification.

[DPOE] uses the 48-bit IEEE 802.3 MAC Address of the D-ONU as the equipment identifier.

The MAC address is used during the DHCP process (in the BOOTP chaddr field) to identify the consumer of the configuration file (the cable modem in [DOCSIS] and the vCM in [DPOE]). Because the provisioning consumer and the provisioned equipment are the same, it is not necessary, though it may be desirable, to use the MAC address in the cable modem configuration file.

CMIM

In some use cases it is also necessary to identify the specific port to which a user is being attached. For example, one port on a multi-port D-ONU may provide Internet access to subscriber-A and another port may provide Layer-2 VPN service to another subscriber. Within the configuration file, the CMIM TLV is used to associate the port to the service flow.

<u>Classification</u>

Classification refers to the process of inspecting incoming packets or frames and associating them with an action based on their content. There are two possible classifier actions in [DOCSIS] and [DPOE] – packets can be dropped or they can be forwarded into a service flow.

Classification occurs independently in the upstream and downstream direction; therefore the DOCSIS provisioning dictionary defines two different classifier types. The classifier TLVs are containers for a rich set of classification criteria.

Classification Criteria

Classification criteria are available for all IPv4 and IPv6 headers, TCP and UDP headers, Ethernet MAC headers and MPLS headers. The complete list of available classification criteria are cataloged in Annex C of [D3-MULPI]. Some examples of particular interest to providing business services are:

CMIM – Specifies the access equipment's physical port on which user traffic is expected to enter the network. Used in upstream classification.

IEEE 802.1ad C-VID – Specifies the customer-supplied VLAN ID (in Provider Bridging format) to use for classification. Typically (but not necessarily) used in upstream classifiers.

IEEE 802.1ad C-PCP – Specifies the customer-supplied CoS markings (in Provider Bridging format) to use for classification. Typically (but not necessarily) used in upstream classifiers.

IEEE 802.1Q VLAN-ID – Specifies the customer-supplied VLAN ID (in the pre-802.1ad format) to use for classification. Typically (but not necessarily) used in upstream classifiers.

IEEE 802.1P User Priority – Specifies the customer-supplied CoS markings (in the pre-802.1ad format) to use for classification. Typically (but not necessarily) used in upstream classifiers.

IPv4 ToS – Specifies the IPv4 ToS markings to match in classification. Typically (but not necessarily) used in upstream classifiers.

IPv6 Flow Label – Specifies the IPv46 flow label to match in classification. Typically (but not necessarily) used in upstream classifiers.

IPv6 Traffic Class – Specifies the IPv6 traffic class markings to match in classification. Typically (but not necessarily) used in upstream classifiers.

MPLS Label – Specifies the MPLS label to match in classification. Typically (but not necessarily) used in downstream classification.

MPLS Traffic Class – Specifies the MPLS Traffic Class (also known as the EXP bits) to match in classification. Typically (but not necessarily) used in downstream classification.

Service Flow Description

A Service Flows is the construct that describes the treatment of a specific classified flow of packets as it is transmitted across the access network. The service flow also maintains the association of the packets to the Forwarder instance in the access aggregation system. Treatment includes the application of bandwidth profiles and could include traffic coloring and re-marking.

Bandwidth Profile

The encodings of most interest in the service flow description are those that describe the *Bandwidth Profile*. The bandwidth profile can take one of two forms. The DOCSIS QoS profile is the most commonly used and can contain settings for Traffic Priority, Maximum Sustained Traffic Rate, Maximum Sustained Traffic Rate, Maximum Traffic Burst, Minimum Reserved Traffic Rate and others.

The second form for a bandwidth profile is the *Metro-Ethernet Service Profile* (MESP). The MESP is based on [MEF10] and contains settings for Committed Information Rate (CIR), Committed Burst Size (CBS), Excess Information Rate (EIR), and Excess Burst Size (EBS) and the color mode.

Forwarder Description

Currently there are three Forwarder types in [DOCSIS], [DPOE], and [L2VPN]: IPv4 and IPv6 forwarders, MPLS forwarders, and layer-2 forwarders (typically 802.1Q-based bridges).

IPv4 and IPv6 Forwarders

In the DOCSIS TLV dictionary there are no TLVs to describe an IPv4 or IPv6 forwarder. Rather, IPv4 and IPv6 forwarders are assumed to be pre-configured and are the default forwarder on the system. That is to say, a service flow will be associated with an IPv4 or IPv6 forwarder unless there is an alternate association specified in the downloaded configuration file. It is possible that multiple IPv4 and IPv6 forwarders exist, therefore it would be necessary to specify a Service Class Name or Attribute Mask TLV to steer the service flow to the intended forwarder instance.

Layer-2 Forwarders

Layer-2 Forwarders are intended to simply pass Ethernet frames received at the CMCI to an external entity using MAC-layer bridging rules. To accomplish this, the forwarder requires a description that simply contains an NSI Encapsulation to specify the layer-2 attributes. Some key TLVs available to configure layer-2 forwarders are:

IEEE 802.1Q Encapsulation – This encapsulation setting defines a layer-2 forwarder with 802.1Q tagging.

IEEE 802.1ad Encapsulation – This encapsulation setting defines a layer-2 forwarder with Provider Bridge tagging.

IEEE 802.1ah Encapsulation – This encapsulation setting defines a layer-2 forwarder with Provider Backbone Bridge tagging.

MPLS Forwarders

MPLS Forwarders are intended to encapsulate into MPLS packets arbitrary frames or packets received through an associated service flow. The forwarders can be configured to create explicitly defined pseudowires or they can be configured with parameters to enable BGP auto-discovery, BGP signaling or LDP signaling.

To create an MPLS Forwarder the NSI Encapsulation must contain the proper combination of the following TLVs:

MPLS Pseudowire ID – Defines the primary pseudowire identifier.

MPLS Peer IP address – Specifies the IP address of the MPLS PW peer.

Pseudowire Type – Specifies whether the forwarder is a member of a VPLS or (for point-to-point pseudowires), whether the PW is in Ethernet tagged mode or Raw mode.

L2TPv3 Forwarders

Similar to MPLS forwarders, L2TPv3 Forwarders encapsulate arbitrary frames or packets into an L2TPv3 tunnel.

End-to-End Service Setup

When dynamic pseudowire signaling (LDP or BGP) and auto-discovery (BGP) are used, it is necessary to provide the forwarder the necessary information to perform these functions.

Attachment Group ID – Specifies the Attachment Group ID (AGI) used in dynamic MPLS and L2TPv3 signaling of point-to-point pseudowires.

Source/Target Attachment Group ID – Specifies the Source and/or Target AGI for dynamic signaling of point-to-point pseudowires.

BGP VPNID – Uniquely identifies the service (VPN) to the auto-discovery protocols. This value is unique to the service, but common among all attachments to the service.

Route Distinguisher – Specifies one or more route distinguishers (RD) to be associated with the service and advertised by this particular attachment (via BGP) to the service.

Route Target (import/export) – Specifies a list of route targets (RT) to be imported from and/or exported to the BGP database.

Supplemental Service Configurations

The TLV dictionary categories discussed above are concerned with establishing connectivity between the user's equipment and the forwarder. Subscribers often require more than simple connectivity, though. The DOCSIS TLV dictionary accommodates supplemental configurations, also.

Two features of recent interest to operator's Commercial Services business are timing and Service OAM. The DOCSIS TLV dictionary contains the constructs necessary to configure distribution of the [1588v2] Precision Time Protocol across the access network. The dictionary also contains the constructs to define ITU Y.1731 Maintenance Entity Groups (MEG) [Y1731].

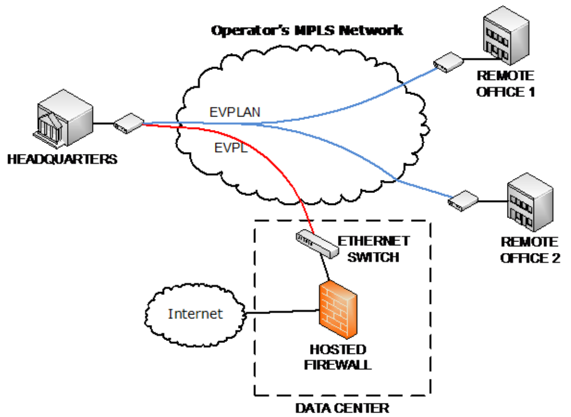


Figure 3 - Example Customer Network Configuration

DOCSIS AND SDN IN CONCERT

The early discussion in this paper described the overall architecture and the key elements of that architecture. Later, the discussion moved to requirements of the service provisioning API and the capabilities of DOCSIS-style provisioning.

What about the practical reality of this combination when asked to provision real services being offered today by cable operators? Given a non-trivial use case (in other words, NOT high-speed Internet access), how would an SDN architecture using DOCSIS-style provisioning function?

Consider a typical Business Services customer that wishes to connect several sites using Ethernet over a Layer-2 VPN and attach an Internet access service to the same VPN. The proposed topology is shown in Figure 3.

In this network, the cable operator would need to provision four (4) end points and a firewall instance. The four end points are

1. D-ONU (assume a DPoE connection) at the Headquarters
2. Cable modem at Remote Office 1
3. D-ONU at Remote Office 2
4. Ethernet Switch Port in the operator's data center

The service-specific provisioning process begins with the business systems requesting service for the four end points and the hosted firewall. Following the SDN provisioning model described earlier the business systems would describe each site to the Service Provisioning API as follows:

EVPLAN at Headquarters

ATTRIBUTE	VALUE
SERVICE ID	CUSTOMER_1_SERVICE_1
SERVICE TYPE	L2VPN
SERVICE ROLE	ROOT
ACCESS EQ'T ID	<HQ D-ONU MAC>
ACCESS PORT ID	Port 1
TRAFFIC SPEC	User-Supplied C-TAG = 99
UPSTREAM BW	50Mbps
DOWNSTREAM BW	25Mbps

EVPLAN at Remote Site 1

ATTRIBUTE	VALUE
SERVICE ID	CUSTOMER_1_SERVICE_1
SERVICE TYPE	L2VPN
SERVICE ROLE	ROOT
ACCESS EQ'T ID	<SITE 1 CM MAC>
ACCESS PORT ID	Port 1
TRAFFIC SPEC	User-Supplied C-TAG = 99
UPSTREAM BW	5Mbps
DOWNSTREAM BW	25Mbps

EVPLAN at Remote Site 2

ATTRIBUTE	VALUE
SERVICE ID	CUSTOMER_1_SERVICE_1
SERVICE TYPE	L2VPN
SERVICE ROLE	ROOT
ACCESS EQ'T ID	<SITE 2 D-ONU MAC>
ACCESS PORT ID	Port 1
TRAFFIC SPEC	User-Supplied C-TAG = 99
UPSTREAM BW	5Mbps
DOWNSTREAM BW	25Mbps

EVPL at Data Center

ATTRIBUTE	VALUE
SERVICE ID	CUSTOMER_1_SERVICE_2
SERVICE TYPE	L2VPN
SERVICE ROLE	NODE
ACCESS EQ'T ID	<DC SWITCH MAC >
ACCESS PORT ID	Port 42
TRAFFIC SPEC	User-Supplied C-TAG = 35
UPSTREAM BW	100 Mbps
DOWNSTREAM BW	10 Mbps

EVPL at Headquarters

ATTRIBUTE	VALUE
SERVICE ID	CUSTOMER_1_SERVICE_2
SERVICE TYPE	L2VPN
SERVICE ROLE	NODE
ACCESS EQ'T ID	<HQ D-ONU MAC >
ACCESS PORT ID	Port 1
TRAFFIC SPEC	User-Supplied C-TAG = 35
UPSTREAM BW	10 Mbps
DOWNSTREAM BW	100 Mbps

The Service Provisioning API would need to expose a method to request the hosted firewall. This paper does not attempt to define this method or the data model for it because DOCSIS-style provisioning is clearly inappropriate for upper-layer services.

Next in the provisioning process the SDN Orchestrator, would process the service provisioning requests and any available data about the network through the business rules. Of course, by nature, the SDN Orchestrator would know the configuration and capabilities of the network and incorporate that knowledge in the processing. The output of this processing would be a set of decisions about how to configure the network to support the requested services.

The SDN Orchestrator would then use this set of decisions to push an abstract configuration set through the Network Provisioning Interface (NPI). The NPI would then translate that abstract configuration into a specific set of parameters to be fed into the DOCSIS Provisioning Interface.

The output of the NPI would be a set of DOCSIS configuration files with one crafted to match each site and the overall service configuration. We assume here that the

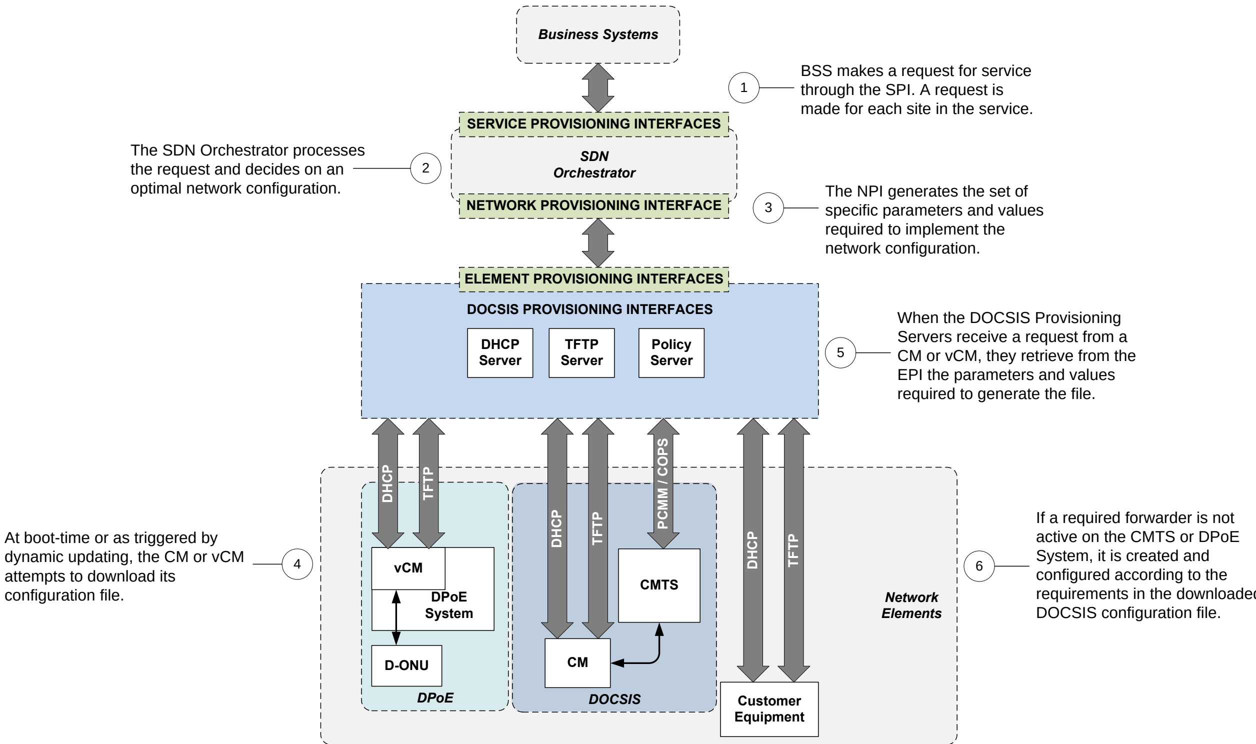


Figure 4 - Provisioning Flow through the SDN+DOCSIS Architecture

operator's network is configured to support VPLS and VPWS using explicitly defined peers. Using this assumption and the input provided through the Service Provisioning API, a set of DOCSIS configuration files can be constructed.

At boot-time the CM or vCM executes the procedures defined in [DOCSIS] or [DPOE], respectively. These procedures include DHCP to obtain an IP address and TFTP to download the configuration file. When the DHCP server or the TFTP server receive a request from the CM or vCM, they will retrieve from the Element Provisioning Interface (or the EPI will provide to them prior to the CM's attempt) the parameters needed to construct a response (DHCP) or the configuration file (TFTP) that will accomplish the intended network configuration.

A dynamic update could be signaled, whether via PCMM or the dynamic file update method defined in [DPOE]. In this case, the policy server (if PCMM) or the DHCP and TFTP servers (if dynamic file update) will retrieve the updated parameters from the EPI and respond back to the CM or vCM.

Once the configuration file is downloaded (or the dynamic update is processed), the CM registers the configuration with the CMTS. The CMTS (or DPoE System) examines the configuration to determine which forwarder to which the defined service flow(s) should be attached. If the forwarder does not exist, then the CMTS will create the forwarder according to the parameters contained in the configuration file.

An example configuration file for the D-ONU located at the headquarters is given below.

```
#BEGIN
Network Access = 1
# Constructs for EVPLAN
US Classifier
 SF Ref = 1
 802.1ad Classifier
   CVID = 99
```

```
DS Classifier
 SF Ref = 2
 802.1ad Classifier
   SVID = 1100
 L2VPN Encoding
   Vendor ID = FFFFFF
   VPNID = "CUSTOMER_1_SERVICE_1"

US Service Flow
 Ref = 1
 QOS ParamSetType = 7
 Min Reserved Rate = 50000000
 Max Sustained Rate = 50000000
 L2VPN Encoding
   Vendor ID = FFFFFF
   VPNID = "CUSTOMER_1_SERVICE_1"

DS Service Flow
 Ref = 2
 QOS ParamSetType = 7
 Min Reserved Rate = 25000000
 Max Sustained Rate = 25000000

L2VPN Encoding
  Vendor ID = FFFFFF
  VPNID = "CUSTOMER_1_SERVICE_1"
  L2VPN Mode = Encapsulation
  NSI Encapsulation Type = MPLS
Pseudowire
  MPLS Pseudowire Type = VPLS
  MPLS Pseudowire ID = 650001100
  MPLS Peer IP Address =
IPv4:192.168.100.1
  MPLS Peer IP Address =
IPv4:192.168.101.1
# end EVPLAN
# Constructs for EVPL
US Classifier
 SF Ref = 3
 802.1ad Classifier
   CVID = 35

DS Classifier
 SF Ref = 4
 802.1ad Classifier
   SVID = 1101
 L2VPN Encoding
   Vendor ID = FFFFFF
   VPNID = "CUSTOMER_1_SERVICE_2"

US Service Flow
 Ref = 3
 QOS ParamSetType = 7
 Min Reserved Rate = 10000000
 Max Sustained Rate = 10000000
 L2VPN Encoding
   Vendor ID = FFFFFF
   VPNID = "CUSTOMER_1_SERVICE_2"

DS Service Flow
 Ref = 4
 QOS ParamSetType = 7
 Min Reserved Rate = 100000000
 Max Sustained Rate = 100000000

L2VPN Encoding
 Vendor ID = FFFFFF
 VPNID = "CUSTOMER_1_SERVICE_2"
 L2VPN Mode = Encapsulation
 NSI Encapsulation Type = MPLS
Pseudowire
 MPLS Pseudowire Type = Ethernet Raw
Mode
 MPLS Pseudowire ID = 650001101
 MPLS Peer IP Address =
IPv4:192.168.102.1
# end EVPL
# EOF
```

The flow through the SDN architecture is depicted in Figure 4.

GAPS, CHALLENGES AND THE FUTURE

DOCSIS/DPOE can be used to provision services, such as L2VPN or Internet Access, that are commonly provided on an MSO's network. Most services can be provisioned via DOCSIS in such a way that no configuration beyond what is gleaned from the DOCSIS configuration file is necessary in the upstream network devices. In this model, the endpoint needs to know very little about the network between it and its remote-side destination(s). Likewise, the network needs to maintain very little state about the underlying service and its topology.

However, there are significant gaps in the support for provisioning more advanced services like L3VPN.

To support L3VPN, additional TLVs would need to be added to the DOCSIS provisioning dictionary to define

- L3VPN topologies (for example, Hub and spoke, partial mesh, NNI/Extranet)
- PE-CE routing protocol configuration
- Route policies needed to ensure the correct routes are announced, filtered, and tagged

Additional gaps exist when considering how to define more complex and chained services, such as providing inline firewall services, or providing access to cloud services from within a VPN. Higher layer services like these are examples of services where the service definition may extend beyond attachment circuits. In L3VPN, elements in the network provide a Layer 3 routed topology, participate in the routing protocol, and require more state exchange between the end point and the network, so in some ways the DOCSIS provisioning model may be inadequate.

One might reasonably ask why the existing framework around DOCSIS/DPOE could not simply be extended to include additional service provisioning since it's intended to be extensible. There is a point at which adding into the DOCSIS provisioning dictionary the TLVs necessary to represent more complex service definitions reaches diminishing returns, especially when considering the need to provision non-network services.

It may be conceivable that the DOCSIS-style configuration file and dictionary could represent virtually any service one might wish to describe. However, when multiple service elements are involved, especially if one or more are separate from the access termination, the constructs of a DOCSIS configuration file are likely to be less rich than modern configuration languages like NETCONF/YANG ([RFC 6020], [RFC 6241]), OpenFlow, or vendor-specific element management APIs.

Further, the DOCSIS method of delivering configuration files is not supported by many systems. While it could be adapted, one is forced to consider whether the effort is deserved over other available methods. In fact, the proposed architecture assumes that there are multiple methods at the Element Provisioning layer and that the SDN Orchestrator would generate the "glue" to chain individual element configurations into an end-to-end service.

The lack of support for DOCSIS-style provisioning in elements beyond cable modems, CMTSes, DPoE Systems and vCMs highlights another gap. Today [DOCSIS] is a monolithic specification. Included in that specification is the MAC and PHY requirements necessary to transmit data over an HFC network and service definitions and provisioning for all of the potential services that use that transmission technology.

It is not reasonable to expect a network equipment manufacturer to understand the entirety of [DOCSIS] if they do not support DOCSIS MAC/PHY interfaces. Yet it is difficult work to identify the specific elements of [DOCSIS] that are required to support DOCSIS-based provisioning without supporting the data transmission protocol.

[DPOE] is a step in the direction of separating DOCSIS provisioning from DOCSIS data transmission, but it accomplishes this not by extracting the provisioning elements from the specifications to make them standalone, but by emulating many of the MAC-layer elements of [DOCSIS] such that the protocol continues to work as-is. This approach works very well for most access-layer network technologies, but extending this to support a wider range of services would require that DOCSIS-style provisioning be separated into a specification of its own.

CONCLUSION

In the combined SDN+DOCSIS model presented here, the service definition and topology are abstracted from the physical access and the devices used to provide the service. This abstraction allows for maximum flexibility in building a provisioning system that is agnostic to the access technologies being used.

As MSOs consider the future of their networks and services, deploying an architecture such as the one proposed in this document allows them to build on existing capabilities while keeping flexibility to evolve the methods used to manage the network and service offerings. Manual provisioning of commercial services that cannot be provisioned via the existing DOCSIS infrastructure today is not a scalable and sustainable strategy.

DOCSIS-based provisioning systems manage hundreds of millions of devices around the world and process thousands of transactions per day. It is a tested and proven system. As cable operators look to expand into the $130 Billion Commercial Services market they will need to find ways to distinguish themselves over competitive operators.

DOCSIS-based provisioning is more than capable of provisioning the transport services offered by MSOs today. The potential exists to extend the DOCSIS TLV dictionary to support additional transport services. However, adding support for upper-layer services like cloud-computing into the DOCSIS provisioning framework requires careful evaluation and probably makes no sense.

By leveraging an SDN model with DOCSIS-based provisioning, the cable operator can reduce their time to bring new services to market, eliminate manual and error-prone configuration of network equipment, and increase the rate at which new customers can be added to the network. Especially in the context of complex Commercial Business offerings, the cable operator is able to reduce the service provisioning process to a transaction rather than a complex series of steps involving multiple systems and humans. These factors will allow the MSO to offer better-quality services at lower cost – making the cable operator a more desirable network provider.

[DPOE-MULPI] DOCSIS® Provisioning of EPON Specifications, DPoE™ MAC and Upper Layer Protocols Requirements, DPoE-SP-MULPIv1.0, Cable Television Laboratories, Inc.

[L2VPN] Data-Over-Cable Service Interface Specifications, Business Services over DOCSIS, Layer 2 Virtual Private Networks, CM-SP-L2VPN, Cable Television Laboratories, Inc.

[DPOE] Refers to the collective set of DPoE Specifications (all versions)

[D3-MULPI] Data-Over-Cable Service Interface Specifications DOCSIS 3.0, MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.0, Cable Television Laboratories, Inc.

[DOCSIS] Refers to the collective set of DOCSIS and L2VPN Specifications (all versions)

[8021D] IEEE Std. 802.1D-2004, IEEE Standard for Local and Metropolitan Area Networks – Media Access Control (MAC) Bridges, June 2004.

[8021Q] IEEE Std. 802.1Q-2011, IEEE Standard for Local and Metropolitan Area Networks – Media Access Control (MAC) Bridges and Virtual Bridge Local Area Networks, August 2011.

[1588v2] IEEE Std. 1588™-2008, IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems.

[MEF10] MEF Technical Specification MEF 10.2, Ethernet Services Attributes Phase 2, Metro-Ethernet Forum, 27 Oct. 2009

[Y1731] ITU-T Recommendation Y.1731 (07/11), OAM functions and mechanisms for Ethernet based networks

[OLIVERDEMO] http://btreport.net/2013/10/oliver-demos-sdn-docsis-provisioning/

[EXPO2012] N. Nadiraju, K. Noll, "Transforming 'DOCSIS Provisioning of EPON' to 'DOCSIS Provisioning of Everything'", SCTE Cable-Tec EXPO Proceedings, October 2012.

[RFC 6241] IETF RFC 6241, Network Configuration Protocol (NETCONF), June 2011

[RFC 6020] IETF RFC 6020, YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF), October 2010

Study of Wi-Fi for In-Home Video Streaming

Alireza Babaei, Neeharika Allanki, Vikas Sarawat
CableLabs

Abstract

With the introduction of 802.11ac and the promise of gigabit Wi-Fi, it seems to be, at least theoretically, possible to simultaneously stream multiple HD quality videos over the Wi-Fi network. This opens up opportunities for cable operators and their customers to distribute cable video to subscriber devices throughout the homes over Wi-Fi networks. This provides an attractive value-add to both MSOs and their customers.

In this paper, our goal is to evaluate the feasibility of using Wi-Fi for wireless distribution of HD quality cable video under different circumstances or configurations. Our results indicate that although Wi-Fi performance can not be guaranteed in all circumstances, it is generally possible to stream multiple HD videos if certain conditions, including conditions on radio configuration, interference, and signal attenuation are met. We present our test observations for 802.11n and 802.11ac Wi-Fi radios with different configurations and provide recommendations for video streaming over operator-managed Wi-Fi networks.

ACRONYMNS

ACS	Auto Channel Selection
AP	Access Point
COAM	Customer Owned and Managed
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
DBDC	Dual Band Dual Concurrent
DBM	Dynamic Bandwidth Management
DTCP	Digital Transmission Content Protection
FCC	Federal Communications Commission
GoP	Group of Picture
HEW	High Efficiency Wi-Fi
IEEE	Institute of Electrical and Electronics Engineers
ISM	Industrial, Scientific and Medical
MCS	Modulation and Coding Scheme
MIMO	Multiple Input Multiple Output
MoCA	Multimedia Over Coax Alliance
MPEG	Moving Pictures Expert Group
PER	Packet Error Rate
PLC	Power Line Communications
PLR	Packet Loss Ratio
QAM	Quadrature Amplitude Modulation
RRM	Radio Resource Management
RSSI	Received Signal Strength Indicator
SDM	Spatial Division Multiplexing
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio
SON	Self Organizing Networks
TCP	Transmission Control Protocol
TDLS	Tunnel Direct Link Setup
TxBF	Transmit Beamforming
UDP	User Datagram Protocol
UNII	Unlicensed National Information Infrastructure
WHDMI	Wireless High Definition Multimedia Interface
WLAN	Wireless Local Area Networks

INTRODUCTION

Consumers want to watch cable video services throughout their homes, but neither the subscriber nor service provider wants the expense and inconvenience of running new cables. Additionally, both subscribers and MSOs would like to avoid the cost of additional set-tops by leveraging Customer Owned and Managed (COAM) devices such as smartphones, tablets and smart TVs where Wi-Fi capabilities are ubiquitous. Wi-Fi offers the opportunity for the service provider to deliver video streaming throughout the home without the expense of new cabling.

Over the last decade Wi-Fi performance has improved exponentially. The latest Wi-Fi standard – 802.11ac – promises support for speeds greater than 1 Gbps. Many Wi-Fi products, including 802.11n, support Multiple Input Multiple Output (MIMO), Transmit Beamforming and operations in 5 GHz spectrum. These technologies promise greater reliability and even better performance. A number of Wi-Fi silicon vendors (e.g., Broadcom, Celeno, Qualcomm, Quantenna, etc.) are working on optimizing Wi-Fi silicon for in-home high definition video streaming.

To evaluate Wi-Fi technology for in-home distribution of Full HD cable video, CableLabs conducted Wi-Fi performance and video quality measurement tests on multiple Wi-Fi products including 802.11n and 802.11ac. The tests were conducted at CableLabs' Louisville facility and in homes of different sizes and construction materials in Colorado and on the East Coast.

This paper first provides a technical overview of wireless technologies enabling in-home video streaming. Following that the paper provides a discussion on challenges for using Wi-Fi for in-home cable video streaming. Subsequently, the paper provides observations from testing and recommendations for operators considering Wi-Fi for video streaming. Finally, the evolution of Wi-Fi networks and the closing thoughts are included in the conclusion section.

HOME MULTIMEDIA WIRELESS NETWORKS

This section provides a technical overview of wireless technologies enabling in-home video streaming.

WirelessHD: The WirelessHD defines a wireless protocol that enables consumer devices to create a wireless video area network. The WirelessHD uses 60 GHz Frequency band. Unlike the Wireless Gigabit Alliance (WiGig) technology, it does not include an option to fall back to 5 GHz band. The indoor coverage range is about 10 meters, which is adequate for video streaming between two devices in the same or the next room. The technology supports Transmit Beamforming and data transmission rates of up to 28 Gbps.

Additionally, it includes support for 3D content and 4K resolution including HDCP 2.0 and DTCP for content protection. WirelessHD products from a number of vendors such as Panasonic, Sony, and LG are available today. Silicon Image is the primary silicon vendor for WirelessHD. The competing technologies include WiGig and Wireless Home Digital Interface. The primary use for WirelessHD is the delivery of high quality, uncompressed A/V content. The picture below shows the logo for WiHD devices.

WHDMI: The Wireless High Definition Multimedia Interface (WHDMI) enables wireless delivery of uncompressed High Definition Television Vision. Unlike the WirelessHD, it uses 5 GHz frequency band and does not include support for 60 GHz band. The indoor coverage range is about 30 meters, which is about the same as 802.11ac. While the WirelessHD is only good for video streaming between two devices in the same or next room, the WHDMI claims support for video streaming throughout the home. The technology supports 20 MHz and 40 MHz channel bandwidth to support up to 1.5 Gbps for uncompressed 1080i and 720p, and up to 3 Gbps for uncompressed 1080p, respectively.
Additionally, it supports capabilities to prioritize the most visually significant bits of a video stream. WHDI products from a number of vendors such as Hitachi, Motorola, Samsung, Sharp, Sony, HP, and LG are available today. Amimon is the primary silicon vendor for WHDI. The competing technologies include WiGig and WiHD. The primary use for WHDI is the delivery of high quality, uncompressed A/V content. The picture below shows the logo for WHDI devices.

WHDI™

WiGig: The WiGig technology offers short-range multi-gigabit connections for a wide variety of applications including video, audio, and data, while the WHDI and WiHD focus is on delivering high quality uncompressed video. The following is a list of applications on which WFA is focusing:

- WiGig Display Extension
- WiGig Serial Extension
- WiGig Bus Extension
- WiGig SD Extension

The WiGig technology is the basis of the IEEE 802.11ad amendment and supports Beamforming and data rates up to 7 Gbps in a 60 GHz frequency band. Many WiGig products are also expected to support Wi-Fi, along with mechanisms for smooth handovers from 60 GHz to 2.4 GHz and 5 GHz band. Similar to the WiHD, the indoor coverage range is about 10 meters, which is adequate for communication between two devices in the same or next room.
A number of vendors, including Atheros, Marvell and Broadcom, Dell, Intel, Panasonic and Samsung, are working with the WFA in the development of technology and certification testing program. The WFA currently expects to launch the WiGig certification program in 2015. The competing technologies include WHDI and WiHD. The picture below shows the logo that WiGig certified products are expected to display.

WiGig™
CERTIFIED
by Wi-Fi Alliance

802.11n and 802.11ac: Both 802.11n and 802.11ac technologies support enough throughput to support in-home HD video streaming. 802.11ac is the current generation Wi-Fi technology, and it supports some features that were not part of the 802.11n standard. The table below provides a highlight of some of the differences between 802.11n and 802.11ac.

FEATURES	802.11N	802.11AC
Frequency Band	2.4 or 5 GHz	5 GHz only
Channel Bandwidth	20, 40 MHz	20, 40, 80, 160, 80+80 MHz
Modulation & Coding Scheme	64 QAM	256 QAM
Spatial Streams	Up to 4	Up to 8
Transmit Beamforming	Optional	Standardized
Max Throughput	600 Mbps	3.2 Gbps
MU-MIMO	No	Yes
Availability	Available for some time now	First generation available now

In addition to the features in the previous table, 802.11ac also includes support for features such as Dynamic Bandwidth Management, which can be very handy in mitigating interference and improving spectral efficiency. This feature allows an AP to dynamically select channel bandwidth to each client on a frame-to-frame basis.

The first generation 802.11ac products support only 20, 40 and 80 MHz channel bandwidth. The current FCC spectrum rules do not allow for a continuous and homogenous 160 MHz channel. Channel bandwidth of 80 MHz+80 MHz and 160 MHz are expected in the second-generation 802.11ac products. Support for MU-MIMO and Dynamic Bandwidth Management are also expected in the second-generation 802.11ac products.

Tunnel Direct Link Setup (TDLS): TDLS allows network-connected client devices to create a secure, direct link to transfer data more efficiently. The client devices first establish a control channel between them through the AP. The control channel is then used to negotiate parameters (e.g., channel) for the direct link. APs are not required to support any new functionality for two TDLS compliant devices to negotiate a direct link.
TDLS offers multiple benefits, including efficient data transmission between client devices by removing the AP from the communication link. Use of a direct communication channel also allows the client to negotiate capabilities independent of the AP. For example, clients can choose a wider channel, efficient modulation scheme, and a security and channel that are more suitable for direct link between the client devices.

TDLS devices, communicating with each other over a direct link, are also allowed to maintain full access to the Wi-Fi network simultaneously, which, for example, allows the client device to stream video to another device in the home over the direct link; and at the same time allows the user to surf internet via connectivity to the AP. If the TDLS direct link is switched to another channel, the stations periodically switch back to the home channel to maintain connectivity with the Wi-Fi network.
The WFA has certified multiple products for TDLS, including Broadcom and Marvel. TDLS is based on IEEE 802.11z, and is one of the optional features of Miracast (Wi-Fi Display).

Wi-Fi Direct: Wi-Fi Direct allows Wi-Fi client devices to connect directly without use of an AP. Unlike TDLS, Wi-Fi client devices are not required to be connected to an AP to establish a Wi-Fi Direct link. Wi-Fi Direct also includes support for device and service discovery. Wi-Fi Direct devices can establish a one-to-one connection, or a group of several Wi-Fi Direct devices can connect simultaneously.

Wi-Fi Direct offers multiple benefits, such as ease of use and immediate utility; enables applications such as printing by establishing a peer to peer connection between the Wi-Fi Direct enabled printer and client device; content sharing between two Wi-Fi Direct enabled devices; and displaying content from one Wi-Fi Direct device to another without requiring any Wi-Fi network infrastructure. Wi-Fi Direct certifies products, which implement technology defined in the WFA Peer-to-Peer Technical Specification. The WFA has certified multiple products for Wi-Fi Direct. As of 2012, there are over 1100 Wi-Fi Direct certified products. Wi-Fi Direct is the core transport mechanism for Miracast (Wi-Fi Display).

Miracast: Miracast provides a seamless display of content between devices using Wi-Fi Direct as the transport mechanism. Miracast also includes optional support TDLS as a transport mechanism.

The key features supported in Miracast include device and service discovery, connection establishment and management, security and content protection, and content transmission optimization. Similar to Wi-Fi Direct and TDLS, Miracast is client functionality and does not require updates to AP devices.

Primary use cases for Miracast are screen mirroring and video streaming. Miracast certifies products which implement technology defined in the Wi-Fi Display Technical Specification. As of this writing many devices (e.g., Smart phones) have been certified for Miracast.

Airplay: AirPlay is an Apple proprietary technology. It enables iTunes and other media systems to use local area networks to stream audio and video to Apple TV or other AirPlay-enabled sound systems or remote speakers. AirPlay is built on bonjour technology, which uses Multicast DNS (mDNS) and DNS-based Service Discovery (DNS-SD).

AirPlay includes support for media protocols such as RTSP, RTP, RTCP and HTTP Live Streaming. The supported audio format includes AAC and MP3. The supported video format is H.264. AirPlay also includes support for FairPlay – another Apple proprietary technology - for DRM and link Protection.

AirPlay is incompatible with DLNA in many ways. For example, the service discovery and link protocols used by the two technologies are different. DLNA also supports many more options for video and audio format.

AirPlay has a limited objective: media networking within Apple's closed ecosystem. Apple builds nearly all of the hardware and software, with the exception of audio-streaming components for music players. With these limited goals, AirPlay requires a simple set of protocols and media formats. Testing interoperability and usability of all the combinations and permutations takes very little time.

Digital Living Network Alliance (DLNA): DLNA is a generalized approach to media networking, designed to work with all media devices – methodologies which interconnect everything from everyone. In contrast to Airplay, DLNA incorporates every possible media format. Universal Plug and Play (UPnP) provides the principal underlying structure for DLNA. The UPnP Forum was formed "to enable device-to-device interoperability and facilitate easier and better home networking." Nearly a thousand companies are now members of UPnP Forum. The UPnP Forum develops and publishes UPnP Device Architectures that "define how to use IP to communicate between devices" and Device Control Protocols, which define specific services between devices.

DLNA divides devices into two broad categories: "home network devices" (essentially anything that is plugged into an electrical outlet and provides at least Ethernet networking) and "mobile handheld devices" (anything that runs on batteries and uses Wi-Fi). The mandatory and optional requirements for these two categories are quite different. Home network devices are required to support only JPEG images, LPCM audio, and MPEG2 video. Everything else is optional. Mobile handheld devices are required to support more formats: JPEG images, MP3 and AAC LC audio, and MPEG4 AVC video.

DLNA uses DTCP-IP for link protection. DTCP (Digital Transmission Content Protection) is a widely- accepted mechanism to protect high-value digital media such as movies and network videos when they are transferred between devices – such as between a digital set-top box and a TV. DTCP-IP (DTCP for Internet Protocol) is an extension of DTCP for protected transmission over IP networks – such as between a PC and a digital

TV, or over a network link between a cable gateway and a remote digital TV.

POTENTIAL CHALLENGES OF VIDEO OVER WI-FI

This section provides a discussion on factors that can have an impact on video service delivery using Wi-Fi.

Video is known to be a demanding application with strict requirements on throughput, delay, packet loss and jitter. One reason for stringent requirements is the way video is compressed. The common objective of all the video compression methods is to reduce redundancy in the spatial and temporal domains. Corresponding to spatial and temporal redundancy, intra-frame and inter-frame compression have been applied in video coding algorithms. Video packets can be classified as inter-frame (I frame) and intra-frame (P and B frames) packets. Intra-frame packets serve as reference for other inter-frame packets and their loss leads to error propagation to adjacent packets. Video transmission over the error-prone wireless channels is, therefore, inherently a challenging task.

One other unique characteristic of video is that it exposes the deficiencies in the network in an abrupt manner. In other words, unlike other applications, there may not be a smooth transition from good quality video to the poor quality video, and if the network cannot guarantee some minimum thresholds on any or a combination of throughput, delay, jitter or packet loss, the video may not be intelligible and its quality may drop immediately.

In the following sections, the main challenges of video transmission over Wi-Fi will be discussed in more detail.

Throughput: Video applications require some minimum throughput depending on the type of coding method used for compression of video. The more advanced video coding algorithms compress the video with a fewer number of bits and, as such, require lower data-rate transmission media. However, mobility and temporal variation of wireless channels lead to orders of magnitude fluctuation of throughput for wireless clients. As a client moves away from the AP with which it is associated, the Wi-Fi signal experiences higher path loss and therefore the client will receive smaller average throughput. In addition, wireless signals experience significant multipath in indoor environments. Even for a fixed position of the Wi-Fi client, the received signal strength may vary significantly over time due to movement of objects and variation in multipath. As a result, during the lifetime of a video session, it is possible that the client's throughput is below or above the minimum requirement. This presents a challenge to the conventional QoS approaches like admission control. While the client can receive sufficient throughput for reliable video transmission at the time of admission to network, there is no guarantee that its throughput will remain above the threshold all the time.

Delay and Jitter: Delay and jitter can be particularly important for interactive real-time video applications. These applications typically use UDP transport protocol to avoid packet retransmissions and meet the delay requirements. That said, UDP is a connectionless transport protocol that can lead to increased packet loss, which can also lead to poor video quality. There can be, therefore, a tradeoff between delay and packet loss for these applications. Due to delay and delay variation (jitter), playback buffers may be needed at the client and AP to smooth out the video when packets arrive after their playout times. Delay and jitter may also have an impact on the buffer size at the client and AP.

Packet loss: As described earlier, a sequence of compressed video frames, called group of picture (GoP), consists of both intra-frame

packets and inter-frame packets. The inter-frame packets act as reference for decoding other packets in the GoP. Loss of inter-frame packets can lead to error propagation which can significantly degrade the video quality. Consequently, for the same packet loss ratio (PLR), video quality can degrade more noticeably compared to other applications.
802.11 networks, on the other hand, are inherently prone to packet loss. One reason is the use of the CSMA/CA random access method in these networks. Although CSMA/CA is based on carrier sensing and collision avoidance, collisions and packet loss cannot be completely avoided due to the distributed nature of this protocol. CSMA/CA uses a retransmission mechanism to resend the lost packets. In addition, the retransmission mechanism can introduce additional delay and jitter.

The other contributor to packet loss is the short-term fast fading which results in rapid fluctuation in received signal strength. The Modulation and Coding Scheme (MCS) chosen by the AP must be adapted corresponding to the variation in the received signal-to-interference-plus-noise ratio (SINR) to keep the bit error and packet error rate below an acceptable level. However, it is possible that the rate adaptation algorithm, chosen by the AP, is not able to keep up with fast channel variation, and some packets may be lost in this transition. In addition, the rate adaptation algorithms are proprietary algorithms chosen by the AP vendors. The more aggressive rate adaptation algorithms can also lead to higher packet loss.

Unlicensed spectrum: Wi-Fi operates in the unlicensed 2.4 GHz Industrial, Scientific and Medical (ISM) band and 5 GHz in the Unlicensed National Information Infrastructure (UNII) band. The ISM unlicensed band is shared by other communications devices and appliances including Bluetooth, Zigbee and microwave ovens, which generate interference to Wi-Fi devices. Wi-Fi devices also share UNII bands with 5 GHZ cordless phones; however the UNII bands are much less crowded compared to the ISM bands. These non-Wi-Fi devices generate interference and can degrade the Wi-Fi throughput and packet loss performance.

Interference: Unlike cellular networks, where frequency planning and interference management is an integral part of the network design, in-home Wi-Fi networks are generally not methodically planned and coordinated. It is not uncommon to have multiple co-channel APs operate in close proximity. Both co-channel and adjacent channel interference can substantially degrade the Wi-Fi performance.

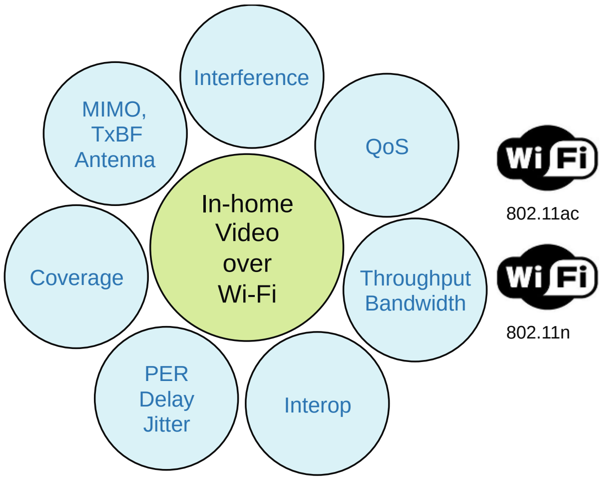
Coverage: The video requirements on throughput, delay and jitter introduce constraints on the coverage of Wi-Fi inside a residential area with acceptable video quality. The farther a client is from the AP, the lower its achievable throughput and potentially the higher the packet delay and jitter will be experienced by the client. The situation is more severe in 5 GHz compared to 2.4 GHz, as path loss for 5 GHz signals is larger for a given distance. In addition, for most of the construction material, 5 GHz signals experience more attenuation compared to 2.4 GHz signals.

The following section provides an overview of observations from the Wi-Fi testing done a CableLabs.

OBSERVATIONS FROM VIDEO OVER WI-FI TESTING

The observations in this section are based on the Wi-Fi and video quality measurement tests on three Wi-Fi products, including 802.11n and 802.11ac. Wi-Fi silicon vendors provided both the Wi-Fi Access Point (AP) and client (STA) for testing.

The tests were conducted in CableLabs' Louisville facility and multiple houses of different sizes and construction material (five in Colorado and six on the US East Coast). Wi-Fi coverage data was collected from an additional 25 Colorado houses to supplement the Wi-Fi performance data. As shown in the figure below, the Wi-Fi performance was measured using a number of metrics.

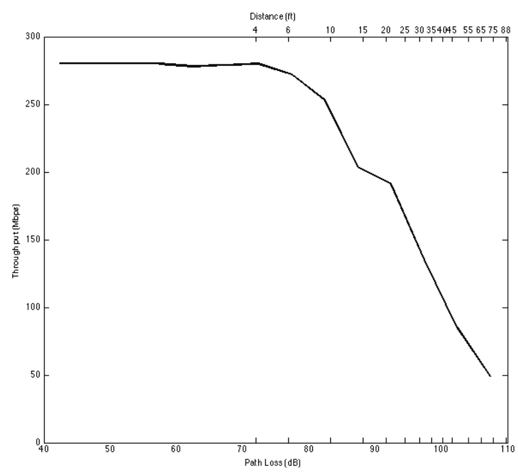


The test results indicate that Wi-Fi can reliably transport HD video in the home; but Wi-Fi network performance is highly dependent on a number of variables, including construction materials, distance between AP and Client, level and type of interference, Multiple Input Multiple Output (MIMO) configuration, Transmit Beamforming, antenna orientation, RF spectrum, background traffic, and device capabilities.

Based on the testing, some of the important considerations and observations for using Wi-Fi for in-home video distribution include:

- Interference can be a significant issue and AP location and channelization should be set to avoid it. Not only co-channel, but also adjacent channel and alternate channel interference are significant issues when the two Wi-Fi APs are less than 20 and 10 feet apart respectively. Interference from 5.8 GHz cordless phones also impacts Wi-Fi performance significantly.

- Wi-Fi signals in the 5 GHz band provide excellent coverage in homes with drywall panel walls. Analysis shows that three MPEG-2 streams can be successfully transmitted to three Wi-Fi clients up to 80 feet from the AP in a home with drywall construction. Wi-Fi signals in the 5 GHz band offer more limited coverage in houses with brick walls (or concrete floors). In these houses, HD Video streaming is possible if there is only one wall between the AP and clients. Wi-Fi signal attenuation is too high for two or more brick walls to reliably support HD video streaming. As the graph below shows, in a drywall house Wi-Fi is capable of serving 60 Mbps up to ~80 feet away from the AP. There are two X-axes in the graph - one at the bottom and the other at the top of the chart. The X-axis at the bottom of the chart is for pathloss and the X-axis at the top shows the distance between AP and the client. The Y-axis shows the TCP throughput in Mbps.

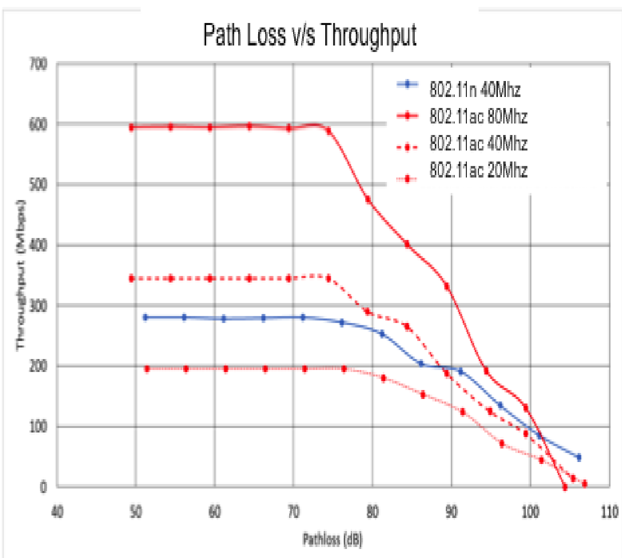


- Wi-Fi networks with the same vendor's AP and clients offer better performance than a Wi-Fi network with multi-vendor Wi-Fi products.

- Support for WMM and airtime fairness is useful on an AP and clients, if the same band and channel are used for both video as well as other services (e.g., data). Some of the APs were successfully able to prioritize video traffic in the presence of congestion from data traffic once the

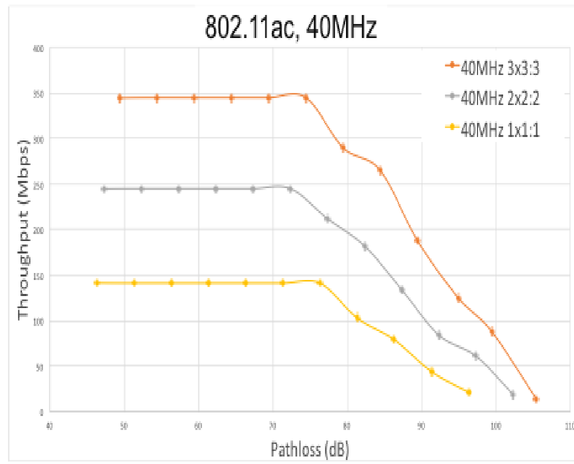
WMM tags were added to the video traffic.

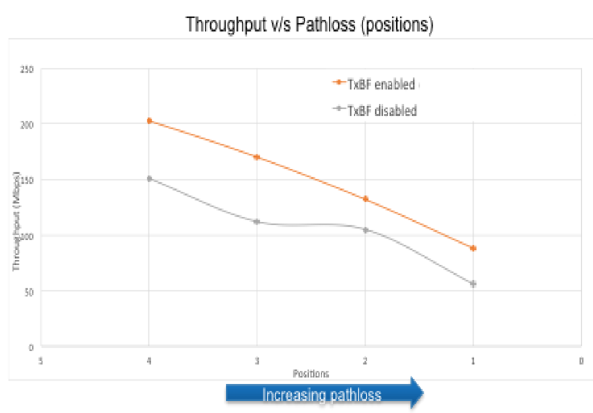
- Both 802.11ac and 802.11n technologies are capable of supporting in-home video streaming. 802.11ac can deliver higher throughput than 802.11n as a result of the support for 80 MHz channel and 256-QAM. This advantage is more obvious when Wi-Fi clients are at close range to the Wi-Fi AP. The performance of 802.11ac and 802.11n technologies is comparable at longer ranges. (e.g., around < -70 dBm RSSI). The following graph, *Path Loss v/s Throughput*, shows average TCP throughput measured in a conducted environment for 802.11n and 802.11ac AP and Client using 20, 40 and 80 MHz channels.



- The Packet Error Rate (PER) performance for some Wi-Fi solutions is consistently low, while for others PER performance varied from one test run to another.

- Transmit Beamforming (TxBF) and spatial division multiplexing improved wireless system performance. Increasing the number of spatial streams increases the Wi-Fi system throughput; however, the relative gain in throughput is less as the number of spatial streams increases. For example, the gain in throughput as a result of going from 2 spatial streams to 3 spatial streams is less than the gain in throughput as a result of going from 1 spatial stream to 2 spatial streams. TxBF is not standardized in 802.11n, resulting in a lack of interoperability. TxBF is standardized in 802.11ac, but wide scale interoperability across multiple vendor products needs to be verified.





- Wi-Fi utilizes unlicensed spectrum that is not solely under the control of operators. Numerous devices with a variety of technologies may utilize the spectrum. Compared to shielded environments such as the HFC network, video over Wi-Fi may be subject to more frequent radio disturbances that are not possible to be completely mitigated by the operator.

OPERATOR GUIDELINES AND RECOMMENDATIONS

This section provides guidelines and recommendations for the operators considering deployment of Wi-Fi for video streaming.

802.11ac vs. 802.11n: 802.11ac delivers higher throughput than 802.11n, as a result of the support for 80 MHz channels and 256-QAM. This advantage is more obvious when Wi-Fi clients are at close range to the Wi-Fi AP. The throughput performance of the two technologies is comparable at long range (e.g., < -70 dBm RSSI). While either 802.11n or 802.11ac can be used for video streaming, 802.11ac is the current generation Wi-Fi technology, and supports some features that were not part of the 802.11n standard. 802.11ac also includes support for features such as Dynamic Bandwidth Management, which can be very handy in mitigating interference and improving spectral efficiency. This feature allows an AP to dynamically choose channel bandwidth to each client on a frame-to-frame basis.

The first generation 802.11ac products support only 20, 40 and 80 MHz channel bandwidth. The current FCC spectrum rules do not allow for a 160 MHz channel. Channel bandwidth of 80 MHz+80 MHz and 160 MHz are expected in the second-generation 802.11ac products. Support for MU-MIMO and Dynamic Bandwidth Management are also expected in the second-generation 802.11ac products.

Installation Consideration: Wi-Fi performance can be affected by a number of factors, including construction material of the house, location of the AP, and interference from other Wi-Fi networks using the same/adjacent/alternate channel as well as non-Wi-Fi sources such as 5 GHz cordless phones. An optimum placement of the AP in the customer house makes a big difference in Wi-Fi signal coverage and performance. MSOs deploying Wi-Fi for video streaming should consider development of best practices for in-home Wi-Fi installation and for educating technicians. The following paragraphs provide recommendations and guidelines for in-home Wi-Fi installation.

APs used in residential deployments normally come equipped with omnidirectional antennas with minimal to no directional antenna gain. For equal coverage in each direction, AP should be placed in a central location in the house. In a multi-storied house, the technician should place the AP on the middle floor. While placement of the AP in the middle of the house is normally a good strategy, there is value in developing a test tool that can identify optimum location for the AP in a house since some factors such as high density of furniture in one location of the house can make the central location non-optimum.

Frequent movement of the AP by customers is also undesirable and should be avoided since antenna orientation and placement can impact performance. At a minimum, the technician should place the client and AP devices in correct orientation (i.e., a wireless bridge that was intended for vertical (“standing”) position, should not be installed horizontally). The location for Wi-Fi devices should be chosen such that they are out of high foot traffic and the children’s play area.

For AP and client devices with external antennas, manufacturer-provided instructions should be followed to properly configure the antennas before leaving the house. It may also be helpful to explain the proper antenna orientation to the customer.

With the AP in place, the MSOs should conduct a Wi-Fi signal coverage survey around the house. If the signal strength at the TV locations is insufficient to support HD video streaming, the AP location should then be adjusted until “appropriate” signal strength (e.g., RSSI) is available at all TV locations. In addition to taking signal strength measurements, the technician should also conduct a throughput test to verify that enough capacity is available to support simultaneous video streaming to all TVs connected to the Wi-Fi network.

One method for throughput measurement is to first identify the number of TVs connected wirelessly, second, take signal strength measurement at each TV location, and then measure network throughput to a client supporting the least Wi-Fi signal strength. If the measured throughput to this client is greater than the throughput required for each video stream times the number of TVs plus a margin of 20%, then the Wi-Fi network can be considered good for delivering video in that house. Changes in the environment (e.g., interference, door closing) can still interrupt the video streaming but a built-in 20% margin should provide some protection. If the Wi-Fi performance is not sufficient (in some parts of the house) for video streaming, operators should consider deploying multiple APs or use other Wi-Fi extension technologies such as Multimedia over Coax (MOCA) or Power Line Communications (PLC) or a combination thereof.

Currently, CableLabs is working on a new project called Future Home Networks; as a part of that, we are taking a look at various home network technologies, including G.hn , PLC etc.

It is also recommended that the technician take following measurements during installation for later review and use by the engineering and customer support team to determine what has changed in the environment over time.

- Collect Wi-Fi coverage data in each room of the house
- Number of other Wi-Fi networks seen from the house, including their signal strength, channel size, channel number, frequency band, MAC address of the APs, number and type of active clients.
- SNR and RSSI at each video client
- Identify legacy Wi-Fi APs and client devices in the home and neighborhood

MSOs should choose a Wi-Fi channel to avoid co-channel, adjacent channel and alternate channel interference. If such choice is not available, to minimize the impact of Adjacent and Alternate channel interference, the technician should first find the location of other APs in the house and neighborhood and place the AP and IP STB in such a manner that it is at least 20 feet away (assuming open air) from the other nearby APs. If it is not feasible to measure the distance then the technician should verify that the signal strength of adjacent channel is no more than 6dB higher than that of the wanted signal. Similarly, the technician should verify that the signal strength of the alternate channel is no more than 36dB higher than that of the wanted signal.

Wi-Fi signals attenuate differently through different materials. Furniture and accessories built with metal can cause significant attenuation. The technician should avoid placing Wi-Fi devices directly behind computers, monitors, TVs or other metal obstructions.

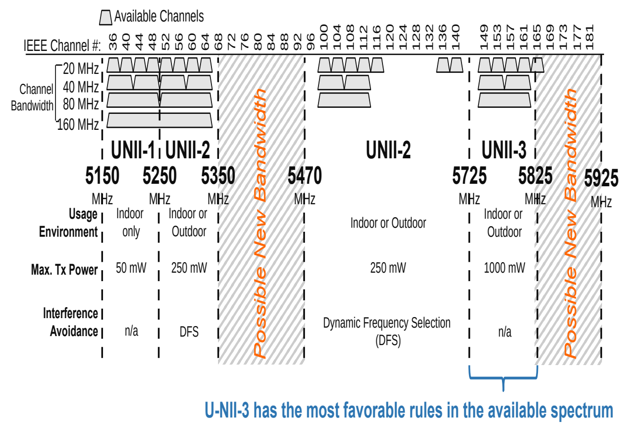
Several vendors also advocated that Wi-Fi clients used for video streaming should also include LED lights to indicate the health of the wireless link. For example, a green light could indicate excellent signal strength for video streaming, yellow could indicate border line signal strength for video streaming and red could indicate poor signal strength to support any video streaming. Further work is needed to define exact mapping between signal strength (e.g., RSSI) and color of the light. The LED could be a simple method for consumers to understand the signal strength and help with customer care.

Wi-Fi Band and Channel: Although the 2.4 GHz band promises better coverage, for video over Wi-Fi, use of the 5 GHz band and a 40 MHz channel is recommended. The 2.4 GHz band is typically too crowded to utilize 40 MHz channels. There are more 40 MHz channels available in the 5 GHz band than in the 2.4 GHz band. The wider channel bandwidth of 40 MHz provides greater capacity to deliver video services. Additionally, there is potential availability of more channels in 5 GHz band. NCTA, CableLabs, and other operators and suppliers are advocating for more unlicensed spectrum and improved rules for Wi-Fi in the 5 GHz band.

CableLabs recently submitted a paper, “Toward Expanded Wi-Fi Access in the 5 GHz Band” to FCC. This paper analyzes the interference risk to certain incumbent services posed by expanding Wi-Fi access, with a particular focus on two sub-bands known as UNII-1 and UNII-4 [4].

As shown in the figure below, 5 GHz band supports multiple sub-bands – UNII-1, UNII 2, UNII 2E and UNII-3 - for Wi-Fi use. Each of these sub-bands enforces unique requirements on devices.

- UNII-3: Devices using UNII-3 band are allowed to transmit using maximum EIRP of 36 dBm (IR=30 dBm, antenna gain=6 dBi)
- UNII-2 and UNII-2E: Devices using UNII-2 and UNII 2E band are allowed to transmit using maximum EIRP of 30 dBm (IR=24 dBm, antenna gain=6 dBi). Additionally, devices working in UNII-2 and UNII-2E band are expected to vacate this band upon detecting any radar activity. After vacating these bands as a result of radar detection, devices are allowed to come back to UNII-2 and UNII-2E bands only after scanning these bands for no radar activity for at least a minute.
- UNII-1: Devices using UNII-1 band are allowed to transmit using maximum EIRP of 23 dBm (IR=17 dBm, antenna gain=6dBi).



For video over Wi-Fi, UNII-2 and UNII-2E are excellent options for medium size homes (up to 4000 sq. ft.) built using drywall panel walls in locations with no radar activity. Use of UNII-2 and UNII-2E is suggested since it offers interference protection from retail Wi-Fi APs working in the 5 GHz band. The retail Wi-Fi APs are not expected to support UNII-2 and UNII-2E bands due to the expense involved in obtaining DFS and FCC certification . UNII-1 should be a good option for use in buildings with small apartments and studios. Use of UNII-3 band is suggested for large (4000 sq. ft. or larger) homes with decent gap (10 to 20 feet to avoid interference from adjacent and alternate channels respectively) between houses and also in places with regular radar activity.

As 802.11ac products start to penetrate the residential market in 2014, the use of the 5 GHz band for Wi-Fi is going to increase, resulting in worsening interference. Consequently, the increased use of 5 GHz band may result in improved interference in 2.4 GHz band. Because of the changing dynamics and need for reliable service, MSOs should consider deployment of Dual Band Dual Concurrent (DBDC) radio APs for residential deployments. The video over Wi-Fi clients (e.g., IP STB) should also support both bands.

Wireless Link Margin: Indoor wireless communication is a very dynamic and multipath rich environment. Environmental changes in such as opening and closing of doors, people and pets moving around, change in AP or client location, and interference are some of the factors that can impact wireless network performance. When designing an in-home Wi-Fi network for video streaming, MSOs should consider a wireless link margin to protect against the uncontrollable changes in the environment.

It is recommended to use a minimum of 5dB margin to factor for these environmental changes. This margin should be included regardless of type and size of home. For example, if a Wi-Fi solution in a lab environment is capable of streaming three MPEG-2 streams to clients with an RSSI of -90dBm, then for real deployments all video clients should be placed in locations supporting -85dBm or better RSSI. Additional margin for various factors discussed in the following below will help improve the robustness of the in-home Wi-Fi network for video streaming.

- Wireless link margin to accommodate for changes in antenna orientation on AP and client. In the worst case tester observed a throughput variation of as much as 30% with changes in AP antenna orientation. This variation is different for different products.
- Video over Wi-Fi networks using 5 GHz spectrum should be designed to avoid legacy devices (802.11a) on the same channel as used by video over Wi-Fi devices. If legacy device must use the same band as video over Wi-Fi devices, a healthy wireless link margin should be included in the design. Further study is required to find the exact number as the margin will vary on factors such as number of active legacy devices, and whether the network supports airtime fairness and WMM.
- Wireless link margin to accommodate for background traffic if the same band and channel are used for both video and data services.
- Implementation loss (e.g., different product casing and design may cause different amount of attenuation). Several vendors indicated that, in their experience, implementation loss can be as much as 5dB.
- Interference (e.g., different source, amount and type of interference will require different amount of wireless link margin).

Airtime Fairness and WMM QoS: Wireless Multimedia (WMM) offers four priority levels by using different minimum and maximum back-off slots, allowing some applications better transmit opportunity than the others. Airtime fairness prevents slow clients from slowing down the fast clients by “fairly” allocating the airtime to clients.

Support for WMM and airtime fairness is recommended on AP and IP STBs, if the same band and channel are used for both video as well as other services (e.g., data,). Additionally, the management of Wi-Fi resources allocated to non-video clients, including slower (as a result of poor link quality or low MCS) and legacy clients is critical.

While WMM is standardized in IEEE 802.11e, airtime fairness is vendor proprietary. For consistent customer experience, the cable industry should consider defining requirements for airtime fairness algorithm and contributing to wireless standards bodies.

For near term video over Wi-Fi deployments, operators should use the default WMM parameter set as defined in IEEE 802.11e. Additionally, the video, data and voice traffic should be tagged as suggested below to assign video higher priority than data.

- Voice- 0xE0 (Decimal TOS precedence value= 6,7)
- Video- 0xB8 (Decimal TOS precedence value= 5,4)
- HTTP- 0x20 (Decimal TOS precedence value=1)

In the long run, the cable industry should consider research to identify an optimum WMM parameter set for in-home video over Wi-Fi use case.

Some vendors offer the capability to provide higher priority to video traffic without having operators to tag the traffic. Vendors claim to prioritize traffic by identifying the video using proprietary traffic signature methods. One example of this technology is “streamboost” from Qualcomm. This technology appears useful and very promising. As part of the Future Home Networks project, CableLabs plans to analyze and test some of these technologies and provide recommendations to the operators.

Interoperability: The basic interoperability exists – different vendor clients seamlessly attach to and send traffic through other vendor APs. However, the Wi-Fi network with same vendor AP and clients offers better performance than the Wi-Fi network with multi-vendor Wi-Fi products.
Interoperability between multi-vendor 802.11ac products is not as seamless as it is with 802.11n products. In at least one case, clients from one vendor were unsuccessful when associating with a different vendor AP. We expect this to improve with WFA certification maturity.

For initial video over Wi-Fi deployments, operators who consider a multi vendor network should consider testing products for interoperability. The cable industry should work with vendors to improve performance in a multi-vendor environment. The cable industry should also consider sharing interoperability test results with industry organizations such as the Wireless Broadband Alliance and the Wi-Fi Alliance. For example, a joint test plan could be proposed based on MSO requirements and vendor recommendations to improve interoperability.

Interference Management: Wi-Fi in 5 GHz has comparatively fewer sources of interference than Wi-Fi in 2.4 GHz in today’s deployments. Interference is a significant factor; we expect it to increase in the 5 GHz bands, and it should be avoided and managed. Major issues are not only co-channel, but also adjacent channel and alternate channel interference. Video over Wi-Fi devices should be placed at least 20 ft. away (assuming open air environment) from the other nearby APs. Nearby APs should be configured to use the other available channels.

Some of the cordless phones available today use the same spectrum as Wi-Fi in 5 GHz. Interference from these phones on the Wi-Fi network can be significant. Operators should identify channels used by 5.8 GHz cordless phones in each house and avoid them for Wi-Fi use. To achieve this, MSO technicians should have access to survey tools that not only identify interference but also the source and type of interference. If a subscriber installs cordless phones after the initial Wi-Fi installation, Wi-Fi features such as ACS should be able to detect interference and move Wi-Fi operations to a better channel.

Sources for interference are many and dynamic in nature. To manage interference, operators should consider the following additional tools:

- Support for Automatic Channel Selection (ACS) at boot up and during operation
- Support for Dual Band Dual Concurrent (DBDC) with support for multiband steering
- Support for Radio Resource Management (RRM) and Self Organizing Networks (SON)
- Site survey and record the interference environment at installation time to help with troubleshooting the post install environmental changes that degrade performance.

Automatic Channel Selection: APs, supporting ACS constantly sense the presence and amount of interference around them. APs then use this information to select and use a channel with "better" operating conditions.
Since the channel conditions can change with time, it is recommended that the AP should be capable of performing Automatic Channel Selection at boot-up and during run-time. The channel selection must be done carefully with consideration to a number of factors, including:

- The transmit power of each band. Different sub-bands within 5 GHz band have different transmit power requirements. Simply moving from high power band, as a result of increased interference, to a low power band can introduce coverage challenges.
- APs are required to scan the UNII-2 and UNII-2E band for at least 60 seconds for radar activity before the APs can use these channels again. APs should be thoroughly tested to make sure they don't move out of UNII-2 and UNII-2E bands a result of false positive radar detection.
- APs should also support background scanning of these bands for the presence of radar without leaving the current channel of operation and potentially affecting services.

CableLabs performed basic ACS tests on three vendors and found that all the vendors support ACS at both boot-up and run-time. CableLabs recommends further testing of this feature to verify proper operation in more complex scenarios. For example, ACS operation in the presence of different amounts and types of interference in multiple channels (e.g., 5.8 GHz cordless phones) should be considered for future testing.

Dual Band Dual Concurrent (DBDC): Wi-Fi APs could be classified into four categories based on the frequency band in which they operate.

- 2.4 GHz only: Supports 2.4 GHz band only
- 5 GHz only: Supports 5 GHz band only
- Dual Band switchable (DB switchable): Supports both 2.4 and 5 GHz bands, but not concurrently
- Dual Band Dual Concurrent (DBDC): Supports both 2.4 and 5 GHz bands concurrently

DBDC would allow operators to use separate bands for video and data services. For example, use of 5 GHz for video and 2.4 GHz for data and voice. With DBDC, CableLabs also suggests support for multiband steering, which, for example, allows operators to load balance between bands, keep slower devices on non-video band and move video services and devices to a band that's less occupied. Several Wi-Fi AP vendors claim support for multiband steering capabilities using proprietary methods.

Multiband steering is an active work item in Wi-Fi Alliance (WFA), and CableLabs actively follows and contributes to the WFA multiband steering working group.

Radio Resource Management (RRM): In addition to ACS, DBDC and site survey, operators should also consider the use of RRM/SON for interference mitigation and performance improvement, especially in dense Wi-Fi deployments.

Wi-Fi networks may be large in scale, comprising of hundreds of thousands or millions of operator managed APs. Self-organizing methods are required for the efficient management of the Wi-Fi resources with large numbers of APs. Wi-Fi SON approaches can include techniques supported by each AP for immediate response to air interface conditions. Wi-Fi SON approaches can also include placing a centralized SON servers in the cloud or network that provides a high level management of specific parameters based upon a wider view of the Wi-Fi network that may be available to individual APs. The goal of the RRM/SON is to provide operators with a centralized Wi-Fi SON control based on a wide view of the Wi-Fi access network, which consists of Wireless Controllers as well as standalone APs from different vendors.

Spatial Division Multiplexing and Transmit Beamforming: Spatial Division Multiplexing (SDM) allows an AP to send multiple streams of data simultaneously to a client using multiple antennas. This results in better throughput at the client.

Transmit Beamforming (TxBF) allows an AP to concentrate its signal energy at the client location. The AP does this by sending the same signal from multiple antennas and carefully controlling the phase of the transmitted signal from each antenna. This results in better signal to noice ratio (SNR) and throughput at the client, and can reduce interference across the network.

Video over Wi-Fi application benefits from the support and use of both TxBF and SDM. Vendors' use proprietary algorithms to select between SDM and TxBF based on the channel conditions to each client. APs tested were capable of intelligent antenna resource allocation using some antennas (e.g., 2 out of 4) for TxBF while simultaneously using the remaining for SDM. Two transmitters cannot be used to perform both SDM and TxBF simultaneously.

TxBF and SDM improved wireless system performance when enabled. Increasing the number of spatial streams increases the Wi-Fi system throughput; however, the relative gain in throughput is less as the number of spatial streams goes up. For example, the gain in throughput as a result of going from 2 spatial streams to 3 spatial streams is less than the gain in throughput as a result of going from 1 spatial stream to 2 spatial streams (T2ss-T1ss > T3ss-T2ss > T4ss-T3ss).

TxBF is not standardized in 802.11n, resulting in a lack of TxBF interoperability. TxBF is standardized in 802.11ac, but wide-scale interoperability across multiple vendor products needs to be verified.

Key Performance Indicators and Network Management: Wireless is a very dynamic environment. In order to proactively manage the network, operators should be able to keep track of the following key performance indicators and take appropriate proactive actions to prevent impact on services.

- Number of currently associated clients and type (Current and historical)
- Current channel of operation
- Quality of wireless link to each client, including Downstream throughput (avg., peak), Upstream throughput (avg., peak), Packet Error Rate (PER), Type of client (e.g., iPhone, IP Set top), MAC address, RSSI,) the highest supported Wi-Fi version by the device, SNR, Number of spatial streams currently being used for transmission, Modulation and Coding

Scheme (MCS), Channel size for each client, Connection state

- Level and type of interference in different channels of a Wi-Fi bands (including current channel)
- Number of channel change events, for example as a result of Automatic Channel Selection
- Channel switch time
- Reason for client device disassociation (e.g., legacy device not supported, SNR below threshold)
- Error events of wireless link state for a client (e.g., IP set-top box) goes below operator defined values
- Maximum Transmit Power
- Channel Utilization (airtime percentage)
- Wi-Fi Carrier Sense threshold used
- Noise floor
- Number of radar detection instances
- Amount of time Wi-Fi service is affected when AP was trying to switch to DFS channels (UNII-2 and -2E)
- Number of packet re-transmissions
- Number of Forward Error Correction (FEC) events: Number of Un-errored FEC Code-words, Number of Correctable FEC Code-words, Number of Uncorrectable FEC Code-words.

Wi-Fi EVOLUTION AND FUTURE WORK

The convergence of voice, video and data, along with the evolution of HD and 3D video, is driving the need for increased throughput connectivity throughout the home, while assuring a high level of reliability and sustained performance. With the proliferation of a number of handheld devices per home, it becomes stringent for any access technology to deliver high speed throughput rates delivery in a reliable and contiguous manner.
Wi-Fi has a strong deployment base with a diverse and healthy vendor ecosystem. A lot of newer technologies have been stemming up in the past couple of years to help support the ever-increasing thirst for bandwidth. For example, High Efficiency Wi-Fi (HEW), Phase 2 features of 802.11ac, Cloud-based AP co-ordination functionality, utilization of 802.11ad for faster throughput rates (at a distance less than 20m), multiband steering, etc. A number of standard bodies (like IEEE, WBA and WFA) are actively working on standardizing these features. CableLabs is actively participating and contributing in the WFA, WBA and IEEE to move the Wi-Fi technology and interoperability forward.

Phase 2 features of 802.11ac

The following table shows the features of 802.11ac phase 2 and the corresponding benefits.

Features	Projected Benefit
Dynamic Bandwidth Management	• Interference Mitigation • Improved Spectrum Efficiency
Mu-MIMO	• Reduced Jitter • Improved System throughput • 1 4x4 STA≈ 4 1x1 STAs • Highly useful for video streaming
160Mhz/80+80Mhz	• Improved throughput
8x8 MIMO	• Improved throughput • Improved coverage (3db more)

Vendors are actively working on supporting these features. Wi-Fi devices with support for phase 2 features are expected to start certification later this year.

High Efficiency Wi-Fi (HEW)

With dense AP deployments becoming a norm nowadays, IEEE 802.11 HEW group aims to improve efficiency in the use of spectrum resources and achieve a very substantial increase in the real-world throughput. IEEE is working on enhancing the PHY and MAC to support real time applications by improving the power efficiency for battery powered devices.

Carrier Grade Wi-Fi

IEEE, WFA and WBA have active work items that address Carrier Wi-Fi, which is a recent industry movement to promote operator requirements for managed Wi-Fi networks. Carrier Wi-Fi scenarios target dense AP deployments with many device associations, Community Wi-Fi where public and private SSID's coexist on the same AP, Transparent Mobile data off load as well as Customer Experience, Device Requirements and Network Management on par with Licensed Cellular Technologies. CableLabs is working on defining proper device behavior which consists of Minimum Performance Requirements, Minimum Standards Compliance and Interoperability for the Wi-Fi APs and clients as a part of Carrier Grade Wi-Fi.

CONCLUSIONS

Our test results indicate that Wi-Fi can reliably transport HD video in the home; but Wi-Fi network performance is highly dependent on a number of variables, including construction materials, distance between the AP and client, level and type of interference (co-channel, adjacent and alternate), Multiple Input Multiple Output (MIMO) configuration, Transmit Beamforming, Spatial multiplexing, antenna orientation, RF spectrum, background traffic, QoS settings on the APs, and device capabilities.

Wi-Fi provides the convenience for consumers to have access to their subscribed video content without being restricted to predefined locations. On the other hand, Wi-Fi utilizes unlicensed spectrum that is not solely under the control of operators. Numerous devices with a variety of technologies like Bluetooth, cordless phones, microwave ovens, etc., may utilize the spectrum. Compared to shielded environments such as the HFC network, Wi-Fi may be subject to more frequent radio disturbances that may not always be completely mitigable by the operator. On the other hand, with video becoming an increasingly important application over WLANs, the Wi-Fi technology trend seems to be cognizant of this fact. Phase 2 features of 802.11ac, High Efficiency WLAN standardization in IEEE, Cloud-based AP co-ordination functionality, utilization of 802.11ad for faster throughput rates (at a distance less than 20m) and Multiband steering are example evolution areas in Wi-Fi which enhance the video streaming performance.

REFERENCES

[1]: NIST Construction Automation Program, Report No. 3, Electromagnetic Signal Attenuation in Construction Materials, available in: *http://fire.nist.gov/bfrlpubs/build97/PDF/b97123.pdf*

[2] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "*Image quality assessment: From error visibility to structural similarity*," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[3] Guojun Jin and Brian L. Tierney, "System Capability Effects on Algorithms for Network Bandwidth Measurement" Distributed Systems Department Lawrence Berkeley National Laboratory. Available here: http://conferences.sigcomm.org/imc/2003/papers/p314-jin11111111.pdf

[4] Rob Alderfer, CableLabs, Dirk Grunwald and Kenneth Baker, University of Colorado "Toward Expanded Wi-Fi Access in the 5 GHz Band". Available here: http://apps.fcc.gov/ecfs/document/view?id=7520933307

SYSTEMS FOR OFDM BASED DIGITAL SIGNAL ASSESSMENTS FOR CATV NETWORKS

Rob Thompson, David Urban, Rob Howald — Comcast Cable
Satheesh Angaiah, David Geeter — Electro Rent Corporation
Murthy Upmaka — Agilent

Abstract

CableLabs has opened a new chapter in CATV digital communications with its latest installment of DOCSIS® 3.1. New DOCSIS® 3.1 capable products are expected to start showing up in 2015 time frame. Meanwhile, cable operators are looking into ways to start kicking the tires of OFDM-based PHY now by employing readily available simulation tools to produce DOCSIS® 3.1 PHY signals and subject them to a variety of simulated channel conditions.

Cable operators may also be aware of how some of these powerful simulation tools can be combined with signal generation and analysis equipment to produce DOCSIS® 3.1 signals that may be used to get an early look at how CATV systems will react to the new DOCSIS® 3.1 PHY. Simulation systems may be created to allow for deeper-dive into supported parameter sets, driving toward identification of optimal modulation profile settings, such as choosing appropriate cyclic prefix (CP) for expected channel conditions, or evaluating windowing parameter impacts to adjacent SC-QAM performance.

The authors wish to provide an initial reference point associated with porting DOCSIS® 3.1 PHY simulations into RF signal stimuli for more detailed analysis of how this new PHY performs across laboratory-based CATV systems. The authors expect to include both test topologies and new data resulting from these enhanced simulation capabilities, enabling others to contribute and add to this exciting new chapter in CATV digital communications.

BACKGROUND

This paper provides an introduction to software and equipment capable of testing DOCSIS®3.1 physical layer (PHY) performance. DOCSIS 3.1 is an essential component of a comprehensive network evolution strategy, delivering capacity increases through improvements in bandwidth efficiency and the addition of spectrum, and enabling the network to deliver long term capacity growth and best-in class data speeds [3].

Readers may be relieved to know that even though DOCSIS®3.1 certified products may not be available until sometime in 2015, there is opportunity now to start evaluating the specification by generating Orthogonal Frequency Division Multiplexed (OFDM) waveforms not just with software simulations, but also with hardware.

Many Arbitrary Waveform Generator (ARB) products are currently available, some of which are capable of generating the 192 MHz OFDM waveforms required by DOCSIS®3.1 downstream. Likewise, there are a variety of devices capable of digitizing large amounts of spectrum so that Vector Signal Analysis (VSA) software may demodulate and quantify fidelity via metrics including Error Vector Magnitude (EVM) or Modulation Error Ratio (MER).

The good news is that these are relatively mature products already being leveraged by the wireless industry for development activities associated with other standards bodies including Long Term Evolution (LTE) and IEEE 802.11ac (Wi-Fi®). Of course, mileage may vary given the many features included in the DOCSIS®3.1 specifications that may not get 100% coverage with custom OFDM waveform generation tools, like the unique upstream pilot patterns. Therefore, now is a good time to evaluate exactly what can be leveraged and what additional specification coverage can be made available going forward.

This paper provides insights of how adequate specification coverage is available to get up and running with some very basic PHY parameters including the following:

1. 4K, 8K IDFT size
2. Continuous Pilots
3. Up to 192 MHz Bandwidth
4. Downstream Cyclic Prefix
5. Downstream Roll-Off Prefix
6. Up to 4096-QAM Modulation

The authors would like to show how the above DOCSIS®3.1 test capability enables some preliminary evaluation of the DOCSIS®3.1 specifications to not only prove that this capability exists now, but hopefully inspire others to join the cause and leverage these benefits to help ensure DOCSIS®3.1 rollouts become the most successful in DOCSIS history.

The most obvious benefit is in verifying DOCSIS®3.1 requirements, not just when certifying DOCSIS®3.1 capable products, but in evaluating the requirements against channel model assumptions [4,5]. This is valuable work that could be done now to identify holes in the current requirement definitions so that issues may be resolved prior to a massive deployment. Many of the channel conditions can be simulated easily within a laboratory environment or field trials. DOCSIS®3.1 PHY can be exercised against the channel model conditions to ensure the PHY behaves as expected.

A less obvious benefit would be in identifying optimum PHY profiles. What is meant by optimum may vary among MSOs. It is the opinion of this paper's authors that the optimum profile is one that reliably provides the greatest capacity for the majority of users. To elaborate further, similar modulation profiles have already been created in DOCSIS®3.0 to reliably deliver optimum capacity of ATDMA channels in the upstream. There are many features available in DOCSIS®3.0 including pre-equalization, interleaving, Reed-Solomon (RS) coding, and preamble length. All these parameters can be tweaked against common channel conditions to reliably optimize capacity. DOCSIS®3.1 is a little different in that there are many more parameters available for tweaking and the opportunity is now to get a sense of which parameter profiles apply for the most relevant set of channel conditions.

Another less obvious benefit is in the identification of any potential operational challenges, ideally prior to mass deployment.

Fortunately, the authors will demonstrate how much of this due-diligence may commence now. This paper will highlight some prospective areas for investigation, like interoperation with legacy services. Specifically, how the various windowing parameters will impact adjacent DOCSIS®3.0 signals is something the authors will cover later in this paper.

The authors hope that this material will stimulate much-needed discussion on this topic and that many operational investigations and results will be brought forth so that our industry may support a quick convergence upon a sound DOCSIS®3.1 readiness strategy, an initiative already underway within the SCTE.

OFDM MEASUREMENTS

Windowing Power Spectral Density (PSD)

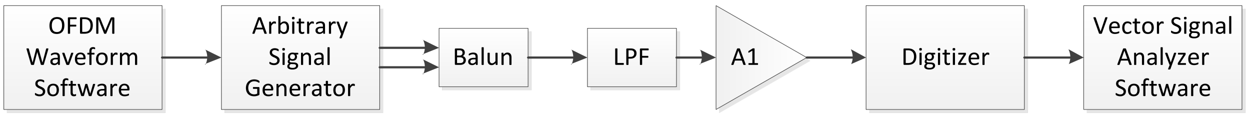


Figure 1 - Windowing Parameter Evaluation Test Topology

To evaluate the effects of windowing, Figure 1 test topology was used. OFDM waveform software is used to create the in-phase, (I) and quadrature (Q) data. This waveform is downloaded into the ARB for playback in the frequency domain. The digitizer bandpass samples the frequency domain content to recover the waveform and the results are aggregated into the vector signal analyzer software for a variety of measurements including MER, constellation plots, RF levels, etc.

There are five different Roll-Off Prefix (RP) settings available via DOCSIS®3.1 [1], which has been provided here in Figure 2 for reference. The test topology of Figure 1 was used to generate and analyze OFDM waveforms for the five different windowing settings from Figure 2. A 192 MHz, 4096-QAM modulated waveform was configured with a Cyclic Prefix (CP) setting of 2.5 µs. Both 4K, and 8K FFT waveforms were generated in this exercise. The maximum number of continuous pilots, M = 120, was used for all waveforms. M may vary between 48 and 120, which represents the number of continuous pilots that will occur at the same frequency location for all OFDM symbols. Their role is to assist in receiver synchronization.

Table 7–35 - Roll-Off Prefix (RP) Values

Roll-Off Period (µs)	Roll-Off Period Samples (N_{cp})
0	0
0.3125	64
0.625	128
0.9375	192
1.25	256

Figure 2 – DOCSIS®3.1 Downstream Windowing Specification

Amplitude traces were captured for all 10 waveforms and plotted onto a single chart for both the 4K, and 8K FFTs, illustrated in Figure 3 and Figure 4 respectively. The effects of windowing may be observed in these traces where the larger windowing settings provide a sharper edge in the frequency domain, resulting in more useful subcarriers. However, these settings come at the expense of reduced efficiency in the time domain. Finding the optimum windowing parameters will essentially be a trade-off between acceptable adjacent channel performance and DOCSIS® 3.1 capacity.

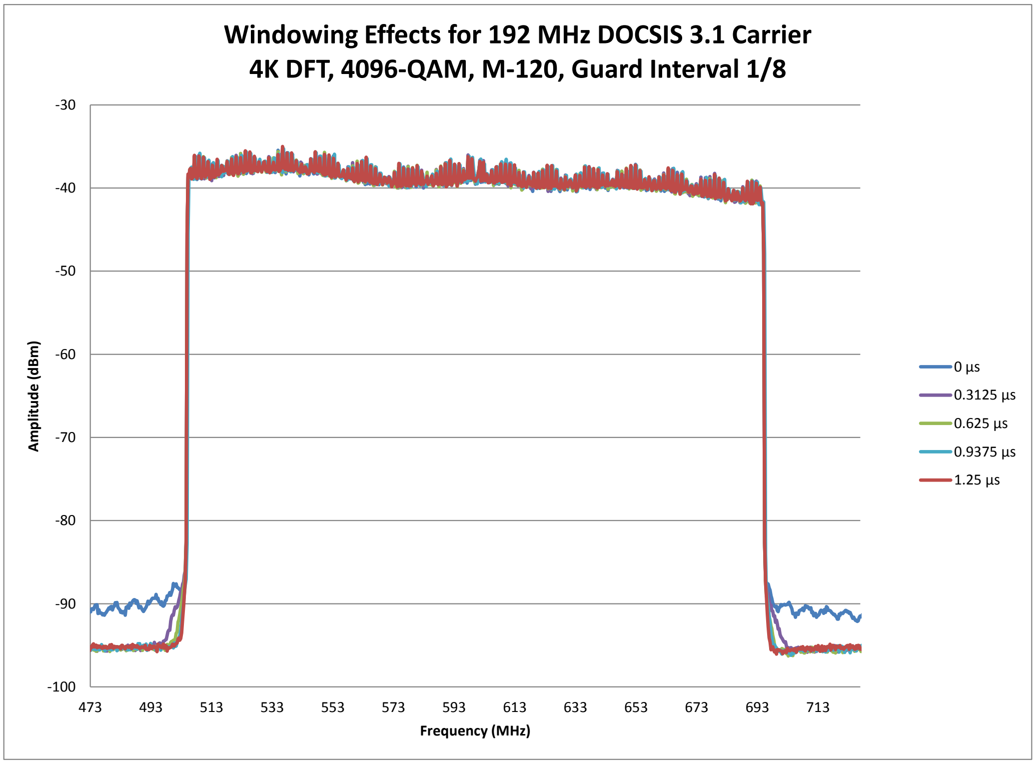


Figure 3 - Amplitude Traces, Using 4K FFT, for DOCSIS®3.1 Windowing

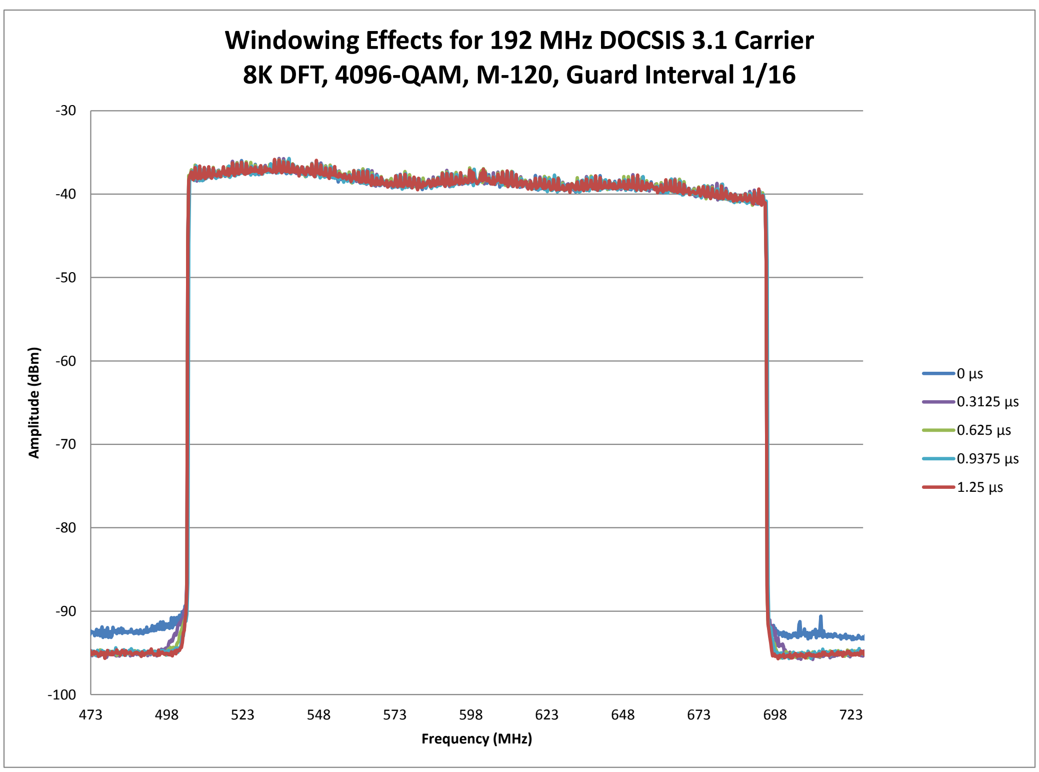


Figure 4 - Amplitude Traces, Using 8K FFT, for DOCSIS®3.1 Windowing

Adjacent Channel Performance

To assess adjacent channel performance, power measurements were made in both the DOCSIS®3.1 band, as well as six adjacent DOCSIS®3.0 bands for either side of the DOCSIS®3.1 signal. Many spectrum analyzers facilitate adjacent channel power (ACP) measurements similar to what has been illustrated in Figure 5.

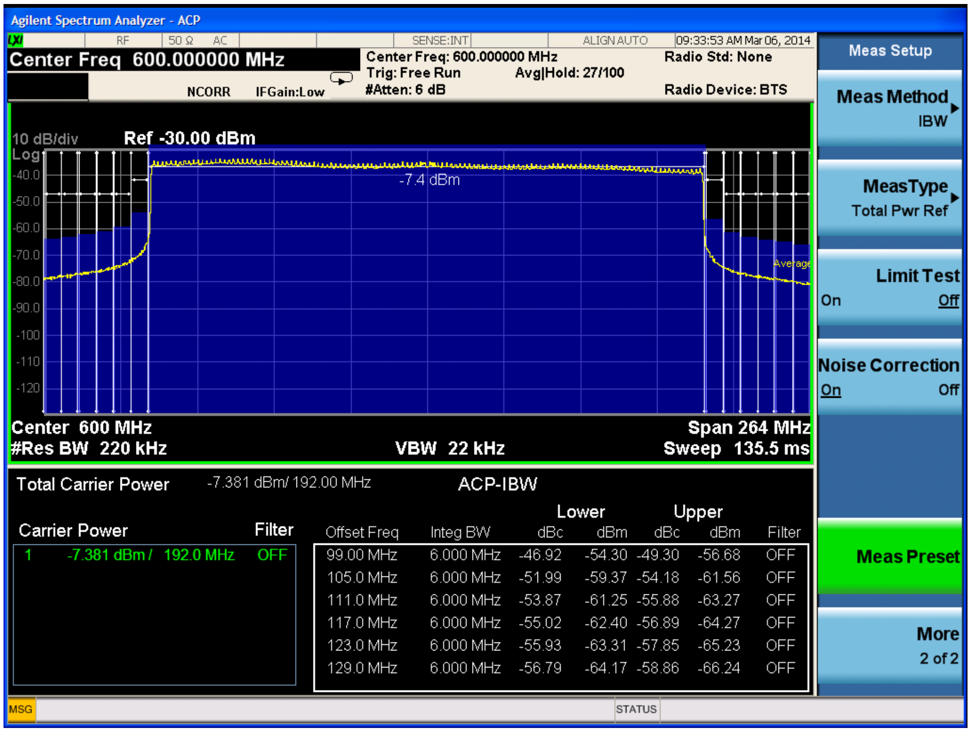


Figure 5 - ACP for 4K FFT, M = 120, CP = 2.5 µs, RP = 0 µs

Figure 6 summarizes the ACP for the 12 adjacent DOCSIS®3.0 carriers using 6 MHz bandwidth. The difference in power between 192 MHz DOCSIS®3.1 and a 6 MHz DOCSIS®3.0 bandwidths is $10*\log_{10}(192/6) =$ 15 dB. It can be seen how the ACP has the greatest impact to the next adjacent 6 MHz slot. There is also appreciably higher adjacent power when no windowing, RP = 0 µs, is used.

Windowing Effects via Relative Adjacent Channel Power - 6 MHz DOCSIS® 3.0 (dBc) Relative (dBmV) to 192 MHz DOCSIS® 3.1 Band Centered at F_c 4K FFT, 4096-QAM, M = 120, CP = 2.5 µs												
RP (µs)	F_c-129 MHz	F_c-123 MHz	F_c-117 MHz	F_c-111 MHz	F_c-105 MHz	F_c-99 MHz	F_c	F_c+99 MHz	F_c+105 MHz	F_c+111 MHz	F_c+117 MHz	F_c+123 MHz	F_c+129 MHz
0	-56.79	-55.93	-55.02	-53.87	-51.99	-46.92	39.61	-49.30	-54.18	-55.88	-56.89	-57.85	-58.86
0.3125	-73.22	-73.18	-73.10	-73.09	-72.22	-51.11	39.73	-53.83	-73.35	-73.83	-73.64	-73.61	-73.71
0.625	-73.01	-73.01	-72.98	-72.96	-72.91	-55.36	39.66	-58.56	-73.75	-73.50	-73.39	-73.31	-73.46
0.9375	-73.16	-73.17	-73.15	-73.09	-73.11	-59.30	39.71	-62.23	-74.08	-73.84	-73.75	-73.70	-73.71
1.25	-73.06	-73.01	-72.99	-73.05	-72.98	-62.96	39.69	-65.64	-73.82	-73.59	-73.49	-73.51	-73.49

Windowing Effects via Relative Adjacent Channel Power - 6 MHz DOCSIS® 3.0 (dBc) Relative (dBmV) to 192 MHz DOCSIS® 3.1 Band Centered at Fc 8K FFT, 4096-QAM, M = 120, CP = 2.5 µs													
RP (µs)	Fc-129 MHz	Fc-123 MHz	Fc-117 MHz	Fc-111 MHz	Fc-105 MHz	Fc-99 MHz	Fc	Fc+99 MHz	Fc+105 MHz	Fc+111 MHz	Fc+117 MHz	Fc+123 MHz	Fc+129 MHz
0	-58.24	-57.46	-56.42	-55.41	-53.37	-48.39	40.07	-51.80	-56.22	-57.88	-58.69	-59.73	-60.59
0.3125	-73.25	-73.25	-73.13	-73.07	-72.55	-53.68	40.18	-56.59	-73.63	-73.76	-73.58	-73.67	-73.68
0.625	-73.08	-72.97	-73.05	-72.98	-73.02	-58.54	40.20	-61.28	-73.82	-73.63	-73.54	-73.46	-73.48
0.9375	-72.78	-72.74	-72.69	-72.69	-72.69	-62.17	40.04	-64.78	-73.41	-73.17	-73.04	-73.02	-73.11
1.25	-73.17	-73.08	-72.94	-73.07	-73.01	-65.47	40.13	-68.23	-73.75	-73.53	-73.39	-73.43	-73.52

Figure 6 - ACP Summary Using Both 4K and 8K FFT DOCSIS®3.1

Legacy DOCSIS®3.0 channels were coupled into the test topology as shown in Figure 7. Minimum Loss Pads (MLPs) were used to combine legacy loading using 75 Ω impedance devices. A variable attenuator was used to adjust relative levels of legacy loading. An additional amplifier was used to compensate for the losses associated with the MLPs and passive devices.

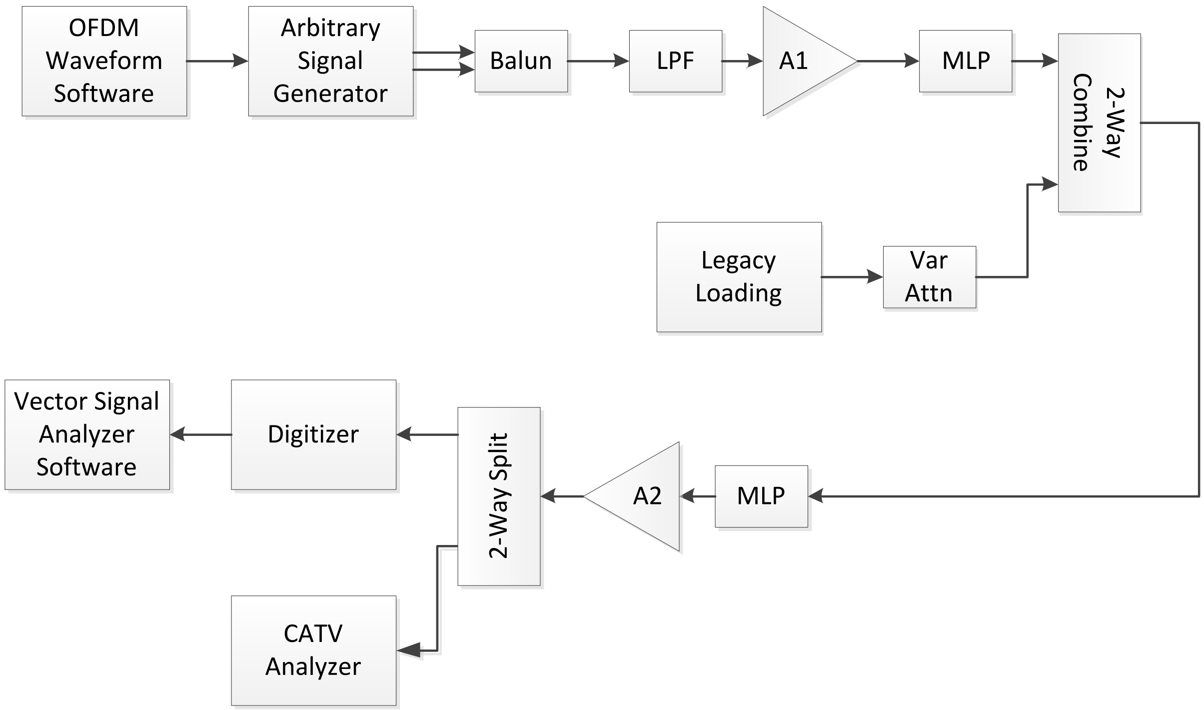


Figure 7 –DOCSIS®3.1 and DOCSIS®3.0 Coexistence Test Topology

Fidelity assessments, via Error Vector Magnitude (EVM) of the DOCSIS®3.1 signal, were made using Vector Signal Analysis software (VSA), shown in Figure 8, while a CATV Analyzer was used to measure Modulation Error Ratio (MER) impact on DOCSIS®3.0 signals using 256-QAM. The EVM of the DOCSIS®3.1 signals without the legacy loading was -53.9 dB for 4K FFT and -54.1 dB for 8K FFT.

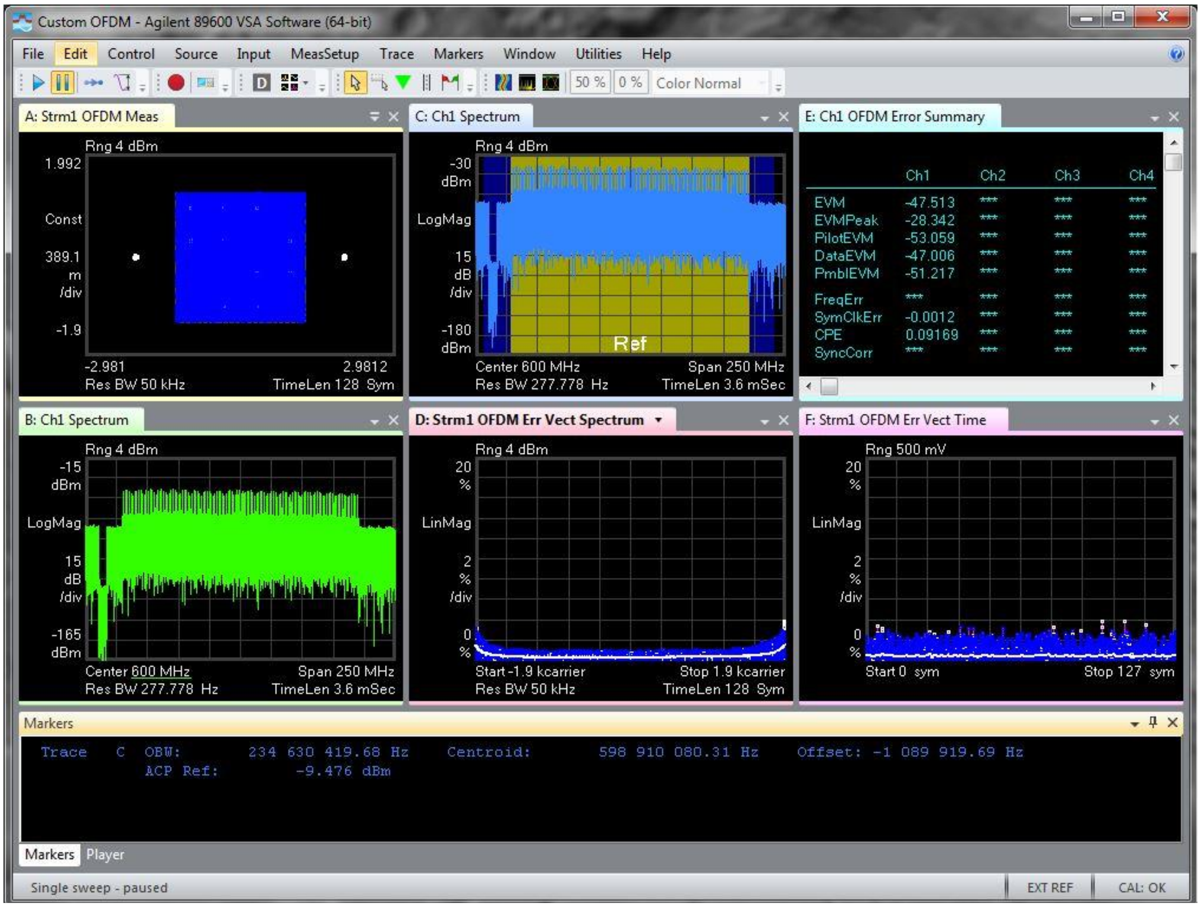


Figure 8 - VSA EVM Measurement of DOCSIS®3.1

The variable attenuator was used to adjust the relative levels of the DOCSIS®3.0 signals such that they were 0, -3, -6, and -9 dB with respect to the DOCSIS®3.1 signal. Previously, we had discussed that the DOCSIS®3.0 levels must be 15 dB lower than the DOCSIS®3.1 signal to ensure 0 dBc on a power-per-hertz basis. The resultant fidelity measurements are summarized for two different windowing settings, RP = 0.3125 and 1.25 µs in Figure 9. As expected, both DOCSIS®3.1 and 1st adjacent 3.0 signals are appreciably impacted, where the RP = 0.3125 µs has a more severe effect on adjacent DOCSIS®3.0 signals than the sharper edge produced by RP = 1.25 µs.

1st, 2nd Adjacent 6 MHz DOCSIS® 3.0 MER (dB) and 192 MHz DOCSIS® 3.1 EVM, Centered at F_c 4K FFT, 4096-QAM, M = 120, CP = 2.5 µs, RP = 0.3125 µs					
Legacy RF Level (dBc)	F_c-105 MHz	F_c-99 MHz	F_c	F_c+99 MHz	F_c+105 MHz
0	48.5	39.3	-42.0	40.8	48.0
-3	47.6	37.8	-43.4	39.0	47.3
-6	46.7	36.2	-45.5	37.4	46.4
-9	45.3	34.2	-47.5	35.3	45.0

1st, 2nd Adjacent 6 MHz DOCSIS® 3.0 MER (dB) and 192 MHz DOCSIS® 3.1 EVM, Centered at F_c 4K FFT, 4096-QAM, M = 120, CP = 2.5 µs, RP = 1.25 µs					
Legacy RF Level (dBc)	F_c-105 MHz	F_c-99 MHz	F_c	F_c+99 MHz	F_c+105 MHz
0	48.5	47.5	-41.6	47.7	48.0
-3	47.8	46.1	-43.4	46.7	47.5
-6	47.1	44.6	-45.3	45.3	46.6
-9	45.6	42.3	-47.2	43.4	45.3

1st, 2nd Adjacent 6 MHz DOCSIS® 3.0 MER (dB) and 192 MHz DOCSIS® 3.1 EVM, Centered at F_c 8K FFT, 4096-QAM, M = 120, CP = 2.5 µs, RP = 0.3125 µs					
Legacy RF Level (dBc)	F_c-105 MHz	F_c-99 MHz	F_c	F_c+99 MHz	F_c+105 MHz
0	48.4	41.4	-43.5	42.9	48.0
-3	47.8	40.0	-46.6	41.5	47.4
-6	46.1	37.2	-47.8	40.0	46.5
-9	45.0	36.1	-49.3	37.8	44.8

1st, 2nd Adjacent 6 MHz DOCSIS® 3.0 MER (dB) and 192 MHz DOCSIS® 3.1 EVM, Centered at F_c 8K FFT, 4096-QAM, M = 120, CP = 2.5 µs, RP = 1.25 µs					
Legacy RF Level (dBc)	F_c-105 MHz	F_c-99 MHz	F_c	F_c+99 MHz	F_c+105 MHz
0	47.7	46.9	-44.1	47.3	47.5
-3	46.8	45.3	-46.3	46.0	46.5
-6	45.6	43.5	-48.2	44.4	45.4
-9	43.4	41.1	-49.8	42.2	43.6

Figure 9 – DOCSIS®3.1 and DOCSIS®3.0 Coexistence Fidelity for 0, -3, -6, and -9 dBc (c = DOCSIS®3.1)

The majority of tests were performed with a CP = 2.5 µs, while all windowing settings were being evaluated. This may not make sense in a typical deployment scenario where a sharp windowing setting may negate the effectiveness of the CP to mitigate channel impairments, such as micro-reflections. A more typical configuration may leverage different CP settings for each windowing setting. In support of this anticipated flexibility, a single test reducing the CP = 1.25 µs for a given windowing setting, RP = 0.3125 µs had a negligible impact on adjacent channel performance, based on the 192 MHz, 4K FFT, 4096-QAM, and M=120 DOCSIS®3.1 signal. This makes sense given that CP does not play a role in shaping the DOCSIS®3.1 signal. Perhaps a more appropriate test may be to evaluate CP effectiveness under varying windowing settings while introducing a constant channel impairment in order to better understand optimal combinations of CP and RP relative to specific channel conditions.

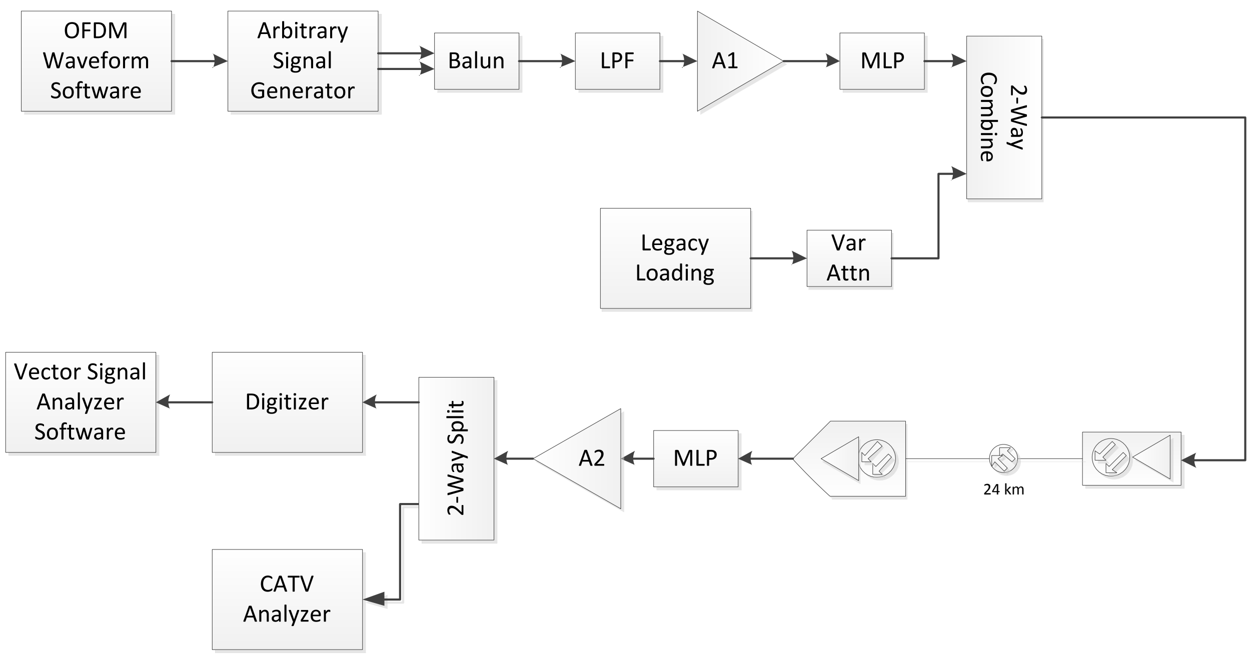


Figure 10 – Optical Link Test Topology

Optical Link Performance

A 1310 nm optical transmitter and node were added to the test topology in Figure 10, with 24 km of fiber in between. The node launch amplifier supported 1 GHz passband, but was only loaded to 700 MHz. The optical link performance was aligned for 696 MHz loading, which included 6 MHz DOCSIS®3.0 signals between 108 MHz and 504 MHz and a 192 MHz DOCSIS®3.1 signal at 600 MHz. Initially, the DOCSIS®3.0 and DOCSIS®3.1 signals were equivalent on a per-hertz basis. The DOCSIS®3.1 signal level was adjusted while fidelity measurements were made on all signals.

Adjacent 6 MHz DOCSIS® 3.0 MER (dB) and 192 MHz DOCSIS® 3.1 EVM 4K FFT, 4096-QAM, M = 120, CP = 2.5 µs, RP = 1.25 µs					
D3.1 RF Level (dBc)	F_c-117 MHz	F_c-111 MHz	F_c-105 MHz	F_c-99 MHz	F_c
0	41.7	41.9	42.0	42.0	-40.0
-3	45.3	45.3	45.3	45.5	-36.7
3	41.8	41.7	41.5	41.0	-41.5
6	35.0	34.6	34.3	33.8	-

Figure 11 - Optical Link Fidelity Summary

A performance summary of the optical link has been included in Figure 11. The 192 MHz DOCSIS®3.1 signal represents approximately 1/3 of the loading. The RF level adjustments reveal that the loading is already starting off too close to the peak performance of the optical link, where 3 dB increase in RF level does not yield a like increase in EVM, and 6 dB drives the optical link into compression. Furthermore, these higher fidelity measurements will require correction factors, specifically regarding the instrument contribution to the measurements provided. The average 4K FFT instrument performance was EVM = -53 dB, which means the instrument could contribute approximately 0.5 dB of measurement error associated with a measurement MER = 43 dB. The practical limitations of today's hardware is something to keep in mind when trying to validate DOCSIS®3.1 requirements.

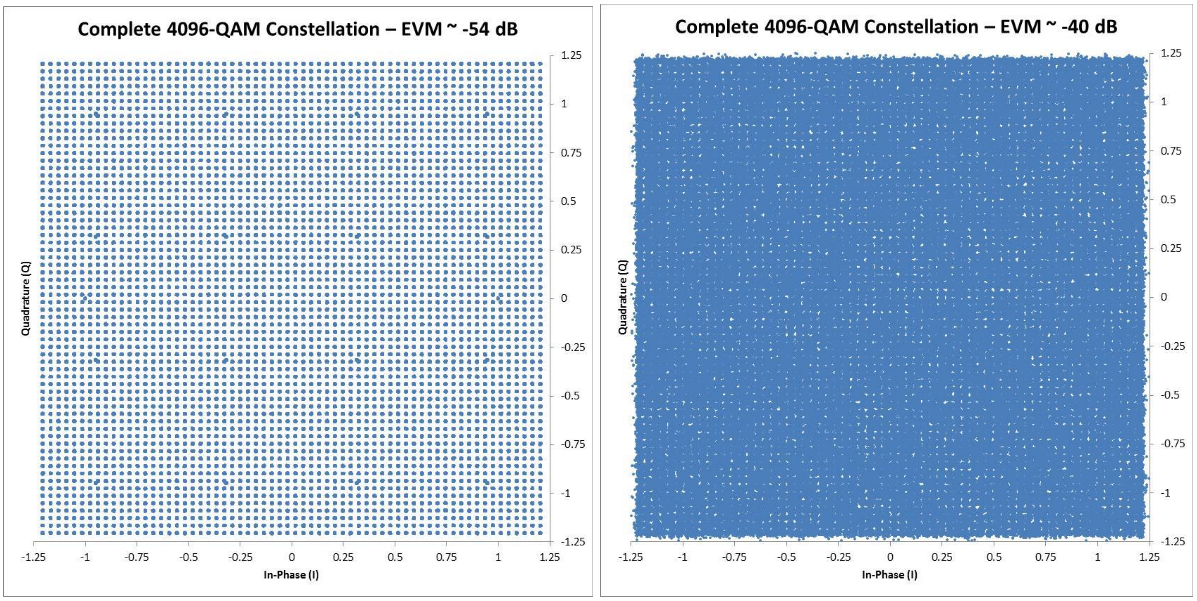


Figure 12 - Full 4096-QAM Constellation, Before and After Optical Link

The measured constellations, both before and after the optical link have been provided in Figure 12. The complete 4096-QAM constellation is difficult to view in its entirety, especially when performance is degraded. Analyzing constellations this large will require some modification of the display to facilitate easy impairment diagnosis traditionally used for lower order modulations.

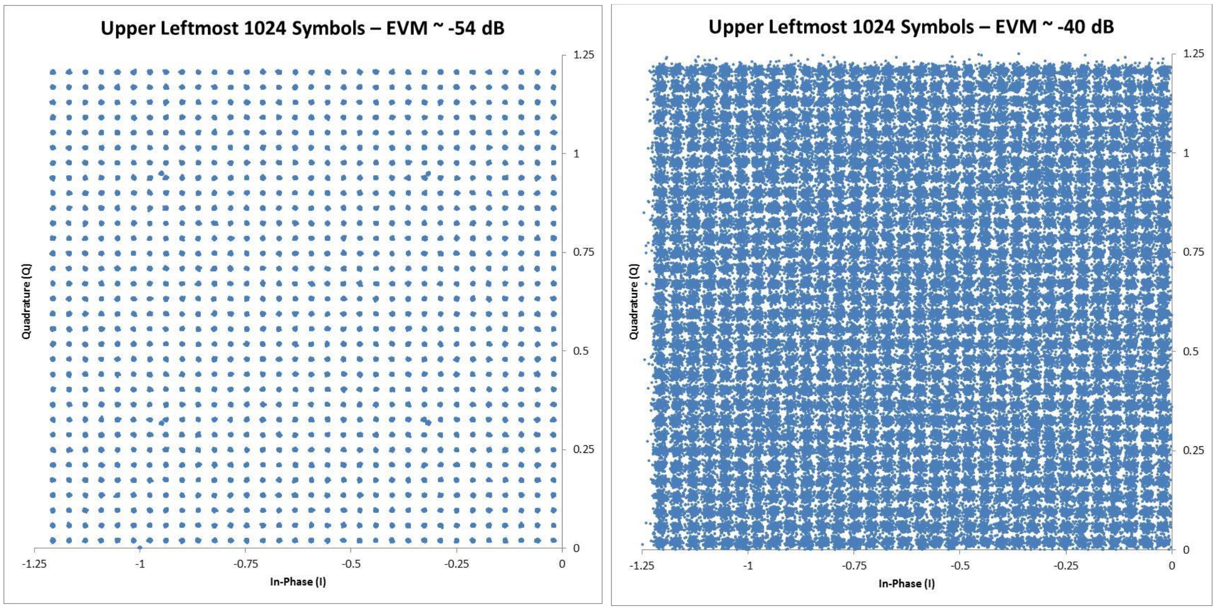


Figure 13 - Upper Leftmost 1024 Symbols of 4096-QAM Constellation, Before and After Optical Link

Zooming into the upper leftmost quadrant, shown in Figure 13, does provide a slightly improved perspective on performance, but still may be challenging for detecting unique impairment effects like phase noise.

Perhaps a better solution would be to color a group of symbols, such as 16 or 64 symbols, with a color associated with a specific value of MER, then repeating the process for the entire constellation assigning a color for each sub-group of symbols. The outcome would be a constellation heat map of sorts, where a single glance would easily distinguish the uniform effect of noise versus the non-uniform effect of phase noise or compression, where the outermost constellation points are impacted more severely than the any of the other symbols.

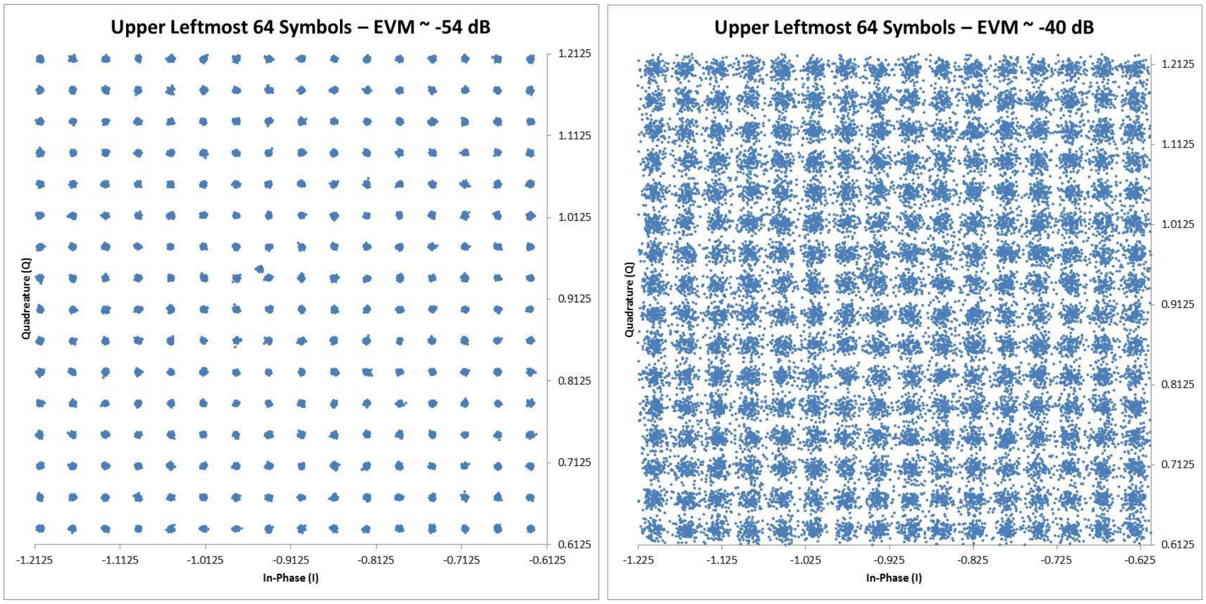


Figure 14 - Upper Leftmost 64 Symbols for 4096-QAM Constellation, Before and After Optical Link

Figure 14 shows only the upper leftmost 64 symbols. At this resolution the effects of the optical link are observed on the right constellation as uniformly impacting all symbols. As a check, the innermost 64 QAM symbols were also included in Figure 15 to verify similar conditions at the center of the constellation. At approximately EVM = -40 dB, uncoded BER = 1.12E-3 where appreciable LDPC error correction would be consumed.

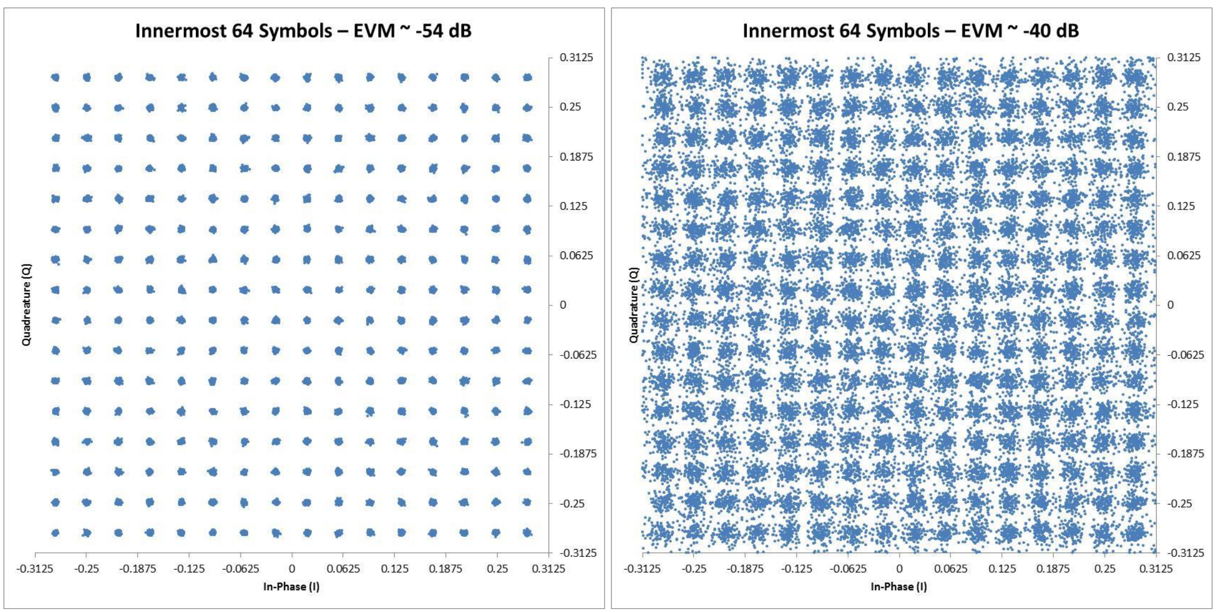


Figure 15 - Innermost 64 Symbols for 4096-QAM Constellation, Before and After Optical Link

Still, EVM = -40 dB does have promise for supporting 4096-QAM using LDPC coding. The DOCSIS 3.1 requirement for the cable modem is that it must meet low packet error criteria for a CNR of 41 dB as long as the input level is -6 dBmV. If we were to apply similar rules used for supporting nearly error-free 256-QAM via [2], where SNR = 34 dB, for BER = 1E-8, then 4096-QAM would require SNR = 46 dB to achieve similar fidelity. If it is possible to achieve the full 10 dB of coding gain predicted in AWGN with DOCSIS 3.1 LDPC, then the link may measure error free after decoding. However, 4096-QAM modulation may still pose challenges operationally; given there would be approximately 4 dB of margin available for the network performance to breathe, and the various non-AWGN impairments contribute more significantly as QAM order increases.

Table 1 - Experimental Setup Materials List

Device	Description	Vendor
M8190A	Arbitrary Waveform Generator	Agilent
M9703A	Digitizer	Agilent
89601B	Vector Signal Analyzer SW	Agilent
N9030A	PXA Spectrum Analyzer	Agilent
N6152A	Digital Cable TV X-series Application	Agilent
M9099	Waveform Creator Application SW	Agilent
5310A	Balun	Picosecond PulseLabs
ZX60-33LN	Amplifier (A1)	Minicircuits
ZX60-2514M	Amplifier (A2)	Minicircuits
SLP-1200	Low Pass Filter	Minicrcuits
GX2-LM1000B	Downstream Optical Transmitter	Motorola
SG4000 Node	Downstream Optical Receiver Node Launch Amplifier	Motorola
SMF-28	24 km Optical Fiber Spool	Corning
E6000	DOCSIS®3.0 Cable Modem Termination System (CMTS)	ARRIS

Test Equipment

Table 1 provides a list of the equipment used in the experimental setups described. An Agilent M8190A 14-bit, 8Gsa/s ARB was used to generate a 192 MHz BW DOCSIS3.1 signal in conjunction with the M9099 Waveform Creator Application Software. The IQ data is digitally up-converted in HW using the Digital Upconversion (DUC) mode, which gives the best signal quality in the desired frequency range.

An Agilent M9703A 12-bit digitizer was used to acquire the wide bandwidth signals with optimized dynamic range from the DUT at a sampling rate of 3.2GS/s.

Agilent's 89601B VSA SW was used to demodulate the DOCSIS 3.1 Signals. The Custom OFDM Modulation Analysis (Option BHF) mode provides the flexibility to perform time and frequency selective measurements over all subcarriers and symbols and report metrics such as EVM.

An Agilent N9030A PXA Spectrum Analyzer in conjunction with the Digital Cable TV X-series Application (N6152A) was used to measure MER of QAM carriers.

SUMMARY

The purpose of this paper was to provide an introduction to DOCSIS®3.1 test capability. This technology is available today and operators and solution providers could begin assessing a variety of DOCSIS®3.1 issues. Ideally this work would identify and resolve potential problems long before a massive deployment of DOCSIS®3.1 devices into customer homes.

Windowing settings will impact adjacent channel performance, in particular the 1st adjacent channel, and to a lesser extent the 2nd adjacent for non-zero windowing settings. How much impact will be dependent upon the window setting, and relative operating levels. Additional relief may come in the form of excluded subcarriers, lower order modulation on the data sub carriers located towards the edges (Mixed Modulation), or enhanced robustness of adjacent SC-QAMs. For example, only video signals with a longer interleaver depth may be used at the 1st adjacent carrier to a DOCSIS®3.1 signal.

Our testing suggests that an N+0 architecture can support 4096-QAM without any optimization of the signal profile or of the sample network, , when supported by LDPC coding. 4096-QAM, with its 12 bits per symbol, represents a significant gain in channel efficiency, it is 50% more efficient than 256-QAM. The question is, can the potential capacity be mined effectively by optimally leveraging DOCSIS®3.1 functionality. This needs to be explored more deeply so that operators can leverage the most optimal DOCSIS®3.1 PHY configurations possible for their specific network operating scenarios.

Our intent was merely to prove that the industry can be exploring optimal DOCSIS®3.1 deployment scenarios now. There are many tests, beyond what has been described in this paper that will help operators and solution providers in this regard. The following test case scenarios are a short-term to-do list for the authors to pursue after this paper has been written.

MoCA Coexistence

This testing examines issues with localizing DOCSIS®3.1 signaling adjacent to MoCA signaling within the home. Depending on the home network, MoCA signal levels may be running appreciably higher than DOCSIS®3.1. Identification of optimal center frequency placement of both would be provided.

CP Effectiveness

This testing examines the effectiveness of CP in the presence of varying linear distortion, as well capturing any impact of varying windowing settings. Identification of optimal CP, RP combinations would be provided.

Upstream Spectrum Allocation Implications

Examines any differences in PSD with equivalent legacy upstream DOCSIS®3.0 operating in 54-85 MHz. This energy may ingress into downstream receivers located within the home via poor isolation splitter and co-located DOCSIS®3.1 devices. If there is a higher expectation for ingress energy, then additional steps may need to be taken to protect legacy home devices, such as notch filters to block additional harmful RF. This work would identify any delta in ingress potential for a representative set of DOCSIS®3.1 configurations.

Mixed Modulation of Data Subcarriers

The testing explored in this paper exclusively uses uniform modulation on all data subcarriers. Mixed Modulation scenarios could be explored where data sub carriers situated near the adjacent SC-QAM carriers can be modulated at lower modulation rate to ensure error free performance.

Modulation Level Guidance

Optimal modulation level guidance for a diverse set of HFC optical and RF scenarios and channel conditions. Measured data support for theoretical expectations would be provided.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support of Neil Spence of Comcast. His assistance with many of the laboratory setups, configurations, and interconnections with legacy services and HFC equipment is very much appreciated.

REFERENCES

[1] Data-Over-Cable Service Interface Specifications, Physical Layer Specification v3.1, CM-SP-PHYv3.1-I01-131029, October 29, 2013, Cable Television Laboratories, Inc., http://www.cablelabs.com/specs/

[2] Howald, Robert, "The Grown-Up Potential of a Teenage PHY", NCTA Cable Show, Boston, MA, May 2013

[3] Howald, Robert, "Growth Architectures: Built to Last, Built to Launch," NCTA Cable Show, Los Angeles, CA, May 2014.

[4]http://www.ieee802.org/3/bn/public/mar13/howald_3bn_01_0313.pdf

[5]http://www.ieee802.org/3/bn/public/mar13/howald_3bn_02_0313.pdf

TAKING THE CUSTOMERS VIEW OF TROUBLE RESOLUTION "PEERING OVER THE WALLS"

Jeff Brooks, Deborah Horan, Gary Cunha, Mike Gordish
ARRIS Assurance Solutions
Charles H. Patti, Ph.D., Ronald Rizzuto, Ph.D.
Daniels College of Business, University of Denver

Abstract

The continued encroachment of competitive service offerings has driven subscriber satisfaction into the arena of completive differentiation. The necessity of service differentiation as a competitive advantage drives the necessity to apply Total Quality Management practices such as customer focus and process analysis and redesign, advanced statistical methods and measurement to the subscriber trouble resolution process. These tools have been widely used to generate local optimums at each stage of escalation of customer trouble reports. However, the practice of ishikawa's root cause analysis and solution analysis and industrial engineering practices of end-to-end process design has not been widely applied.

This paper will look at the theoretical improvements that could be achieved with a system that begins with system process documentation and process redesign to achieve outage preventive management, leading to customer experience improvements.

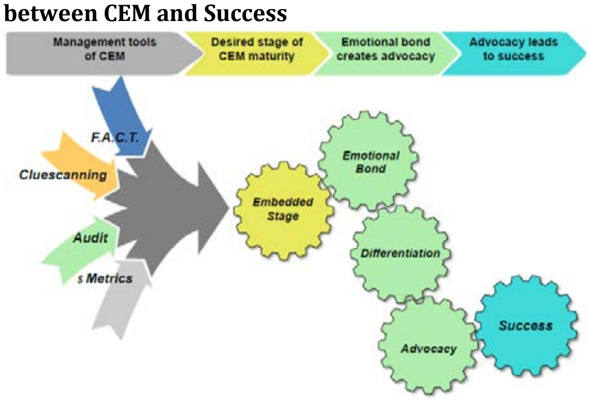
INTRODUCTION AND PURPOSE

Each year, there are over 10,000 new business books published in the United States alone. That's hundreds of millions of words giving advice on a range of business topics—from leadership, to survival tactics, to marketing. Every now and then, one of these business ideas rings so true that we change the way we think about—and conduct—business. Customer Experience Management (CEM) is one of those rare business ideas that is so compelling that it has pushed itself through the maze of business rhetoric and risen to the forefront of business thinking. When The Experience Economy was published in 1999 (Pine & Gilmore, 1999), it became a best seller, but managers didn't immediately embrace the ideas and change their ways of doing business. Powerful ideas often take time to penetrate the way we think and how we lead our companies. Today, fifteen years after the publication of The Experience Economy, CEM is still maturing into a mainstream business framework. But, much has been accomplished during this period. CEM constructs are emerging, ideas about the obstacles to full-scale CEM implementations are being discussed, and methods for gaining greater insight into how customers feel about themselves and the brands they experience are gaining acceptance. We now know that CEM is penetrating business leadership and practice. Companies from every sector of the economy now have CEM positions. Ideas about 'maturity levels' of CEM within organizations are being discussed. There is a lively debate about the components of CEM. And, CEM is being applied to all business sectors (B2B, not-for-profit, etc.) and across all stakeholder groups (See Firms of Endearment, Sosodia et al, 2007). This paper uses the CEM ideas and business tools we have developed over the past several years as the context for describing the implementation and measurement of a CE initiative in the cable industry. The ability of a business idea such as CEM to gain full acceptance is largely dependent on the ability of management to demonstrate the financial payoff of investments in the idea. The cooperation between Doctors Patti, Rizzuto and ARRIS Assurance Solutions has resulted in the isolation of the financial benefits of a CE investment.

Creating Value through CEM

As shown in Exhibit 1, our view of how CEM leads to profitability is a four-step process that starts with understanding how to use the CEM tools that we and others have developed over the past several years. Those tools include F.A.C.T. (the four components of CEM); Cluescanning™ (a method for understanding how well a company is delivering the customer's desired experience; the CEM Audit (a method for discovering the 'experience gap'—the difference between the desired experience and the delivered experience); and CEM Metrics (the tools used to measure the impact of CEM investments).

Exhibit 1: Framework for Understanding the Link between CEM and Success

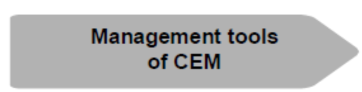


Source: Original material developed by the authors.

The second step in the process is to reach the most desired stage of CEM Maturity (Embedded Stage). We developed the concept of CEM Maturity by first acknowledging that full-scale adoption of CEM within an organization is a journey. Below we explain more about the CEM Maturity concept and how it affects the emotional bond between customer and company.

The third step embraces the idea that emotional bond between customer and company is the strongest point of brand differentiation. Further, meaningful differentiation is what leads to advocacy—the state in which customers willingly help recommend your brand.

The results of advocacy, increased sales, market share, profitability, and other financial measures, are the ideal measure of success (step 4). While most companies struggle to identify the financial payoff of CEM investments, we are strong proponents of such measures, and indeed the case history described in this paper illustrates how this connection can be made. In summary, we recommend the adoption of the framework in Exhibit 1 because: (1) it shows how to discuss CEM in terms of a process that leads to success, and (2) it embraces all of the main developments of CEM thinking.



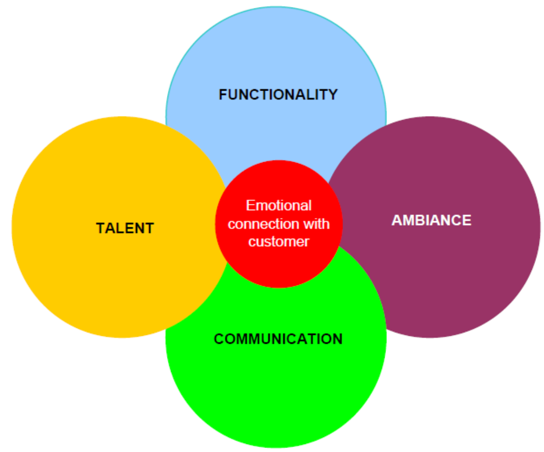
MANAGEMENT TOOLS OF CEM

1. F.A.C.T.—the components of CEM

Functionality. Long considered the single, most important reason that buyers work with specific sellers, Functionality has to do with a company's core product/service offering. This includes how effective the product/service is; how well it performs its functions; how much value it provides to the customer; the R&D capability of the seller; the financial terms offered; the quality and availability of service; and any number of other tangible attributes the seller provides. In the hotel industry, for example, hotels are typically evaluated on the basis of the comfort and cleanliness of their rooms, their reservation system, price, and other on-property amenities that surround the core product.

Ambiance. Although we most often associate the relevance of ambiance (look and feel) with consumer goods and services (e.g., restaurants, hotels, department stores, airlines, etc.), ambiance is also important in the B2B sector where it includes appearance of company representatives, appearance and convenience of office space, product packaging (where applicable), and any other aspect that conveys an appealing sensory feeling. For example, some manufacturing companies take great pride in their factories, including the Bang & Olufsen factory in Denmark, the Serta factory in Illinois, the BMW and Audi factories in Germany, and many others that understand the relevance of look and feel to the experience of customers, employees, and other stakeholders.

Exhibit 2: The F.A.C.T components of Customer Experience



Source: Original material developed by the authors.

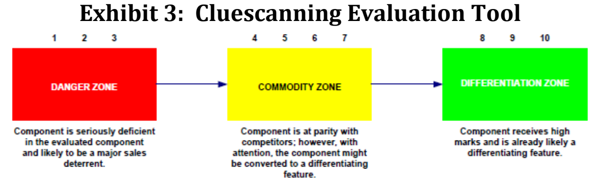
Communication. We live in world of message overload. Over 90 percent of U.S. adults own cell phones and 61 percent of U.S. adults own smartphones (Marketing Land). We talk, we text, we use social media, we are exposed to thousands of advertising messages each day, and we increasingly do business in non-face-to-face media, e.g., online, call centers, etc. A popular statistic about communication is that the majority of it (estimated between 70-90%) is non-verbal (Mehrabian). Further, it is estimated that within the next five years, 85% of a customer's relationship with a business will be done without face-to-face interaction (Gartner). At the same time, nearly 90% of customers quit doing business with a company because of a bad customer experience (Tierney), and increasingly that bad experience is through one or more forms of communication. Examples include, not answering phone calls, long wait times, incomplete or unsatisfactory responses, and inappropriate language. The Communication component of the F.A.C.T. definition of CEM is about the what, when, where, and how of messages between buyer and seller.

Talent. The rise of communication technology has fueled even more interest in the effects of people and culture on customer experience. Increasingly, business is in the hands of employees who interact with customers. Being knowledgeable, having the ability to listen to customers, anticipating customer problems and desires, and focusing on problem solving (first-contact resolution) are all characteristics of the Talent (people and culture) component of CEM. For example, employees at retailers such as Nordstrom's, REI, Container Store, and Whole Foods Market possess all of these characteristics. The other dimension of Talent is the organization's culture. Customer experiences—and then ultimately, the development of a positive emotional bond with customers—are enhanced when the organization is founded on the idea that sustainable profits are rooted in a 'customer-first' culture. If you listen to Gary Kelly, CEO of Southwest Airlines, he speaks of building an airline that is first devoted to giving customers what they want and to build trust and authenticity.

2. Cluescanning™

In his book, Clued In, Lou Carbone talks about the development of a way to measure how well a company is performing in the delivery of F.A.C.T. Carbone calls this process Cluescanning™, a method of self-evaluating the quality of delivering on the four F.A.C.T. components of customer experience. This method involves company representatives or others external to the company developing a keen sense of observation of the four dimensions. Then, the evaluators apply their observational skills to evaluate and score each dimension on its delivery of the dimensions. This diagnostic tool is particularly helpful when trying to assess current performance and then formulate action plans for improvement. Exhibit 3 is an adaption of Carbone's Cluescanning™ scale. This scale can be used by having the evaluator conduct a comprehensive, objective analysis of all four CEM components. Once that is completed for each component, the company can identify where improvements need to be made and which of the four components offers the best possibility to create a differential advantage in the marketplace. The idea is to try to move the evaluations of all four components to the right, toward the differentiation zone. However, it is rare for any company to achieve scores of 8-10 in all four components. Still knowing strengths and weakness and becoming committed to ongoing improvements puts an organization on the path to delivering the best possible customer experience.

Exhibit 3: Cluescanning Evaluation Tool



Source: Adapted from Clued in, by L. Carbone

3. The CEM Audit

The idea of an audit is not new, and it is applied in many fields. An audit—the practice of verifying information—is as old as the idea of counting and record keeping. Investor Words (www.investorwords.com) defines the concept of audit as, "An examination and verification of a company's financial and accounting records and supporting documents by a professional, such as a certified public accountant." In the U.K., the Medical Act of

1858 launched the auditing of medical practices (www.bristol-inquiry.org.uk). Business Publication Audits (BPA) has been verifying the circulation of business magazines and trade show exhibits for decades, thus providing its customers with assurances of the accuracy of data (www.bpaww.com). And, the idea of a marketing audit has been around for over fifty years (Kotler et al). The idea of an audit is simple enough—an objective process of verifying information. The value of an audit lies in its characteristics of thoroughness, standard procedures, objectivity, and reporting. When conducted well, an audit provides a report card for its sponsor, helping the organization understand its strength and weaknesses, thus indicating how to improve performance.

The CEM Audit is relatively new. It is a natural extension of the audit idea to an emerging concept, CEM. While we see no shortage of customer satisfaction surveys as well as more formal reports on brand performance, e.g., J.D. Power and Associates (www.jdpower.com), Consumer Reports (www.consumerreports.org), the idea of extending the audit idea to CEM is recent. This means that the formal procedures for conducting a CEM Audit are not firmly established. Instead, those who work in the CEM area are beginning to put forth their ideas of CEM audit components and the process for conducting the CEM audit. In Clued In, Lou Carbone identifies three components of a CEM Audit (identifying customer needs, deconstructing the experience from the customer's point of view, and determining the 'quality gap'). Smith and Wheeler developed a 20-item 'Branded Customer Experience Assessment' tool (Smith and Wheeler). In their 1999 book, The Experience Economy, Pine and Gilmore suggest a less formal auditing procedure that they call, "Shifting Up the Progression of Economic Value"(Pine & Gilmore) And, Bernd Schmitt provides a five-step framework for producing a winning CEM program (Schmitt).

The idea behind all of these approaches is the same: (1) identify the gap between what customers see as the most relevant experience and what brands are delivering, and (2) develop and deliver a plan to exceed customer expectations for an experience that truly separates the brand from competition. By doing this, a brand creates value that is manifested in customer advocacy. Originally articulated in a classic business strategy book by Ansoff in 1965 (Ansoff), gap analysis widely used today to identify differences between a current state and a desired state. For example, in one of our studies of nearly forty companies across a wide spectrum of industry sectors, we found that nearly 80% of the companies surveyed believe they are delivering good service while only 28% of their customers believe they received good service. In recent years, this gap has been repeatedly demonstrated to be extremely important to the financial well being of the companies. Here are just a few of the findings from various attempts to document the importance of customer experience and one of its key aspects, customer service:

Companies that prioritize the customer experience generate 60% higher profits than their competitors (Murphy) 81% of consumers are more likely to give a company repeated business after good service (Murphy) 60% of Americans businesses haven't increased their focus on providing good customer service, up from 55% in 2010 (American Express).

Concepts and Processes for Conducting the CEM Audit

There are many ways to get a close feel for the desired and delivered customer experience. Traditionally, surveys are conducted, asking customers a variety of questions about their levels of satisfaction on a variety of brand/service attributes. In recent years, this technique has been criticized for a number of reasons, including low response rates, inability to capture what customers truly want as an experience, the reliability of the data, and the inability to provide a deep understanding of the relationship between the customer experience and brand value. A number of new techniques have emerged, including the ZMET technique (Coulter & Zaltman; Zaltman & Coulter).

Outline for CEM Audit

Exhibit 4 illustrates the flow of the customer experience audit process, starting with an understanding of the desired customer experience (using focus groups, surveys, and ZMET); measuring the delivered experience (using Cluescanning™ as described above). From the information collected in steps 1 and 2, the organization can make decisions about what new or revised experiences should be created or revised. In step 4, the organization uses all normal channels of interacting with customers and prospects to deliver the new and/or revised experiences. Finally, in step 5, the organization measures the impact of the new and/or revised experiences through surveys, Net Promoter Score (Reichheld), and various financial measures. Creating a CEM program out of this sequence often leads to misallocation of human and financial resources. For example, starting at step 2 without first knowing what experiences customers and prospect desire is likely to be inefficient and ineffective.

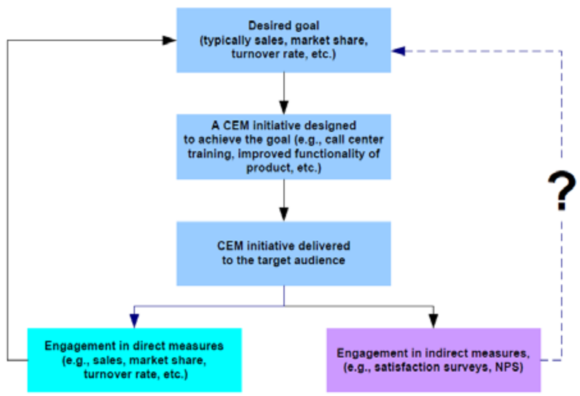
Exhibit 4: The CEM Audit Process

Source: Original material developed by the authors.

4. Measuring CEM's Impact

One of the objectives of this paper is to bring clarity to the measurement of CEM initiatives. Because CEM was first conceived as another name for customer service, early measurements were focused on various types of customer satisfaction surveys. This was followed by widespread adoption in several of industries of the Net Promoter Score (Reichheld), a measure of customer advocacy. Other measurements include customer effort (Dixon et al) and proprietary measures that combine satisfaction, effort, and advocacy. While all of these approaches can add value to understanding CEM's value to the organization, we advocate the use of measures that are directly tied to financial outcomes, e.g., sales, market share, customer turnover rate, customer lifetime value, etc. The reliance on indirect measures is the traditional approach because the challenges to measure financial results are significant. When there are many, and often uncontrollable, variables that affect the relationship between investment and desired outcome, there is a tendency to use indirect measure and assume that the indirect measure will affect the desired outcome. See Exhibit 5 below.

Exhibit 5: Indirect vs Direct Measures of CEM Initiatives



Source: Original material developed by the authors.

The above Exhibit illustrates the case for the use of direct measures. While it's often challenging to control all of the variables intervening between the desired goal and the measurement of impact, we believe that this challenge is no more difficult than trying to establish the relationship between indirect measures and the desired goal. In this paper, we illustrate how CEM initiatives can be directly related to financial outcomes.

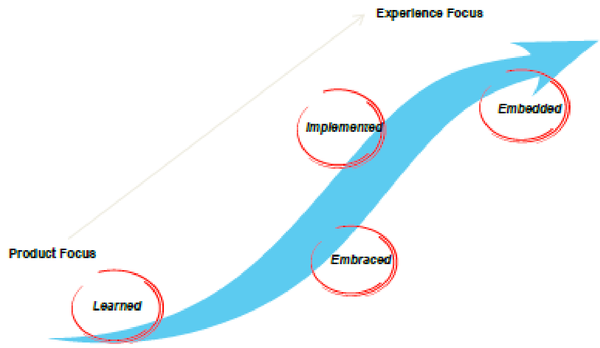
THE CEM MATURITY CURVE

In his 2011 article in Forbes, Steven Denning observed,

> *Changing an organization's culture is one of the most difficult leadership challenges. That's because an organization's culture comprises an interlocking set of goals, roles, processes, values, communications practices, attitudes and assumptions.*

Clearly, the full-scale adoption of a 'customer first'/customer experience corporate culture takes time, commitment, and energy by everyone in the organization. During the past several years, we've worked with over thirty companies on CE initiatives—observing their business models and organizational structure, evaluating their CE efforts, and helping them understand the obstacles to full-scale adoption of a CE culture. A major outcome of working with these companies is the development of the Customer Experience Maturity Curve (shown in Exhibit 6 below). The Curve is based on the idea that adoption of any culture change is a journey that has an identifiable pattern. In the case of CEM, this pattern consists of four distinct phases—Learned, Embraced, Implemented, and Embedded. Basic characteristics of each phase are summarized in Exhibit 7. As the organization moves toward the Embedded state increasing levels of knowledge about CEM are required, e.g., complete understanding of how to use Cluescanning™, the CEM Audit, and financial measures of impact. Also, the 'silo mentality' is replaced by a completely integrated approach to serving customers. The CEM Maturity Curve is useful in terms of understanding the current location of where the organization is on the curve and what steps are required to move up the curve.

Exhibit 6: CEM Maturity Curve



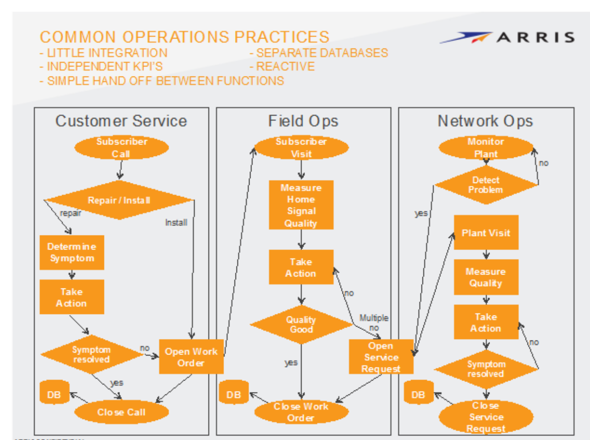
Source: Original material developed by authors.

Exhibit 7: Stage, characteristics, and orientation of CEM adoption

Stage	Charactertistics	Orientation
Learned	organization has acquired a minimum level of information about CEM and is familiar with the concept and some of the management processes associated with CEM. However, little has been done to implement CEM as an operating framework.	Product
Embraced	Organization has accepted the CEM idea as an operating framework; however, CEM is not yet implemented in the organization.	Product
Implemented	Organization is engaged in practicing selective aspects of CEM practices, but CEM is not fully integrated throughout the organization.	Customer
Embedded	Every member of the organization is committed to building emotional bonds with customers. Concepts and processes such as F.A.C.T, Cluescnning™ CEM Audit, and Metrics are part of the core fiber of the organization.	Customer

Source: Original material developed by the authors.

In a manual operational center, the silo affect among all departments is in full affect. The departments work within their specific silo and do not understand the actions within other departments that occur throughout a day and negatively impact the customer experience. Work orders incorrectly coded, forced scheduled times, missing timeframes, lack of customer facing time and manual field processes can all degrade overall operational efficiencies and produce a negative impact on subscriber experience. A comprehensive Field Service Management (FSM) tool integrated with Network and Service Management (NMS) tools helps build a cross-functional organization by providing valuable department-specific information in an automated fashion and helps mainstream the data flow within the FSM tool. By linking cross-functional departments and automating the delivery of the right information to each, organizations can drive efficiencies, improve customer experiences and reduce costs.



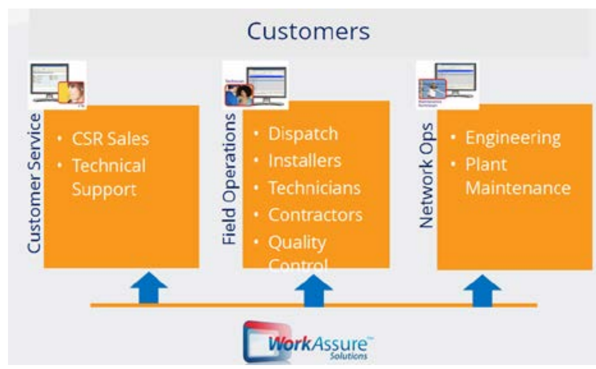
BEFORE

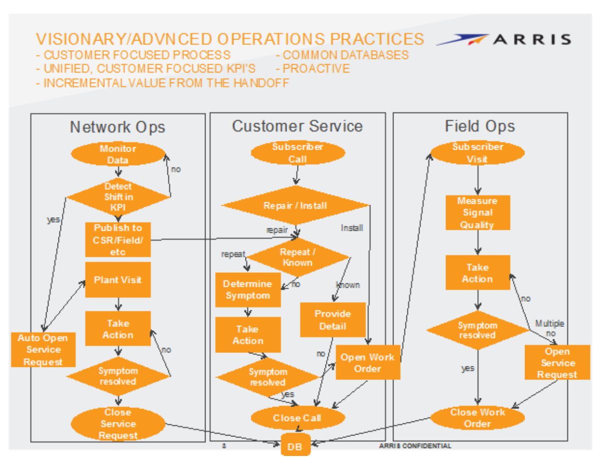
This integration between the FSM and NMS solutions impacts more than just technicians and dispatchers. It helps build a seamless cross-functional organization by incorporating an automated process enablement feature set that keys on actionable events that occur in the field. This process generates notifications across the organizations that are important to that specific group based on near real-time events. Often this integration does not exist and all steps involving other departments are completed in a manual process that invoked human error prone situations, which could impact the end subscriber or entire network of subscribers if work is not completed correctly.

For example – During the day, subscribers have been trained from previous experience to always call the customer service center asking the infamous question – “Where is my Tech?” A comprehensive modern FSM tool can provide the proper details to that end subscriber based on real actionable events through an IVR solution - without impacting the customer service department. This allows the customer service department to focus on generating new customer opportunities and retaining the current customer base, while reducing the wait times for subscribers who need to reach a representative.

These advanced FSM tools when integrated with the NMS and other Business and Operational Support Systems BSS/OSS do much more than deploy technicians in the field based on a scheduled work order pool. They also empower the technicians to complete the majority of their tasks without communication with the back office personnel, providing more face time with customers, and placing less strain on the back office resources. Additionally, during the completion of work orders, technicians are required to collect specific details related to the subscriber home network and access network telemetry. This information is valuable in proactively identifying possible network related problems that may require engineering or plant maintenance resources. Using the home and network telemetry information, the integrated FSM system invokes an automated process that creates a maintenance related order and automatically routes it to the next available plant technician to research and resolve the problem - before additional subscribers make a problem related phone call. This process improves subscriber satisfaction by preventing additional service issues, while allowing the Service Provider to identify network issues using their normal workforce during day-to-day activities, which helps keep operational costs low.

Building a well-organized cross-functional operational center is vital to providing a quality product and service to the end subscriber.





AFTER

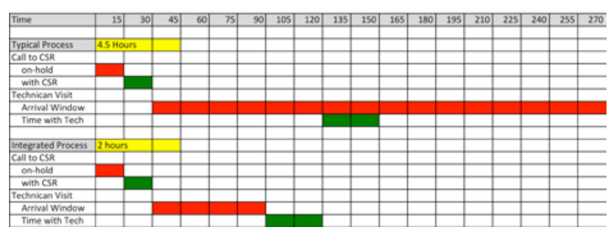
Emotional bond creates advocacy

CUSTOMER EMOTIONS LEAD TO ADVOCACY

The customer emotion—differentiation—advocacy sequence shown in Exhibit 1 is a powerful, but often neglected, approach to developing brand advocacy. Most organizations believe that customers become brand advocates because of superior product characteristics or outstanding service. These functional aspects of a brand are indeed important. But, the strongest driver of advocacy is the collection of positive emotions that customers feel about themselves when they purchase or use certain brands. We all want to feel proud, happy, content, smart, admired, pampered, important, and any number of other positive emotions. We want this in all aspects of our lives, including through the purchases we make. When we feel this way—and we can associate it with a brand—we become brand advocates. It's more important and relevant for customer to feel good about themselves than it is for them to feel good about the brand. Organizations increase the opportunities for brand advocacy—and all of its benefits (e.g., lower turnover rate, lower cost to acquire new customers, increased customer lifetime value, etc.) when they deliver a total and positive experience. This importance of emotions is underscored by the results of a recent study on customer experiences that concluded, "…Instead, people's feelings about a company often depend on the company's ability to gauge customer emotions, which account for more than half the typical customer experience." (The Huffington Post)

Considering technicians are often the face of the cable operation, it is very important for them to be empowered with a solution that allows them to be efficient and focus more attention on the end subscriber for all type of work being done in the field. Modern FSM systems must incorporate a dynamic routing algorithm that efficiently routes work orders throughout the day using the changes to the work that occur through out the day. The Service Providers are now able to provide more user friendly timeframes, reducing the four or two-hour windows down to one hour timeframes and even providing exact timeframe selections. The dynamic routing logic often has the added benefit of reducing drive times. The dynamic routing logic and reduction in none work related activities, allows the technician to spend more time with the customer to ensure that their issues are resolved successfully.

Customer Effort for a Single Trouble Call



Further the integration of FSM systems to the NMS tools will assure that the technician will install new services or resolve problems correctly the first time, immediately reducing subscriber frustrations associated with repeat trouble calls.. Field technicians are empowered with tools that help quickly identify any concerns related to the services the customer has purchased. For example, before a technician even arrives at a residence, a house check feature pulls key statistics about the installed components to proactively identify possible home or network related problems that may impact the installation or repair to be performed upon arrival. In addition, during the installation process of a complex service offering, a task checklist feature to ensure that the technician does not overlook a key installation step and conducts the steps in the proper sequence should be implemented. A history of all work orders should be kept and presented to the technician, allowing the technician to conduct a trend analysis resolution path based on both historical and real time polls, speeding the troubleshooting process and providing more time to answer customer questions.

Enabling the call center: hearts, minds, and resolution

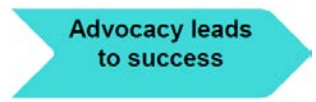
The ultimate goal of the NMS platform is to proactively detect service impairments and drive resolution before a subscriber ever experiences an issue. That said, there will always be the need to carefully manage customer calls to the call center since, as mentioned above, industry trends indicate that within the next five years, 85% of a customer's relationship with a business will be via remote interaction. (Gartner) The NMS system should be designed to help manage the subscriber call experience. If the a subscriber calls, there are several dynamics already in motion:

- Subscriber is already frustrated by the need to call
- Subscriber home network is increasingly complex and subscriber may feel ignorant or isolated in their lack of understanding
- Subscriber may have had a poor call experience with a different product/service
- The call agent is remote, making it difficult to project empathy and understanding to subscriber

Legacy support methods typically required the subscriber to try verbally representing the situation to the call agent, right down to which lights were blinking on the home device. The Today's NMS tools use a variety of telemetry protocols to provided full visibility into the subscriber home network and enable the customer/technical service agent (CSR/TSR) to service the entire subscriber account, not simply react to issues the subscriber identifies. Even before the subscriber calls, the NMS has been monitoring systems and measuring subscriber quality of experience across multiple services. While the call agent is first connected with the subscriber, the system can automatically inventory and check the home network without asking the subscriber to do anything. To support this, the NMS provides a normalized view of the home, despite what make, model, or technology is involved, whether subscriber devices are managed through SNMP, TR-069, or supporting Internet, voice, Wi-Fi, or video parameters. The CSR sees a similar view, colorized to reflect issues. Since all home devices are visible and initial status quickly checked, the CSR can begin by reviewing the installed devices with the subscriber. For example, the CSR may say "I see you have a cable modem with Wi-Fi, an MTA, and a Digital set top box, and all main devices are online and responding, including the game system and tablet you have connected to Wi-Fi. I see you've been experiencing Internet degradation since 7am today. Let me look more closely at your Internet service since I see something out of tolerance that may be causing the issue you're experiencing." Through this interaction, there is a level of credibility, knowledge, control, and assurance that this business can help.

Legacy support methods typically focused on a specific device and its telemetry in isolation. Historically, call agents were not empowered to resolve issues themselves, only lead the subscriber through a series of troubleshooting steps followed by escalation to a more technical group. For in-house issues, the NMS provides tools to enable the call agent to resolve many of the most common call drivers. For example, for "no dial tone" voice issues, the call agent may reset a line on the MTA. For Wi-Fi challenges, the call agent may reset a password, setup security, or set the Wi-Fi gateway to use a channel with less interference. Outside the home: The NMS platform provides a broader view of the entire network giving service to this subscriber, with the intelligence to quickly determine the next step. There will be scenarios where something bigger is going on and the final resolution will not take place on the initial subscriber call. When this occurs, it's important that the call agent be empowered to recognize this quickly, in a repeatable and systematic way, so they can focus on managing the customer's expectations. Therefore, service degradations caused by a neighborhood network outage or bandwidth capacity challenge are flagged to the call agent. In these cases, neither rolling a truck to the house nor using the in-house tools will help resolve the problem, so this visibility enables the agent to avoid deploying the wrong resource, and provides a platform for accurate communication with subscribers.

When a subscriber calls, whether the underlying cause is related to the malfunction or misconfiguration of in-house devices or a larger plant or capacity issue, the NMS must be designed to provide visibility across the network and services without segmentation by the current organizational structure with in the Service Provider in order to deliver a high-quality experience for subscribers and increase their satisfaction and loyalty.



ADVOCACY LEADS TO SUCCESS

As suggested above, there are significant financial benefits to investing in customer experience programs that lead to advocacy. These benefits include lowering the cost to acquire customers. When current companies advocate your brand, the investment in customer acquisition is lower. And, while the costs to acquire customers vary widely by industry, this cost can be over $300 (Safko). Advocates tend to stay with the brands they buy longer, thus extending their lifetime value. And, they tend to spend more because they are brand loyal and they enjoy the positive emotions they receive through engaging with their advocate brands. The strongest argument for the relevance of advocacy is the writing of F. Reichheld, the force behind the Net Promoter Score (NPS)—an advocacy measurement tool. In Reichheld's article, "The One Number You Need to Grow," Reichheld puts forth the case that measuring advocacy is the key to understanding a brand's financial growth potential.

CEM Case Study

Given the CEM context described above, we present a case study that provides: 1) an illustration of the application of some of the CEM concepts and tools discussed above to the cable telecommunications industry, 2) a discussion of some the difficulties encountered in measuring the impact of CEM investments, and 3) lessons learned for future case studies applications.

1. Background for Case Study

As a backdrop to discussing the case study, it is important to point out a key difference between the services provided by the cable telecommunications business and the products/services of other businesses. Specifically, cable services like video/high speed data/telephone are 'always on' or continuous use services. Hence, a key element of the 'functionality' of cable telecommunications is the ongoing delivery of the service in addition to the fulfillment at the time of purchase. This is in contrast to the 'functionality' of a discrete service like buying a cup of coffee. When consumers purchase coffee, they have no expectation as to the 'functionality' of the product beyond the consumption of the coffee and the associated 'buzz'/wake up effect afterwards.

As a consequence of the continuous use nature of cable services, cable telecommunications companies need to provide technical repair and service operations.

2. The Problem for Cable Operators

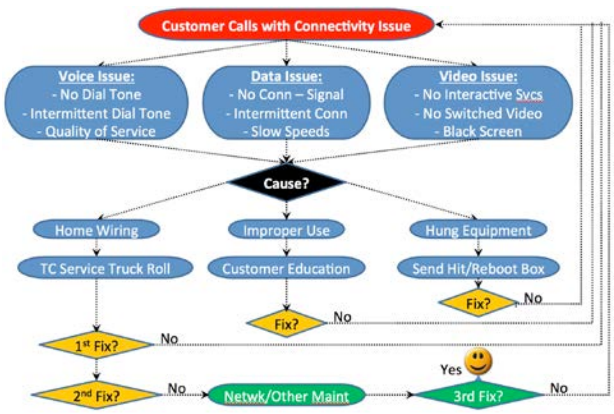
Before a cable operator can solve the customer's problem, they need to determine the root cause of the problem. The challenge, however, is that cable operators are faced with a voluminous amount of data generated from disparate sources that must be evaluated in the process of diagnosing and solving a subscriber's technical problem. These data sources include, but are not limited to, the following: tech support customer calls, premises truck rolls, maintenance actions, weighted telemetry, network operations center (NOC) alarms, poor/failed calls, voluntary disconnects and customer credits. The key challenge for operators with this vast amount of data is that there is no way to systematically consider and correlate all of this data into actionable information. The result of this challenge is that technical problem solutions are more trial and error in character to subscribers and less customer friendly.

The 'experience gap' for the customer

The 'trial and error' nature of problem diagnosis and solutions is illustrated in Exhibit 8. The Exhibit depicts the sequence of events that occurs when a cable subscriber calls with a valid hybrid fiber coaxial (HFC) connectivity or back-office (e.g., provisioning) issue that has yet to be detected by the Service Provider. In this case, there is a problem in the plant affecting the subscriber experience. When there is a problem in the service delivery infrastructure, it manifests with subscribers calling about a number of different connectivity issues encompassing voice, data and video examples.

These issues can be manifestations of connectivity problems that may actually reside in the plant somewhere. So, if there is bad ingress, bad power levels, or high error rates in the HFC infrastructure, the subscriber could experience any of these issues.

Exhibit 8: Call Center, Service and Maintenance Life Cycle of Subs with valid 'Plant Connectivity' or 'Back Office' Issues



Source: Developed by the authors with contributions from Rev2. (see www.rev2.com)

Let's look at the flow of the subscriber call in Exhibit 8 in sequence. Across the second row the call center service representative (CSR) is faced with the challenge of categorizing each call based solely on the customer's report. It is straightforward for the CSR to categorize a call in the area of voice, data or video. But there is no way for the CSR to recognize that the issues might be related. To illustrate, in the diamond-shaped box called "Cause," the CSR has to determine what he/she thinks is the cause of the problem. This must be done quickly and based only on the information by the subscriber/caller. On the right hand side of Exhibit 8, the CSR has decided that the equipment needs to be reset because it is hung (not responding) or needs to have services re-authorized/refreshed. To remedy this, the CSR can send a hit to the box or reboot the modem. Sometimes the remedy will make the box come back online. However, if the problem is actually in the plant and not in the modem, it will rear its ugly head again. As a result, the customer calls back again.

Another possible resolution is the middle path. The CSR could determine that the modem has been improperly used and, after another reboot, could recommend a resolution around customer education—teaching the customer how to reboot the modem, how to clean out the cache on the browser, or even how to troubleshoot some of the wiring.

But if we go back to the assumption that the problem is really in the plant, no kind of customer education is going to fix the problem. The customer will eventually call back. At some point, we need to follow the left-hand path where the CSR decides that the problem is in the home and a truck is sent to the home. When the service technician visits the home, he/she will probably perform one or more of the following five resolutions: replace the modem or set-top box, run a new drop, run new inside wiring, change a splitter configuration, or troubleshoot the ground block. If this first visit to the home fails to resolve the problem, then a second truck roll to the home may be required. If this next home visit fails to resolve the problem, then the problem will likely be referred to network maintenance.

This trial and error process is a double negative for Service Providers. First, it increases the cost of technical operations as a result of the unnecessary telephone calls and truck rolls. By identifying and resolving issues in a timely manner, Service Providers can lower customer care costs and, thereby, increase the customer lifetime value (CLV) of customers. Second, it degrades the quality of the customer experience since it requires greater effort from the customer and the customer is not receiving the desired experience. Ultimately, customers may become frustrated with their Service Provider and will disconnect and find another provider. All of these outcomes are contrary to the CE framework presented earlier—that is, the Functionality of the product has been compromised and the customer is receiving far less of the desired experience. The 'experience gap' is widened and if Cluescanning™ were conducted, the Functionality component would score quite low.

3. The Proposed Solution

The proposed solution for this a voluminous amount of data generated from disparate sources and the resultant trial and error problem diagnosis is the utilization of an operational tool that included the following elements:

- Data warehouse with risk analytics software and visualization capabilities
- Introduction of three additional inputs: connectivity calls, failed telemetry, and trouble calls/plant maintenance
- Correlation of all of the data helps prioritize problem areas and identify the appropriate remedy, i.e., reboot, customer education, residential truck roll, maintenance truck roll, etc.

The NMS platform must be designed specifically to tackle these large impactful challenges. Through constant automated surveillance of the network, a data warehouse acts to establish a historical record. Modern technologies and techniques designed for Big Data are leveraged for storing and processing large amounts of information, even from 3rd party business systems. Through these mechanisms, the entire BSS/OSS system can be monitored for current outages, predicted network issues, service anomalies, and cross-system impact - and prioritized for business and financial risk. The goal is to automate system surveillance and analytics in a repeatable and predictable way, modeling what a network engineer or business analyst would do manually, to provide shortest path to resolution for the largest number of subscribers.

Active outages require special attention. The pressure to detect current service issues may push network analysts to roll trucks before knowing the true cause and location of the issue. This approach reduces resources available for other, potentially more impactful, anomalies. A NMS system designed with the ability to correlate root cause will transforming the way operators are managing outages, resulting in fewer and more focused plant truck rolls, the prevention of residential trouble call truck rolls when related to an outage, and opportunity for IVR call intercept in the call center for known outages. As operators become more proactive in their plant maintenance, the value of extinguishing issues before they turn into an active outage becomes apparent. Advanced analytics techniques are used to predict issues, target the fix location, and thus prevent subscribers from ever experiencing an outage. Global operators are still following a simple scheduled "inspect and tighten" campaign where, despite the state of the health of that network segment, plant elements are inspected. While some level of scheduled inspection may always be useful, the proactive NMS platform leverages advanced technology to show operators where their plant maintenance resources will be most impactful.

Determining business risk is much more than simply counting the number of affected subscribers during an outage. By looking beyond network telemetry and including other OSS/BSS indicators, a clear picture of business risk is available. For example, business risk may be highest or subscribers may be more at risk of “churning out” if service in an area is newly launched, has a high concentration of business-class customers, has a history of chronic (subtle repeated) networked issues, has a pattern of high call center call volume, or is in an area with heavy competition (cable vs. telco vs. satellite). Operators may also look to consider the cost of components/maintenance involved, where proactively maintaining plant elements is far cheaper than waiting until they require replacement, or the service location is in a challenging area requiring special equipment or teams to service. The NMS platform platform must supports this risk analysis, which provides the critical input for business executives to use when making important investment decisions. This risk analysis also helps managing the top priority maintenance activities with limited resources.

Forward thinking Service Providers are examining new ways to not only leverage information, but to make it actionable for maximum customer satisfaction, loyalty and business advantage. To this end, an integrated FSM and NMS ecosystem provides full end-to-end automation of plant maintenance management. The NMS system must automatically discovers and monitor the network, detect current and future network issues, and measure and prioritize the business risk. Verified plant issues can then be published to the FSM for automatic work order routing to plant technicians, who are electronically notified of a critical work order while they are already in the field. This workflow is further enhanced by an FSM tool that with the intelligence to assign task using multiple variables including the locations, skillsets, schedules and equipment of all available technicians.

Economic Proposition

The annual cost of the operational tool and the additional process involved in integrating the tool plus more maintenance trucks being deployed should be offset by savings garnered from less residential trouble calls (TCs), fewer connectivity calls to the call center, and reduced CPE (consumer premise equipment) cost (i.e., unnecessary discard of CPE equipment).

	Typical Process					Proactive Process			
Measure		Qty	Cost per	Mo. Rev per	Annual Net Value	% Improvement	Qty	Annual Net Value	ESTIMATED IMPROVEMENT
Total Subscribers		1,000,000		$100	$1,200,000,000	0%	1,000,000	$1,200,000,000	**$0**
Total Calls/Month	100%	1,000,000				1%	990,100		
New/Upgrade Service	5%	50,000	-$100	$50	$120,000,000	5%	52,475	$125,940,000	**$5,940,000**
Billing Related	30%	300,000	-$25	$0	-$90,000,000	0%	300,000	-$90,000,000	**$0**
Other	30%	300,000	-$25	$0	-$90,000,000	0%	300,000	-$90,000,000	**$0**
Where's my Tech	5%	50,000	-$25	$0	-$15,000,000	0%	50,000	-$15,000,000	**$0**
Technical Support	30%	300,000	-$25	$0	-$90,000,000	4%	287,625	-$86,287,500	**$3,712,500**
remote resolution	75%	225,000	-$25	$0	-$67,500,000	0%	225,000	-$67,500,000	$0
Require Truck Roll	25%	75,000	-$100	$0	-$90,000,000	17%	62,625	-$75,150,000	**$14,850,000**
Repeat Truck Roll	10%	7500	-$100	$0	-$9,000,000	50%	3,750	-$4,500,000	**$4,500,000**
Avoidable	15%	11250	-$100	$0	-$13,500,000	50%	5,625	-$6,750,000	**$6,750,000**
Referred to Plant	5%	3750	-$100	$0	-$4,500,000	80%	750	-$900,000	**$3,600,000**
Other	70%	52500	-$100	$0	-$63,000,000	0%	52,500	-$63,000,000	**$0**
Trouble Calls Converted to new/upgrade service						20%			
TOTAL ANNUAL VALUE					$945,000,000			$969,502,500	**$24,502,500**

CEM Impact

The improvements in the diagnosis and repair of customer problems will improve key indicators of customer satisfaction:

> Connectivity calls | Repeat calls | NTF (no trouble found) calls | Repeat TCs | Churn | Network availability | Customer satisfaction scores |

Over time, these improvements will increase the emotional bond between the customer and the cable operator which would lead to customer advocacy, customer loyalty and an increase in the Lifetime Value of the Customer Lifetime (CLV).

4. Case Study Results

As noted above, the ideal metrics for measuring the financial payoff for investments in CEM are the typical ROI (return on investment) metrics as Payback period, Net Present Value, and Internal Rate of Return. Unfortunately, these tools could not be utilized since there was not sufficient data to do a before-and-after cost/benefit comparison. In this case study situation, the decision was made to begin using this operational tool without documenting the before situation. Ideally, this before case would have included such benchmarks as the following:

- Phone Calls (annual numbers)
- Truck Rolls (annual numbers)
- Repeat Truck Rolls (annual numbers)
- Voluntary Disconnects (annual numbers)
- Technical Phone Calls per Customer per year
- Customer Effort (minutes) per year
- Total Operating Costs per Customer per year
- Net Promoter Score
- CLV (customer lifetime value)

By collecting this data on the front end, the change in both operational costs as well as CEM metrics can be documented so that the cable operator can gauge the financial and CEM benefits of the investment in the operational tool.

In this case study, we found that sometimes it is difficult to accurately measure before-and-after costs because the operational environment is dynamic. That is, the volume of phone calls and maintenance truck rolls are dependent on numerous factors such as weather, marketing promotions, plant repairs, etc. As a result, before-and-after cost comparisons may not provide absolute documentation as to the financial payoff for the CEM investment. Given this reality, cable operators may need to focus on indirect measures of financial payoff. Examples of indirect metrics of performance include repeat truck calls, number of voluntary disconnects, customer effort, and churn. Although documented changes in these operational and CEM metrics do not guarantee that there is a financial payoff, they do provide strong anecdotal evidence of a positive ROI.

5. Lessons Learned

There are three key lessons learned from this case study with respect to determining the financial payoff from a CEM investment. These lessons are:

- Determine how you are going to measure results on the front end.
- Include CEM metrics in the data set along with operational metrics.
- Remember CEM is an integrated effort. Hence, consider costs/benefits beyond field service management.

Every organization has the opportunity to close the experience gap, and in this case example, the experience gap could be closed through an investment in technology that directly affects the Functionality component of CEM. In other situations, the gap is closed by improvements in Ambiance, Communication, or Talent. The overall point is that to achieve any of the more advanced stages on the Customer Experience Maturity Curve; an organization must become committed to understanding the gap and then making a sound financial decision about CEM investments. This case illustrates how such an investment can be measured to show its benefit to subscribers and to the organization.

REFERENCES

American Express® Global Customer Service Barometer, 2013.

Ansoff, H. Igor, *Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion,* McGraw-Hill, New York, New York, 1965.

www.bristol-inquiry.org.uk

www.bpaww.com

Carbone, Lewis P., Clued In: *How to Keep Customers Coming Back Again and Again,* Prentice Hall, Upper Saddle, New Jersey, 2004.

www.consumerreports.org

Coulter, R. and G. Zaltman, "Using the Zaltman Metaphor Elicitation Technique to Understand Brand Images," *Advances in Consumer Research,* 1994,

Denning, Steve, "How do you change an organizational culture?" *Forbes,* July 23, 2011.

Dixon, Matthew, Nick Toman, and Rick DeLisi, *The Effortless Experience,* Conference Executive Board, Washington, D.C., 2012.

Gartner Predicts, February 2013. (http://mktimages.gartner.com/pv_obj_cache/pv_obj_id_92DB197B681D1A1003F701CB30B626FF3B111B00/filename/)

The Huffington Post, http://www.huffingtonpost.com/2011/09/20/the-top-10-companies-with-most-admired-customer-experience_n_972027.html

Investor Words (www.investorwords.com)

www.jdpower.com

Kotler, P., W. Gregor, and W. Rodgers, III, "The Marketing Audit Comes of Age," *Sloan Management Review,* Winter 1989.

Marketing Land, http://marketingland.com/pew-61-percent-in-us-now-have-smartphones-46966

Mehrabian A., "SilentMessages", (http://www.nonverbalgroup.com/2011/08/how-much-of-communication-is-really-nonverbal/)

Murphy, Emmett C., and Mark Murphy, *Leading on the Edge of Chaos: The 10 Critical Elements for Success in Volatile Times*, Prentice-Hall, Upper Saddle, NJ, 2002.

Pine, B. Joseph and James H. Gilmore, *The Experience Economy: Work is Theatre and Every Business a Stage,* Harvard Business School Press, Boston, Massachusetts, 1999.

Reichheld, F., "The One Number You Need to Grow," *Harvard Business Review,* December 2003.

www.rev2.com

Safko, Lon, "How much did that new customer cost you?" *Entrepreneur,* January 14, 2013,

Schmitt, Bernd H., *Customer Experience Management,* John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.

Smith, S. and J. Wheeler, *Managing the Customer Experience,* Financial Times Prentice Hall, Upper Saddle, N.J., 2002

Tierney, Jim, "Most Consumers Will Drop a Company Due to Bad Customer Service Experience," *Daily News,* Aug 13 2013 (http://loyalty360.org/resources/article/most-consumers-will-drop-a-company-due-to-bad-customer-service-experience#sthash.ojiV3uIP.dpuf)

Zaltman, G. and R. Coulter, "Seeing the Voice of the Customer: Metaphor-Based Advertising Research," *Journal of Advertising Research,* July/August 1995).

THE IMPORTANCE OF AIR TIME ALLOCATION IN WI-FI QUALITY OF SERVICE

Eli Baruch
ARRIS

Abstract

As Cable and Telco service providers strive to own the subscriber's home experience by deploying ever more sophisticated Wireless Gateways and multi-service offerings; wireless technology plays a key enabling role in the way consumers use the various services. With the proliferation of Multi-SSID Wireless gateways that offer multiple services such as an in-home private network, Home Security and Appliance Monitoring, Public Hotspots and Video over Wi-Fi; the need to monitor and assure basic service levels and an overall Quality of Service (QoS) becomes essential to the user's quality of experience.

Through a series of real-life tests, this paper will show how these competing services can impact the home user's wireless experience, in both single user and multi user environment. Also demonstrated, is how the shortcomings of existing wireless systems, the limits of existing WMM, and the lack of a proper QoS monitoring and enforcement mechanisms can degrade the user experience and create severe subscriber retention problems. This paper will show that the critical resource on the wireless network is not raw bandwidth management but rather comprehensive management of the time allotted on the air interface, or Air Time Management. Lastly a variety of options will be discussed illustrating methods of avoiding delivery of a poor user experience and guaranteeing an acceptable basic service level.

WIRELESS QUALITY OF SERVICE

Wireless 802.11 technologies were designed as a method of extending LAN-type service over the air. As such, they were seen as an extension of Ethernet LAN services, for which QoS usually did not play a major role in the end user's experience. Due to the nature of the wireless medium, and in anticipation of multiple types of traffic using the air interface, a basic QoS mechanism was defined in the 802.11e standard and adopted by the Wi-Fi Alliance as part of their certification and interoperability program under the name of Wi-Fi Multimedia (WMM)[1].

WMM defines four Access Categories (AC):

- VoIP—Very low throughput with highest priority and strict latency requirements
- Video—High priority with latency requirements
- Best Effort—Low priority
- Background—Lowest priority

Incoming traffic is tagged and assigned one of four different priorities. Individual packets are then directed to one of four internal queues and prioritized according to the AC into which they fall. Packets from higher priority ACs are transmitted with a smaller inter-frame space and a smaller random back-off window, which allows transmission to the wireless medium with less delay on average.

The WMM mechanism thus provides statistical priority for winning access to the air interface. The algorithm used by a WMM enabled Access Point is probabilistic and depends on two timing parameters that vary for each AC:

1. The minimum interframe space or Arbitrary InterFrame Space Number (AIFSN), and
2. The Contention Window (CW), a random backoff wait time.

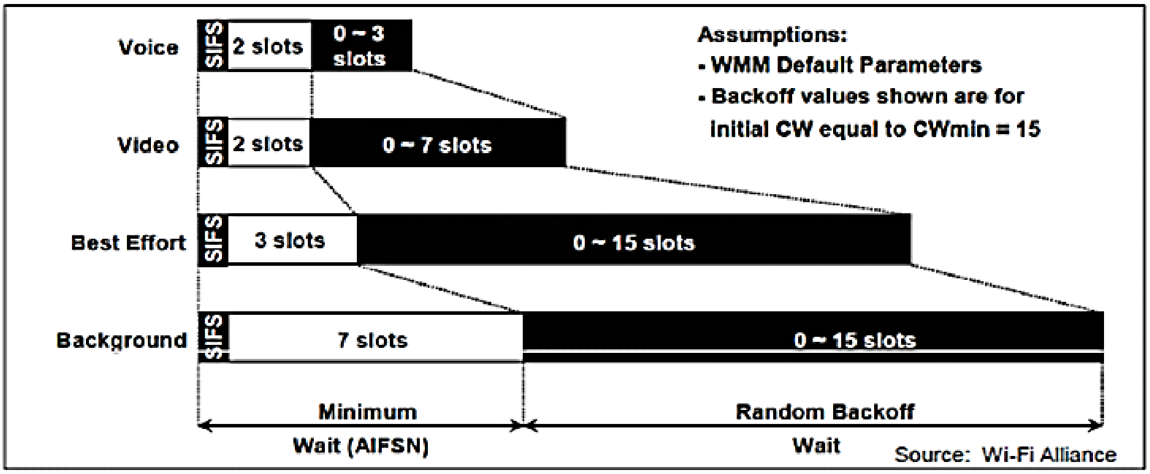


Figure 1: Wireless Multimedia – wireless QoS mechanism

After each collision, the CW is doubled until a maximum value, which also depends on the AC, is reached. As frames with the highest AC tend to have the lowest backoff value, they are more likely to be transmitted. However, little consideration was given to a multi-SSID scenario, in which different services may use different SSIDs to indicate the expectation of different levels of service. For example, a home security service may use a very limited amount of bandwidth. The user, however, expects that bandwidth to be available whenever needed. On the other hand, a hotspot service may have inconsistent bandwidth requirements, but the user does not expect such service to degrade or compete with the wireless home network's bandwidth.

Level of Service Expectations

A basic minimum level of service is expected from any Internet service provider. To fulfill this expectation, most service providers manage their High Speed Internet services with a dynamic set of DOCSIS service flows or using IP Differentiated Services protocols. These mechanisms govern how bandwidth is allocated to different services and users. Such schemes distribute the maximum overall bandwidth to and from the home by dynamically allocating bandwidth between preferred and best-effort services. The DOCSIS service flow mechanism, however, only governs the traffic allocation to and from the cable modem or wireless gateway. This mechanism does not enforce any priority over the wireless network. A fundamental question to be addressed is the actual workings of the intuitive picture depicted in Figure 2: Theoretical behavior of Dynamic Service Flow.

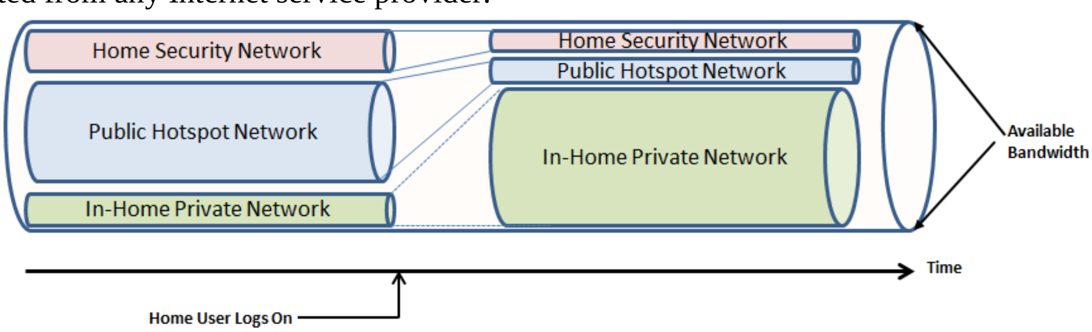


Figure 2: Theoretical behavior of Dynamic Service Flow

UNDERSTANDING AIR TIME

As a wireless signal propagates through the air, its signal strength and the ratio of signal to noise as seen by the receiving end diminishes as a function of distance, attenuation (e.g. obstacles , Multipath, and reflections), and temporal or other interference. The minimum received signal power level required to achieve a sufficient signal-to-noise ratio (SNR) is called “receive sensitivity.” If the received signal power level falls below the receive sensitivity required for a given data rate, communication at that data rate becomes unreliable. A common manifestation of such unreliable communication is video stream buffering or freezing, as well as dual-band switchable clients (e.g. iPad) switching between 5GHz and 2.4GHz bands.

In order to maintain a reasonable SNR, the transmitter will modify the Modulation and Coding Scheme (MCS) by changing the modulation profile, the error correction scheme, and the number of spatial streams that send traffic to the receiver. Without going into a deep technical description, we can simply state the higher the SNR, the better the performance. Higher order QAM are not as robust against noise or other degradations as lower-order QAM. Using higher-order QAM without increasing the bit error rate requires a higher signal-to-noise ratio (SNR). The table in
Figure 3: 802.11n MCS index and data rate
 lists the different Modulation and Coding Scheme index used by a 3x3:3 802.11n wireless device.

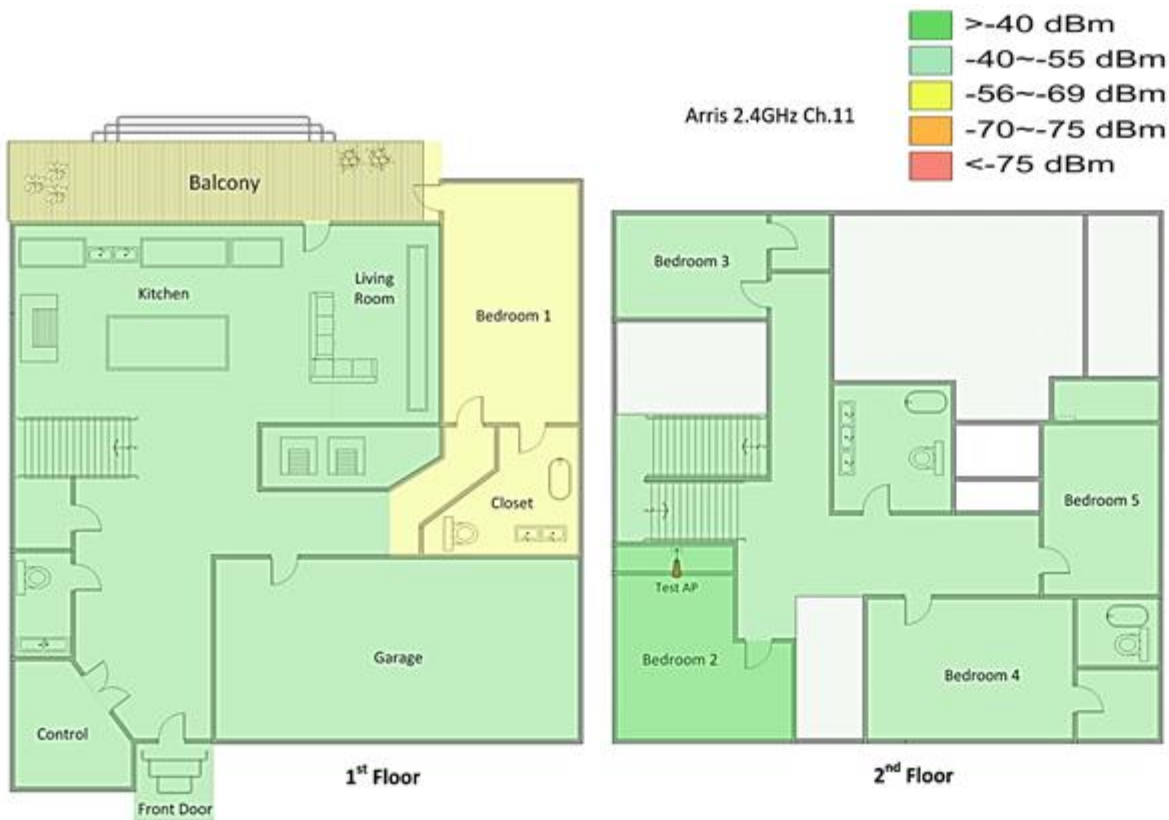
MCS	Modulation	Coding	Spatial Streams	Theoretical Data Rate (Mbps) GI = 800ns		MCS	Modulation	Coding	Spatial Streams	Theoretical Data Rate (Mbps) GI = 800ns	
				20MHz	40MHz					20MHz	40MHz
0	BPSK	1/2	1	6.5	13.5	12	16-QAM	3/4	2	78.0	162.0
1	QPSK	1/2	1	13.0	27.0	13	64-QAM	2/3	2	104.0	216.0
2	QPSK	3/4	1	19.5	40.5	14	64-QAM	3/4	2	117.0	243.0
3	16-QAM	1/2	1	26.0	54.0	15	64-QAM	5/6	2	130.0	270.0
4	16-QAM	3/4	1	39.0	81.0	16	BPSK	1/2	3	19.5	40.5
5	64-QAM	2/3	1	52.0	108.0	17	QPSK	1/2	3	39.0	81.0
6	64-QAM	3/4	1	58.5	121.5	18	QPSK	3/4	3	58.5	121.5
7	64-QAM	5/6	1	65.0	135.0	19	16-QAM	1/2	3	78.0	162.0
8	BPSK	1/2	2	13.0	27.0	20	16-QAM	3/4	3	117.0	243.0
9	QPSK	1/2	2	26.0	54.0	21	64-QAM	2/3	3	156.0	324.0
10	QPSK	3/4	2	39.0	81.0	22	64-QAM	3/4	3	175.5	364.5
11	16-QAM	1/2	2	52.0	108.0	23	64-QAM	5/6	3	195.0	405.0

Figure 3: 802.11n MCS index and data rate

A wireless gateway/access point will change the MCS used to communicate with different clients. In other words, the same client, depending on its location relative to the access point, may use a different MCS setting and therefore have a different level of data throughput at different locations. The further the client is from the access point; the fewer number of bits it will transmit in a given amount of Air Time. Moreover, the more time is required for communicating with one client, means that less time—and hence less overall bandwidth—remains for others using the same access point. It is important to note that the MCS index by which the access point will communicate with a given client is not fixed. As a wireless environment changes, an access point may switch MCS while transmitting to the same client based on SNR and Packet Error Rate (PER).

In order to illustrate the receive sensitivity distribution based on location, some use a color coded map, also known as a Heat Map, that shows the receive levels seen by a client from the access point. Such maps, although not always a direct representation of the bandwidth one might expect at different locations, give some indication of the behavior to expect from close and remote clients.

Figure 4: Example of receive levels “heat map”



The Impact of Location on Performance

To demonstrate the effect that a lower MCS due to distance has on the amount of Air Time that a client uses, we conducted several tests in a test house. The test house layout is depicted in**Error! Reference source not found.**.

We measured the traffic sent from the Access Point (AP) located at the main floor to a single client in location 9, the reception room on the main floor. We placed the client in location 9 10 feet away from the AP and in line of sight.

We also measured the traffic sent from the AP to location 6, the bathroom on the upper floor located at the edge of the house. Location 6 was 55 feet and 5 walls away from the AP.

We measured the throughput of downstream traffic from the AP to the client and the response time that the TCP connection needed to acknowledge the transfer of a fixed block of TCP packets. The AP we used was an ARRIS Wireless Gateway with Dual Band Concurrent 3x3:3 802.11N radios. We conducted all measurements using the 2.4GHz band.

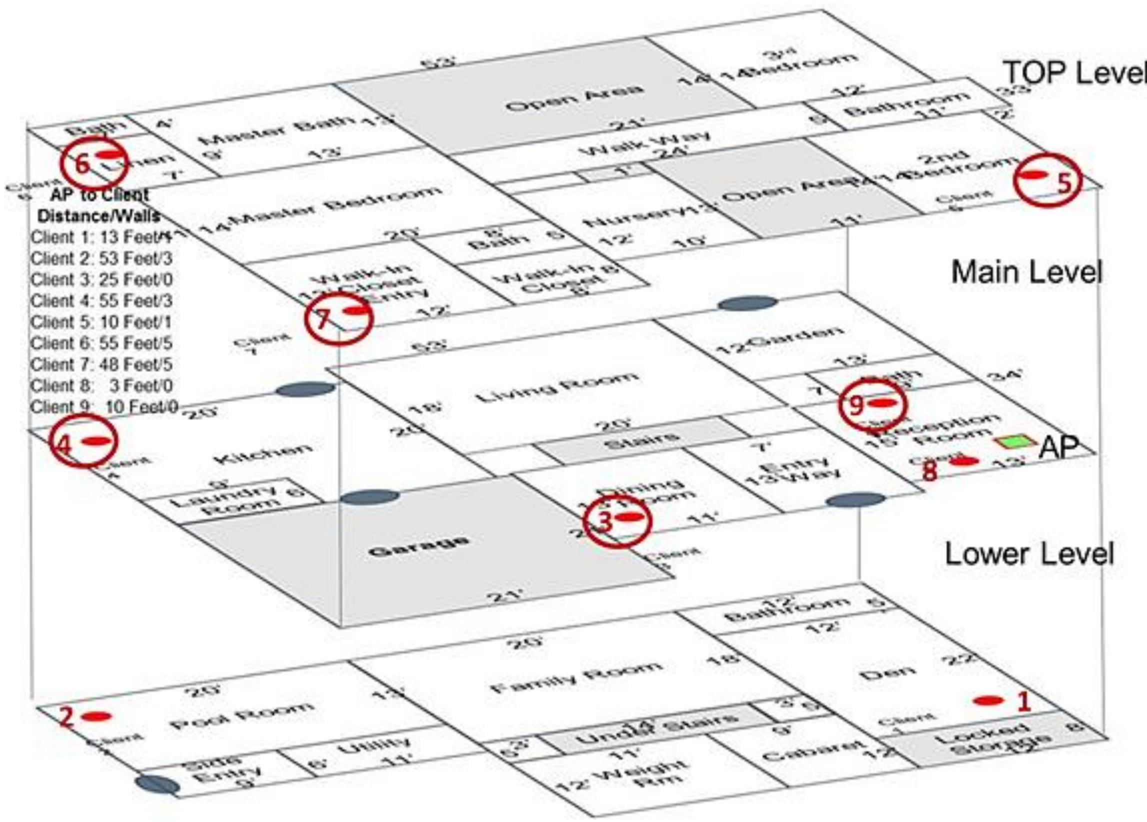


Figure 5: Wi-Fi test house client location map

The client in location 9 is close to the AP, which used MCS=22 to communicate with it. As result, it enjoyed good connection speeds, averaging 96.37 Mb/s, and fast response times, averaging of 0.834 sec.

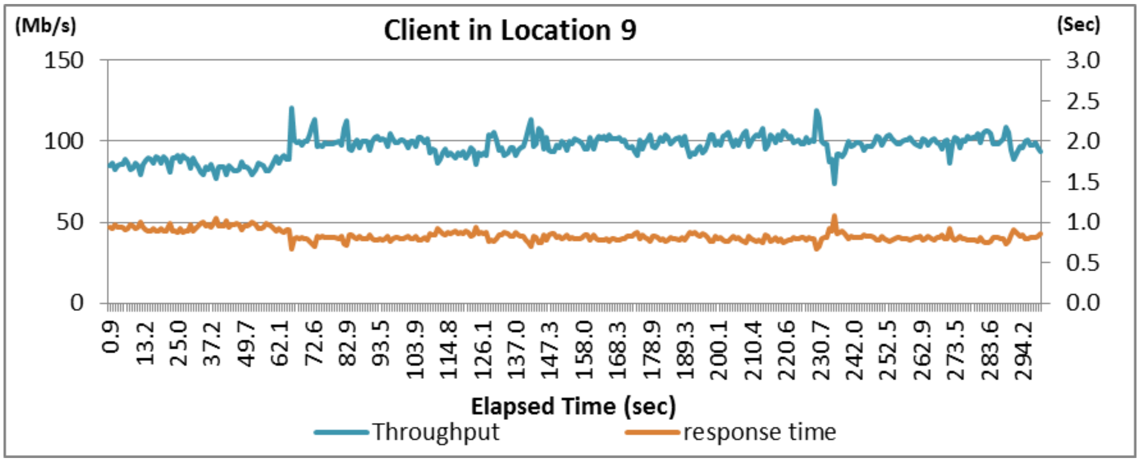


Figure 6: Throughput and response time measured in location

The client in location 6 started temporarily with MCS=17, continued for a while with MCS=19, and finished with MCS=20. As result, it experienced lower throughput than Client 9, with an average speed of 58.75Mb/s, and experienced longer response times, averaging 1.378 sec. In other words, in order to send the same amount of traffic, the client in location 6 used **40%** more time to complete the data transfer.

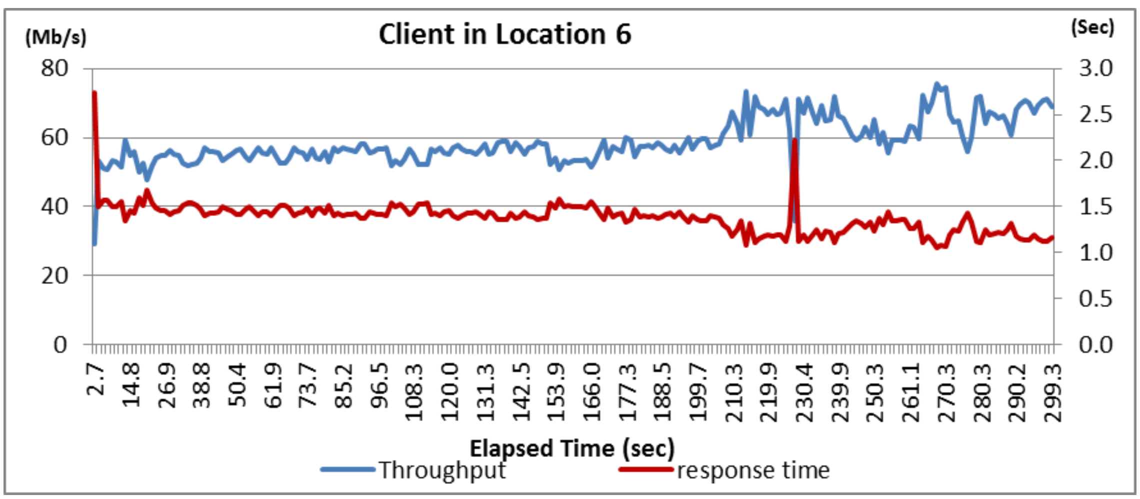


Figure 7: Throughput and response time measured in location 6

MULTI-CLIENT ENVIRONMENT

Our statistics show that, on average, there are six active wireless devices in a home that share access to the home wireless network **Error! Reference source not found.**. Since those different clients may reside in different rooms in the house, and at different distances from the AP, the overall aggregated available bandwidth in a multi-client environment is reduced slightly in comparison with a single device.

Assuming a typical package of 50Mb/s WAN-side Internet service, here is an example of test results for bandwidth usage and bandwidth distribution from amongst one to six wirelessly connected 802.11n devices in the home. In the example shown in Figure 8: Bandwidth usage in multi-client environment , all of the devices are located at similar distances from the house's main AP.
Note that even at similar distances and using the 802.11n standard, without any slower 802.11g or 802.11b clients on the network, there may still be differences in overall bandwidth between devices because of differences in their wireless characteristics (e.g., the number of receive and transmit antennas, power levels, etc.). The overall aggregated bandwidth used by the varying number of devices, however, is affected only slightly by the number of active devices. As one would expect, the more devices that are actively connected, the lower the overall bandwidth shared among all devices—including the individual bandwidth allocated per device.

Although when multiple clients are attached to the same AP, the overall available time on the wireless medium for data transmission is also reduced compared to a single device. The way the access point choses to distribute the Air Time significantly impacts the individual client data throughput and may as well impact the overall aggregated data throughput. In some cases, an AP may use a "fairness" algorithm to allow for the aggregated bandwidth to be allocated fairly between the clients. Such "fairness" algorithms may use a packet base round robin between clients. The drawback of this approach is that with TCP type of IP connection the client using the higher rate MCS will end up consuming the majority of the available bandwidth to the point that it may starve the other clients. Other algorithms may try to take into account time-based information such as Wi-Fi frames retransmissions and MCS rate.

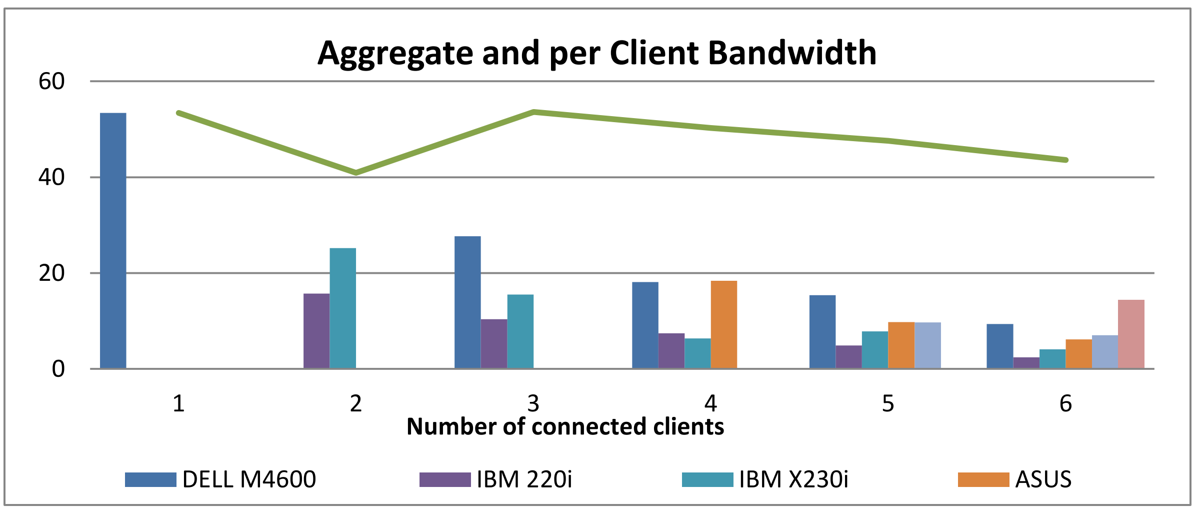


Figure 8: Bandwidth usage in multi-client environment

The example below shows test results of 10 clients connected to two different access points. Each access point uses a different algorithm to allocate Air Time between clients. Although the aggregated bandwidth consumed by the 10 clients in both tests was almost identical, 36Mb/s, note how differences in Air Time "fairness" algorithm manifests in each client's bandwidth.

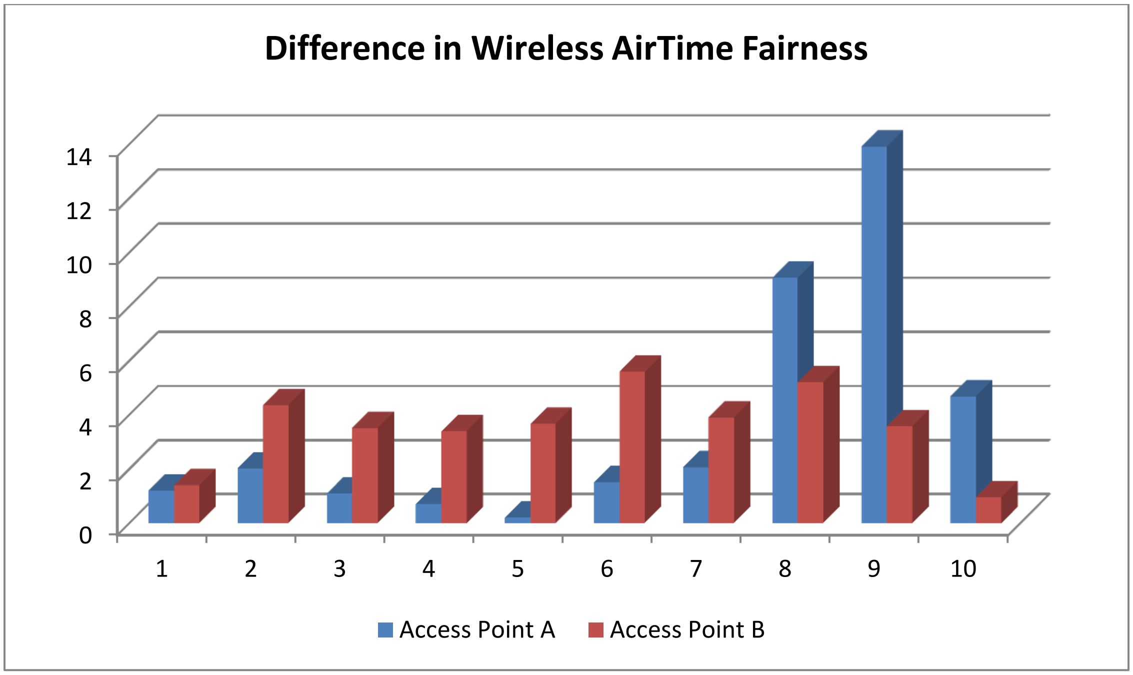


Figure 9: Bandwidth usage in multi-client environment

In the example shown in Figure 110 and Figure 121, six clients were placed at locations 3, 4, 5, 6, 7, and 9 in the test house. In order to eliminate the variability of the clients' Wi-Fi characteristics, the same version of client was used in each location. The AP was an ARRIS Wireless Gateway with a 2x2:2 2.4GHz 802.11N radio.

Although the six clients were located in different rooms, the results show a relatively even distribution of overall data throughput between the different clients, with the three clients closest to the AP—3, 5 and 9—enjoying higher throughput. As one can see from the graphs below, the bandwidth per client varies over time.

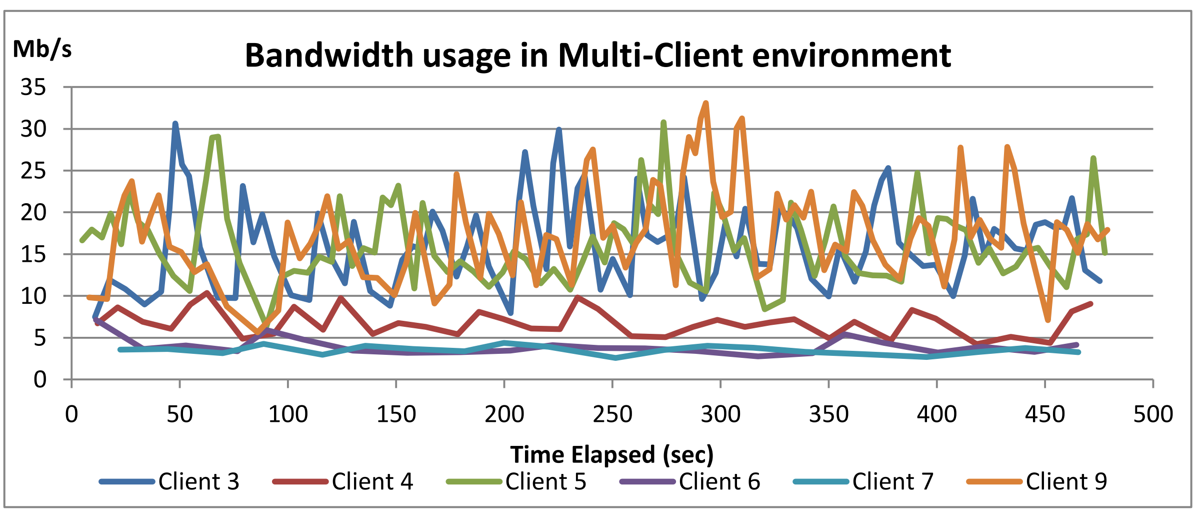


Figure 10: Bandwidth usage in multi-client environment

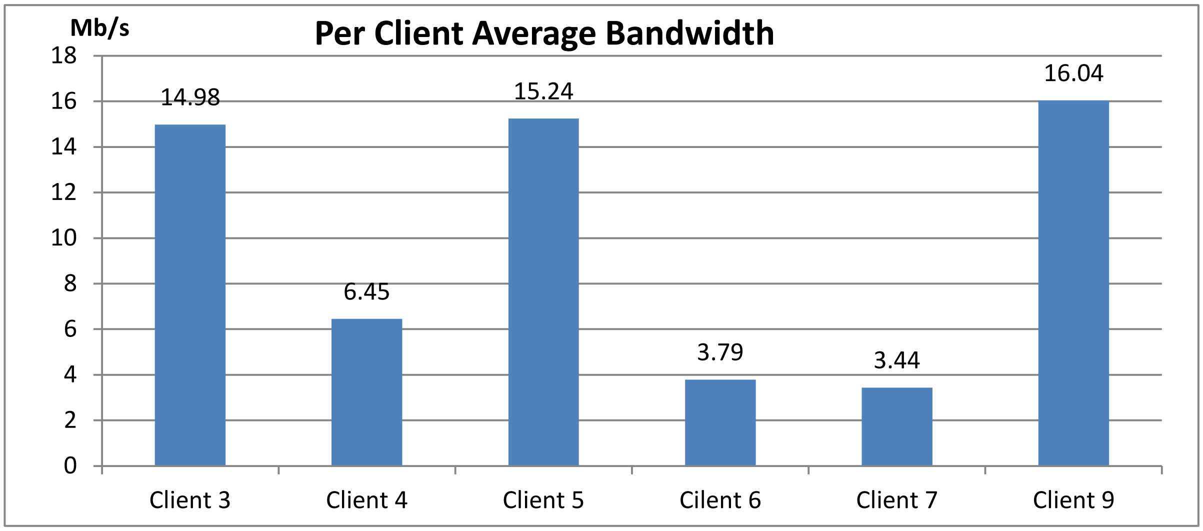


Figure 11: Average bandwidth usage in multi-client environment

	Location 3	Location 4	Location 5	Location 6	Location 7	Location 9
Average Bandwidth (Mb/s)	14.98	6.45	15.24	3.79	3.44	16.04
Average Response Time (Sec)	5.34	12.40	5.25	21.11	23.27	4.99
Percentage of Air Time Used	7.38%	17.14%	7.26%	29.16%	32.16%	6.90%

Figure 12: Air Time distribution in multi-client

If we take a closer look at the over-the-air allocation of time, the remote clients with lower MCS values take significantly more Air Time to deliver the same number of byte

THE IMPACT OF WIRELESS HOTSPOT SERVICE

Most service providers want to augment new and existing outdoor wireless deployments with home hotspot services using always-on, multi-SSID wireless gateways. The active hotspot serves Wi-Fi roaming users and/or offloads cellular traffic from the 3G/4G network onto the Wireless network and into the service provider's DOCSIS or IP backhaul networks.

Although some of the users connected to the hotspot SSID may be houseguests, they are more commonly roamers who are situated outside the house, but still within the wireless coverage range of the home access point. By the very nature of being outside, those roaming devices may be at the fringes of the home's wireless coverage and therefore would use a lower speed MCS compared to devices in the home.

As the number of roaming clients connected to the home hotspot increases, the impact on the existing clients in the home becomes more noticeable. These distant clients use a large portion of Air Time, which causes the in-home clients' bandwidth to decrease noticeably.

In order to demonstrate the impact of home hotspot roamers on in-home clients, the bandwidth and response time of Client 9 were monitored after situating two Wi-Fi clients just outside the house. These clients connected to the home's hotspot service and streamed an HD YouTube video. The home hotspot clients started streaming the movies approximately four minutes after test began.

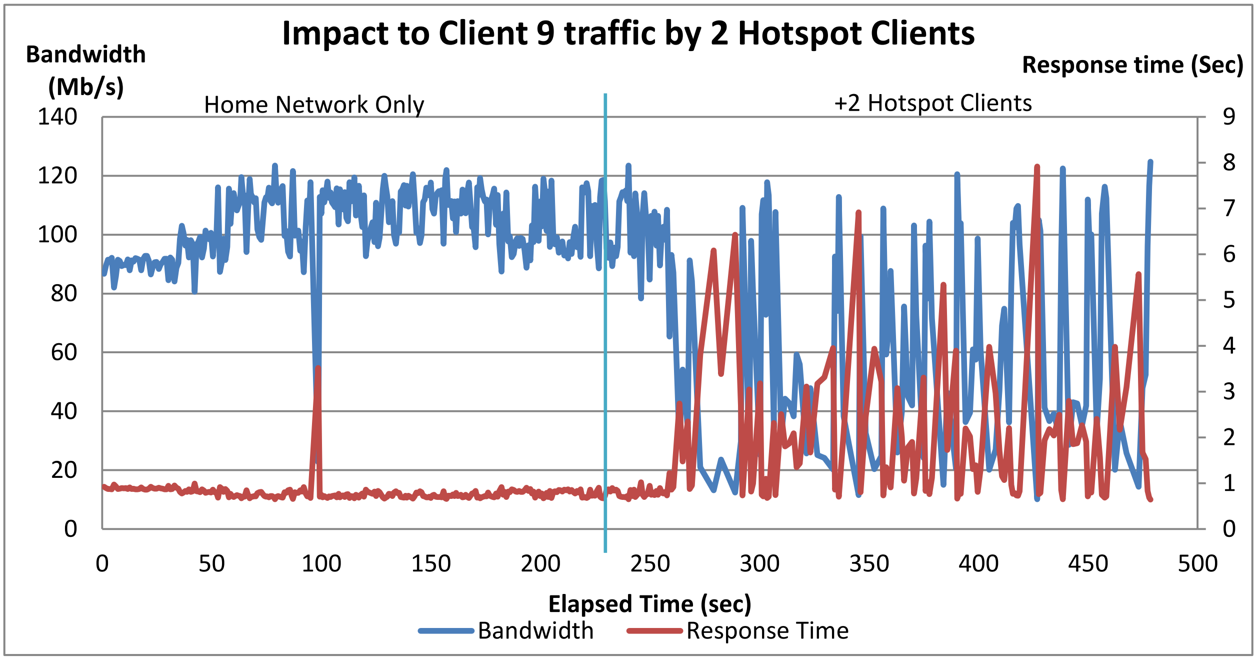


Figure 13: Client 9 bandwidth and response time (Hotspot traffic starts 4 min into the test)

Another test was done with six home clients along with the two hotspot Wi-Fi clients just outside the house streaming HD YouTube videos. As in the previous test, the home hotspot clients started the streaming of the movies approximately 4 minutes after test began. This time, in an attempt to preempt the impact of the outside Hotspot clients, WMM was enabled and the home traffic was assigned to the Video Access Category thus giving the home clients the highest priority. Please note, the YouTube traffic towards the Hotspot clients was not manipulated by the access point. In other words any DSCP marking used by the origin server was carried over and mapped into the relevant WMM queue.

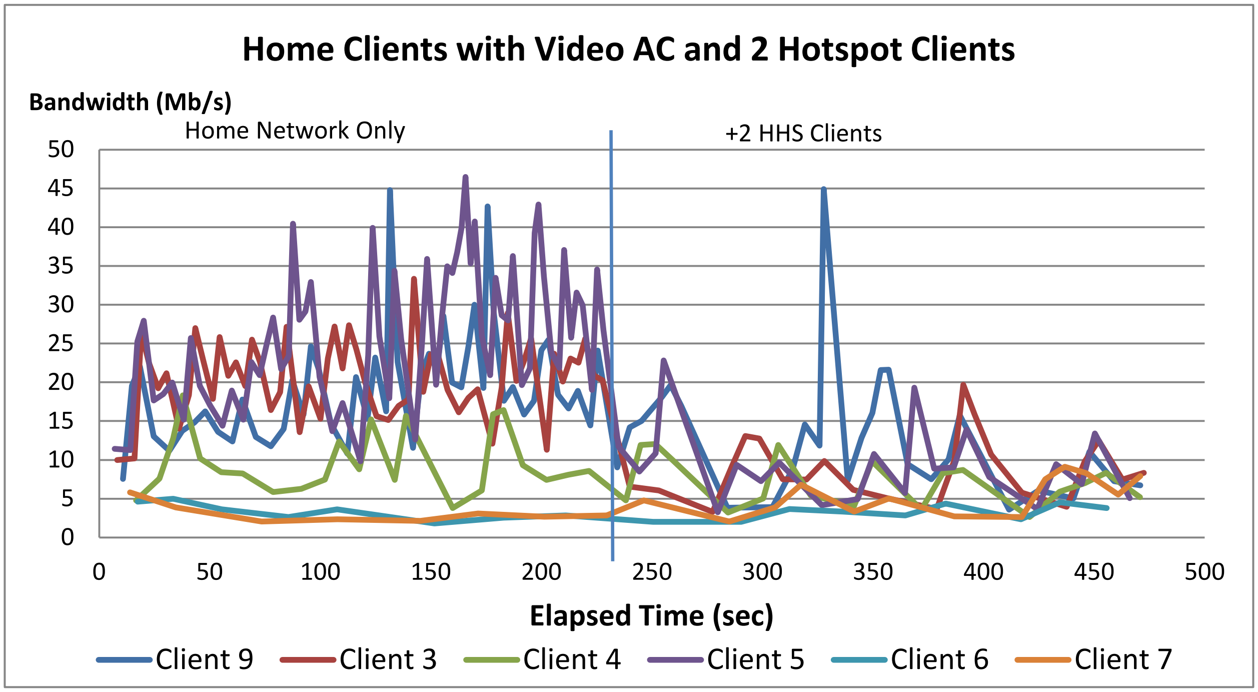


Figure 14: 6 home clients bandwidth
(Hotspot traffic starts 4 min into the test)

The impact of the outside clients was very noticeable, even with WMM applied. As seen by the chart below, the bandwidth of the home clients was reduced when the hotspot clients were active. Surprisingly, for some clients, the use of WMM priority actually resulted in lower bandwidth when compared to the client bandwidth without the use of WMM.

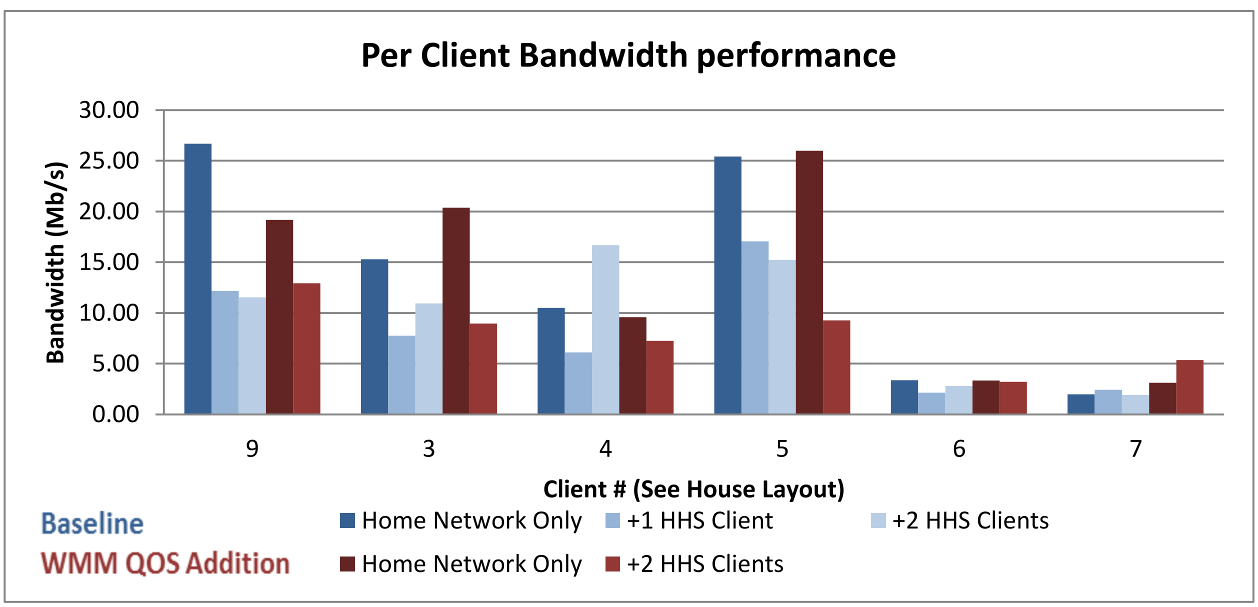


Figure 15: Per client bandwidth when hotspot clients are active with and without WMM

Bandwidth (Mb/s)					
Client	Baseline (no WMM)			With WMM QOS	
	Home Network Only	+1 HHS Client	+2 HHS Clients	Home Network Only	+2 HHS Clients
9	26.69	12.15	11.55	19.19	12.93
3	15.29	7.75	10.94	20.37	8.95
4	10.48	6.10	16.68	9.57	7.24
5	25.43	17.05	15.21	26.01	9.27
6	3.36	2.14	2.81	3.34	3.21
7	1.96	2.40	1.92	3.09	5.35
Aggregate	83.21	47.60	59.10	81.58	46.95

Figure 16: Comparison of per client bandwidth
when hotspot clients are active with and without WMM

Response Time (sec)					
Client	Baseline (no WMM)			With WMM QOS	
	Home Network Only	+1 HHS Client	+2 HHS Clients	Home Network Only	+2 HHS Clients
9	3.44	9.55	10.00	4.70	8.88
3	5.85	14.77	12.96	4.28	11.08
4	8.63	16.80	10.51	9.93	13.68
5	3.58	7.84	10.03	3.51	11.01
6	26.00	41.23	35.14	26.36	27.19
7	42.27	39.62	44.63	28.68	18.57
Average	14.96	21.64	20.54	12.91	15.07

Figure 17: Comparison of per client response time
when hotspot clients are active with and without WMM

As seen by this example a simple WMM mechanism that is based on extended backoff times has little positive impact on the bandwidth available to the home clients. The impact of remote clients with low MCS settings cannot be overcome by simply applying WMM. Customers experiencing such in-home Wi-Fi performance degradation will most likely have poor quality of experience, which may result in service calls to the service provider to “fix the Wi-Fi problem”.

The Impact of Home Security

A modern home security service usually involves the use of cameras that are located at the edges of the subscriber home and connected wirelessly to the home Access Point. Such devices and services are becoming more common than ever as found by our survey.(2) Other home security sensors may also be deployed, and most of them will probably be wirelessly connected. Although such cameras and sensors are always on, they usually send very low bit rate pictures or other information to a monitoring portal. At any time, however, the user may choose to view a live streaming feed from one or more of the connected cameras, which will require much more bandwidth to deliver the video. Security services therefore demand a changing level of allocated bandwidth in order to fulfill the promise of service. They also may require preferential treatment of these specific clients at the expense of other clients on the same in-home network. One of the key challenges here is the fact that the traffic from these IP cameras and sensors is upstream towards the access point. Any access point downstream WMM or other fairness mechanism does not apply, since the clients transmit upstream towards the access point based on the collisions and backoff mechanism inherent to Wi-Fi.

SOLVING THE IN-HOME WIRELSS QUALITY OF EXPERIENCE: MANAGING AIR TIME

As we have demonstrated in this paper, remote clients can impact a home network's overall performance severely. Lack of a proper quality of service monitoring mechanism can degrade the user experience and create severe subscriber retention problems. Standard Wi-Fi tools, such as WMM Access Categories, cannot mitigate the impact these clients have on network performance. A more sophisticated scheme that takes each client's time consumption of the shared air interface into consideration is needed. To provide the expected level of service, service providers must apply specific scheduling algorithms coupled with higher level application logic.

In order to minimize the impact of roaming hotspot clients, it is important to limit not only the number of such associated devices, but to also cap the amount of time they are allowed to use the air interface. One way of controlling Air Time is by allocating time based on SSID. For example, an AP might allocate only 10% of the total Air Time to the clients associated with the home hotspot SSID, while clients attached to the home SSID can enjoy 90% of the Air Time. The differentiation based on SSID, and the type of service that is associated with each SSID, is especially crucial when the service provider wants to deliver high-definition video over Wi-Fi.

In addition, an admission control mechanism for allowing or blocking slow clients (based on low bit-rate MCS) connection to the hotspot SSID is needed. Service providers can incorporate such mechanisms as part of the handshake with the associating client (e.g., RADIUS authentication of the requesting client). The AP is aware of its resources in terms of overall bandwidth and number of active clients. When the wireless media is congested, an AP may decide to disassociate or block a hotspot client from gaining access to the air interface.

In the case of Home Security, where specific, known clients need extra bandwidth, the AP should be configured to use the air interface in a manner that meets their minimum needed bandwidth. These clients may need preferred treatment by the Air Time scheduler depending on the overall network usage at that time. A pre-defined minimum bandwidth configuration may be needed to ensure the video streams these clients send arrive intact and with minimum delay. Different services require different solutions and logic; a dynamic Air Time management scheme should allow the service provider to allocate the Air Time resource between types of services associated with different SSIDs while differentiating between clients associated with the same SSID.

References

(1) Wi-Fi Multimedia™ Technical Specification Version 1.0 Wi-Fi Alliance

(2) Managing Cable TV Migration to IP – Part 1 Series Introduction: Market Drivers and Tech Challenges, S. Howe C. Cheevers, ARRIS

UltraHD, HEVC and Higher Fidelity Video
Why It's Not Just Pixel Density Anymore

Indra Laksono, Cindy Guo
Vixs Systems

Abstract

We present the advantages of 10-bit video over the traditional 8-bit video compression. Results on both objective and subjective video quality are discussed along with some highlights of implementation issues for anyone looking at a IC implementation of a 10-bit HEVC decoder. 10-bit video compression can achieve higher coding efficiency while delivering visible quality improvement.

INTRODUCTION

The deployment of digital video is facilitated by the development of modern video coding standards such as MPEG-2, H.264/MPEG-4 Advanced Video Coding (AVC), and mostly recently, High Efficiency Video Coding (HEVC) otherwise known as H.265. Each generation of adopted video coding standards has offered a target bit rate savings of at least 50%. This enables broadcast industry to deliver video distribution of resolutions from 720x480 or 720x576 (SD), to the current standard of 1920x1080 (HD), and now to 3840x2160 (UltraHD).

As pixel density increases from SD to HD/UltraHD, the basic unit of video is still YUV 4:2:0, with chroma channel subsampled, and each colour components are 8 bits per sample. This so-called "True Colour" system can support more than 16 million colours. While some study suggests that human vision system can only discriminate up to 10 million colours [1], banding artifacts (also known as contouring) can be visible in areas with low chrominance variation as shown in Fig. 1 [2]. The prominence of banding artifacts in distributed video today suggests that the dynamic range of colour representation chosen today in distributed video still has gaps that are discernible to the human eye. Furthermore, with increase of display sizes and with display's wider colour gamut and higher dynamic range, visual artifacts of 8-bit video are exposed more easily.

The limitations of 8-bit video were recognized by the International Telecommunication Union (ITU). In August 2012, ITU released its recommendation for UltraHD TV, commonly known as ITU-R BT.2020. This recommendation is not just a change in the colour space conversion matrix coefficients; it specifies that the bit depth of each colour component will be 10 bits or 12 bits. Taking their cue from ITU, the Joint Collaborative Team on Video Coding (JVT-VC) added a consumer oriented 10-bit profile, named as Main10, in October 2012 [3]. This is notable as this is the first time a major video codec standard, targeted for wide consumer adoption has formally allowed a bit depth higher than 8 bits in the first ratified release.

This paper will discuss the various aspects of 10-bit digital video implementations vs. 8-bit, studying visual differences, bandwidth

Fig. 1: A screen shot from the introduction sequence of "House of Cards". Only luma channel is shown here.

requirements, and some IC implementation issues of supporting 10 bits. Moving from 8-bit colour representation to 10 bits will quadruple the dynamic range of each colour component, and provide a 64 times increase in total dynamic range of different colours. This shift can be accomplished with little incremental cost while delivering a visible tangible improvement to the end user. The resulting implementation will move us from 16 million colours to over 1000 million.

The traditional use of 8 bits for colour representation provides us a 24-bit colour space commonly known as True Colour. Due to the limitation of rounding to the component bit depth, mapping between RGB and YUV colour spaces are not generally reversible. For example, given all 24-bit RGB triplets, only 24% of them can be exactly recovered from converting to 8-bit YUV and back to 8-bit RGB using Rec.709. Because of clipping, most RGB triplets are off by ±1. However, increasing bit depth can improve the density of representation of colours and brightness throughout the entire range of colours. When converting between YUV 10-bit and RGB 8-bit, the recovery rate is 100% because of the extra precision. Hence there are always measurable advantages of higher bit depth in video fidelity.

Colour representations of 30 bits and higher are collectively known as Deep Colour in TV parlance and is available widely in 1080p TVs sold today.

HIGHER FIDELITY VIDEO

Moving from 8-bit video compression to 10-bit poses several challenges and the primary concern is on coding efficiency. Coding efficiency is measured by the balance between video quality and bit rate. As 10 bits per pixel (bpp) is 25% more than 8 bpp, to get the same subjective or objective video quality, one might question whether delivering 10 bits video would require more data bandwidth. Another concern is on the implementation. For software encoders and decoders, 8 bpp fits nicely into bytes; whereas 10 bpp requires either data packing or doubling memory requirements. On the ASIC side, there are implementation issues on memory bandwidth usage.

In this section, we examine various technical challenges of 10-bit video processing.

Coding Efficiency

We choose 4 test sequences, which are summarized in Table 1, to test coding efficiency. All sequences are progressive with YUV 4:2:0 colour format and 10 bpp. 8-bit YUVs are generated by rounding up the last two least significant bits (LSB) of the 10-bit sequences. All test materials are coded with HM12.0 reference encoder [4] using the random access configuration but with only 4 B frames (instead of 8) per GOP. Each test sequence is encoded at 4 different bit rates, using quantization parameters (QP) 20, 26, 32 and 38.

Coding efficiency is evaluated subjectively with rate-distortion curves and peak signal-to-noise ratio (PSNR) is chosen as the distortion measure. Fig. 2 shows the rate-distortion curves for the 4 streams, in which Y-PSNR is plotted as a function of the average bit rate. We only consider luma PSNR because luminance channel is the most important one.

Table 1: Test Sequence Used

Sequence	Resolution	Length (frames)	Source
SteamTrain	2560x1600	250	ITU
Netuba	2560x1600	250	ITU
Tears of Steel	1920x1080	250	mango.blender.org/download
Sintel	1920x1080	300	www.sintel.org/download

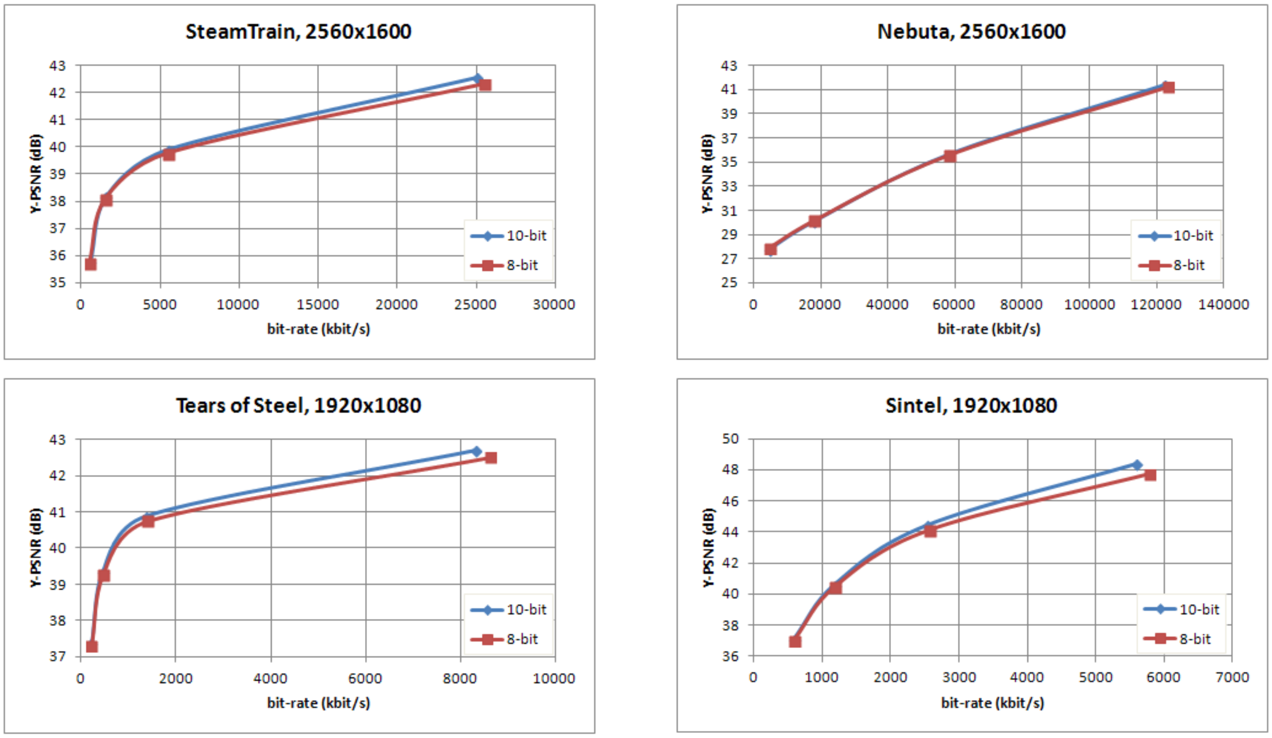


Fig. 2: Rate-distortion plots for test sequences

Results in Fig. 2 indicate that at lower QP and higher bit rate, 10-bit coding provides increased coding efficiency while maintaining the same quality. At higher QP and lower bit rate, the coding performance of 8-bit and 10-bit is similar.

The increased coding efficiency at higher bit rate was also reported in [5]. One of the reasons is that 10-bit pixels are better correlated than 8-bit pixels, and thus can be predicted more easily. This characteristic is well suited for transform based coding techniques such as AVC and HEVC. At low bit rate, on the other hand, the information in both 8-bit and 10-bit sources are heavily reduced and therefore the coding efficiency is about the same.

Video Quality

Evaluating objective video quality is a difficult subject. People often have different preferences when looking at video streams. Moreover, in our experiences, the PSNR difference between the 10-bit and 8-bit streams, which is less than 0.7dB, is not visible to most untrained eyes. Therefore, we judge video quality objectively using frame-by-frame comparison.

Fig. 3 shows the side-by-side comparison of a scene from the sequence “Sintel”. To show the quality of the video one would see on a 10-bit display, the following process is applied to the luma channel in order to generate this figure.

1) From the decoded 10-bit and 8-bit images, a sub frame of size 512x512 pixels is cut from the top left cover.

2) Pixels from the 10-bit image range from 138 to 278. They are normalized so that they map to 0 to 255. The output from this step is the 10-bit image Fig. 3(b).

3) Pixels from the 8-bit image range from 34 to 70. They are also mapped from 0 to 255. The output from this step is the 8-bit image Fig. 3(c).

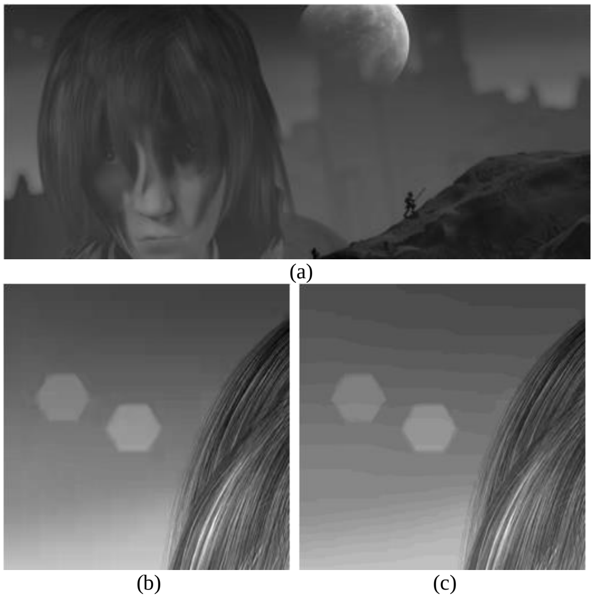


Fig. 3: Side-by-side comparison of 10-bit and 8-bit images: (a) 8-bit original frame, (b) decoded 10-bit image, and (c) decoded 8-bit image.

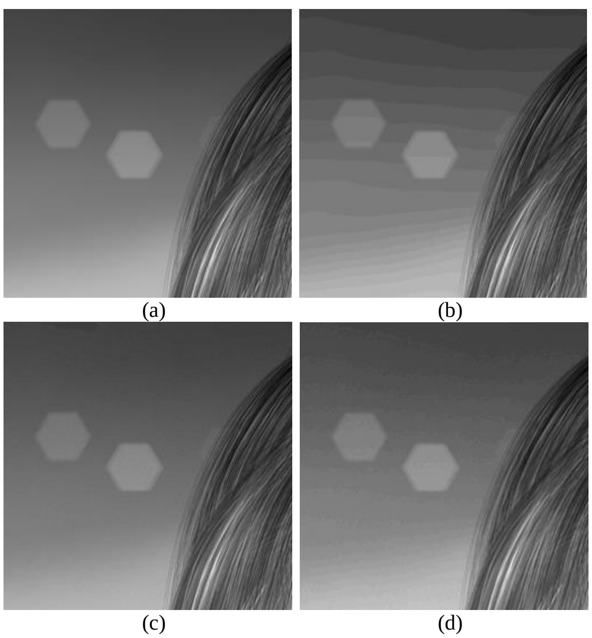


Fig. 4: (a) 10-bit original “Sintel” frame, (b) 8-bit frame generated by rounding the LSBs, (c) 8-bit frame generated using Floyd-Steinberg dithering, (d) lowpass filtered frame in Fig. 4 (c).

One can see that 10-bit compression removes contouring artifact noticeable in an 8-bit video. This is because 10-bit video can provide a higher dynamic range than that of 8 bits, and thus the transition from darker areas to brighter areas can be a lot smoother. This observation is also supported by [5] and [6].

<u>Contouring Countermeasures</u>

It is generally accepted in compression labs that when a 10-bit or higher digital master exists, one way to reduce contouring is to dither the source 10-bit content before down sampling to 8 bits (Fig. 4 (c)). This is the current remedy widely used in the field. However this approach has two drawbacks.

1) The particular choice of the dithering algorithm introduces distortions to the original content and requires more human intervention in the encoding process.

2) In broadcast video, to hit compression targets, we often see noise reduction filters applied to images as part of the compression process. This noise reduction filter tends to undo much of the dithering introduced and would reintroduce contouring (Fig. 4 (d)).

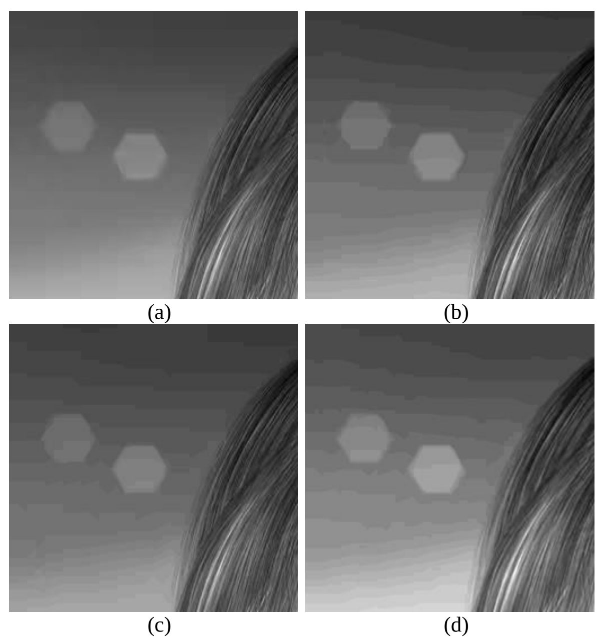
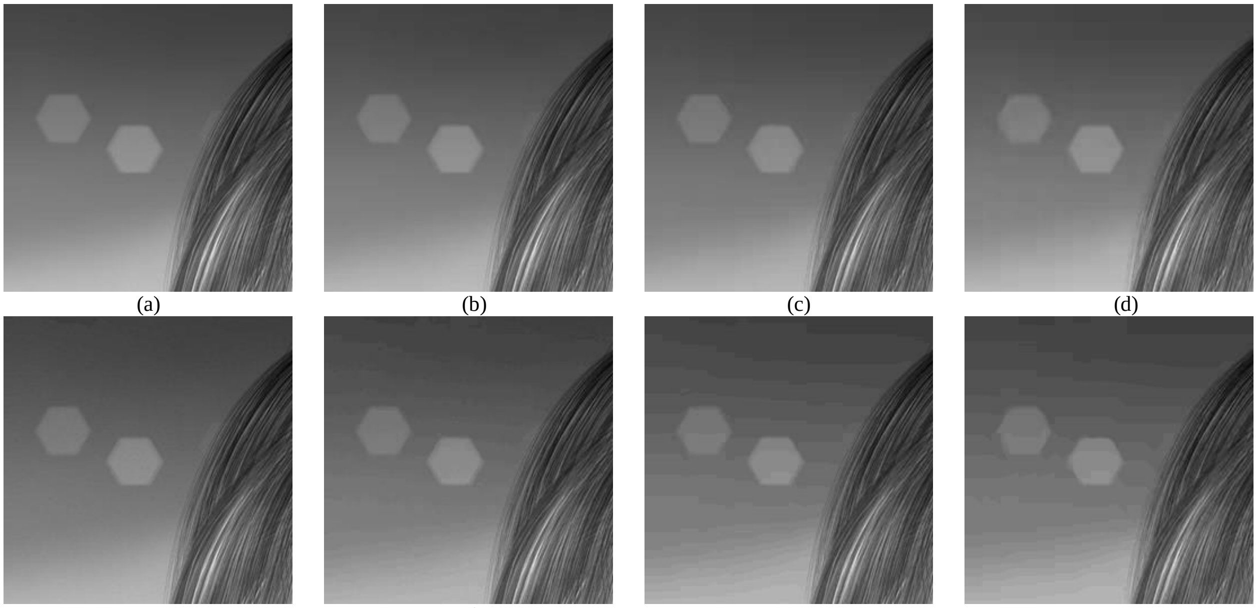


Fig. 5: (a) decoded 10-bit “Sintel” frame, (b) decoded 8-bit frame generated by rounding the LSBs, (c) decoded 8-bit frame generated using Floyd-Steinberg dithering, (d) decoded noise filtered frame in Fig. 4 (c).



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 6: (a) ~ (d) decoded 10-bit “Sintel” frame with QPs 5, 10, 15, and 20, (e) ~ (h) decoded dithered 8-bit “Sintel” frame with QPs 5, 10, 15 and 20.

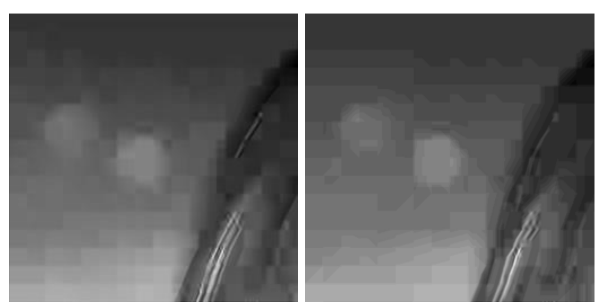
Fig. 4 (a) shows a sub frame of the original 10-bit “Sintel” image. When a simple rounding method is used, banding is visible as seen in Fig. 4 (b). Fig. 4 (c) is an 8-bit frame generated from the original using Floyd-Steinberg dithering. Colour banding is not visible here. However, after noise filtering, traces of contouring reappeared in Fig. 4 (d).

Fig. 5 is the set of decoded frames from the original pictures in Fig. 4. The QP used to encode these frame is 22 for Fig. 5 (a), (b) and (c), which gives the final bit rate of the sequence of about 5.5Mbps. The lowpass filtered picture, Fig. 5 (d), can achieve similar bit rate with a slightly lower QP of 19. It is clear that contouring is visible in all the 8-bit images.

Quantization and Contouring

Rate control in video compression (after considerations of tools used and within the same implementation) is largely determined by the quantization, which we call QP. We further publish our results to show how quantization values behave for both 10-bit and 8-bit video.

Fig. 6 shows the same decoded “Sintel” frame at QPs equal to 5, 10, 15, and 20 for both 10-bit and dithered 8-bit pictures. Banding starts to become visible for 8-bit compression when QP is 10, whereas it remains unnoticeable for 10-bit compression at QP equals 20. Given that a QP increase of 6 roughly halves the bit rate of a stream, this means that no artifacts are seen on the 10-bit sequence even though its bit rate is about a quarter of that of the 8-bit sequence. Even at very low bit rate, 500 Kbit/s and QP equal to 38, the 8-bit sequence still shows some contouring (Fig. 7 (b)). By comparison, the 10-bit sequence mostly exhibits blockiness, which is a common compression artifact at low bit rate (Fig. 7 (a)).



(a) (b)

Fig. 7: (a) decoded 10-bit “Sintel” frame at QP 38, (b) decoded 8-bit “Sintel” frame at QP 38.

Memory Bandwidth Usage

The increase of coding efficiency and better video quality do not come for free. Implementing 10-bit video encoder and decoder has many difficulties. A critical subject in a video encoder or decoder design is memory bandwidth usage. This is because memory access is often the slowest part of a video coder pipeline due to the limit of the memory technology. In designing HEVC, some memory bandwidth considerations were put forward so that HEVC does not require much more bandwidth than AVC, even though HEVC uses an 8-tap interpolation filter for luma pixels [7]. Nevertheless, supporting a simplistic straight forward 10-bit video processing still requires memory bandwidth increase of about 25% over 8-bit (in theory). The caching of pixel blocks to avoid refetching previously fetched pixel data clearly can help, even though some judicious tuning is required to get more optimal cache management to make this work better.

One way to save on memory bandwidth is to reduce memory access when fetching reference pixels in motion compensation (MC). Since reference pixels usually exhibit strong spatial correlation, it is possible to compress the reference pixels before they are written to the frame buffer. Hence, memory bandwidth can be reduced for both memory read and write.

A simple lossless compression algorithm based on JPEG-LS is implemented in order to monitor memory bandwidth usage for an HEVC decoder [8]. MC cache is two-way set-associative with a size of 256x256 pixels.

Fig. 8 shows the bandwidth usage for the "Sintel" stream. Without reference picture compression, 10-bit stream requires about 21% more bandwidth than the 8-bit stream on average. *When compression is enabled, the 10-bit stream only needs about 4% more bandwidth on average than the 8bit stream.*

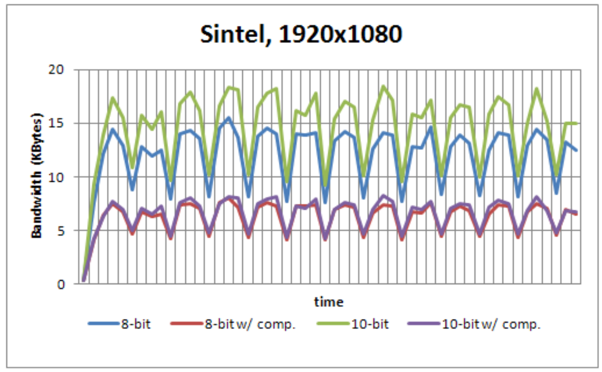


Fig. 8: Memory bandwidth measurement for 8-bit and 10-bit streams with and without reference picture compression

We also note that by turning on reference picture compression alone saves memory bandwidth usage by about 50% for the "Sintel" stream. It shows that this technique has great potential in supporting future codecs and profiles that include 12-bit and 14-bit video and YUV 4:2:2 and 4:4:4 formats.

It is understandable that bandwidth savings with compression is highly dependent on the content of the pictures. In the case of random noise, compression provides no savings at all. However, in most cases, bandwidth savings with compression is significant. We strongly suggest though that normal video for live content and animation are generally known to behave well with modern compression algorithms, since such codecs were targeted at a wealth of such content and improvements were designed and improved upon the successes of behavior with previous algorithms.

Most significant finding here is that we can support lossless 10-bit frame buffer compression with an incremental memory bandwidth increase of 4% over 8-bit frame buffer compression (when the raw traditional approach would require 25% more memory bandwidth). This is clear indication to us that supporting higher dynamic range video should strongly consider frame buffer compression as a key architectural focus of building IC codecs for both encode and decode. The result will carry over to 12bit implementations as well.

CONCLUSIONS

In this paper, we have discussed various aspect of 10-bit video processing. We have shown that comparing to 8-bit video, 10-bit has increased coding efficiency because 10-bit pixels are well suited for transform based video coding techniques. Furthermore, because of its higher dynamic range, 10-bit video can reduce banding or contouring artifacts visible in an 8-bit video, especially in areas with low chrominance variation. We have examined one particular aspect of 10-bit encoder and decoder implementation, namely memory bandwidth usage. By using reference picture compression, we are able to limit bandwidth increase because of 10-bit to about 4%. We believe that with minimal cost increase, broadcasters should be able to deliver 10-bit content and provide visible quality improvement to the consumers.

REFERENCES

[1] D. B. Judd and G. Wyszecki, *Colour in Business, Science and Industry*. Wiley Series in Pure and Applied Optics (3rd ed.). New York: Wiley-Interscience, 1975

[2] http://reddoorreview.com/house-of-cards-review/

[3] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand, *High Efficiency Video Coding (HEVC) Text Specification Draft 10*, document JCTVC-L1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC). Jan. 2013.

[4] K. McCann, B. Bross, W.-J. Han, I. K. Kim, K. Sugimoto, G. J. Sullivan, *High Efficiency Video Coding (HEVC) Test Model 12 (HM 12) Encoder Description*, document JCTVC-N1002, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC). Aug. 2013.

[5] P. Larbier, *Keep Video Quality Pristine throughout the Prodcution Process: 4:2:2 10-bit AVC/H.264 encoding*, http://x264.nl/x264/10bit_03-422_10_bit_pristine_video_quality.pdf

[6] A. Duenas, A. Malamy et al, *On a 10-bit consumer-oriented profile in High Efficiency Video Coding (HEVC)*, document JCTVC-K0109, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC). Oct. 2012.

[7] T. Hellman, W. Wan, *Motion Compensation Restrictions to Alleviate Memory Bandwidth Concerns for High Resolution Video*, document JCTVC-H0441, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC). Jan. 2012.

[8] M. J. Weinberger and G. Seroussi, *The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS*, Proceedings of the IEEE Data Compression Conference, Snowbird, Utah, Apr. 1996.

Unicast or Multicast for IP Video? Yes!

John Horrobin, Yoav Schreiber
Cisco Systems

Abstract

Cable operators are leading the industry with innovative cloud-based video solutions such as TV Everywhere and Cloud DVR. These solutions are expected to play a prominent role as operators aim to deliver more content to more devices and ultimately transition to all-IP with full-scale, managed IP Video services. With current-generation IPTV solutions, operators have the choice of using unicast or multicast to deliver Linear TV services, and multicast has proven to be very efficient in existing IPTV deployments. However, cloud-based video solutions currently rely exclusively on unicast delivery, which can have a dramatic impact on the cable access network as the number of subscribers served from the cloud grows. Thus operators need to carefully consider the network impact of migrating to cloud-based video solutions, and look for ways to optimize the network efficiency.

Using example use cases based on field data and our own assumptions, we illustrate the potential for as much as 80% savings in CMTS capacity and HFC spectrum using multicast delivery for real-time viewing and in-home DVR recording. While moving the DVR functionality to the cloud can increase the network capacity required for time-shifted viewing by 40% in our example, it can also significantly decrease the control plane traffic load on the CMTS during prime time. Even with cDVR solutions in place, we expect multicast will continue to be a valuable tool for cable operators in serving non-cDVR subscribers. Hybrid DVR solutions can offer the best of in-home DVR and cDVR solutions if the business case challenges can be overcome.

Many cable operators have deployed TV Everywhere solutions as an overlay to their existing digital cable systems. As operators migrate to all-IP, there are clear benefits to using a common platform for delivering both managed and unmanaged IP Video services to all devices. The TV Everywhere solutions deployed to date can serve as a foundation for the common platform to support all IP Video services in the future. Since these solutions typically employ ABR streaming, the development and deployment of multicast ABR video transport solutions will enable operators to leverage the benefits of multicast as they migrate to a common infrastructure and client for supporting all IP Video services.

While the optimal solution for delivering full-scale, managed IP Video services will be unique to each operator, it will undoubtedly involve both multicast and unicast delivery approaches.

INTRODUCTION

Cable operators are responding to their subscribers' seemingly insatiable appetite for compelling TV programming on their schedule and platform of choice. The proliferation of Set-top Boxes (STB's) with integrated Digital Video Recorder (DVR) functionality and TV Everywhere services are just two examples of the cable industry's innovations. New solutions such as cloud-based DVR (cDVR) promise to keep the cable industry at the forefront of the highly competitive Service Provider market. However, one must carefully compare the impact of in-home and cloud-based solutions on the network infrastructure when choosing between these disparate approaches, and look for ways to leverage the strengths of each approach when deployed together.

In this paper, we will analyze the network requirements for supporting both real-time and time-shifted viewing of Linear TV programming on cable access networks in three scenarios: non-DVR STB, in-home DVR, and cDVR. Our focus is on full-scale, managed IP Video solutions in which all Linear TV programming is delivered via the DOCSIS access network; however, the methodology and results can be applicable in assessing network requirements for traditional QAM-based digital video solutions as well as hybrid QAM + IP Video solutions.

For the purposes of this paper, Linear TV viewing is considered to be either real-time, or time-shifted. Real-time viewing refers to direct playout of the programming from a STB or other rendering device as it is delivered across the cable network, with no user control beyond "tuning" to the TV channel. Time-shifted viewing refers to user-controlled playout from an in-home or network-based storage device, such as a STB with DVR capability, a cDVR system, or a cable operator's service such as Time Warner Cable's Look Back® and Start Over® services. When properly equipped, a user may shift between real-time viewing and time-shifted viewing of the same programming during the broadcast window. Also, note that time-shifted viewing may occur during the broadcast window, or post-broadcast.

An important distinction between real-time and time-shifted viewing in IP Video solutions is that programming viewed in real-time may be delivered via unicast or multicast, while time-shifted programming must be delivered via unicast. Multicast delivery can result in significant savings in the HFC spectrum and CMTS capacity required to support Linear TV services. We will illustrate the savings of multicast and the network impact of offering time-shifted viewing options via in-home DVR and cDVR solutions.

Analyzing the network impact of Linear TV viewership on customer-owned-and-maintained (COAM) devices such as PC/laptops, tablets, game consoles, smart TV's and smartphones, is outside the scope of this paper. However, we will highlight how multicast ABR video transport can enable operators to migrate to a common infrastructure and client to support both STB's and COAM devices.

NON-DVR STB USE CASE

In this scenario, all Linear TV viewership via STB's is in real-time. The percentage of STB's that are receiving Linear TV at a given time is referred to as the STB concurrency. A number of factors affect STB concurrency, including the time of day, day of week, programming choices and popularity, demographics, and even weather and local events. Cable operators must consider the impact of these and any other relevant factors to predict the STB concurrency in each of their cable systems. In a recent study of field data from a cable operator's network (see [1]), we observed STB concurrency in the range of 50–60% during prime time. This is consistent with some prior studies, and we will assume 60% peak concurrency in our illustrative examples below.

Another key factor in estimating the network impact of Linear TV is the efficiency of multicast delivery. As we explained in [1], multicast gain is the average number of viewers per Linear TV channel (i.e. per-channel concurrency) within a service group (SG). The multicast gain that can be achieved depends on a variety of factors, most importantly the number of viewers per

SG, the number of TV channels offered, and the popularity of the programming. In our examples below, we will assume a multicast gain of 5, which is consistent with the gain reported in [1] for service groups of comparable size.

Table 1 provides an illustrative example of the methodology for estimating the HFC spectrum and CMTS downstream (DS) channel capacity required to support multicast delivery of Linear TV services to STB's. This example assumes all Linear TV programming delivered to STB's is encoded as HDTV with a constant bit rate of 8 Mbps. Based on our assumptions, a cable operator would need to allocate 13.7 CMTS DS channels and associated HFC spectrum for each SG to support Linear TV services (assuming an all-switched approach…refer to [1] for more information on all-switched and broadcast approaches).

Parameter	Value
Homes passed per Service Group	500
Cable TV Take Rate	40%
Cable TV Subscribers per SG	200
STB's per Subscriber	2.5
STB's per SG	500
Linear TV STB Concurrency	60%
Linear TV Viewers per SG	300
Multicast Gain	5
Unique Linear TV Streams per SG	60
Linear TV Stream Bit Rate (Mbps)	8
CMTS DS Channel Capacity (Mbps)	35
CMTS DS Channels per SG	13.7

Table 1. Non-DVR STB Use Case

IN-HOME DVR USE CASE

According to the latest Nielsen report on TV viewership [2], roughly 50% of TV households have some type of DVR device. The DVR functionality may be integrated in the STB provided by the service provider, or may be a stand-alone device connected to a STB (or gateway). In either case, the DVR records Linear TV programming received from the network according to the subscriber's input.

From the cable network perspective, the DVR recording appears as real-time viewership. In an IPTV system utilizing IP multicast for delivering the Linear TV programming, the DVR client sends an IGMP (or MLD) Join message to receive the Linear TV program(s) the subscriber has scheduled to be recorded. This is indistinguishable from an IGMP/MLD Join sent by an IPTV STB when a subscriber "tunes" to the same Linear TV program. If the DVR and STB are on the same DOCSIS bonding group, they will receive the same multicast video flow (i.e. the CMTS will send one copy of the Linear TV program on the DOCSIS bonding group).

If the Linear TV programming being recorded by DVR's is the same programming being viewed in real-time by subscribers, then no additional network capacity is required to support the DVR's. In this case, DVR recording increases the multicast gain, since more devices receive a given multicast video flow. However, if the programming being recorded is not the same as that being viewed in real-time, then DVR recording will increase the amount of network capacity required to support the Linear TV services. Although we do not have empirical evidence on the specific programming being recorded during prime time, our presumption is that subscribers with DVR's predominantly record the popular programming that is also viewed in real-time by subscribers without DVR's, and thus there is no material impact on cable network capacity for DVR recording during prime time.

According to a recent study on DVR usage [3], 120% of DVR devices record Linear TV programming during prime time. Exceeding 100% concurrency is possible since most

DVR devices are able to record multiple programs simultaneously. This high level of concurrency is indicative of the tendency for DVR users to record a large volume of content (ED: how full is your DVR?) so they have a wide variety of recorded programming to choose from whenever they watch TV.

Using the DVR penetration and recording concurrency data cited above, we can estimate the number of DVR recordings to be expected during prime time, and the resulting multicast gain, in our example use case. See Table 2 for the methodology and results.

Parameter	Value
Cable TV Subscribers per SG*	200
DVR Take Rate	50%
DVR's per SG	100
DVR Recording Concurrency	120%
DVR Recordings per SG	120
Linear TV Viewers per SG (STB's)*	300
Linear TV Viewers + Recordings per SG	420
Unique Linear TV Streams per SG*	60
Multicast Gain	7

Table 2. In-home DVR Use Case

*See Table 1 for derivation

Assuming the DVR's in our example use case are recording the same Linear TV programming that is being viewed in real-time by subscribers without DVR's, the same number of unique Linear TV streams are required per SG as in the non-DVR STB use case. By dividing the total number of real-time viewers and DVR recordings by the number of unique streams per SG, we can calculate the multicast gain for the in-home DVR case. As shown in Table 2, the multicast gain is 7 when serving both STB's and DVR's, which represents a 40% increase in multicast efficiency compared to the non-DVR STB use case. Figure 1 depicts how in-home DVR recording increases multicast efficiency since more end points are served with each multicast video flow.

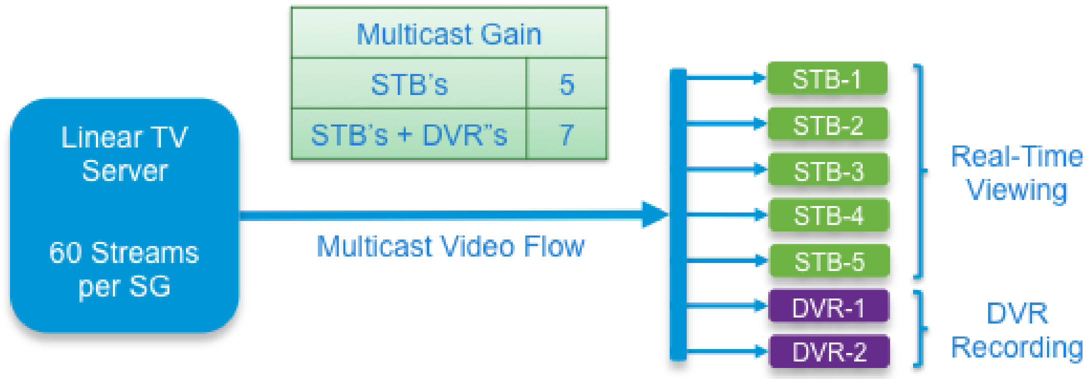


Figure 1: Multicast Gain with STB's and In-Home DVR's

Our example also illustrates the impact that DVR recording can have on Linear TV concurrency. If we assume all DVR's are separate from the STB's, then the overall concurrency of STB+DVR devices in the in-home DVR use case is 70% (420 total viewers + recordings / 600 STB's + DVR's), compared to 60% in the non-DVR STB case. If all DVR's are integrated with the STB, then the overall concurrency is 84% (420 total viewers + recordings / 500 STB's). These examples illustrate that in-home DVR recording can be a significant contributor to the overall Linear TV concurrency.

Although in-home DVR recording may not impact the network capacity required for Linear TV, it does increase the control plane traffic (i.e. IGMP/MLD and DBC messaging) that the CMTS must support, particularly during prime time when a high percentage of DVR's are scheduled to record at the top of the hour. Referring to our example use case, if 10% of the Linear TV viewers using STB's change the channel at the top of the hour, then the CMTS will receive 30 IGMP/MLD Join messages simultaneously. If half of the DVR's that are scheduled to record at the top of the hour initiate a channel change, the CMTS could receive an additional 60 IGMP/MLD Join messages, representing a 200% increase. Clearly the synchronized timing of channel change requests from pre-programmed DVR's can dramatically increase the control plane traffic load on the CMTS during prime time.

CLOUD DVR USE CASE

Cloud DVR solutions move the DVR recording and playout functions to the network, and enable remote access from clients on a variety of platforms. With cDVR, all recording of Linear TV content takes place in the cloud, and thus requires no capacity on the cable access network. However, unlike the in-home DVR case in which all time-shifted viewing is served from the DVR and thus has no impact on the cable network, all time-shifted viewing in the cDVR case is supported from the cloud, and thus requires capacity on the cable access network. Since the cDVR solution utilizes unicast delivery (i.e. a unique video stream per user), the amount of network capacity required is directly proportional to the number of cDVR subscribers viewing time-shifted content at any given time.

In order to estimate the impact of cDVR on the cable network, we refer again to the recent study of DVR usage described in [3]. According to that study, 20% of DVR's are utilized for time-shifted viewing during prime time (not including stream control of Live TV programming). We assume the take rate for the cDVR service is the same as DVR service, since subscribers should not care if the DVR functionality is supported in the home or the cloud as long as the user experience is equivalent. However, given that cDVR service is accessible from any STB in the home (which is not the case with an in-home DVR unless it is a whole-home DVR), we will assume the concurrency of time-shifted viewing among cDVR subscribers is slightly higher than in-home DVR subscribers. Finally, since we are focusing on Linear TV viewership on STB's, we will assume the same encode rate for cDVR service as we did for real-time viewing and DVR recording (8 Mbps). As shown in Table 3, we estimate that 5.7 CMTS DS channels and HFC spectrum is required to support the time-shifted viewing from a cDVR solution in our example use case.

Parameter	Value
Cable TV Subscribers per SG*	200
cDVR Take Rate	50%
cDVR Subscribers per SG	100
cDVR Time-Shifted Viewing Concurrency	25%
cDVR Time-Shifted Viewings per SG	25
Linear TV Stream Bit Rate (Mbps)	8
CMTS DS Channel Capacity (Mbps)	35
CMTS DS Channels per SG	5.7

Table 3. cDVR Use Case

*See Table 1 for derivation

In addition to the time-shifted viewing estimated in Table 3, there is also real-time viewing by non-cDVR subscribers (and cDVR subscribers who choose real-time viewing). Assuming that time-shifted viewing does not materially impact the STB concurrency or multicast gain, we can estimate that 13.7 DS channels are required for real-time viewing (per Table 1). Thus a total of 19.4 CMTS DS channels are required per SG for Linear TV service (13.7 for real-time viewing plus 5.7 for time-shifted viewing). Thus, based on our example use cases, cDVR solutions can increase the network capacity and HFC spectrum required for Linear TV services by roughly 40% compared to in-home DVR solutions.

Since the cDVR solution performs all recording in the cloud, the CMTS does not receive control plane traffic associated with channel changes for DVR recordings. Given the propensity of DVR users to record a large volume of content, especially during prime time, we expect cDVR solutions to offload a significant amount of control plane traffic from the CMTS. While actual results will depend on a number of factors, our illustrative example described previously indicates a 67% reduction in control plane traffic load on the CMTS if the DVR recording function is moved to the cloud.

HYBRID DVR USE CASE

The primary benefit of cDVR solutions when compared to in-home DVR solutions is the avoidance of costly DVR functionality and storage in the STB and the higher maintenance costs of DVR STB's compared to non-DVR STB's. With that in mind, it seems counter-intuitive to consider a "hybrid" DVR use case in which operators deploy both in-home DVR and cDVR solutions. However, two scenarios may merit consideration for such an approach. Both scenarios rely on the assumption that the popular Linear TV programming will be delivered via multicast for real-time viewing by non-DVR subscribers (or by DVR subscribers that choose real-time viewing in some instances). If that is the case, then having an in-home DVR capability that can provide the stream control functionality (for a limited duration) could mitigate the scenario in which a large percentage of subscribers simultaneously access stream control from a cDVR solution, resulting in a large spike in cDVR traffic on the cable network. This in-home DVR capability could be limited to stream control only, with a relatively small cache, and all recording and storage supported by the cDVR solution.

The second scenario is the ability to push the popular Linear TV content to in-home DVR devices with a moderately sized cache, and avoid the cost of storing the popular Linear TV programming in the cloud (i.e. licensing fees and storage capacity). In addition, recording the most popular programming in-home has virtually no impact to the cable network, yet would significantly reduce the network capacity required to support time-shifted viewing of the popular programming from the cloud.

If the hybrid DVR approach is not feasible, a "pseudo-hybrid" solution can be envisioned in which the cDVR service falls back to multicast if the cable network is congested and cannot provide sufficient capacity to support the stream control features from the cDVR solution.

MULTICAST ABR VIDEO TRANSPORT

Many cable operators have deployed TV Everywhere solutions to enable their subscribers to access Linear TV and On-demand services on a variety of customer-owned-and-maintained (COAM) devices such as PC/laptops, tablets, game consoles, smart TV's and smartphones. In most cases, the TV Everywhere platform has been deployed as an overlay to the existing digital cable platform, and designed to deliver unmanaged video services to COAM devices. As operators migrate to all-IP, there are clear benefits to using a common platform for delivering both managed and unmanaged IP Video services to operator-owned STB's as well as COAM devices. The TV Everywhere solutions deployed to date can serve as a foundation for the common platform to support all IP Video services in the future.

The TV Everywhere solutions typically employ Adaptive Bit Rate (ABR) video streaming, which dynamically adapts to network conditions to deliver the best possible user experience without provisioning guaranteed quality of service. The ABR video streams are delivered via IP unicast with a bi-directional TCP/IP connection between the ABR client and server so the client can continually monitor the network throughput and request the appropriate ABR video profile from the server as the throughput fluctuates. As we described in [1], using unicast delivery for all Linear TV services in a full-scale IP Video system is not feasible due to the exorbitant CMTS capacity and HFC spectrum required. Hence, an effort is underway within the cable industry to define and develop a multicast ABR video transport solution to enable operators to utilize multicast delivery for Linear TV services with an ABR-based IP Video system.

To describe the operation of a multicast ABR video transport system, let us first review the operation of multicast delivery in existing IPTV deployments, as depicted in Figure 2. A multicast server receives the Linear TV programming from the source, and outputs a multicast video flow labeled with a unique multicast group address. The multicast client residing on the IPTV STB, which has a priori knowledge of the multicast group address for each Linear TV program available to the user, initiates IGMP or MLD Join/Leave messages to receive the multicast video flow based on user input (or DVR recording activity if so equipped). The CMTS processes the IGMP/MLD messages in accordance with the DOCSIS 3.0 specification, sends PIM messages to receive the specified multicast video flow from the multicast server, and replicates the flow to the downstream interface on which the cable modem/gateway is connected. The CMTS then sends DOCSIS 3.0 (D30) management messages to the cable modem/gateway, instructing it to forward the multicast video flow to the IPTV STB.

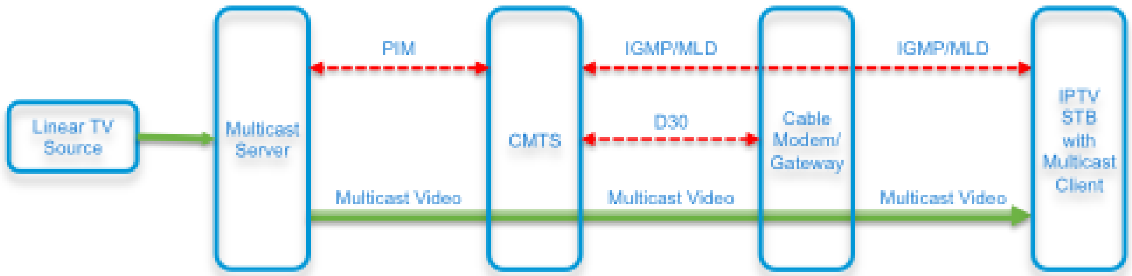


Figure 2: Multicast IPTV System

In a multicast ABR video transport system, an ABR client residing on the IPTV STB or COAM device sends HTTP GET messages to the ABR video server to request a specific Linear TV program (following standard ABR video system operation). However, in this case, a client on the cable modem/gateway intercepts the HTTP GET messages and, with a priori knowledge of the programming available from the multicast server, sends IGMP or MLD Join/Leave messages to receive the requisite multicast ABR video flow from the multicast server. The multicast server is specially equipped to fetch the Linear TV stream from the ABR video server (using HTTP) based on the IGMP/MLD messages received from the client, and encapsulate the ABR video streams into multicast flows. The CMTS follows the same process described above to receive the multicast ABR flows from the multicast server and forward the flow to the cable modem/gateway. The modem/gateway then converts the multicast flow back to unicast for delivery to the IPTV STB or COAM device.

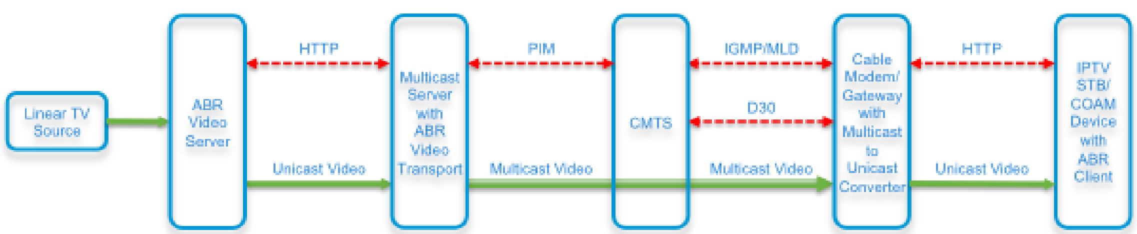


Figure 3: Multicast ABR Video Transport System

The use of multicast delivery in the cable network is transparent to the ABR client and server. If the requested Linear TV program is not available from the multicast server, the system falls back to standard ABR video system operation in which the modem/gateway client simply forwards the HTTP GET from the ABR client, and the

Linear TV service is delivered via unicast. The multicast ABR video transport system is also compatible with in-home DVR and cDVR solutions. Hybrid DVR functionality can also be supported with local caching on the gateway to support stream control, for example.

When deploying a multicast ABR video transport system, cable operators will need to select which Linear TV programs, and which ABR profile(s) for each program, to make available from the multicast server. These choices will affect the multicast efficiency that can be achieved. If the popular Linear TV programming is available from the multicast server, and all STB's and DVR's access the same ABR profile, then the multicast efficiency can match that of standard IPTV multicast systems. In fact, it is possible to achieve even higher efficiency if a multicast flow can serve not only STB's and DVR's, but also COAM devices.

Although not shown in Figure 3, an optional enhancement to the multicast ABR video transport system is the use of a reliable multicast transport protocol. This would enable the multicast server to be notified of lost packets, and re-transmit those packets (within a re-transmit window) in order to improve the quality of experience.

CONCLUSION

Using example use cases, we illustrated a few key takeaways for cable operators to consider in assessing the network requirements for delivering full-scale, managed IP Video services:

1. Multicast is essential for delivering popular Linear TV programming in prime time. Multicast efficiency can reduce the CMTS capacity and HFC spectrum required to support Linear TV services by as much as 80% in our example.
2. In-home DVR scenarios benefit the most from multicast; however, multicast is still a valuable tool in cDVR scenarios where real-time viewing is still prevalent (among non-cDVR subscribers and/or cDVR subscribers who choose real-time viewing in some instances).
3. cDVR increases the network capacity and HFC spectrum required for Linear TV services by 40% in our example; on the other hand, cDVR reduces the multicast control plane traffic load on the CMTS by 67% in our example.
4. Hybrid DVR solutions can enable operators to leverage multicast to record the most popular Linear TV programming in-home with virtually no impact to the network, and record all other programming in the cloud. Hybrid DVR solutions can also avoid the potential for large spikes in cDVR traffic associated with stream control during prime time.
5. Multicast ABR enables operators to continue to leverage the benefits of multicast for Linear TV services as they migrate to a common infrastructure and client for supporting all IP video services across a variety of STB and COAM devices.

The examples provided in this paper are based on field data where available, and on our own assumptions where field data is not available, and thus are for illustrative purposes only. Cable operators are encouraged to adapt the methodology described herein to their own particular environment. While the optimal solution for delivering full-scale, managed IP Video services will be unique to each operator, it will undoubtedly involve both multicast and unicast delivery approaches.

REFERENCES

1. John Horrobin and Gitesh Shah, "Pioneering IPTV in Cable Networks", SCTE Cable-Tec Expo, October, 2013

2. The Nielsen Company, "An Era of Growth: The Cross-Platform Report", March, 2014

3. Carol Ansley and John Ulm, "The Dawn of Cloud-based DVR Services", SCTE Cable-Tec Expo, October, 2013

USING SDN AND NFV FOR INCREASING FEATURE VELOCITY IN A MULTI-VENDOR NETWORK

Jeff Finkelstein, Cox Communications
Alon Bernstein, Cisco
Samir Parikh, Gainspeed

Abstract

As shiny objects of the past decade go, few have achieved the shininess of Software Defined Networking (SDN) and Network Functions Virtualization (NFV). You cannot open a trade publication either soft or hard copy, without seeing these TLAs (three letter acronyms). Every conference has topics related to them. In fact, there are even entire conferences dedicated to these technologies with presenters and vendors discussing what it all means to potential customers and the communication industry at large.

This paper is on a similar path, but more specific to cable and how operators (MSOs) may take advantage of these technologies to achieve two key objectives in their quest to improve services. First, MSOs need to more efficiently scale the network on many levels so as to control OPEX and CAPEX, as well as to fit the necessary equipment into existing facilities. Not only does the equipment need to scale, but so too does the support infrastructure. Power, heating, cooling, physical space, cabling and the like all have to reach a scale that was unimaginable just a few years ago.

The second objective MSO's must accomplish relates to the velocity of service enablement, creation, management and provisioning. The operator community is very good at whiteboarding ideas, producing slides showing how these new services will make things better, and talking about them in meetings and conferences. Where they frequently miss is the speed at which they actually create a product and bring it to market so that it adds revenue to their bottom line. If an MSO could get new products and services out more quickly, and at a lower cost, not only would they be more competitive on a technical level, but they could open new ways of getting ahead of our competition.

The challenge is in finding ways to meet both of these objectives — scaling more efficiently and increasing the velocity of service deployments — without increasing complexity. By using SDN and NFV to drive this innovation, operators can actually simplify operations by decomposing network elements and functions and spreading them across multiple devices. Rather than have a large monolithic architecture we distribute the work across smaller pieces of equipment using simple interfaces to communicate between them. As the industry has learned over time, simplification of technologies helps drive cost out.

In our paper, we will present descriptions and use cases of how MSOs may use SDN and NFV in the headend and at the customer premises to achieve these goals.

INTRODUCTION TO SDN AND NFV

A brief introduction is needed to set the stage for the rest of this paper. SDN and NFV can take on different meanings depending on context, so let's start with the basics.

What is SDN?

Historically, the network and the applications have been managed as discrete entities with no direct interaction outside of someone changing network paths or elements as awareness of an application's needs became apparent. Software Defined Networking (SDN) integrally links these two components. Simply put, the key concept behind Software Defined Networking (SDN) is that the network becomes aware of applications and that applications become aware of the network.

There have been a variety of functional definitions of SDN. A couple that have been more commonly used by cable operators and equipment vendors are:

- Separation of control and forwarding functions
- Centralization of control and distribution of processing/forwarding

The importance of SDN was recognized as the use of cloud services grew and their value was better understood, it became clear that the data-center needed to evolve to effectively support these services. The industry realized the need for a method to connect and control the resources (virtual machines, network capacity, storage capacity) making up the cloud in parallel with the applications running on the network. Specifically the network and cloud resources needed to be able to quickly and reflexively react to which applications were being used and how they were being used. Simultaneously, applications needed to be able to communicate what they needed from the network and react to changes in available resources. All of this needed to happen automatically and in real-time SDN had an enormous impact on how we envision and build data-centers in this era of cloud computing.

What is NFV?

Network Functions Virtualization (NFV) was created to leverage standard IT virtualization to consolidate many different network elements into a cloud architecture that uses standards commercial-off-the-shelf (COTS) hardware. This hardware runs many functions which typically have been handled by purpose built platforms. It is important to note that for some functions, particularly those that involve processing packets at line rate, purpose-built platforms remain a best practice. However, placing other functions on a virtual machine (VM) on COTS hardware can accelerate service deployment and optimize equipment spend.

How do SDN and NFV work together?

SDN and NFV are complementary technologies aimed at streamlining the operation of the network and deployment of new services. They can be view as two pillars in the evolution of managed services networks. NFV-enabled applications interact with an SDN-enabled network infrastructure. NFV streamlines the deployment and scaling of new applications in a cloud environment. SDN streamlines the deployment and scaling of traditional network resources. While they can be deployed together to achieve maximum efficiency, they are independent technologies that can exist without each other.

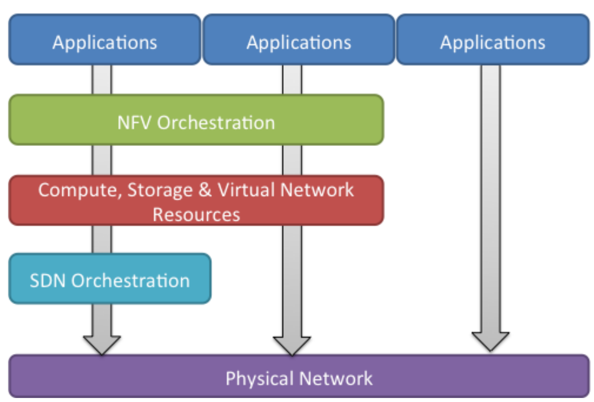
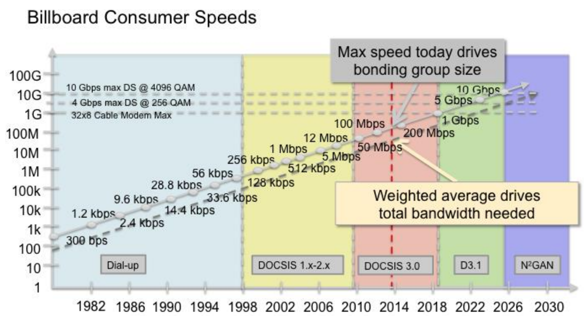


Figure 1: SDN and NFV Layers

Why SDN and NFV?

Graphs showing the increased bandwidth required to meet ever-growing consumer demand are commonplace. In meeting after meeting, MSOs debate why the growth curve cannot possibly continue at the historical CAGR. Yet, it does continue along the existing line and shows no sign of abating.



Graph 1: Bandwidth Growth Curve

Accepting that this growth will continue along its present curve, the challenge of keeping pace with bandwidth requirements relative to operators' CAPEX and OPEX spends becomes evident. To keep up with the ever-increasing bandwidth demands, MSOs must replace current hardware with higher capacity equipment and change the plant to support higher order modulations. This in turn allows operators to compete more efficiently and to replace customer premise equipment with next-generation equipment to deliver a new and improved user experience.

In the past, MSOs would have had an opportunity to add capacity and then reap the benefits, realizing a higher profit margin thanks to capitalization of technology assets. However, the useable lifetime of network assets is decreasing rapidly as we approach the asymptote of the capacity growth curve. Capacity requirements now are simply growing at too great a rate for operators to get ahead of the curve in terms of CAPEX and OPEX.

MSOs must find a new way to increase the velocity of service enablement in order to meet or get ahead of their customers' technology curve. At the same time, operators need to manage the scale of new equipment needed to provide those services while reducing equipment density to stay within the power, heating, cooling, and space availability of existing headend and hub facilities.

The advancements of SDN and NFV provide MSOs with a new paradigm for scaling services and support infrastructures through decomposition of functions.

In the following sections, we will discuss how SDN can enable Network Virtualization and how NFV can enable CPE Virtualization.

SDN AND NETWORK VIRTUALIZATION

SDN can take multiple forms in an operator network. This section will focus on how it enables operators to virtualize the headend.

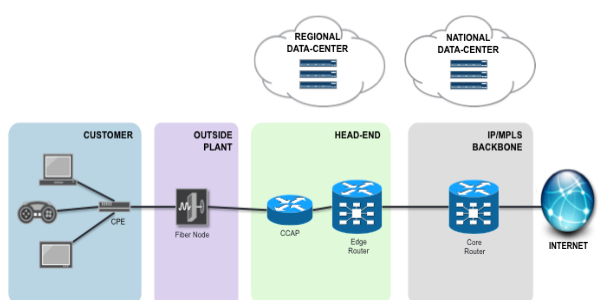


Figure 2: Typical MSO Network Today

As operators scale services, the headend infrastructure (power, cooling, rack space) is coming under increased pressure. The existing architecture of separate video (EQAM) and data (CMTS) platforms, coupled with the combining and splitting networks, occupies a tremendous and growing amount of space and consumes an enormous amount of power. Each new generation of products has provided greater efficiency, but not enough to keep up with demand. Additionally, keeping pace with demand requires operators to churn hardware without maximizing its usable lifespan.

Converged Cable Access Platforms (CCAP), which combines EQAM and CMTS in a single box, are supposed to alleviate this problem. However, evaluating available CCAP products suggests that this is just the next step on the same technology curve.

As previously discussed, MSOs must find a new way to keep up with bandwidth demands on the network, minimize power, cooling and space requirements and prolong equipment's usable life in the network

The key is to build upon CCAP, adding SDN technologies to enhance it, and thereby creating a Virtual CCAP.

This approach allows CCAP to be viewed not as a platform, but rather a collection of base functions that support the services offered by cable operators. These base functions include:

- Cable Control Plane
- IP/MPLS Control and Forwarding Plane
- Subscriber Management
- Video (QAM) processing
- DOCSIS processing
- RF modulation

SDN technologies enable MSOs to move away from the notion that all of these functions must be collocated in a monolithic system centralized in the headend.

In addition, virtualizing these base functions enables operators to leverage best-of-breed products for each function, thereby creating an eco-system which will provide the most feature-rich, reliable and yet lowest cost solution.

SDN technologies allow creativity with regards to where in the network and how these functions are implemented. This paper examines the Virtual CCAP architecture that enables operators to collapse layers/functions without changing their overall network operational model.

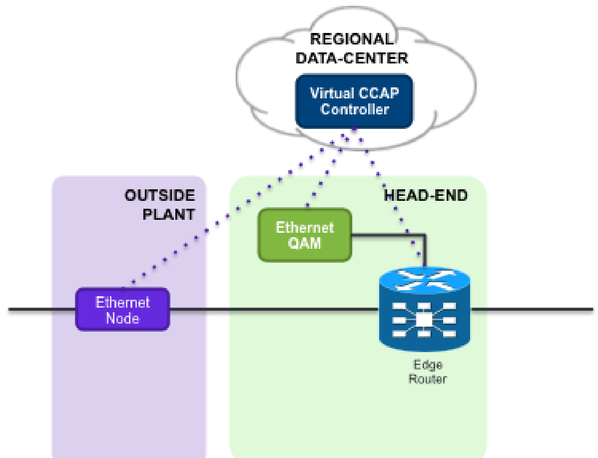


Figure 2: Virtual CCAP Architecture

No changes are required to the CPE in the home, the services delivered to the end customer, or the back-office systems (provisioning, billing, etc.).

Virtual CCAP Components

The Virtual CCAP is composed of the following components:

1. Virtual CCAP Controller: This is a software application in the datacenter that runs the Cable Control Plane and interfaces with existing back-office systems. It is the brains of the Virtual CCAP and is responsible for all control and management of the system. It orchestrates the entire solution, tying all of the pieces together to effectively operate as a single CCAP device.

2. Edge Router: This is the router that is already deployed in the headend. It simply takes on subscriber management functions on top of the functions it already performs (i.e., IP/MPLS control and forwarding plane)

3. Ethernet QAM: This is a new category of edge QAM devices that output a fully groomed MPTS for digital video as a multicast stream on an Ethernet interface as opposed to an RF port.

4. Ethernet Node: This is a new category of fiber nodes that handles all DOCSIS and RF processing. This allows analog transport to be replaced with standard Ethernet.

Benefits of NFV- and SDN-Based Architecture

The benefits of adopting a Virtual CCAP architecture can be broadly categorized as CAPEX and OPEX reduction.

CAPEX Reduction

Operators can achieve significant CAPEX savings with a Virtual CCAP architecture. Savings come from a number of fronts:

1. Replacing legacy analog optical transport with standards based Ethernet transport: 10 Gigabit Ethernet is the lowest cost per bit transport solution available today. It benefits from massive economies of scale from use in data centers and telecommunications infrastructure across wireless, wireline and cable operators.

2. Eliminating the physical CCAP system and leveraging existing headend equipment: Equipment reduction and use of existing gear_lowers the cost of the Virtual CCAP solution.

3. Increasing overall system scalability: The scalability of the solution (number of Ethernet Nodes supported per edge Router, full spectrum agility of the Ethernet Node) provides a much lower cost point as operators add new service group (SG) to the network and scale DOCSIS services on new and existing SGs.

OPEX Reduction

1. Eliminating manual tuning of RF and physical RF combining and splitting: By distributing RF generation to the node and removing it from the headend, spectrum allocation changes can be made quickly and cheaply in software. The manual steps required today to balance RF power levels from the headend towards the transmit laser and in the node between the optical receiver and coax are eliminated. This saves hours and hours of planning and implementation. In addition, physical RF combing and splitting is eliminated and replaced by a "digital" combiner that is controlled entirely by software.

2. Reducing power, space and cooling requirements: By eliminating the physical CCAP, the power and space requirements in the headend are dramatically reduced. This has a follow-on benefit of lowering the cooling capacity required, further increasing the OPEX savings. Virtualizing the physical CCAP also eliminates the need to augment facilities as the network grows.

3. Optimizing use of plant: By pushing DOCSIS (MAC and PHY) processing out of the headend into the node, the need for a tight timing relationship between the headend and node is eliminated. Furthermore, the use of Ethernet transport makes it cost-effective to have longer fiber paths between the headend

and the node. Add to that the reduced footprint that results from virtualization and operators can now consolidate smaller hub facilities into larger, central headends.

4. Leveraging existing management systems: The Virtual CCAP is managed just like a physical CCAP. Control and management flows through the centralized Virtual CCAP controller which now manages up to 400 SGs as opposed to the 64 SGs of a traditional CCAP system.

5. Improving service quality and customer satisfaction: Digital transport to the node increases SNR, improving service performance. This will lower the number of trouble tickets and truck-rolls, resulting in significant savings on an ongoing basis and increased customer satisfaction

Virtual CCAP enables operators to accelerate service velocity and keep up with customer bandwidth demands while concurrently lowering CAPEX and OPEX, all without impacting existing operational systems and processes.

Once this foundation is in place, new and interesting applications can be rapidly delivered to the customer premises by leveraging NFV technologies.

CPE VIRTUALIZATION

In addition to applying SDN and NFV to the headend, an operator can apply these technologies to virtualize the CPE. Virtualizing CPE enables certain capabilities, traditionally implemented in the home to be implemented in the operator's cloud infrastructure.

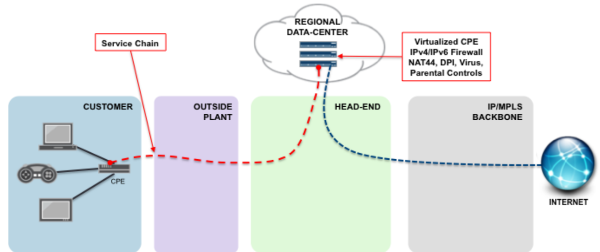


Figure 3: Virtual CPE

Virtualization Candidates

CPE functions can be categorized under the following building blocks:

Networking

1. Layer 1 Physical access media data/control

2. Layer 2/3 networking functions such as routing

3. Upper layer functions such as Network address translation (NAT), firewall, and Deep Packet Inspection (DPI).

It is possible to "virtualize" the control plane or the packet-by-packet handling of the data plane for each of these functions.

Home Appliances

Some of the appliances within the home can be virtualized as well. An example already being deployed in some cable networks is the network video recorder. A household's recorded TV programs would be stored in the cloud instead of on a hard drive in the home. The same way a home backup system can be moved to the cloud. Another example would be the rendering of TV guides, games and other graphic applications in the cloud instead of the home device.

Moving to the Cloud

Which of these functions should be moved away from the home and into a virtualized cloud service must be determined on a case-by-case basis by performing a cost-benefit analysis.

This section will explore the value of CPE virtualization for a couple of specific use cases.

Parental Controls

Parental control refers to the ability to block access to certain URLs. Performing the task of comparing URLs against a large black list is fairly CPU-intensive and thus well-suited for an NFV-type of application rather than standard packet processing.

A virtualized version of parental control would perform the URL filtering in the cloud instead of running it on individual computers or in the home gateway. Updating the list of blocked sites is easier to manage centrally.

Firewalls

Firewalls are network security appliances meant to protect the home network from external security threats. A virtualized firewall would run in the cloud, inspect all packet streams and block those that are a security risk. As with the parental control example, managing the list of security threats centrally is easier and more reliable then having it distributed to a home gateway of consumer devices.

Benefits of Virtualization

The business benefits of CPE virtualization can be categorized as improving time to market, CAPEX and OPEX reduction and system simplification.

Faster Time to Market

1. Accelerating software development: Software development for a virtualized application is less constrained than in a typical embedded system. This not only reduces the time needed to develop apps, but also to add new features as there are more development tools, fewer memory constraints, faster CPUs and a wider pool of engineers capable of programming in a standard environment.

2. Not relying on specific hardware: With virtual applications, there is no need to wait for equipment orders to arrive and hardware to be installed. New services can be launched on the existing server infrastructure.

3. Reducing testing time: Testing of a virtual application can take place in a production environment by starting with a small-scale implementation and then rolling it out as the software stabilizes. In total, this results in shorter qualification cycles.

4. Eliminating dependency on devices in the home: From a logistical standpoint, home devices are not easily replaced. Consequently, new features that depend on an updated home device take years to roll out. With a virtualized environment, this complexity and delay is removed since the dependency on the home is eliminated.

Reduced CAPEX

1. Reuse of resources: The same compute resources can be used for multiple applications, thereby maximizing utilization and reducing the need for additional resources. For example, during high traffic daytime hours, firewall services require a lot of resources.

However, when traffic drops at night, the freed up resources can be used, for back-up services.

2. Statistical multiplexing of resources: Historically, operators could statistically multiplex bandwidth in order to maximize utilization. With NFV operators can also_statistically multiplex the resources needed for a function. Furthermore, the ability to "cloud burst" into other "clouds" when required capacity exceeds available resources allows operators to engineer the network for average usage rather than peak usage.

3. Lowering CPE costs: Because services are enabled and delivered from a virtual machine in the network, they can be enabled for existing customers without replacing operational CPE.

Lower OPEX

1. Reducing manual operations: Because SDN and NFV can facilitate network automation, many operations that previously required manual intervention, including new service deployment, can be automated.

2. Minimizing CPE software upgrades: For every function that runs on the virtual CPE, the cost and complexity of managing CPE software upgrades is eliminated.

3. Improving service quality and reducing support costs: Deploying a unified service that operates independently of the hardware platform deployed at the customer premise leads to better service and less support.

4. Dynamically scaling and optimizing use of resources: CPE consumes power even if it is not forwarding packets. A virtual machine can be completely turned off when a function is not in use and its resources can be reallocated to another function. This enables dynamic scaling of resources, optimizing power, space and cooling.

Operational Simplification

The operational implications are closely related to the OPEX reduction, however it is worthwhile to spell them out:

1. Eliminating the need to manage software versions: CPE functions run as virtualized network functions leverage a single service model. This eliminates the need to deploy and manage software versions across the typically varied CPE footprint and streamlines service roll-out.

2. Speeding error identification and resolution: When software is deployed as a system-wide resource, troubleshooting, isolating and remedying a defect is simpler and quicker.

3. Standardizing infrastructure across multiple applications: Since CPE functions are being run on centralized platforms, operators can standardize the infrastructure (hardware, software, tools), reducing complexity and improving reliability.

4. Increasing customer satisfaction: The home environment is typically a jungle of improperly plugged cables and equipment tucked away without proper air circulation. By moving features/functions into the operator cloud, the complexity of in-home systems is reduced, which will result in fewer customer calls due to improper installs.

Open Issues

The benefits of using SDN and NFV to virtualize CPE functions are considerable. There are, however, a couple of issues which must still be addressed.

1. Requires greater technical skills and knowledge: While the overall operation is simplified by using NFV to perform CPE tasks, a more skilled work force is needed to maintain the system. For example, connecting an Ethernet cable between two physical devices does not require a high degree of skill. However, operating the virtualized network functions requires substantial technical knowledge and a higher skill level.

2. Requires greater security: A cloud environment requires a higher level of security as some of the benefits of NFV could backfire in the event of a security breach.

3. Requires more careful data center planning: A data center that is too heavily oriented around video distribution might need to be upgraded to support virtual CPE applications. Geographic placement of the data center will also come into play for latency reasons. The closer the data center is located relative to the customer, the better response times would be – especially compared to companies that might provide virtual CPE services over the top.

CONCLUSION

In this paper we have explained what SDN and NFV are, how they may be used both in the access network and at the customer premise, what they achieve from a CAPEX and OPEX perspective, and the benefits in terms of resource management.

We can now see how SDN and NFV can work together to help scale the network to levels never before considered. This is achieved by breaking the functional components of devices into software, placing these software elements into a virtual machine environment, and having each use resources only as needed. This helps operators optimize their network spend by using COTS hardware and managing this equipment using an abstraction layer under the control of the orchestration layer and hypervisor.

Real-world examples of how this works are rapidly appearing in labs, proof-of-concept trials, field trials, and in actual deployments. The authors of this paper believe this is the beginning of a new era of cable technologies unlike any we have seen to date. Operators are no longer required to add new hardware except when existing equipment runs out of computing or packet processing resources. We invite you to join the many efforts underway at CableLabs, IEEE, IETF, ETSI, and other standards organizations to help develop the specifications, use cases and models needed to drive these exciting new technologies into daily use so we may all benefit from their adoption.

Virtualized Software Transcoding for Cloud TV Services

Yasser F. Syed Ph.D., Comcast Distinguished Engineer
Weidong Mao Ph.D., Comcast Senior Fellow & IEEE Fellow

Abstract

This paper will examine what will be needed for Cloud TV video services to utilize a virtualized software platform for orchestrating distributed transcoding workflows.

There are many encoding constructs that are now becoming realizable in this environment. Until recently, customized IC and hardware-based designs had been a practical approach to getting the job done using general-purpose servers. But these general processors are getting faster, more efficient, and more cooperative in multi-processor configuration groups.

Encoding/transcoding technologies, and especially those designed for Multiple Bit Rate (MBR) technologies using H.264/ Advanced Video Coding (AVC), as well as High Efficiency Video Coding (HEVC), are increasingly designed to become more flexible, adaptable, and scalable.

Furthermore, encoding/transcoding quality continues to increase, resulting in higher density growth at a much lower cost using general-purpose servers.

This paper discusses how transcoding workflows can be adjusted, so as to to take advantage of new server architectures, better networking, and distributed storage -- through a concept called a Transcoding Resource Manager.

OVERVIEW OF MULTI-SERVICE CLOUD TV ARCHITECTURE

Though all video services may look the same to a viewer, in actuality, cloud-enabled services differ from traditional service infrastructures in several ways. Those differences attributable to cloud-based video are contributing to increased service velocity, operational efficiencies, and infrastructure utilization. This will increase as the footprint for cloud-enabled services grows.

For the video industry, candidate services for cloud-based delivery can include:

- *Linear* - Supplying live broadcast & cable TV Channels, as well as live event programming across multiple device platforms
- *VOD (Video On Demand)*- Offering TV programs, movies, and other content on a subscription or transactional basis across multiple device platforms
- *cDVR (Cloud DVR)*- Providing services to record linear TV programs & TV series into a digital locker, to be viewed across multiple device platforms
- *cUI (Cloud UI)*- Providing single search & navigation services across all device platforms
- *EST (Electronic Sell Through)*- Providing a service to purchase TV programs, movies, and other content with storage in a digital locker and viewing across multiple device platforms.

To implement services in a cloud infrastructure, three basic building blocks are required: Computing, networking, and storage. Cloud gives the ability to process information, move it around in a timely manner, and to store the results for later use. This needs to happen both at a macro level (across a network) and at a micro level (within a device). To make the cloud platform optimized and realizable, the infrastructure needs to be suited to handle not just one type of information, but different kinds of information, and in a concurrent manner, in order to shift resources where needed.

The infrastructure to do this includes a hosted set of networked, general purpose servers located in a collection of Data Centers that are tied together through an IP network physically running over fiber, Ethernet, coax, and sometimes even wireless mediums.

Storage, in this sense, is a federated architecture, with large, centralized server farms, CDNs, and caching edge servers that implement the file migration strategies that assure data is available at the right place and time. The servers, in turn, house processors that can be used in collective, re-combinable sub groups, so as to perform tasks that vary from simple to complex.

The networking element allows data to be transferred within and between servers, as well as from network ingress points to network egress points, in an expedited manner, and thus making the task meaningful.

Consumers are already showing their enthusiasm for the benefits of increased bandwidth capacity and speed. The number of devices that can handle cable services has expanded recently, from set top boxes to PCs, game consoles, tablets, connected TVs, and cell phones. A DVR service no longer requires a hard drive in the home, because the content can reside in the cloud. A new UI or feature can be rolled out without testing exhaustively on each device platform, and the UI or feature can be designed to exhibit the same "look and feel" or "experience" on every screen. This can be accomplished by handling most of the processing in the cloud, with only rendering happening on the client device.

Peak hours for requested content can be scaled to handle larger demands through load balancing on processors, so as to better handle requests, and to strategically copy/migrate content closer to the viewer, which localizes bandwidth demands.

To achieve this performance, processors need not only to be fast, but also to work in a multi-core configuration. Bandwidth on both the server and on the core network needs to increase, and has -- from 10 Gbps to 100 Gbps. Now, 1 Tbps speeds are being approached [2]. Similarly, storage and accessibility to storage needs to improve its block accessibility for read/ write fetches, so as to create a federated storage/caching architecture with bulk access and retrieval of data.

These types of improvements will also lend themselves well to other tasks beyond just delivering services to a client/customer. In this paper, we will examine how content transcoding workflows, through virtualization, can adapt to and take advantage of a cloud infrastructure.

Virtualization, in a software transcoder sense, works by executing content workflows in a way that is agnostic to specific hardware, processing assignment, storage location or resource capacity. The operator simply interacts with the task/workflows with the monitoring and rendered results presented to them on an interface device. Using the cloud for virtualization means that resources through the network can be dynamically allocated to transcoding workflows, according to demands and schedule.

DESCRIPTION OF TRANSCODING WORKFLOWS

In early cable Linear/VOD workflows, content was targeted to be played back on a STB device through a VOD streamer or a

QAM-based live linear feed. For trusted content providers/ aggregators/ broadcasters, these distributable assets could be created and delivered to the service provider according to specifications including CableLabs ADI/VOD [5] and the SCTE's video specs [11]. This was good for scheduled material that would often come as a content refresh of a VOD catalog, or as live linear content targeted for the just the STB and no additional networked services.

For resource planning, the capacity of the transcoding system needed to be based on a peak transcoding rate covering file transcoding equipment for VOD and the recording of live feeds. That's because the peak transcoding rate is more dependent on the time constraints of processing the material than on catalog volume [4].

Another issue was getting the mezzanine assets or contribution feeds to the transcoding systems. Traditionally, this was done through a satellite pitcher/catcher infrastructure, or, more recently, with IP distribution delivery taking place over FTP-like connections (e.g. Aspera).

To get expanded Linear/VOD services to the STB, transcoding resources need to be continually increased to keep up with demand. Today, the demand for transcoding resources is exponentially exploding due to:

- VOD services being expanded to COAM (Customer Owned & Managed) retail devices (Tablets, PC, Cell Phones)[1]
- Turn-around demands getting tighter (to hours instead of days/weeks, C3/C7)
- Expanded Current Program content (over 200 TV shows for Fall 2013! [12])
- Expanded Backlog content
- Cloud DVR Services based on linear feeds
- EST (Electronic Sell Through) Assets w/ DLNA
- Higher Resolution/ Frame rate [1080p, UHD, 60 Fps] versions for content
- More diverse and increasing set of content providers

To handle this expanded scale of encodings, content workflows need to evolve to be faster. They also need to consider integrating priority of the job as part of the content workflows. This integration of job priority could be categorized into 4 content workflow types:

1. Just-in-Time (JIT) Content Workflow

A JIT content workflow requires a quick-turnaround service (immediate, minutes, hours). Immediate workflows may be for linear live services, and recordings of linear services. Examples of a "minutes & hours workflow" are highest priority VOD transcoding services, such as live events, or unscheduled VOD assets that require a mezzanine/contribution asset as a source.

2. Near-Time (NT) Content Workflow

A NT content workflow could be a high priority transcode for unbroadcasted material to handle live turn-around events (e.g. non-broadcasted Olympics events), broadcasted events that require a transcode, C3/C7 re-encodes, and repairs to pre-existing VOD assets.

3. Catalog Content Workflow

A catalog content workflow could handle scheduled assets that are a part of a VOD content catalogue refresh, EST offering, or catalog migration. This is typically not a high priority transcode unless a VOD presentation time is quickly approaching.

[1] To address CO&AM devices, adaptive streaming technologies generate 5-10 new encodes for each content asset. Each then needs to be wrapped in 1-3 DRMs and packaging technologies, dependent on device and delivery agreements.

4. Assembly Content Workflow

Assembly Content Workflows do not require additional transcoding processes but may need things like JIT packaging, manifest generation, manifest conditioning, or dynamic ad replacement. This may be initiated by a customer request, and ties in with using the Cloud for delivery of the service to the viewer.

How can a transcoding service take advantage of being virtualized in a cloud-based infrastructure to handle a large scale of content workflows? The benefits of being virtualized are multiple: 1) a service instance can be created or destroyed based upon transcoding demand, 2) a service instance needn't be dependent upon dedicated hardware but general purpose servers, and 3) a service instance can be moved anywhere in the infrastructure, for network bandwidth or storage optimizations. In the near term, it can handle more of the lower priority workflows and reduce the anticipated site-based transcoding equipment needed. In addition, cloud-based resources can be shared with other task-based services to reduce transcoding resources.

What is needed to evolve in this direction? First, a move to using software based transcoding, followed by higher bandwidth for movement of data at both the network level and server level. Lastly, a more defined breakdown of transcoding processes into more granular, operable atomic units.

REALIZING A SOFTWARE TRANSCODER ON GENERAL PURPOSE SERVERS

Software video transcoding requires a significant amount of calculations around the stages of pre-processing, transcoding, and wrapping/packaging. Pre-processing functions are things like decoding, pre-filtering, cropping, scaling, and watermarking. Transcoding processes include transforms (DCT/Integer), quantization, block-based operations, motion search/prediction, and entropy operations. Wrapping/packing functions exist to prepare the content for streaming, as well as chunking/manifest creation for adaptive streaming, or file formats wrapped in MXF. Some of these operations are coefficient based and can scale up to the number of sampled pixels. Others are more blocked-based, repeatable operations. Some other operations don't happen at this scale and require more sequentially-based threads. The types of operations for video encoding are basically a combination of sequential processing, quick paralizeable calculations, and fetching/putting from fast cache memory.

Video transcoding can be implemented in the following architectures with each type having tradeoffs in the areas of programmability (new code), platform rework (redesign based on server hardware type), quality (is it fixed or can it continuously improve?), costs per stream, cloud fit (deployability in Cloud infrastructure), density (rack compactness), and power [See Table 1].

Architecture	Program mability	Platform Rework	Quality	Costs/ stream	Cloud Fit	Density	Power
Hardware Based	Hardware Update	None	Fixed	High	Low	High	Low
Software w/CPU (single/multi-core)	Flexible	Low/High	Improve	Low	Mid/High	Low/Mid	Mid/High
Software w/ CPU & GPU	Flexible	High	Improve	Low	High	High	Mid

Table 1:Video Transcoding Architectures & Tradeoffs

Hardware architectures can be the most efficient and integrated for high-density transcoding -- but because of their permanence, specialization, and high costs, they represent a difficult fit into cloud architectures. The higher costs are attributable to specialization, and an inability to drive costs down through volume. Also, transcoding hardware may not be optimized for doing other task, which is a key to unlocking cloud architecture potential.

A general-purpose server that is suitable for multiple tasks, including transcoding, is a stronger option for distributed infrastructures with large volume scalability.

A key component of its effectiveness is the performance of a single CPU. A single CPU is capable of running 1-2 threads, using time division multiplexing. Each thread can run a single set of sequential instructions. The faster the performance of the CPU, the more threads and more complex set of instructions can be used. The limitation on the performance of a single CPU is due to an upper bound on clock frequency, which in turn stems from silicon die limitations.

An alternative is to implement a multi-core architecture. This allows for multiple processors to be put on a single silicon die. When designed properly (optimizing sequential threads across each core), this can lead to higher performance at lower processor speeds, with improved power management and cooling. Early multicore architectures involved 2 cores, but have been evolving to 2/4/8/16/32 cores. For processing purposes, performance optimization may not be directly proportional to number cores, because each individual process needs to be independent but sequentially assembled. An additional factor to consider is the amount of code rework needed to optimize over different types of multi-core architectures, which can vary over type as well as number [6]. Some experiments of multi-core architectures in HEVC studies have seen performance speed-ups of 25 times that of a single core architecture [7].

Pure software encodes can run on this platform, but do not always take full advantage of the multicore architecture -- even though lower level toolkits may provide ways of taking some advantage of this. Implementation can be further accelerated through the use of Graphical Processing Units (GPUs) and fast access memory. GPUs became popular for the computer gaming industry as specialized PC cards used in creating graphics operations, such as textures and shading. GPUs harness a parallel throughput architecture that is suitable for processing many concurrent threads slowly, and can be suitable for some transcoding operations of a highly scalable nature -- such as coefficient level/ pixel-based operations or block-based operations. GPUs need to have a full input pipeline to be most efficient, and this is where fast access to memory is needed, with optimizations in the code to allow for large block fetching. GPUs can exist as a separate chip on a board (e.g. GT 60xx), in which can handle large volumes but also need to be tied into larger bandwidth and faster memory access that is off the die. GPUs and memory can also coexist on the die, but process less volume and memory while optimizing data management between CPU and on-board GPUs. Originally, GPUs were an afterthought on the die for basic computer graphics purposes. If used for video transcoding, the encoding time speeds up, but the picture quality could suffer.

As chip architecture developed, more room became available on the die (to accommodate I/O demands), which allowed the inclusion of more and better GPUs, memory, and other specialized processors, to address the growing importance of customer-facing media applications [8,9]. This has greatly improved the performance of GPUs on the die -- in some cases up to 75 times over the span of 4-5 years [9] . The popularity of this architecture is evident, with 9 out of 10 PCs already shipping CPU and GPU on the same piece of silicon. [8]

Today, specialized processors are built on the die or GPU chip that is dedicated to perform minute operations involved in video coding. This aims to improve quality while increasing processing speed. GPU/CPU architectures designed to share more memory can further increase performance. There are several new specialized media structures used to help accelerate video encoding [such as Quicksync (Intel – Haswell), and VCE logic (Nvidia Radeon)]. These use video middleware toolkits such as Handbrake and OpenCL as efficient access layers to the hardware. As GPU performance improves [See Fig. 1], it increasingly appears that the use of the evolving general-purpose servers that have these capabilities will be an ideal architecture for transcoding operations. [6]

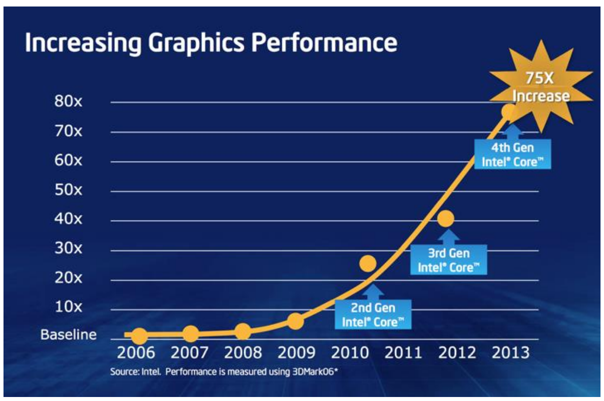


Figure 1: Evolution of Chip Graphics Performance [9]

Software video encoding can be realizable on server architectures, but ultimately needs to be optimized for the multicore architecture that can be accelerated through GPUs, cache memory, and specialized processors using video conversion toolkits. Again, the three components needed for the cloud are computing , networking, and storage [See Fig. 2] -- at a macro and micro level. A multicore architecture with mixed CPUs/GPUs integrates well with this concept [See Fig. 3]. Building servers for this should consider a combination of architecture & capabilities, BOM costs, and data center operational costs.

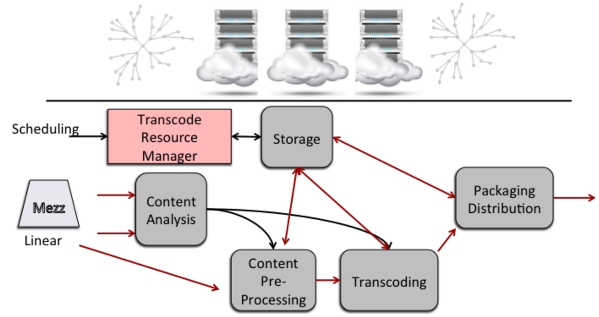


Figure 2: Transcoding Workflows using the Cloud

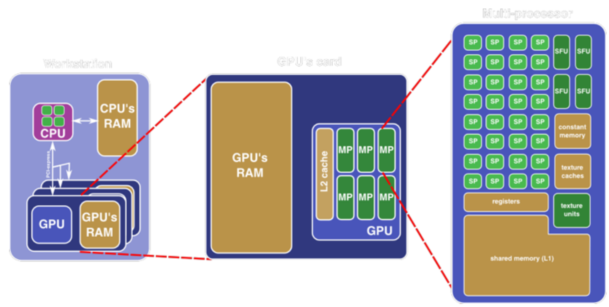


Figure 3: Mixed CPU/GPU architecture [11]

TRADEOFFS OF QUALITY, DENSITY, AND PERFORMANCE

Bringing up a transcoding instance on a general-purpose server allows the flexibility of performing tasks in more than one way. Different approaches may be optimized to converse resources, speed-up transcodes, or improve quality. The advantage of this over dedicated hardware is to be able to look on a larger scale (beyond just a box) to assign and adjust resources (often across different locales) as needed to produce an acceptable output volume.

In dedicated transcoding hardware, some algorithms may be fixed to avoid the complexity of putting multiple approaches in the hardware. An example of this is motion search algorithms. This can affect the accuracy and number of motion vectors generated, which can in turn affect the size of predicted pictures. In dedicated hardware, motion search approaches and range may be limited. Under certain circumstances, using an alternative approach may actually improve picture quality. This opportunity (and others such as dynamic encoding, dithering, and edge enhancement) can now be available using this type of architecture. Another example is the ability to assign cache memory resourced for a single instance. Expanding the amount of on-cache memory and allowing large block fetch movements can optimize the interaction between CPUs and GPUs, but optimization can differ depending on the size of the picture [1].

Lastly, in a software/general server-based environment, more than one instance can now be invoked to handle a task. A "split and stitch" process could use the multiple instances to split the content up, process each section in parallel, then reassemble the compressed stream [See Fig. 4]. This can be helpful when dealing with near time or JIT content workflows.

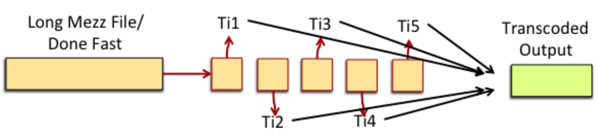


Figure 4: Split & Stitch Workflow

With a more actively managed orchestration of a workflow, there exist many yet-to-be-explored advantages that may be useful for content transcoding workflows using software on general-purpose servers. It does, however, require changes in traditional content workflows.

DESIGNING CONTENT WORKFLOWS FOR THE CLOUD

A. Adjusting Traditional Content Workflows

A traditional approach in transcoding is to make use of "hot folders" to process content in each step of the workflow [See Fig. 5]. Basically, content and/or metadata is put into a storage folder. The transcoding operation would then see the content and start processing it. At the end of the task, it places the content in an output folder, which then may act as an input folder for another task.

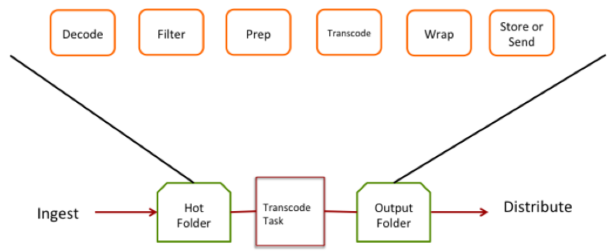


Figure 5: Hot Folder Workflow

There are several items to be aware of when using hot folders:

- Content needs to be processed in the same way (issues arise when the content gets placed in the wrong folder)
- Content needs to be complete before a task is started
- It is difficult to indicate priority, and thus harder to adjust the workflow because of priority
- Resources can idle, due to empty folders, making it all the more difficult to share resources

An alternative approach is to associate tasks with the content as a "job". In this sense, the content is acted upon, rather than placed in a folder. The transcoding instance is not designed to do a dedicated set of tasks, but instead looks to see what task needs to happen on a piece content at the moment. This is similar to the use of a general-purpose server rather than a dedicated hardware device. With an associated ID and metadata (such as due date, aspect ratio, etc.) , the concept of a job can be a powerful concept and can enable:

- Workflows that adjust to content
- Reduction of storage

- Scheduling of resources even before content arrives
- Prioritizing workflow to either handle voluminous or bursty traffic
- Reducing transcoding resources when not needed and applying them to other cloud tasks

Additionally, another change that needs to be made to improve adaptability to cloud and data is to break down tasks into more atomic units. Instead of just transcode, it can be broken down further into operations such as:

- Job verification
- Schedule
- Unwrap
- Verification (syntax, file size, semantic)
- Baseband processing (e.g. Cropping, Scaling, prefiltering)
- Decode
- Watermark, Fingerprint
- Encode
- Delete
- Wrapping, Distribution

Each of these tasks has their own combination of resources and can vary in complexity and time-to-finish. These three factors may not be deterministic, but can be bounded. These tasks can be combined in different manners to adjust the workflow to the types of source content, types of job, available resources, and time constraints. Creating smaller task modules and using a job workflow will allow better processing in the cloud in a resource optimized way.

B. Orchestrating Workflows Using A Transcoding Resource Manager

In order for a virtualized software transcoding instance to operate in a cloud infrastructure, the following are required: A balance of resources, network bandwidth demands, and timely output volume of assets. In order to keep these three aspects in balance, while exploiting the advantages of non-dedicated, general-purpose resources, the concept of a transcoder resource manager needs to exist [See Fig. 6]. This is more than just a workflow manager, because it balances server resources, storage, and bandwidth in a scheduled manner.

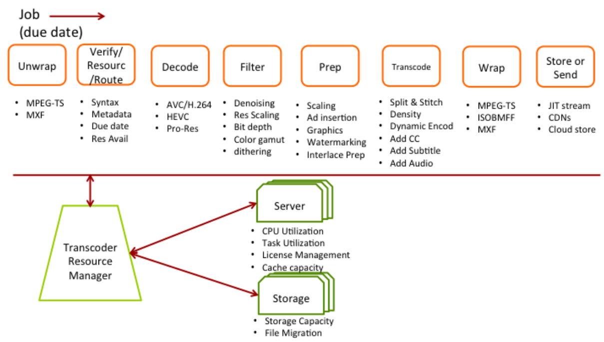


Figure 6: Workflow Using a Transcoder Resource Manager

The transcoder resource manager determines when and where an instance(s) will be created to address a job and when that instance will be destroyed/taken down. It will consider what server(s) to use and where to locate storage. It will schedule jobs based upon priority, type of content workflow and estimated time to completion (JIT, Near-Time, Catalog, Assembly). It will also ensure that jobs will be done agnostic to any types of equipment failures.

Some of the decisions that the transcoder resource manager could make can be exemplified by a mezzanine workflow to create a UHD-1 output as well as a set of MBR (Multi-Bitrate) streams:

- Determine how instances can exist at site(s) closest to the content mezzanine
- Verify source characteristics through metadata and syntax validation
- Create an instance to convert color gamut from BT 2020 to Rec 709
- Convert 10-bit input into an 8-bit version

- Split content into 5 sections and create 30 instances to process each section, for both 4K and MBR outputs
- Determine task assignment on servers using CPU utilization as a guideline
- Record linear programming
- Index file generation
- Manifest generation
- CDN distribution
- Destroy task instances and copies

OPERATIONAL IMPACTS

The basic building blocks for cloud infrastructure (computing, networking, storage) need to support the maximum performance needs of individual tasks for virtualized software transcoding. Cloud-based software and virtualized transcoding platforms will also impose unique operational requirements for the service providers.

These may include: (1) Capacity planning of compute and networking resources to enable simultaneous transcoding streams (dependent on types of services - linear, VOD, cDVR) and mode of transcoding (real-time, just-in-time); (2) Distribution topology of transcoding resources to national or regional data centers, based on attributes of source origination, network connectivity, national or local channels, and content delivery network; (3) Remote operational monitoring of software transcoding configuration, usage, performance, and availability; (4) Redundancy with automatic failover; and, (5) Transcoding software upgrade strategy.

CONCLUSION

Shifting to cloud-based transcoding services will help address the ever-increasing demand for video services with flexible resources. As algorithms continue to be optimized for multi-core architectures, and as networking bandwidth increases, and storage gets more accessible, cloud-based transcoding services can scale to meet the demand without scaling the costs in the same manner. Modification in the content workflows that are needed are: 1) implementing a job-based priority workflow, 2) breaking down the transcoding processes into smaller, atomic tasks, and 3) creating a transcoding resource manager to coordinate tasks and distributed equipment resources.

This approach can open up vast new types of business and operational models, where transcoding resources can be leased rather than purchased, and used to off-load peak transcoding demands. Virtualized software transcoders can create a flexible transcoding process that will be needed as the types and volumes of content for linear, VoD, and cDVR services grow.

REFERENCES

[1] F. Lee, "implementing H.264 encoding algorithms", Embedded Systems, March 2006, pp.34-37.

[2] J. Baumgartner, "Comcast's Unsung Hero: The Network", Multichannel News, Feb. 3rd, 2014, p.2.

[3] M.Ferrell, "Why Comcast Is Winning Subscriber War (for Now)", Multichannel News, Feb 3., 2014, p.29

[4] Conversations with Austin Vrbas, Brian Taft, Chris Cunningham, Feb 2012.

[5] www.cablelabs.com/specifications: MD-SP-ADI1.1, MD-SP-CONTENTv3.0, MD-SP-VOD-CONTENT1.1.

[6] Cyril Kowaliski, A look at hardware video transcoding on the PC: Performance and image quality with black boxes and Open CL. Tech report,., www.techreport.com, July 30th , 2012.

[7] M. Horowitz, F. Kossentini, and H. Tmar, " Informal Subjective Video Quality Comparison Between the eBrisk-UHD HEVC and x264 AVC Encoders", JCTVC-P0158, 16th Meeting of JCT, San Jose , Ca Jan 9-17, 2014.

[8] I. Cutress, and R. Garg, " AMD Kaveri Review: A8-7600 and A10-7850K Tested" AnandTech, Anandtech.com, Jan. 14, 2014

[9] A. L. Shimpi, "Intel Iris Pro 5200 Graphics Review: Core i7-4950HQ Tested" Anandtech, Anandtech.com, June 1, 2013.

[10] www.aurelian.plyer.fr/phd/gpu-programming

[11] www.scte.org/standards: ANSI/SCTE 128-1 2013, ANSI/SCTE 128-2 2013, ANSI/SCTE 54 2009, ANSI/SCTE 35 2013a

[12] tvbythenumbers.zap2it.com /2013/08/01/list-of-2013-fall-tv-show-premires-dates/195183

"Wireless Shootout: Matching Form Factor, Application, Battery Requirement, Data Rates, Range to Wireless Standard

David John Urban
Comcast

Abstract

Wireless standards are developed to meet the needs of many different applications, data rates and form factors. The IEEE 802.11ac standard is optimized for data rates greater than 300 Mbps for ranges of 1 to 10 meters with notebook computers having all day battery life. IEEE 802.15 based Bluetooth low energy is optimized for wearable devices such as watches, heart rate monitors, and stride sensors with data rates of 1 Mbps in extremely small form factors with battery life of weeks, months, and even years. Sensors that are geographically dispersed requiring years of battery life and very minimal data transfer may be best served using a lower frequency band with modulation tailored for low data rates and long range.

This paper discusses these three different use cases and describes how modulation, spectrum, and bandwidth can be matched to the unique requirements of each application and form factor

INTRODUCTION

The first campus wide wireless local area network was built at Carnegie Mellon University, named Wireless Andrew. The project was the brain child of Alex Hills and Marvin Sirbu. The first equipment was installed in 1995, using 915 MHz unlicensed spectrum and direct sequence spread spectrum providing 1 to 2 Mbps data rates. At the time this was orders of magnitude faster than the 19.2 kbps speeds available with other technologies and services.

In 1997, the IEEE standardized WLAN in the 802.11 working group and Wireless Andrew was upgraded to 2.4 GHz conforming to the 802.11 standard so that equipment from multiple vendors could work together. This is described in an entertaining book by Alex Hills that is highly recommended reading[1].

An interesting aspect of the book is Alex's path from tapping out Morse code to fellow hams halfway around the world, to running an AM radio station in Alaska, to installing radio telephones in rural Alaska, to building Wireless Andrew at 915 MHz 1 Mbps and upgrading to 2.4 GHz 11 Mbps. Each progression involved much higher data rate, Morse code to voice broadcast to two way telephone to WLAN. And each progression involved operating at a higher carrier frequency in order to accommodate the higher data rate. And yet with each progression the range decreased from half way around the world, to 50 miles, to several miles, to hundreds of feet.

Today, as we look at a range of products that have different speed and data rate requirements a useful guide is how different parts of the spectrum and different bandwidths and modulations were used for specific applications in the past.

This paper begins by analyzing the latest WLAN standard readily available for the fastest data transfers for a notebook computer. Measurements are reported of the data transfer speeds and range of 802.11ac with an 867 Mbps notebook computer in a single family home. The next section takes a look at the WLAN used in many phones and tablets. This is followed by a discussion of the personal area wireless technology for wearable devices that can pair with a smart phone or tablet. Finally, the paper takes a look at spectrum and physical layer considerations for standalone sensors that require years of battery life and can be dispersed over tens of miles.

WLAN 802.11AC FOR NOTEBOOKS

IEEE 802.11ac is a wireless local area network standard for 5 GHz unlicensed band that is optimized for data transfer speeds greater than 300 Mbps over distances from 1 to 10 meters. Notebook computers using 802.11ac with two stream capability can operate for eight hours powered by a battery. A typical battery type can be Li-Ion Polymer 7.6V, 54.4 Whr, and 7150mAh. So if the battery is pretty much drained after eight hours of heavy use, which means that the average power consumption of the device is around 6.8 watts.

To understand the importance of RF transmit power in the overall budget for power consumption consider a rough estimate of DC power into the final stage transmit power amplifier. The Wi-Fi radio will support two streams of data so two transmit power amplifiers are required. The output power of the notebook computer used for the tests could have an output power of up to +17 dBm in the 5.6 GHz band for an 80 MHz channel according to FCC test reports. The peak to average ratio of an OFDM signal is roughly 10 dB so that the final stage power amplifier must be linear for signal envelopes up to +27 dBm, a half watt of RF power. A standard design linear class A power amplifier may have a DC to RF efficiency of 10 percent, that is 5 watts of DC power are needed to support 0.5 watts of linear RF power. With two transmit chains, the DC power needed for the final power amplifier stages would be 10 watts using a standard class A power amplifier design and 100 percent duty cycle. This will not work since it exceeds the total average power consumption of the whole computer. Class AB power amplifier designs can have DC to RF efficiency as high as 50 percent which would take the final stage power amplifiers average DC power consumption down to 2 watts. This is a little bit closer to a workable amount of battery consumption. It is evident from these calculations that efficient power amplifier design is essential to be able to transmit at the power levels required by 802.11ac. These rough estimates suggest that lowering the duty cycle and output power of the Wi-Fi final stage power amplifier is an important consideration in improving overall battery life.

Devices supporting 450 Mbps have been commercially available for several years using the 802.11n standard. Getting to a 450 Mbps transmit rate with 802.11n requires three streams of data, 64-QAM modulation, a short guard interval, and a 40 MHz channel width. The details are described in full in references [2], [3], [4]. 802.11ac client devices typically use only two antennas compared to the three antennas used by 802.11n. This simplifies the radio circuitry and reduces power consumption. Testing in a single family home has found that it is very rare to get three streams of data during steady transmission. Testing has also shown that two streams of data are quite common at normal usage ranges. The key to increasing the data rate with 802.11ac is the 80 MHz channel width. This doubles the data rate while only increasing the SNR requirement by 3 dB. 802.11ac adds a 256-QAM modulation rate with 5/6 code rate. So what parameters are required to get an 867 Mbps data rate with 802.11ac? The channel width must be 80 MHz. The modulation must be 256-QAM with 5/6 code rate. Two streams of data must be sent. The guard interval must be 400 ns. The transmit rate of 867 Mbps is calculated in equation (1).

$$867Mbps = \frac{(8bps/Hz \cdot (\frac{5}{6}) \cdot 234 subcarriers \cdot 2 streams)}{(3.2\mu s + 400ns)} \quad (1)$$

The subcarrier spacing for all 802.11 OFDM signals is 312.5 kHz. This ensures backwards compatibility for 20 MHz, 40 MHz, and 80 MHz operation. Since the subcarrier spacing is 312.5 kHz the useful symbol time or FFT duration can be calculated as the inverse of the subcarrier frequency spacing which is 3.2 µs. A short guard interval of 400 ns is added to the FFT duration to determine the full symbol time. The FFT size can be calculated by dividing the channel width by subcarrier frequency spacing which for an 80 MHz channel width the FFT size is 256. This means that the 80 MHz channel is subdivided into 256 subcarriers spaced by 312.5 kHz. Of these subcarriers, some are nulled near the RF carrier and at the extreme low and high parts of the spectrum. Some subcarriers are used as pilots. This leaves 234 out of the 256 subcarriers that are used to transmit data as seen in equation (1).

The wavelength of a radio wave is the speed of light divided by the frequency. For our tests the frequency was 5,765 MHz. λ = 0.052 m = 2.047 inches. The path loss between two isotropic antennas in free space is proportional to the distance divided by the wavelength, $L = 4\pi d/\lambda$ as shown in equation (2). The free space path loss at 1 meter is 47.7 dB for a frequency of 5.765 GHz. The free space path loss at 3 meters is 57.2 dB for a frequency of 5.765 GHz. The free space path loss at 10 meters is 67.7 dB for a frequency of 5.765 GHz.

The subcarrier spacing for 802.11 OFDM signals is 312.5 kHz. The same subcarrier spacing is used for 20, 40, and 80 MHz channels. This makes it easy for an 802.11a device that is limited to a 20 MHz channel width to work alongside an 802.11ac device that can operate with an 80 MHz channel width. Since the subcarrier frequency spacing is always the same the FFT size can be calculated as the channel width divided by the subcarrier spacing. For an 80 MHz channel width the number of subcarriers is 256. The 80 MHz channel width is divided into 256 subcarriers with each subcarrier separated in frequency by 312.5 kHz.

Of the 256 subcarriers in an 80 MHz channel a few subcarriers are nulled out at the frequency extremes and in the center in order to make reception easier. By nulling out subcarriers at the center frequency, the receiver does not have to worry about DC offset when converting the signal to baseband. By nulling out subcarriers at the frequency extremes near the upper and lower ends of the channel, the receiver does not have to worry about interference from the spectral spillover of adjacent channels. In addition to the nulled subcarriers, some subcarriers are used as pilots. Pilot subcarriers do not carry data, they send a known pattern. Pilot subcarriers are important for frequency synchronization and channel estimation. By measuring two pilots at two different symbol times, the frequency offset can be determined. The time difference between two symbols is known, this gives us the change in time. So by determining the phase difference between two pilots at two symbol times the frequency offset can be determined as the change in phase over change in time. Likewise, by looking at two pilots within a symbol at different frequencies, the receiver can estimate the frequency response of the channel in between the two pilots and apply a correction. This is called channel estimation. After subtracting out the pilots and the nulled subcarriers from the 256 point FFT, the number of subcarriers remaining that can to be used as data subcarriers is 234.

Each of the 234 data subcarriers are modulated. Modulating a subcarrier means that the magnitude and phase of the subcarrier is adjusted to represent a set of bits. The modulation can be BPSK, QPSK, 16-QAM, 64-QAM, or 256-QAM. For BPSK modulation a 0 bit corresponds to a modulation of $\cos(2\pi ft)$ and a 1 bit corresponds to a modulated signal $-\cos(2\pi ft)$. A zero bit is distinguished from a one bit because the transmitted signals are sent 180 degrees out of phase from each other. In the same fashion QPSK maps two bits 00, 01, 10, 11 into sinusoidal signals with four distinct phases, 45°, 135°, 225°, 315°. For 16-QAM things are slightly more complicated, four bits are mapped into 16 amplitude and phase combinations. Since $\cos(0)=1$ and $\sin(0)=0$ and $\cos(90°)=0$ and $\sin(90°)=1$ with this pattern of the cosine function going to 0 when the sine function has a magnitude of 1 repeating every 90°, it is possible to send $\sin(2\pi ft)$ and $\cos(2\pi ft)$ at the same time. These two signals can each be amplitude modulated and as long as the combination is sampled every 90° they will not interfere with each other. This is called quadrature amplitude modulation, QAM. In the case of 16-QAM, 2 bits amplitude modulate the sine wave and 2 bits amplitude modulate the cosine wave. Likewise, 64-QAM maps 6 bits to 8 possible sine wave amplitudes and 8 possible cosine wave amplitudes.

The modulation that is new in 802.11ac is 256-QAM. 256-QAM maps each subcarrier to 8 bits. 256-QAM requires a signal to noise ratio of 30 dB for a bit error rate of 1 in ten thousand. The highest modulation in 802.11n is 64-QAM so that the maximum number of bits that can be mapped onto a subcarrier using 802.11n is 6 while the maximum for 802.11ac is 8 bits per subcarrier, thus the spectral efficiency of 802.11ac can be 33 percent more than 802.11n.

The bits that map to the data subcarriers are code words. LDPC is an option for 802.11n and 802.11ac.

The forward error correction can be convolutional coding or LDPC. LDPC coding is extremely powerful.

Tests were performed in a single family home in an attempt to see just how fast files could be transferred using 802.11ac Wi-Fi with real devices. The next sections describe these tests. The transmit rate of an 802.11ac signal can be as high as 1300 Mbps. A local server running an Apache2 web server using Ubuntu Linux operating system was connected directly to the wireless router. The wireless router was capable of 1300 Mbps 802.11ac speeds. The web server was connected to the wireless router with a CAT-5 unshielded twisted pair cable with 1 Gbps Ethernet ports at the computer and the router. An objective-C program was written using the CoreWlan class
Library to measure parameters such as signal, noise, transmit rate, channel width, PHY mode. The program was run on a notebook computer with an 802.11ac 867 Mbps client station. The program transferred a 600 MB
file from the server to store on the notebook computer while measuring the time before and after the download and then calculating the average transfer rate. This was done twice with a waiting period of 20 seconds between downloads. The average download speed of the two measurements was then recorded in a test array. These tests were performed at many locations throughout a single family home. Fig. 1 to 9 of this paper records the main results from these battery of tests.

In order to check the validity of the program used to make the measurements of download speed from a local web server, first direct TCP and UDP traffic between two 3by3 802.11ac Wi-Fi adapters was measured. The download speed measured by the program included such factors as web server settings and computer hard drive speeds so it was desired to get a maximum TCP and UDP throughput measurement that did not have the same dependencies.

An 802.11ac 1300 Mbps access point was connected to an 802.11ac 1300 Mbps wireless to 1 Gbps Ethernet adapter in bridge mode. Ubuntu computers running iperf were connected to the 1 Gbps Ethernet ports of the access point and adapter. The wireless link was 3 meters long through one wall. A TCP transfer rate of 512 Mbps was observed using the iperf default client and server setting for Linux. A UDP test was performed over
an eight hour period using a 400 Mbps transfer rate. The test was done overnight so that very little competing traffic would be present. The UDP packet error rate for a 400 Mbps data rate over an eight hour period was measured to be 4.5E-4. Over short time periods there was little packet loss for transfer rates at 500 Mbps and below. The packet loss for UDP transfer rates at 600 Mbps and above was very high, typically 25 percent.

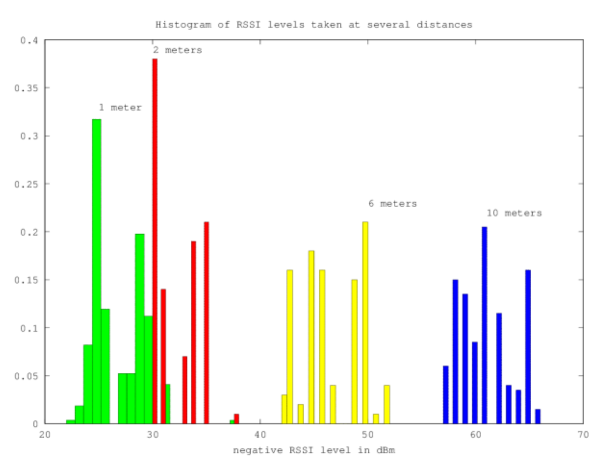


Fig.1 Measurement of RSSI.

Fig.1 shows histograms of the RSSI values measured at distances of 1, 2, 6, and 10 meters between the AP and the STA. The RSSI reading was taken from the notebook computer in between downloads of 600 MB and after a pause of 20 seconds. Typically a hundred to two hundred measurements were made at the same location while the orientation of the notebook computer was adjusted. The RSSI decreases as the distance increases and even at the same location the RSSI can vary. Notice that while the distribution of RSSI at a given location appears to follow the expected Gaussian probability density curve, there seems to be two bell curves at a given location rather than one. A theory to help explain this observation is that receive antenna gain of the notebook computer varies with azimuthal orientation resulting in two distinct RSSI levels. The variation around these two levels indicate that the signal propagation can vary over even a short time period by several dB.

Free space path loss has a $1/r^2$ power density attenuation verses distance relationship while a simple direct ray and ground reflection path loss model has a $1/r^4$ attenuation verses distance relationship. The free space path loss equation is shown in equation (2) where L is the loss, d is the distance between isotropic antennas and λ is the wavelength. These free space and the ground reflection models are represented by the straight lines in Fig.2, the top one free space and the bottom one ground reflection.

$$L = \frac{4\pi d}{\lambda} \qquad (2)$$

The plot indicated that measured levels tend to fall somewhere in between these two models. For a clear line of sight path between the AP and the STA with little in the way of scattering objects, the measurement can be very close to the free space path loss model. When there is an obstructed path and many scattering objects, the two ray model gives a more accurate prediction. While it may seem discouraging that we cannot pin point the receive level any closer than 20 dB with our two models and whichever model we choose an actual measurement is almost certain to differ from our prediction, things are not so bad. We can treat this as a statistical problem which will allow us to define a required fade margin for a given desired availability.

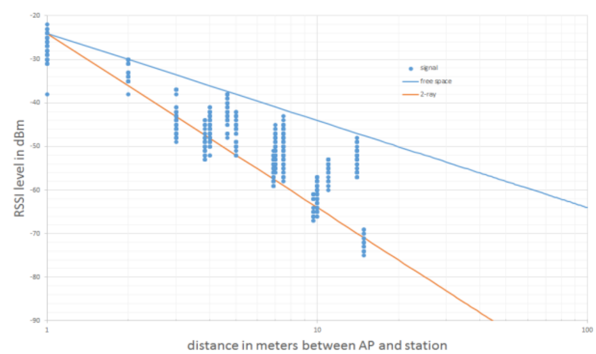


Fig. 2 RSSI vs. Distance.

$$L = A_0 + 10nLog_{10}\left(\frac{d}{d_0}\right) + N_\sigma$$

(3)

Equation (3) shows the formula for a log normal path loss model with free space path loss of A_0 at reference distance d_0 with a path loss exponent of n and the random variable N having a normal probability density function with zero mean and standard deviation

σ. L is the path loss and d is the distance between antennas. We split the difference between free space path loss and the two ray model and select a path loss exponent of 3. This means that we expect the path loss to increase by 30 dB per decade of distance increase and 10 dB per octave of distance increase. This leads us to Fig. 3 which shows a histogram of the difference between measured values of RSSI and predicted values using a logarithmic model with a path loss exponent of 3.

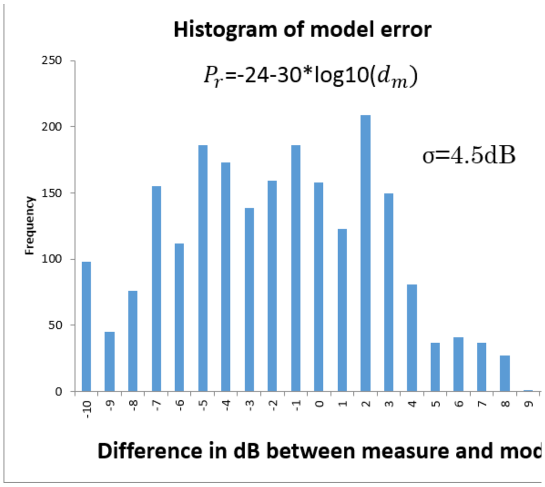


Fig. 3 Histogram of model error.

The model is not very good at predicting an actual measured value of RSSI since errors range from -10 dB to +10 dB. However, the important thing is that we can calculate a standard deviation of this error of 4.5 dB. This allows us to predict that allocating 9 dB fade margin in our link budget calculations will provide 99.8% availability over a range of times and locations. This is based upon the integration of the normal distribution from minus infinity out to two standard deviations above the mean.

This model of radio wave propagation is called a lognormal path loss model. The path loss is modeled as free space path loss at a short reference distance, in this case 1 m. The path loss exponent is applied to calculate the median path loss at a distance away from the reference point, in this case the path loss exponent that best matched measurements was 3. Measurements are expected to have a normal distribution around the median value, in this case the standard deviation was found to be 4.5 dB.

Using the values found in FCC type acceptance reports of the devices used for testing, the transmit power is around 14 dBm with 5 dBi antennas for the transmitter and receiver. The free space path loss at 1 meter for 5,765 MHz is 48 dB. Thus, the receive level at a 1 meter distance is -24 dBm. At 10 meters, the receive level will drop by 30 dB if the path loss exponent is 3, taking the median receive level to -54 dBm at 10 meters. If we add a 9 dB fade margin to account for two standard deviations of variation, then the level is -63 dBm. With an RSSI of -63 dBm, the data rate and reliability of 802.11ac has been shown to be excellent. Thus, up to a 10 meter distance we can expect a 99.8% availability with solid reliability and high throughput. Stations will operate at ranges greater than 10 meters, but there will likely be times when the reliability and throughput is sacrificed.

The guard interval for 802.11 OFDM is either 400 ns or 800 ns. The guard interval of 400 ns is called the short guard interval and the guard interval of 800 ns is called the long guard interval. Surprisingly the guard interval was one of the parameters that was often tweaked in the adaptive modulation process.

The speed of light is 299,792,458 m/s. By applying unit conversions to the speed of light, one can see that it takes light 1.0 ns to travel 1 foot. This is a convenient fact to remember when thinking about communication systems. A time delay of 400 ns corresponds to a radio wave traveling 400 feet, perhaps reflecting back and forth from a wall that is 200 feet away. This is why the short guard interval for an 802.11 WLAN signal is 400 ns; to account for reflections from objects up to 200 feet away. The long guard time is 800 ns which can eliminate intersymbol interference from reflections off of objects up to 400 feet away.

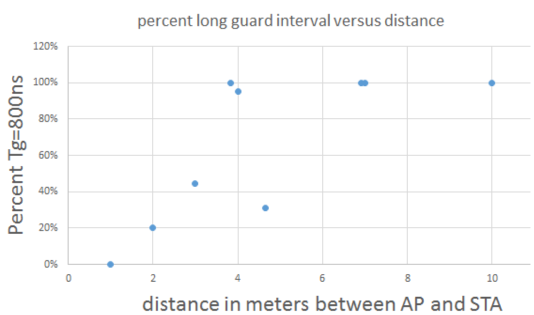


Fig. 4 Guard Interval versus Distance.

A simple analogy from everyday life, helps in understanding the role of a guard interval in OFDM in order to eliminate inter-symbol interference due to reflections. Suppose I have appointments for 1 hour private meetings with Joe at 1 o'clock and with Sue at 2 o'clock. If Joe shows up 15 minutes late and still wishes to meet privately for one hour, I have a problem. From 2 to 2:15 I will have both Joe and Sue in my office and will not be able to have a private conversation with either. If I make the meetings 45 minutes long and I set aside the first fifteen minutes for small talk, then Joe can show up fifteen minutes late, we can have our forty five minute private conversation, we can have small talk with both Joe and Sue in my office from 2:00 to 2:15 and Sue can still have a forty five minute private meeting. While not the most time efficient method, it is a very simple way to prevent conflict without having to worry about the precise arrival times. It works as long as no one shows up later than the designated guard time of fifteen minutes.

Fig. 4 shows the measured guard interval versus the distance between the AP and the STA. This plot indicates that a short guard interval is used a large percentage of the time for distances less than 5 meters. For distances greater than 5 meters the short guard interval is never used. Only at a 1 meter distance is the short guard interval used all of the time. This makes sense since when the STA is very close to the AP then long reflections will experience much more relative attenuation than the direct ray. At longer distances there may not be a direct ray and long reflections will be closer in amplitude to the direct ray.

Fig. 5 shows the percentage that each of the 5 modulation types were reported during the testing. 256-QAM was the modulation used for about 30% of the tests. 64-QAM was used for 58% of the tests. Adaptive modulation is a key feature in 802.11 WLAN. Since the channel conditions change rapidly and widely, sometimes the best modulation is BPSK and sometimes it is 256-QAM. The WLAN radio needs to be able to adjust automatically for the best modulation.

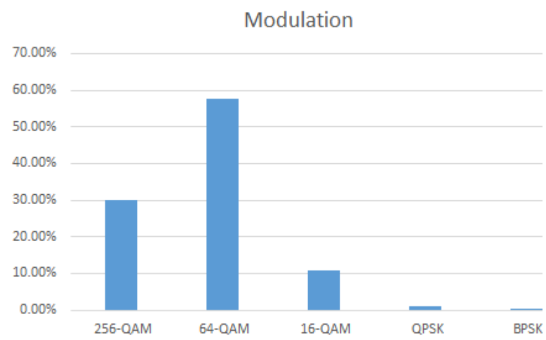


Fig. 5 Modulation Usage of Tests.

While 802.11 has moved on from Alex Hill's direct sequence spread spectrum used by Wireless Andrew at Carnegie Mellon in the 90's to 80 MHz wide 256 point FFT OFDM with a 400 ns guard interval, 256-QAM modulation and LDPC coding, all of the set up and signaling for 802.11 is performed with 802.11b/g signals. This is the beauty of 802.11, it can update to ever higher speeds and still work with older devices since signaling is performed using signals that every device can understand while data to a specific device is tailored to the highest speed that device is capable of utilizing. Thus, a device capable of 2 stream 802.11ac may still be using older style modulation for request to send, clear to send, and acknowledgement messages.

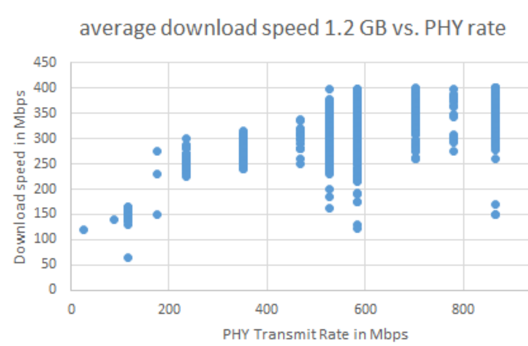


Fig. 6 Download Speed and Transmit Rate.

In addition to signaling needed to set up a packet transfer, each packet has a training sequence and overhead sent along with data symbols. This means that when the transmit rate is 867 Mbps, the actual download speed will be less than 867 Mbps due to training fields, signaling, set up, and other overhead. Fig. 6 shows the download speed measured versus the reported transmit rate. As expected, the general trend is that download rate increases with increasing transmit rate and that the average download speed is less than the transmit rate. We measured around 500 Mbps TCP throughput and 400 Mbps UDP throughput without packet loss using an 802.11ac 1300 Mbps adapter at very close range. The reported speed in Fig. 6 were made using a notebook computer with an 867 Mbps 802.11ac built in Wi-Fi adapter and downloading 600 MB files from a web server on the local area network and then averaging several measurements. During the testing the competing traffic on the local area network was small compared to the greater than 300 Mbps download speeds of the notebook. However, there were other users of the wireless network during measurements and this likely accounts for some of the cases where download speeds were much lower than transmit rate.

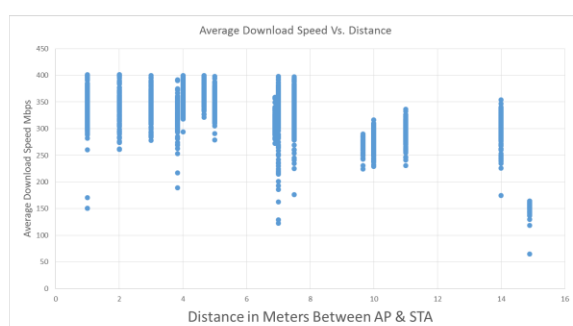


Fig. 7 Download Speed vs. Distance.

Fig. 7 show the average download speed from a local web server to a notebook computer after downloading two 600 MB files as a function of the distance between the AP and the STA. The download speeds were not appreciably different for distances less than 10 meters but did fall off noticeably beyond 14 meters.

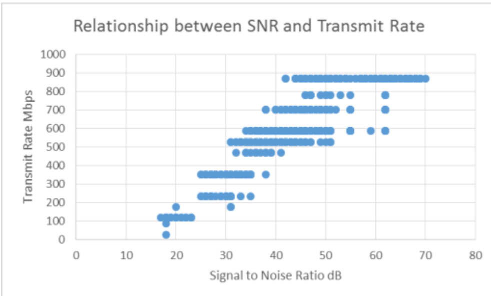


Fig. 8 SNR Vs. Transmit Rate.

Fig. 8 shows the transmit rate reported by the notebook computer as a function of the signal to noise ratio. The signal to noise ratio was calculated from the RSSI and the noise measurement provided by the CoreWlan library of the notebook computer. As expected the higher the signal to noise ratio the higher the transmit rate.

WLAN 802.11N MIMO FOR PHONES AND TABLETS

Many of the latest phones and tablets use 802.11n while moving to two transmit and receive chains for MIMO operation. With two transmit and receive chains these devices are capable of two spatial streams of data sent over the same frequency and at the same time. A SISO device has a single input and a single output; only one transmit chain and one receive chain. A MIMO device has multiple input and multiple output capability; two receive chains and two transmit chains. With a rich multipath environment signal processing is able to invert the 2by2 matrix describing the four paths between the two receive antennas and the two transmit antennas. If the 2by2 channel matrix can be inverted then the two streams of data can be calculated at the receiver. Higher SNR is required in order to separate the multiple streams of data at the receiver.

When comparing the performance of two devices using 802.11n with one device having MIMO capability, there is no change in the maximum bandwidth of 40 MHz and the maximum modulation order of 64-QAM. The only difference is that the MIMO version is capable of transmitting and receiving two spatial streams while the SISO device can only transmit and receive one spatial stream.

The doubling of the data rate can easily be observed by performing a speed test using a broadband connection of 100 Mbps. The 802.11n MIMO device recorded a download speed of 114 Mbps while the 802.11n SISO device measured a download speed of 48 Mbps. Speed test results using standard smart phone and tablet applications vary widely, but in general the MIMO device downloads at twice the data rate compared to the SISO device. This is as expected since the MIMO device can send two streams of data while the SISO device can only send a single stream of data.

From experience while using SISO and MIMO devices, there are additional benefits beyond just two stream capability from the 802.11n MIMO update. These may be chalked up to the multiple transmit and receive chains and perhaps general RF performance such as better noise figure, antenna pattern, AGC.

Watching live TV from a public outdoor Wi-Fi hotspot at a line of sight distance of around 80 meters, the newer tablet with 802.11n MIMO Wi-Fi displays high quality video without losing connection or buffering. The older 802.11n SISO device showed the live video at very low quality and had connection and buffering issues making the application unusable.

WPAN 802.15 BLUETOOTH LOW ENERGY FOR WEARABLES

Bluetooth low energy is designed for personal area network applications, PAN. This means that it is intended for communications between devices around a person, such as wireless headphones to listen to music from a phone in your pocket.

The aim of the Bluetooth low energy specification is to provide connectivity over several meters with small devices having long battery life. The battery life can be prolonged by limiting the amount of energy consumed by the device. The energy is equal to the power integrated over time. Reducing the signal to noise ratio requirements allows the transmit power to be reduced which in turn reduces the energy consumption and elongates the battery life.

Another, even more powerful technique to improve battery life is to turn the device off. Of course, a device that is always turned off is not particularly useful, at least with regards to providing us with sensor data. With a certain amount of data to be sent, it is beneficial to be able to send the data quickly so that the device can be turned off quickly. This leaves us with competing factors, low transmit power reduces energy consumption while the device is on, but on the other hand high transmit power improves the signal to noise ratio which allows higher spectral efficiency which in turn reduces the time that the device needs to be turned on. The right tradeoff between these two depends upon the amount of data to be transmitted, among other things. For transferring GB's of data between a notebook computer and a server, the 867 Mbps 802.11ac signal allows the data transfer to occur in less than a minute. For transferring smaller amounts of data from a wearable device to a phone, the 1 Mbps transmit rate of Bluetooth low energy allows the data to be transferred quickly.

802.11ac is designed for greater than 300 Mbps data rates over ranges of 1 to 10 meters using battery powered computers that are expected to last all day on a single charge. Bluetooth low energy is designed for connectivity over a distance of 1 meter with rates around 1 Mbps using wearable devices powered by coin cell batteries with capacity of 230 mAh.

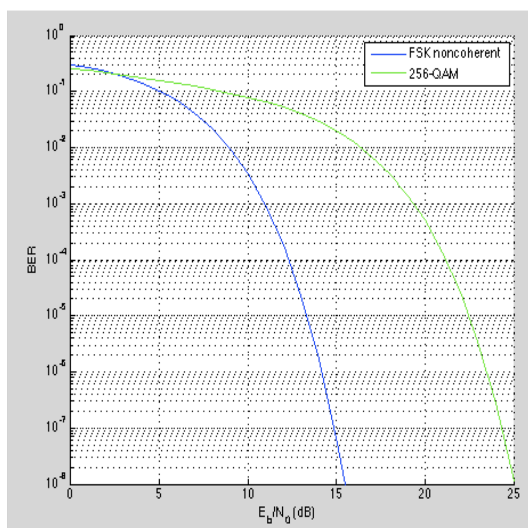


Fig. 9 BLE vs 802.11ac SNR.

The modulations used are tailored to the application and the battery requirements. 802.11ac has modulation rates that can go up to 256-QAM while Bluetooth low energy uses Gaussian frequency shift keying. A comparison of the bit error rate versus energy per bit over noise spectral density, E_b/N_0, is shown in Fig.9. The bit error rate for 256-QAM is 2% for 15 dB E_b/N_0. The number of bits per symbol for 256-QAM is 8 and $10*Log_{10}(8)$=9dB so that the SNR is 24 dB for a 2% error rate for 256-QAM. LDPC coding which is an option for 802.11n and 802.11ac is able to correct for a 2% error rate provided that the code rate and code word length are sufficient.

For FSK using noncoherent detection the bit error rate is 2% for an E_b/N_0 of 9 dB. FSK transmits 1 bit every symbol so that the SNR is also 9 dB for a 2% bit error rate. Bluetooth low energy does not use LDPC coding, however retransmissions of packets allow good throughput even with a 2% bit error rate.

This is illustrated in Fig. 10 which was recorded while listening to music using a Bluetooth headset. The music sounded fine during the test without any disruption. While walking away from the computer sending the Bluetooth transmission the RSSI changed from -50 dBm to below -80 dBm and this resulted in the retransmission rate going from 2% to 8%.

The modulation used for Bluetooth low energy can work at 15 dB lower SNR than the highest order modulation of 802.11ac. Working at a lower SNR requires less transmit power which reduces energy consumption and extends battery life.

A Bluetooth low energy device is required to have a receiver sensitivity of -70 dBm. Bluetooth low energy operates in the 2.4 GHz ISM band. The transmit power of Bluetooth low energy devices ranges from -20 dBm to +10 dBm. There are 40 channels for a BLE device with a center frequency beginning at 2402 MHz and ending at 2480 MHz separated by 2 MHz. The wavelength for 2.4 GHz devices is 4.9 inches (4).

$$\lambda = \frac{c}{f} = \frac{299792458m/s}{2402MHz} = 12.48cm = 4.914in$$

(4)

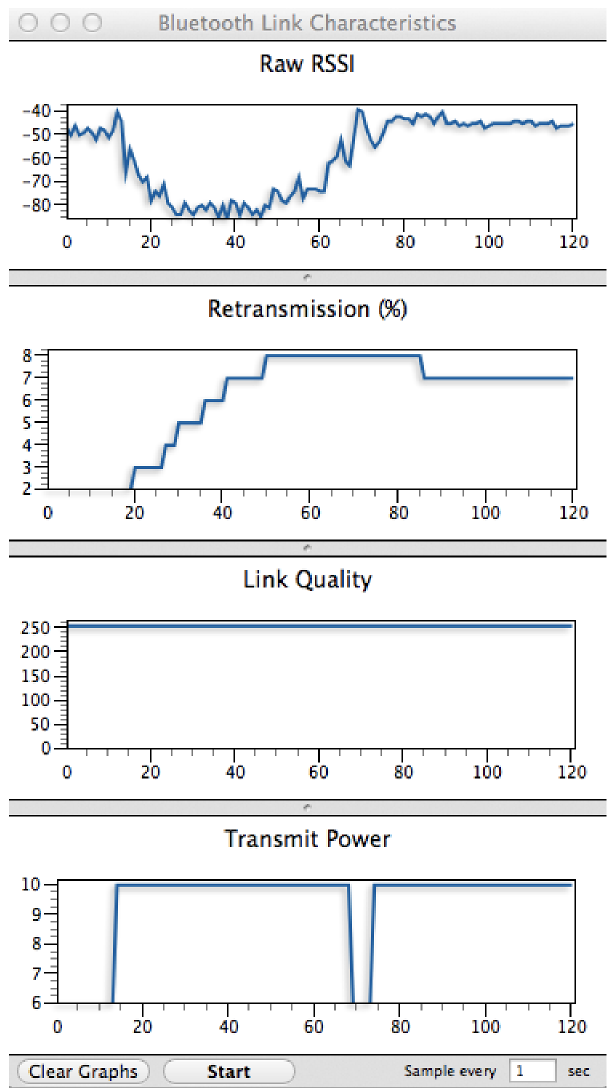


Fig.10 Retransmissions in Bluetooth.

A typical CR2032 button cell battery provides 3 volts with a capacity of 240 mAh. The CR denotes that the “Lithium coin” battery uses a chemical system of Lithium/Manganese Dioxide, Li/MnO_2. The 20 in the part number indicates that the diameter of the battery is 20 mm and the 32 in the part number indicates that the height or thickness of the coin cell battery is 3.2 mm.

The current draw for a Bluetooth low energy device is complicated. The current draw is different for various modes of operation such as sleep, wake up, transmit, and receive. The application note found in reference [8] carefully measured the current draw and time duration for various modes with an oscilloscope. As a rule of thumb, it was determined that the average current draw during continuous operation was 23 µA.

Bluetooth low energy uses Gaussian frequency shift keying modulation. This means that a zero bit is represented by one frequency and a one bit is represented by another. There are 40 channels in the 2.4 GHz band. BLE frequency hops between these channels.

$$h(t) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{t^2}{2\sigma^2}\right)$$

$$H(\omega) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} exp\left(-\frac{t^2}{2\sigma^2}\right) exp\left(-j\omega t\right) dt$$

$$\frac{t^2}{2\sigma^2} + j\omega t - \frac{\omega^2\sigma^2}{2} = \left(\frac{t}{\sqrt{2}\sigma} + \frac{j\omega\sqrt{2}\sigma}{2}\right)^2 = \frac{(t+j\omega\sigma^2)^2}{2\sigma^2}$$

$$H(\omega) = exp\left(-\frac{\omega^2\sigma^2}{2}\right) \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} exp\left(-\frac{t^2}{2\sigma^2}\right) exp\left(-j\omega t\right) exp\left(+\frac{\omega^2\sigma^2}{2}\right) dt$$

$$H(\omega) = exp\left(-\frac{\omega^2\sigma^2}{2}\right) \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} exp\left(-\left[\frac{t^2}{2\sigma^2} + j\omega t - \frac{\omega^2\sigma^2}{2}\right]\right) dt$$

$$H(\omega) = exp\left(-\frac{\omega^2\sigma^2}{2}\right) \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} exp\left(-\frac{(t+j\omega\sigma^2)^2}{2\sigma^2}\right) dt$$

$$H(\omega) = exp\left(-\frac{\omega^2\sigma^2}{2}\right)$$

Fourier transform pair

time domain $\longleftrightarrow$ frequency domain

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \longleftrightarrow e^{-\frac{\omega^2\sigma^2}{2}}$$

B is the 3 dB bandwidth

$$\sqrt{2} = e^{\frac{(2\pi B)^2\sigma^2}{2}}$$

$$ln\sqrt{2} = \frac{1}{2}ln(2) = \frac{1}{2}(2\pi B\sigma)^2$$

$$\sigma = \frac{\sqrt{ln(2)}}{2\pi B}$$

Fig. 10 Gaussian Filter Equation

The equation for the Gaussian filter that is applied to the bit stream of a Bluetooth low energy signal is shown in Fig. 10 along with the derivation of its frequency response. Fig. 11 shows the waveform at the output of the Gaussian filter with a bit stream input.

There are a couple of things to note. First, the Gaussian impulse response has the property that its Fourier transform is also a Gaussian response, both the impulse response and the frequency response follow a bell curve. The parameter σ, the standard deviation, can be calculated from the 3 dB bandwidth of the filter’s frequency response and this term is in the numerator of the impulse response and the denominator of the frequency response. This gives the expected result that a narrow impulse response is required for a wide frequency response and a wide impulse response in time is required for a narrow frequency response.

For Bluetooth low energy the parameter B is set to one half the bit rate, BT=0.5 where T is the symbol period of 1 µs. The 3 dB down point of the Gaussian filter for Bluetooth low energy is 500 kHz. Another thing to note is that a Gaussian filter is not a Nyquist filter. That means that it does not have an impulse response that goes to zero at every symbol time other than its own. A Nyquist filter eliminates intersymbol interference which is critical for high rate modulations such as 256-QAM. In order for a Nyquist filter to work in the presence of multipath, either an adaptive equalizer or a guard interval is required. A Gaussian filter reduces but does not eliminate intersymbol interference. For this reason a Gaussian filter does not support high spectral efficiency modulations. On the plus side, the Gaussian impulse response decays steadily with time so that for modulations such as FSK, there is no need for a guard interval or adaptive equalizer.

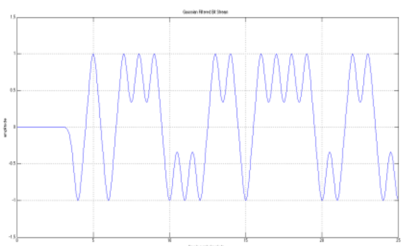


Fig. 11 Gaussian Filter Convolved With Bit Stream.

Standard Bluetooth also uses Gaussian frequency shift keying modulation. The frequency of the carrier is shifted by around 180 kHz to represent a bit. Low energy uses a longer 250 kHz frequency shift. With a 250 kHz frequency shift and a 1 MHz bit rate the two keyed signals are orthogonal. 250 kHz is the smallest frequency deviation for which the two signals are orthogonal. For this reason, this is referred to as minimum shift keying. Imagine a demodulator that splits the signal in two and runs the signal into two mixers. One mixer uses a local oscillator locked to the frequency that represents a bit 0 while the other local oscillator is tuned to the frequency representing a bit 1. The output of the mixer is integrated over the symbol period of 1 µs. If the frequencies match then the output of the mixer will be a DC voltage. If the frequencies do not match then the output of the mixer will have a low frequency component of twice the frequency offset, in this case 500 kHz. The period of a 500 kHz sine wave is 2 µs. The sinusoidal signal will traverse a half cycle which will integrate to zero. The principal of minimum shift keying is illustrated in equation (5) and Fig. 12.

$$\cos(\omega t + \Delta\omega t)\cos(\omega t + \Delta\omega t) = \frac{1}{2} + \cos(2\omega t + 2\Delta\omega t)$$

$$\cos(\omega t + \Delta\omega t)\cos(\omega t - \Delta\omega t) = \frac{1}{2}\cos(2\Delta\omega t) + \cos(2\omega t) \quad (5)$$

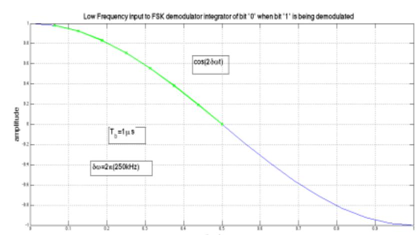


Fig. 12 Principal of minimum shift keying.

Frequency shift keying is used so that the signal can be demodulated at low input levels with poor signal to noise ratio conditions. This reduces the transmit power requirements and improves battery life while reducing the form factor. Bluetooth low energy uses a wider frequency deviation than standard Bluetooth so that the two waveforms are orthogonal. The wider frequency deviation increases the separation requirement between channels from 1 MHz to 2 MHz and reduces the number of channels from 79 channels for standard Bluetooth to 40 channels for Bluetooth low energy. The standard Bluetooth channels are shown in Fig. 13.

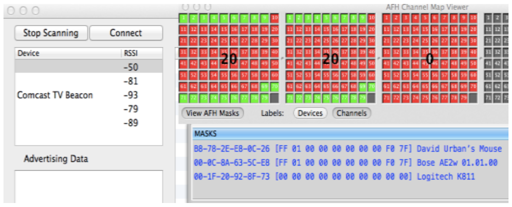


Fig. 13 Bluetooth standard has 79 channels.

GFSK with minimum shift keying is a constant envelope signal so that a linear power amplifier is not required. This helps improve battery life.

WMAN INTERNET OF THINGS LONG RANGE LOW BIT RATE

The Internet of things has been earmarked by many as the next big thing in technology. Although opinions vary about future manifestations of the Internet of things, providing connectivity to things never before imagined to need connectivity is a common element. This includes things that are very far away from Wi-Fi, Bluetooth, and even cellular coverage. Examples are garbage bins, restroom soap dispensers, parking spaces, and all manner of environmental sensors. Other applications include providing wide area connectivity to Bluetooth low energy wearable devices for those times when it is inconvenient to carry your cellular phone.

Geographically dispersed environmental sensors occasionally sending data to a base station require coverage area measured in miles and battery life measured in years. Applications for machine to machine communication along with a standard intended to ideally suit the applications is described in reference [10]. The group developing the radio interface for this application is called Weightless.

Operating in the UHF spectrum from 470-860 MHz gives the base stations greater range than the 2.4 GHz band used by Bluetooth low energy and the 5 GHz band used by 802.11ac. The wavelength at a center frequency of 470 MHz is 63.8 cm which is just over 2 ft. A half wave dipole antenna would be a foot long. At UHF band the antenna needs to be physically larger for an equivalent antenna gain compared to operating in the 2.4 GHz band or 5 GHz band.

The free space path loss at 1 GHz and 1 km is 92.4 dB. This leads to the familiar formula for path loss shown in equation (6). The path loss at 50 km which is 31 miles and 500 MHz is 122 dB.

$$L = 94.2 + 20 \cdot Log_{10}(d_{km}) + 20 \cdot Log_{10}(f_{GHz}) \tag{6}$$

The system link budget is 170 dB. For free space path loss conditions getting a 31 mile coverage range is no problem with such a large link budget.

The uplink channel width is 2 kHz. Using only a 2 kHz bandwidth improves the system link budget. The noise floor of an 80 MHz 802.11ac channel is 46 dB higher than the noise floor of a 2 kHz signal, if both have the same noise figure. The thermal noise spectral density at room temperature is -174 dBm/Hz. So with a 4 dB noise figure receiver and a 2 kHz channel width, the noise floor is -137 dBm. DBPSK, differential binary phase shift keying, is the lowest possible modulation. DBPSK requires a 10.5 dB SNR. A ½ rate convolution code is used for error correction which has 7.5 dB coding gain. A spreading factor of 4 is used to reduce the required signal to noise ratio by 6 dB. With these three factors, the signal can be demodulated successfully at 3 dB below the noise floor, SNR=-3dB. The transmit power is 20 dBm for the uplink device. The antenna gain for the uplink device is -4 dBi and the antenna gain for the base station is 14 dBi. The noise figure of the base station receiver is 4 dB. Using these parameters the system gain is calculated to be 170 dB as shown in equation (7), that is the path attenuation allowed between the two antennas. A measurement of the spectrum showing the 2 kHz uplink channel and the 6 MHz wide downlink channel is shown in Fig. 14. The system uses time division duplex so that both the uplink and downlink use the same spectrum at different times.

$$G_s = P_t + G_t + G_r + 174 - 10 \cdot LOG_{10}(B) - F - \eta = 20 - 4 + 14 + 174 - 10 \cdot LOG_{10}(2e3) - 4 - (-3) = 170dB \quad (7)$$

A benefit of the log-normal path loss model used in an earlier section for 802.11ac propagation is that the model is physics based with three intuitive parameters. The reference distance can be estimated based upon the expected distance for near line of sight transmission. For an indoor WLAN AP this will be around 1 meter. For a 100 foot high cellular base station antenna the reference distance may be around 100 meters. For a 1,000 feet high broadcast tower, the free space reference distance may be around 1 km. When a strong reflection is expected the path loss exponent will be close to 4 and when a strong direct ray is expected the path loss exponent should be close to 2. The standard deviation may be close to 5 dB when obstructions are uniform throughout the coverage area while the standard deviation will be closer to 10 dB when obstructions vary throughout the coverage area.

Consider a deployment from a broadcast tower. Using a path loss exponent of 4 and a reference distance of 1 km and a 10 dB standard deviation, the maximum distance calculated with a log-normal path loss model at 470 MHz with a 20 dB fade margin is 40 km. This is shown in Fig. 15.

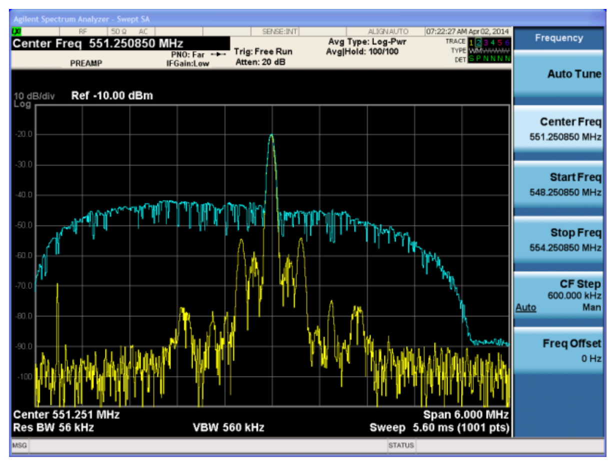


Fig. 14 Spectrum of WMAN IoT Signal.

In order to check the validity of the model, the calculation was compared to the prediction of TV coverage from a website using the Longely-Rice Irregular Terrain Model and the FCC data base. At a distance of 23.5 miles the predicted signal strength was -36 dBm from a broadcast station with an output power of 739 kW and an antenna height above average terrain of 179.9 meters. This would be a path loss of 124 dB, just 5 dB below the free space path loss. In reality, the signal could be picked up with an indoor antenna only after careful positioning and even then it required re-adjusting every once in a while in order to maintain picture quality.

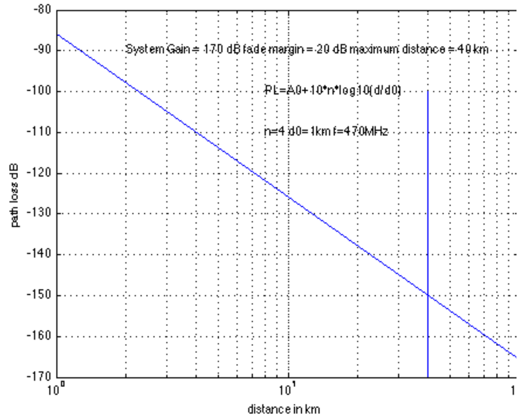


Fig. 15 Broadcast Tower Path Loss

Fig. 15 predicts a receive level at this distance of -62 dBm. In this particular case the log normal path loss model provided a prediction that was more accurate than the terrain based model. For prediction of UHF coverage over long distances the Longely-Rice Irregular Terrain Model and the FCC data base are critical in predicting coverage at a particular location. The log normal path loss model is of little help in predicting the actual coverage at a particular location because it does not account for the topography between the two antennas. In particular, if a hill shadows the RF signal then there is little chance of reception. On the other hand, if one can see the tower with a pair of binoculars then strong signal reception can be expected. The log normal path loss model is useful for statistically analyzing the likelihood of a system working over a range of locations and times.

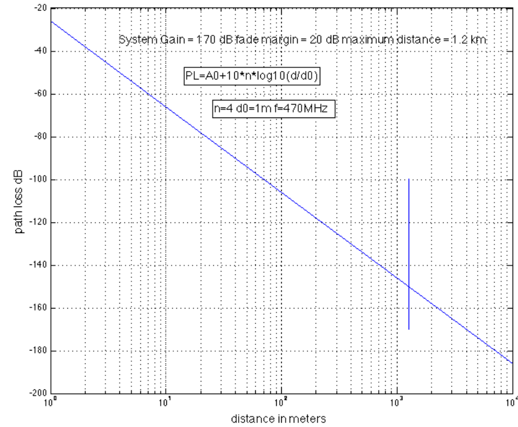


Fig. 16 Strand Mount Path Loss

Fig. 16 show that the log normal path loss model predicts a coverage of 1.2 km for expected parameters from a shorter tower on the order of 10-30 meters. The reference distance used is 1 meter and the path loss exponent is 4. The frequency is 470 MHz and the standard deviation used is 10 dB, thus requiring a 20 dB fade margin. The system gain used is 170 dB.

Note that the uplink uses a 2 kHz channel width with direct sequence spread spectrum having 4 chips. Thus, a 500 Hz rate is spread out to 2 kHz. The modulation is DBPSK with 1 bit per symbol and the code rate of error correction is ½. Thus, the data rate is only a paltry 250 bits per second. For remote sensors, only having to occasionally report small amounts of data back to the base station, this low data rate may be sufficient. The bit rate may be low but the 170 dB system gain makes sure that the range is very long. With sensors placed throughout an area, many devices will not require the full 170 dB system gain. These devices can use higher order modulations. As with Bluetooth low energy and 802.11ac, sometimes a high data rate can allow the device to turn off more quickly and thus preserve battery life.

CONCLUSION

Battery life is an increasingly critical parameter in consumer adoption and satisfaction of electronic communication devices. Great improvements to battery life can be expected in the coming years. Today, a notebook computer, tablet computer, and smart phone are expected to last all day before having to recharge the battery. Smart watches and fitness bands with Bluetooth low energy and Bluetooth headsets can last several days before having to recharge. Wearable devices using Bluetooth low energy can last weeks, months, sometimes almost a year before the coin cell battery needs to be replaced. Great care has been taken in order to achieve these impressive battery lifetimes. Even more attention to reducing energy consumption will be needed to meet the needs and expectations for future devices.

The two counteracting forces at work are lowering the SNR requirement to reduce transmit power and increasing data rate to reduce duty cycle. The data rate needs to be tailored to the size of data that needs to be transmitted, the range of transmission, and the form factor of the device. Computers with 120 GB hard drives need to send large data files. With an 80 MHz channel width, two streams of data, 256-QAM 5/6 LDPC coding, and a 400 ns guard interval, an 802.11ac signal can transmit a 1.2 GB file in 24 seconds. Tests have shown that files can be transferred in about this amount of time over a range of 1 to 10 meters.

Wearable devices using Bluetooth low energy need to send smaller size data files but it is still important that the time it takes to transfer data is short. A 1 Mbps data rate is realized with GFSK modulation having a 250 kHz frequency separation. The 3 dB bandwidth of the Gaussian filter is 500 kHz. No information is carried in the amplitude of the signal so that the constant envelope signal can use nonlinear high efficiency amplifiers. These and other features of Bluetooth low energy are designed to optimize battery life while still meeting the device data transfer requirements.

802.11ac uses the 5 GHz band while Bluetooth low energy uses the 2.4 GHz band. In order to collect data from millions of geographically diverse sensors having very small data transfer needs it may be advantageous to utilize UHF frequencies which have benefits for long range connectivity. By using very narrow band channels, for example 2 kHz uplink channel width, low order modulation with spread spectrum and forward error correction it has been shown that a 170 dB link margin is possible with reasonable transmit powers, noise figures and antenna gains in the 470-860 MHz UHF band. This is a highly coveted band but only a small sliver of it is needed for collecting sensor data. With the UHF band being repacked for broadcast TV and separated into FDD uplink and downlink blocks, this low power application can make use of small sections of spectrum that would otherwise need to be unused and reserved for guard bands. With 170 dB system gain, distance ranges from 40 km to over a km can be expected as antenna height goes from 1000 feet to 18 feet.

This paper has made measurements and analyzed the performance and applicability of many form factors and data rates. The channel width went from 80 MHz for a notebook computer to 2 MHz for a wearable sensor to 2 kHz for a remote sensor. The spectrum, likewise varied from 5 GHz to 2.4 GHz to 500 MHz. The data rate varied from 500 Mbps to 1 Mbps to 250 bps. It has been demonstrated that large size file transfers benefit from very high data transfer speeds using high order modulation, wide channel width and higher frequency spectrum bands. While at the same time applications requiring smaller data file transfers and smaller physical form factor can benefit from lower order modulation, narrower channel width and lower frequency spectrum bands.

REFERENCES

[1] Alex Hills, “Wi-Fi and the Bad Boys of Radio: Dawn of a Wireless Technology”, dog ear Publishing, ISBN: 978-145750-560-7, 2011.
[2] Albano., Ong,I.,Gilson,R.,Urban. “Design of Wireless Home Network”, SCTE CABLE-TEC Expo 11, November 15-17, Atlanta, GA, 2011.
[3] David Urban, “Video Calling Over Wireless Networks”, 2012 Spring Technical Forum, CableLabs, NCTA, SCTE.
[4] Chris Albano, Dinesh Devotta, Dave Urban, “Delivering DOCSIS 3.0 Cable Modem Speeds over the Home Network”, ”, SCTE CABLE-TEC Expo 10, 2010.
[5] Brittain, James E., “Harald T. Friis”, Scanning the Past column, Proceedings of the IEEE, VOL. 83, NO. 12, December 1995, pp. 1674.
[6] Matthew Gast, 802.11ac: A Survival Guide, O’Reilly, ISBN-13: 978-1449343149.
[7] Heydon, Robin, “Bluetooth Low Energy The Developer’s Handbook”, Prentice Hall, 2013, ISBN-13:978:0-13-288836-3.
[8] Kamath, Sandeep & Lindh,Joakim, “Measuring Bluetooth Low Energy Power Consumption, Texas Instruments Application Note AN092.
[9] Masuch,J., Delgado-Restituto,M., “A 190-µW zero-IF GFSK Demodulator With a 4-b Phase-Domain ADC”, IEEE Journal of Solid-State Circuits, Vol. 47, NO. 11, November 2012, pp.2796-2806.
[10] William Webb, “Understanding Weightless; Technology, Equipment, and Network Deployment for M2M Communications in White Space”, Cambridge University Press, 2012, ISBN 978-1-107-02707-7.

ABREVIATIONS

WLAN wireless local area network
WPAN wireless personal are network
WMAN wireless metropolitan area network
OFDM orthogonal frequency division multiple access
AP access point the WLAN wireless router
STA station the WLAN client adapter
GFSK Gaussian frequency shift keying
FSK Frequency shift keying
QAM quadrature amplitude modulation
SNR signal to noise ratio
MIMO multiple input multiple output: antenna techniques used to send multiple streams of data over the same spectrum at the same time
SISO single input single output: with only a single antenna multiple streams of data are not possible
BLE Bluetooth Low Energy
QAM quadrature amplitude modulation
LDPC low density parity check coding

ISBN 0-940272-01-6; 0-940272-08-3; 0-940272-10-5; 0-940272-11-3; 0-940272-12-1; 0-940272-14-8; 0-940272-15-6; 0-940272-16-4; 0-940272-18-0; 0-940272-19-9; 0-940272-20-2; 0-940272-21-0; 0-940272-22-9; 0-940272-23-7; 0-940272-24-5; 0-940272-25-3; 0-940272-26-1; 0-940272-27-X; 0-940272-28-8; 0-940272-29-6; 0-940272-32-6; 0-940272-33-4; 0-940272-34-2; 0-940272-35-0; 0-940272-36-9; 0-940272-37-7; 0-940272-38-5; 0-940272-39-3; 0-940272-40-7; 0-940272-41-5; 0-940272-42-3; 0-940272-43-1; 0-940272-44-X; 0-940272-45-8; 0-940272-46-6; 0-940272-47-4; 0-940272-48-2; 0-940272-49-0; 0-940272-50-4; 0-940272-51-2; 0-940272-52-0; 0-940272-53-9; 0-940272-54-7