DAA Field Deployment, Path to Scaling, and Digital Node Use Cases

A Technical Paper prepared for SCTE•ISBE by

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

Steve Sigman
Director, Access Architecture
Comcast Cable Communications
1701 JFK Blvd – Philadelphia, PA 19103
+1 (609) 685-3480
steve_sigman@cable.comcast.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

1. Introduction

Cable operators have been actively converging video and data services into a common Converged Cable Access Platform (CCAP) platform for a few years now. This trend, which requires an evolution towards newer, more modern, and denser equipment, should free up space in the headend.

However, as the success of high-speed data and video-on-demand services continues its seemingly eternal growth, the evolution of the access network progresses relentlessly towards expanded capacity and ever-smaller service groups. As a result, the spectrum allocated to narrowcast services increases, driving operators to deploy Data Over Cable Service Interface Specification (DOCSIS©) services including 32 SC-QAM (Single Carrier Quadrature Amplitude Modulation) channels and at least a fraction of an OFDM (Orthogonal Frequency Division Multiplexing) channel. In some cases Cable operators have even gone farther, to 36 and even 40 SC-QAM channels, and moving beyond a single OFDM channel onto a second one. In addition, if free spectrum in the network is not available, operators support capacity growth by segmenting service groups into smaller and smaller areas.

These expansion trends result in a continuous growth of headend equipment, which is already starting to exceed the capacity that headend facilities can support.

Therefore, the above trends are now intractably linked to two additional evolutions: distribution of components of the access network and virtualization of the core network functions. Furthermore, the implementation of a Distributed Access Architecture (DAA) is different in places where the plant is being upgraded, such as where a migration to a deeper use of fiber or N+0 is being implemented, versus in Cable networks where the existing plant needs to be segmented or expanded.

This paper will begin by outlining the evolution of service provider networks, and then describe why and how the migration to a distributed architecture is necessary and beneficial. The paper will then expand into features that can be implemented with Distributed Architectures and discuss the topic of Virtualization. Finally, the paper will explore how the implementation of DOCSIS 4.0 could be implemented in Distributed Architecture networks.

The paper is divided into the following sections:

- Cable network evolution, including:
  - Typical HFC networks today and growth projections
  - Advent of DOCSIS 3.1
  - The analog modulated forward link, options for its implementation, and comparison of the various options
- Why and how distributed architectures are useful, covering the following key benefits:
  - HFC network performance improvements by migrating to digital transport, especially to maximize the use of the higher order modulation rates that DOCSIS 3.1 and beyond offers
  - Headend density increases, which is becoming critical as service providers are segmenting service groups more and more, including extending fiber deeper into the plant and implementing passive networks
  - Trunk fiber savings as we move to higher capacity digital links that can be muxed much more than analog links, and
  - The ability to eventually virtualize the remaining upper layers of the CCAP
- Describe a Distributed Access Architecture, including:
  - DAA components
  - Key aspects of DAA interoperability
  - Use Cases
  - Generations and the advent of DOCSIS 4.0

2. Cable Network Evolution

2.1. Typical HFC Networks Today

Most MSO's hybrid fiber-coax (HFC) networks have been designed with an upper spectral boundary of 750 or 860 MHz, while some are designed to support 1 GHz and other newer networks designed to support 1.2 GHz. For the more abundant 750 or 860 MHz networks, if not already fully utilized, it is expected that use of their capacity will be increased to the point of exhaustion This will happen as a result of 1) increased DOCSIS® usage for even faster high-speed data (HSD) service tiers; 2) additional high-definition (HD) programs (for broadcast [BC] and especially narrowcast [NC] services, such as video on demand [VOD] and switched digital video [SDV]); and 3) new service additions such as internet protocol (IP) video and cloud-based digital video recorder (cDVR.)

In recent years the growth in, and demand for, HD programming has resulted in the need for allocation of large numbers of EIA (Electronic Industries Association) channels for HD services, both for BC (Broadcast) and NC (Narrowcast), 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 numbers of them are, including 100 to 150 BC programs.

The above is also applicable to a great extent in systems utilizing SDV technology for content distribution. The difference is that the HD and SD versions of the program are not distributed unless a subscriber 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. All network operators have offered increased service tiers and observed an increased use of HSD service capacity for well over a decade now, as shown in Figure 1, which amounts to a constant year-over-year compounded growth. The applications have changed throughout this time, and the demand has continued to increase at the same relentless rate.

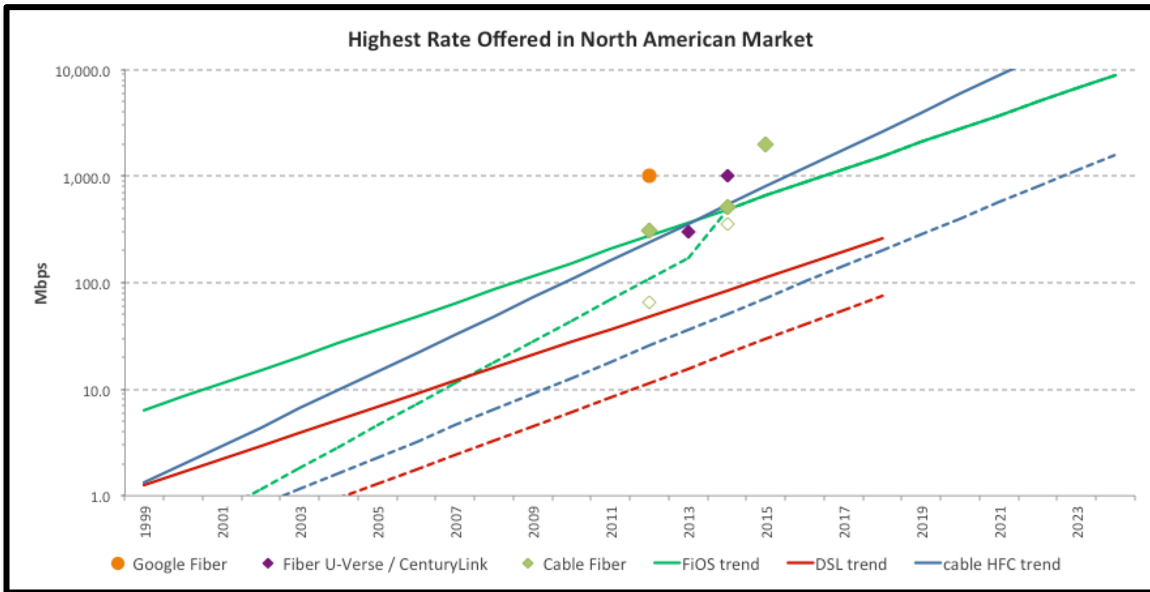

Figure 1: Examples of HSD service tier capacity increase over time

How does this compare to other operator's data services and a longer period? Projecting an operator's HSD service growth back in time to when Internet services started 25+ years ago, services should have been about 100 bps. This coincides with the history of telephone modems from 110 and 300 baud modems from the mid-80s, to 56 Kbps/V.42, into ISDN (Integrated Services Integrated Network) 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 operators in recent years.

2.2. Growth Projections

From all of the above, it follows that, should the usage growth pattern continue at the same rate as in the past, networks will be required to provide >1 Gbps HSD services within the next few years. This growth, coupled with the surge in HD video formats, and more personalized narrowcast services, will result in a significant growth in narrowcast capacity, as shown in Figure 2.

To support this growth, MSOs have deployed bandwidth reclamation tools such as SDV for digital broadcast, digital terminal adapters (DTAs) for analog service reclamation, 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 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 to reduce bandwidth required for existing services. As SDV's role in the network grew, the efficiencies have been even greater, especially as SDV expanded in scope to support niche service introductions that have low initial 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 tier analog customers, such as only the traditional expanded basic subscribers, the move 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 digital-equivalent transport. Using the same comparison method as the above SDV case, this is a >90% efficiency. If extended to the entire analog tier the efficiency gains are very significant.

Despite the availability of these tools, they are not universally applicable. With respect to SDV, in general it is not likely that all broadcast programs will be switched since experience shows that many broadcast programs are constantly viewed by someone in the service group during peak hours, which will leave a large portion of the spectrum still used for broadcast. Similarly, not all analog channels can be removed in the short term due to operational and/or cost constraints.

Additionally, while many 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 compression standards (VCS), and variable bit-rate (VBR) multiplexing. In the case of VCS, coding efficiencies of approximately 50%, depending on implementation and content type, can be obtained with H.264 / MPEG-4 Part 10 . Furthermore, with the release of the H.265 standard in April of 2013, it is possible to achieve a 50% improvement over H.264. The use of VBR promises to result in a capacity efficiency gain of as much as 70% versus constant bit rate (CBR) . The combined gains from using the above approaches for multiple services are even more significant.

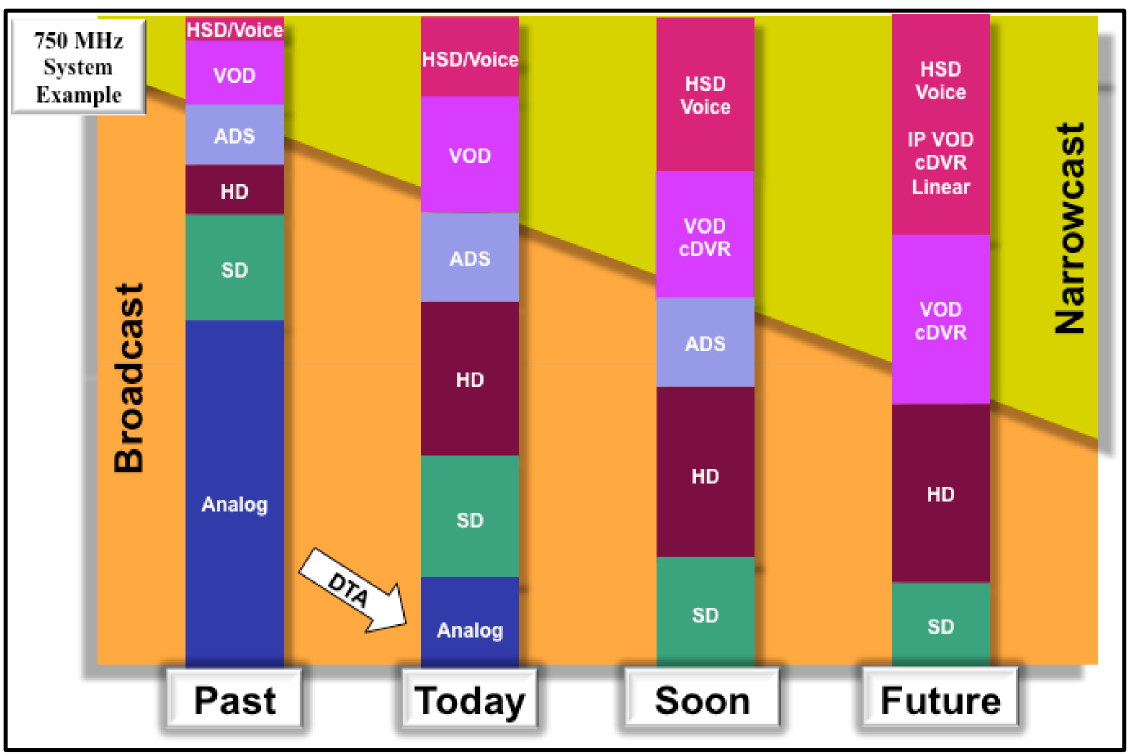

Figure 2: Example of narrowcast service growth over time

However, these are difficult tools to take advantage of, from a network perspective, because on a proportional basis, relatively few legacy set-tops will support all these technical advances, especially

H.265. These tools are more likely to find significant support in equipment designed to handle newer, IP-video based services.

And, this approach will nonetheless require additional capacity from the network. This is especially true when considering that the deployment of these advanced video services will result in an additional simulcast of video programs, at least initially. This is because, realistically, advanced services will not, at least initially, replace the currently deployed service formats

Furthermore, ubiquitous support for such devices would require considerable spectrum if the legacy services are maintained for an extended period-- which is expected, given that legacy devices will continue to be deployed, for some amount of time.. Moreover, this increase in simultaneous use of the more advanced IP video services while maintaining legacy services will be especially impactful over time, as the number of IP video services increases.

All of the above, coupled with the success experienced by MSOs in recent years with business services, homes security, etc., will likely require the deployment of IP capacity beyond what can be supported today. As well, it will require the development of tools for increased spectral efficiency and/or unleashing additional spectrum in the HFC network. The following sections of this paper will enumerate some ways in which this can be achieved.

2.3. The Advent of DOCSIS 3.1

As it has been pretty well advertised in the media, DOCSIS 3.1 development has been quite extensive. Most MSOs announced deployments of DOCSIS 3.1 across their markets, and several operators have even deployed DOCSIS 3.1 throughout their entire footprint.

The key motivation for the 3.1 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 of DOCSIS deployments it was possible to offer Internet services and support its growth with just one downstream DOCSIS channel, over the last 5-10 years services have required many more channels. This is because the year-over-year growth drove service speeds well above the initial levels, to speeds of 50 Mbps, 100 Mbps, and even much higher Mbps tiers today, which can’t be supported by the single channel. Therefore, MSOs deployed multiple DOCSIS channels using DOCSIS 3.0, sometimes using 32 or more channels, and even requiring capacity beyond that supported by DOCSIS 3.0.

To that end, the 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:

    a. The use of more efficient forward error correction (i.e., replacing the older and less efficient Reed-Solomon approach with the far more efficient Low Density Parity Check (LDPC), and

    b. The 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 per Hertz/second of improvement in both the upstream and downstream signal directions, while the use of the new forward error correction approach provides approximately 5 dB better RF performance. The end result is that MSOs are able to transport 1 Gbps of DOCSIS capacity in about 120 MHz of spectrum. For context, doing the same with DOCSIS 3.0, using 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 (Orthogonal Frequency Division Multiple Access)for the upstream should enable MSOs to reduce costs by "packing" more bits in the HFC network more efficiently. As a result, these technologies will likely attract a larger supplier ecosystem, increasing innovation and fueling competition.

3. Enable a simple and orderly transition strategy, both with respect to compatibility with the previous generation of CMTS and CM equipment while simultaneously supporting an expanded spectrum capacity in the HFC network.

Specifically, DOCSIS 3.1 cable modems operate with DOCSIS 2.0 and 3.0 CMTS/CCAP equipment, enabling deployment of DOCSIS 3.1 CPE (Customer Premise Equipment) as soon as available. Similarly, DOCSIS 3.1 CCAPs 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 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.

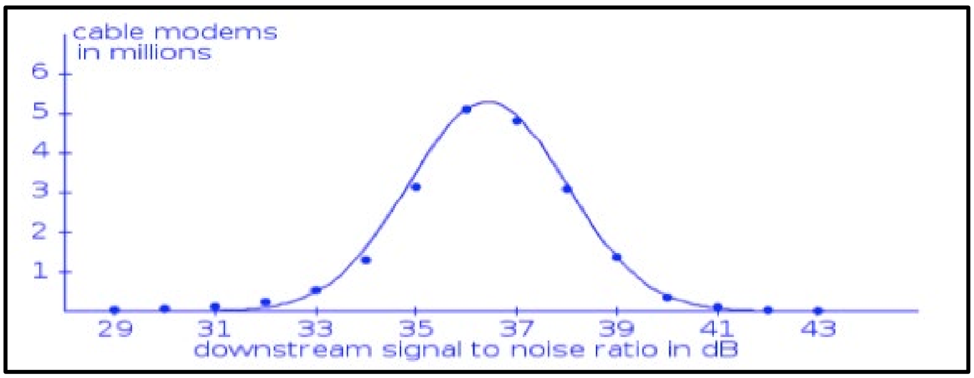

Figure 3: Example of downstream SNR for a large population of cable modems

Figure 3 depicts the downstream signal-to-noise ratio (SNR) as reported by a very large population of cable modems. This data shows that many cable modems will be able to support the high-order modulation profiles included in DOCSIS 3.1.

Modulation	Signal-to-Noise Ratio
512 QAM	27 dB
1024 QAM	30 dB
2048 QAM	33 dB
4096 QAM	36 dB
8196 QAM[1]	39 dB
16384 QAM	42 dB

Table 1: SNR required for DOCSIS 3.1

[1] 8196 QAM and 16384 QAM are included for future consideration in the DOCSIS 3.1 specifications

Assuming an 8/9 LDPC coding ratio, Table 1 shows the required SNR for the modulation rates included in DOCSIS 3.1:

Applying the SNR requirements from Table 1 to the population of modems shown in Figure 1, 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.

2.4. 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, as a means to improve performance, was originally developed with analog modulated lasers and receivers in both signal directions, upstream and downstream.

Over time, the performance of the upstream link was improved by replacing the analog modulation with a digital transport. This 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 needed to transport the downstream spectrum has been considerably higher, reaching and even exceeding 10 gigabits per second.

Because of the above, analog forward links continue to be used. Even though headend equipment is currently capable of launching signals with >47 dB MER performance, which is 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 in the best of the cases.

2.5. Description of Options for Digital Forward Link

As time has gone by, technology evolution and certain developments as described below have enabled options for implementing a digital forward link. These include:

1. Evolution of 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 or completely.

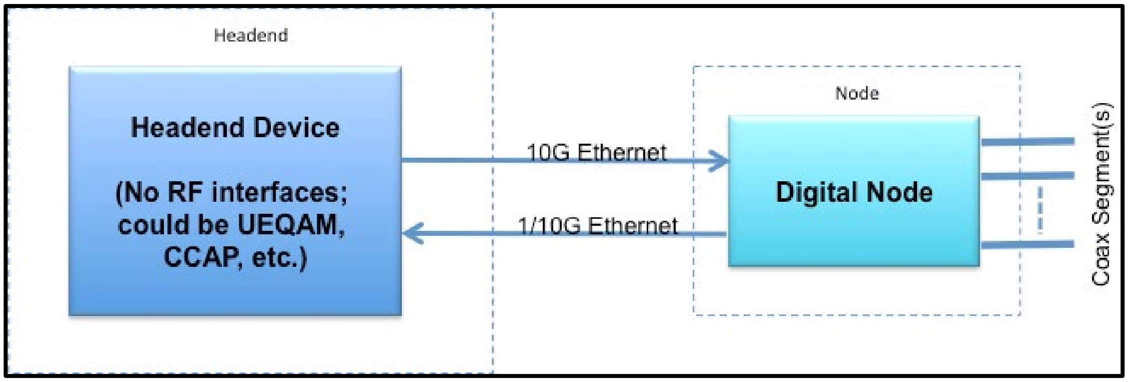

Figure 4: Digital Forward - High-level Architecture

With this technological evolution, it is conceivable to remove the RF combining network, and instead implement it digitally in the edge device, such as the CCAP.

This evolution of the edge headend devices makes it possible to envision several options for digitizing the forward link.

Fundamentally, the migration to a digital forward includes the components included in Figure 4, as follows:

- 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 generates 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

2.6. Comparison of Options for Digital Forward Link

There are pros and cons for each of the options. The following sections outline these tradeoffs.

2.6.1. Option 1: RF remains in the headend

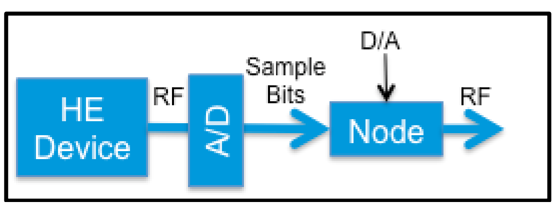

Figure 5: Block diagram for Option 1

- Equivalent to digital return, the RF output from the headend device is digitized, transported digitally, and converted back to RF in the node.
- 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 (Modulation Error Ratio) 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

2.6.2. Option 2: Digital-to-analog conversion is moved to the node

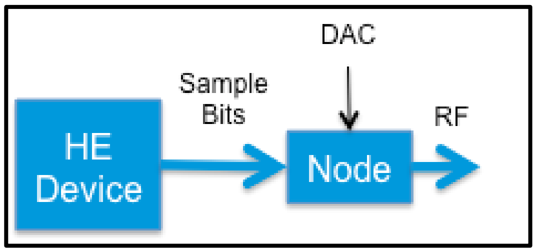

Figure 6: Block diagram for Option 2

- 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

2.6.3. Option 3: Lower PHY is moved to the node

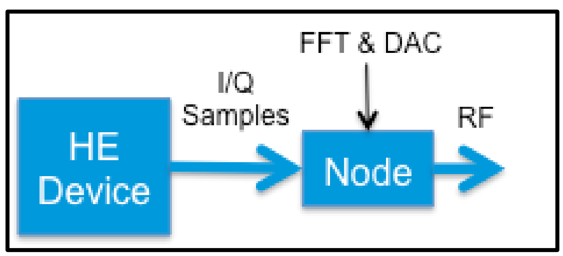

Figure 7: Block diagram for Option 3

- 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 an 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

2.6.4. Option 4: Entire PHY is moved to the node

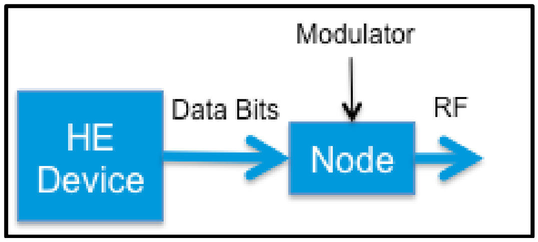

Figure 8: Block diagram for Option 4

- 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

2.6.5. Option 5: Move PHY and MAC to the node

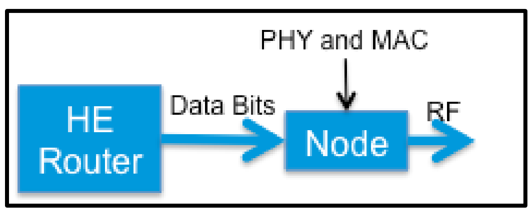

Figure 9: Block diagram for Option 5

- This option puts the most intelligence in the node
- The data rate between the headend and the node is equivalent to the actual data transmitted, except for the addition of ancillary network data
- Same packet-based network benefits as Option 4
- Same highest MER performance as Option 4

Either of the 5 options described above accomplishes a migration away from the analog-modulated forward link and into a new era where the link between the headend and the node becomes a digital link. And, while either of the above approaches accomplishes a migration towards a digital link, over the years options 4 and 5 have received the most attention because of their relative implementation simplicity versus options 1, 2 and 3. We now call these options Remote PHY and Remote

MAC(Media Access Control)-PHY, and we call the devices that implement them RPDs (Remote PHY Devices) and RMDs (Remote MAC-PHY Devices) respectively.

As we migrated towards the implementation of a digital link, and separated either the physical layer in a Remote PHY implementation, or also migrated the MAC in a Remote MAC-PHY implementation, we stepped into the era of the Distributed Architectures. In these Distributed Architectures, the remainder of the CCAP in the headend no longer needs to be implemented in an application-specific hardware design. Instead, the remainder of the CCAP in the headend can be implemented entirely as software running in general purpose compute platforms, which we now call a Virtualized headend platform.

The next sections of this paper will focus on the benefits of a distributed architecture, discuss some of the features of distributed architectures, and outline network evolution strategies.

3. Benefits of Distributed Architectures

There are many benefits from the implementation of Distributed Architectures. The following sections of this paper describe them.

3.1. Improved performance

Improvements on performance are achieved in multiple ways, including:

- Improved SNR characteristics
- Longer link distances
- Higher reliability
- Better use of capacity

As described in the above sections of this paper, one key benefit of Digital Forward Link is the improved performance resulting from the migration from an analog to a digital link. This gain varies depending on the characteristics of the analog link being replaced, but can be generalized as an improvement of 5 dB in 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 the higher order modulations as shown in Figure 3 and Table 1 for more of the cable modems in the network. This will enable significantly higher transport capacity for customers in the HFC network.

In addition, the Digital Forward Link will enable longer distances between the headend and the node. This is because digital interfaces, such as an Ethernet link, are designed to operate over much longer distances while carrying the designated capacity. Extending the distance between the CCAP and the Digital Node would enable 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 the Digital Forward Link is improved reliability of the optical link. It is well known that analog links require periodic maintenance and are subject to the effects of environmental changes. By contrast, Ethernet optical links are far more stable across a wider range of environmental conditions, and require little to no maintenance. The impact of this benefit could be very significant to MSOs.

Finally, the data transmitted through the link can be used more efficiently. One key example of such efficiency is the case where one link is used for multiple remote devices. As shown in Figure 10, one link from the headend CCAP device can be used to transport broadcast services once for multiple remote devices. This is achieved by using multicast addressing, whereby each of the remote devices uses the same lineup for each of the respective service groups. In doing so, a single link from the headend CCAP can be used for all the remote devices without exhausting the transport link capacity.

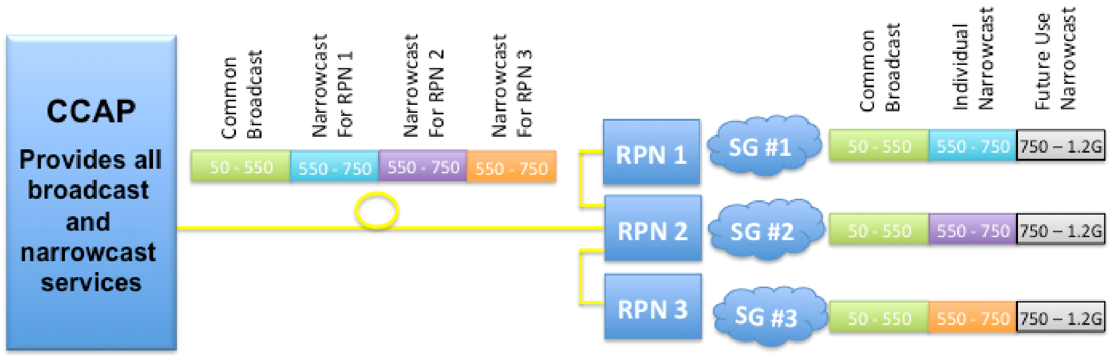

Figure 10: Reuse of broadcast capacity across multiple RPNs

3.2. Increased Headend Equipment Density

The implementation of distributed architectures 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 distributed architecture line card could implement both upstream and downstream. This, in effect, doubles the capacity of a CCAP chassis.

In addition, a typical CCAP downstream line card will house 8 or perhaps 12 RF ports, as defined by the printed circuit board space consumed by the components required for RF modulation, plus the sheer connector spacing required. However, Ethernet connectors can be placed considerably closer to one another, allowing a similar line card to easily house 16 to 24 ports. This additional density gain once again doubles the capacity of a CCAP chassis.

Finally, it is possible to consider “daisy chaining” digital nodes (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, it is possible to generate an RPN “channel lineup” by transmitting the broadcast content once to multiple RPNs. As depicted in Figure 10, the data stream transmitted from the CCAP could contain a single “copy” of the broadcast line-up content, plus individual versions of the narrowcast content for each of the RPNs. The RPNs would then reuse the broadcast lineup content to recreate the individual RPN channel lineup. In this way, each service group served by the CCAP port would contain the same broadcast lineup while allowing for its own unique narrowcast line-up.

Then, as the narrowcast lineup capacity grows over time, CCAP ports would be segmented to support less RPNs, akin to the way service groups are split today to provide more narrowcast capacity as it is required.

As summarized in Table 3 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 can facilitate huge 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 3: Distributed architecture headend density gain

But, just how meaningful is this headend density gain?

Consider: A migration from an HFC architecture with an average of N+5 (meaning an optical-to-RF node followed by 5 cascading amplifiers) to N+0 would require about 10x the number of nodes, and the headend density benefits resulting from the distributed architecture would neutralize the potential increase in CCAP equipment.

It is then quite clear that from a space and power savings, distributed architectures take the benefit of CCAP to a whole new level.

3.3. Integration of HFC and Fiber Services

One of the largest areas of growth for 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 (Passive Optical Network) technologies (either EPON [Ethernet Passive Optical Network] or GPON [Gigabit Passive Optical Network]).

This duality results in the existence of two parallel networks. One of them, the HFC infrastructure, uses fiber from the headend to the node via an analog 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 the RPN, and especially because this digital fiber is based on Ethernet technology, it is possible to collapse both of these networks into a single infrastructure. Even without fully collapsing the Ethernet network better utilization of physical fiber can be enabled by the move to common DWDM wavelengths and multiplexers.

Therefore, the implementation of RPNs with an Ethernet interface between the CCAP and the RPN would make it possible to implement a PON interface at the RPN.

The benefits from this integration include:

- Reduction of the optical link for PON to the distance between the node to the customer premise

  - The typical distance from a node to a customer premise in an N+0 architecture is 1-2 kilometers. This would virtually eliminate any distance limitations for PON, making it possible to implement the largest possible densities.

  - In addition, this shortened distance would enable the use of lower power optics, which can translate into significant savings -- especially for 10 Gbps optics, and 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.

3.4. 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 distributed architectures allow for unsynchronized changes.

Furthermore, ideally the migration to distributed architectures 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 out, distributed architectures enable such gradual, unsynchronized and opportunistic changes in the network. What follows is an overview of the steps and components involved in the migration to distributed architectures.

Starting with the components of the network on both sides of the distributed architectures, 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 distributed architectures, and any additional MIBs for management and/or commands for configuration as needed can be added well before the first distributed architecture 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 distributed architectures, and any enhancements that are made possible through the introduction of distributed architectures would be implemented in CPE equipment that can be introduced before or after the migration to distributed architectures.

The critical portion of the network where changes need to be made are in the headend and the plant.

To begin with, the changes required in the headend are primarily in the CCAP platform. The CCAP architecture was specifically designed to support multiple technologies simultaneously, which makes it possible to install regular RF upstream and downstream line cards and distributed architecture line cards in the same chassis. While some MSOs may choose to deploy a separate CCAP platform for distributed architectures, it is certainly possible to support both types of line cards in the same chassis. Of course, these distributed architecture 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 distributed architecture line cards or nodes.

Turning our attention to the plant, it is similarly possible to migrate regular nodes to distributed architecture 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 distributed architecture.

Figure 11 depicts a single HFC node connected to a CCAP device.

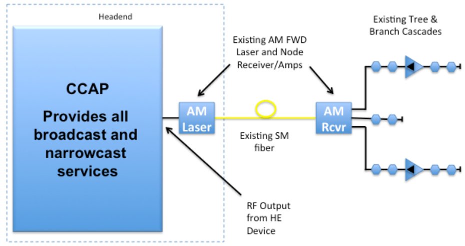

Figure 11: Single traditional HFC node

Figure 12 shows how the HFC node would be converted to RPN while the rest of the HFC network remains unchanged. The distributed architecture line card in the CCAP would have been deployed in the headend in a prior activity, and even the RPN could have been deployed before the day of the cutover. Then on the day of the change the fiber cable could be swung in the headend from one AM laser to the CCAP distributed architecture card, and in the field from the HFC node to the RPN. Of course it is not necessary to perform the migration in such a fashion, but it would be possible if desired.

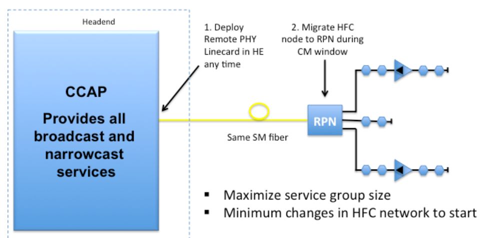

Figure 12: RPN deployment step 1

Figure 13 depicts a possible step 2 in the process, whereby additional RPNs are installed to segment the original service group further. These additional RPNs could be daisy chained from the original RPN 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 RPNs in an N+0 HFC architecture with some turnaround of passive components.

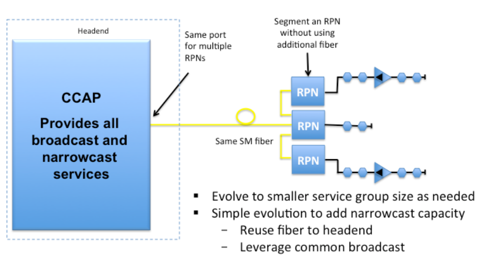

Figure 13: RPN deployment step 2

Figure 14 shows how further segmentation could take place by replacing the remaining amplifiers in the network with RPNs.

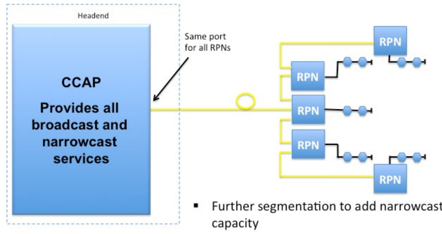

Figure 14: RPN deployment step 3

Figure 15 shows that the RPN service group depicted above is segmented as additional narrowcast capacity is required. In this example, 2 of the RPNs from the distributed architecture service group shown in Figure 14 are split into separate service groups using separate CCAP ports.

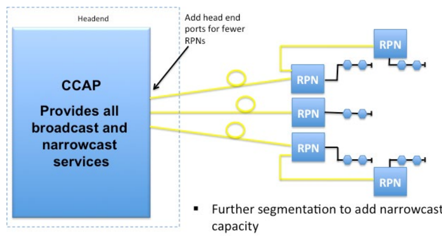

Figure 15: RPN deployment step 4

Eventually each of the RPNs could be connected to an individual CCAP RPN port as shown in Figure 16. This would provide up to 10 Gbps of capacity to each RPN. This could, for example, be desirable to provide both RF and PON services from the RPN.

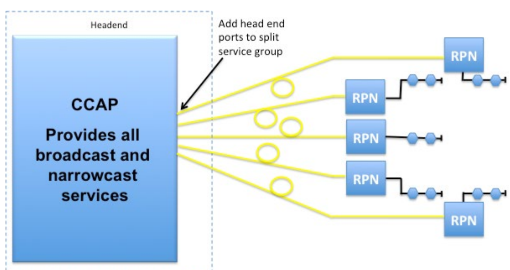

Figure 16: RPN deployment step 5

Similarly, the distributed architecture line card in the CCAP could be upgraded to support even more capacity as such capacity is needed and becomes cost effective. For example, the Ethernet link from the CCAP to the RPN could eventually be upgraded to 40 or 100 Gbps, both of which are already commercially available.

3.5. From Today to Virtual CCAP

As the network has to continue in operation through the transition, virtualizing the CCAP requires careful planning and a sensibly staged process. As with roads, where cars must be kept moving during any lengthy highway reconstruction, in the network customer traffic must continue flowing day after day. In a sense, while road work is visible to car drivers, in a network the modifications remain invisible to the end user.

One way to do so is to migrate individual functions, one at a time. So, one must develop a list of the functions that would be virtualized, and this list would be prioritized, such as on the basis of complexity of implementation and benefit. Those features with the lowest implementation complexity and the highest benefit would be prioritized higher in the list, and consequently implemented first.

In DOCSIS 3.1, one of the functions that would rise to the top of any such list is Modulation Profile Management (MPM). This is because MPM will take time to be implemented by vendors in a CCAP chassis, but implementing externally via virtualization could be quite simple. In the process, its benefit to operators is quite significant since it would enable better efficiencies from DOCSIS 3.1.

Over time, implementing virtualization of the various functions of the CCAP would lead to a completely virtualized CCAP platform. Such a platform would be more easily scalable than CCAP platforms are today, where segmentation of service groups requires the addition of more chassis in a linear relationship fashion.

In addition, and perhaps more importantly, virtualizing the CCAP will enable the development of additional functionality, and improvements to such functionality, to occur much more rapidly than it is possible to do today.

4. DAA Components, Use Cases and Generations

As MSOs move forward with the implementation of Distributed Access Architectures, there are already many useful lessons learned. The following sections of this paper describe three key aspects: DAA components, use cases, and the generational aspects of DAA.

4.1. DAA Components

As the link from the headend to the node is converted from analog modulated forward to digital, using Ethernet as the transport, several approaches can be taken for the implementation of the remaining headend components. In this section we will examine one approach in some detail, for which a key goal is to convert all required components into functionally individual software pieces implemented independently.

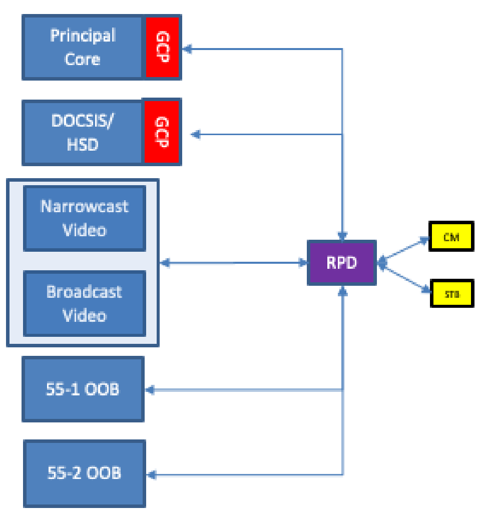

Figure 17: DAA implementation components

4.1.1. Advantages and disadvantages of discrete components

As shown in Figure 17, an implementation approach for DAA is to develop discrete SW components for each of the various DAA components. Some of the advantages for doing so include:

- The implementation can consist of a multi-vendor platform, where each component can be developed by a different party.
- By having smaller functional components their implementation tends to be simpler.
- Time to market also tends to be reduced.

However, implementation of smaller discrete components has its downsides, such as:

- There is an implied requirement to more tightly specify the behavior of each component to ensure that the overall system will operate as intended.
- Interface specifications between the various components is required.
- Management of the various components, including their configuration and upgrade is generally more complicated, and requires more elaborate orchestration.

4.1.2. Key DAA discrete components

The key components depicted in Figure 17 include:

- The Principal Core, which is the first component that the RPD will contact after receiving an IP address.
    - The Principal Core is implemented such that it will configure all the RPD functions except DOCSIS channels and behavior, which will be implemented by the DOCSIS CMTS.
    - As depicted in Figure 17, the Principal Core communicates with the RPD using the GCP protocol, for which it is known as the GCP Principal, or GCPP for short.
    - Included in the GCP Principal Core are all the non-DOCSIS command and control functions for the RPD, including configuration, management and reporting.
- The DOCSIS Core, which is the second component that the RPD will contact in the network.
    - The DOCSIS Core also communicates with the RPD using GCP.
    - The DOCSIS Core provides all configuration, command and control for DOCSIS channels, both downstream and upstream.
- Narrowcast and Broadcast Video engines

    - Implemented as separate components, the Narrowcast and Broadcast Video engines provide all the content services for the various RPDs in the network.
    - Neither the Narrowcast nor the Broadcast Video engines communicate with, nor have knowledge of, the RPDs.
    - Services are configured statically in the Narrowcast and Broadcast Video engines upon their bring-up, and are multicasted to all RPDs, which listen for these services as configured by the GCP Principal.
    - The Narrowcast and Broadcast Video engines could be implemented separately, but they could be operated together as a single functional system.
- Out-of-Band engines or cores
    - The OOB (Out Of Band) functions are implemented separately from the video engines.
    - Given that video systems are implemented using a single encryption and command/control technology, only one (i.e., either SCTE 55-1 or SCTE 55-2) of them is deployed in any one system.
    - The OOB function may or may not implement GCP for communicating with the RPD. When GCP is implemented the OOB server is a Core, and it will configure the OOB downstream and upstream OOB channels in the RPD. However, when GCP is not implemented the OOB server is an engine, and the GCP Principal will configure the downstream and upstream OOB channels.
- Finally, not depicted in Figure 17 is a very important component: the Timing Server.
    - The Timing Server, also known as the Grand Master, provides the critical timing synchronization for all the DAA components.
    - Each of the DAA components will include a Timing Client, which will communicate with the Timing Server to maintain timing synchronization.
    - While timing synchronization is not absolutely critical for video services, it is imperative for DOCSIS service to operate. Therefore, video services may be initiated before timing synchronization is achieved, but DOCSIS services will not.

4.2. Key aspects of DAA interoperability

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

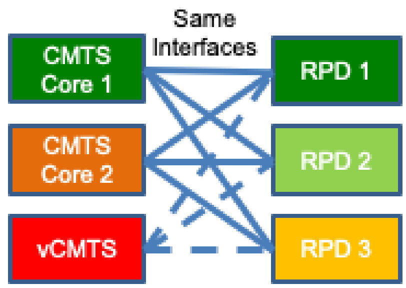

Figure 18: Functional CMTS-RPD interoperability matrix

As depicted in Figure 18, the number of combinations of interoperable components increases geometrically as additional components are added on either side of the interoperability matrix. Having a single CMTS interoperate with multiple vendors' RPDs is complex and requires a lot of careful planning and implementation. If the number of CMTS implementations is increased to 2 or 3, the interoperability complexity doubles and triples respectively.

When considering the overall CCAP system, the complexity to achieve multi-vendor interoperability is even larger. For example, if multiple GCP Principals and/or multiple video engines and/or multiple OOB engines/cores are introduced into the mix, the amount of complexity and work required for testing and interoperability results in increases by orders of magnitude.

Therefore, a multi-vendor RPD deployment coupled with a single headend implementation is a sensible approach to an interoperable DAA ecosystem.

4.3. Use Cases

In the same way as there are different kinds of nodes for different HFC network applications, there are Remote PHY devices with different characteristics that are best suited for each of the specific HFC network use cases. Similarly, while there are use cases for Remote PHY devices in the outside plant, there are also applications for Remote PHY devices in headends, or "inside plant" as it is frequently called, which will have different implementation characteristics. The following sections cover the key scenarios.

4.3.1. Outside plant

The environmental characteristics of Remote PHY devices developed for outside plant make the design of such devices very different than for RPDs developed for inside plant. The key characteristics for outside plant RPDs are as follows:

- Designs must conform to very tight space availability requirements inside of a node enclosure
- RPDs must support an environment where heat dissipation without the use of fans is critical
- Powered from quasi-square wave power supplies used in HFC networks
- Minimize power consumption to the extent possible given the limited amount of power during normal operation and especially during stand-by power mode

In addition, and perhaps more importantly, there are different kinds of nodes for different HFC network applications, which will drive varying designs for RPDs for outside plant, as follows:

- Traditional HFC networks include cascades of multiple amplifiers and cover a plant footprint of a few hundred homes. In such cases the network capacity offered by the Remote PHY device should be maximized, such as including multiple downstream and multiple upstream ports.

- Newer HFC networks are built with less, or even no amplifiers, and are targeted to cover smaller network footprints. In such cases it is not necessary for the RPD to support much more than a single downstream port, with either a single or dual upstream ports.

- Finally, given that scaling is needed as in any other network application, it should be possible to support greater capacity over time to the extent possible. For example, while initial deployment may only require a single downstream and/or upstream, over time service group segmentation may require additional downstream and/or upstream ports in a single node. For that purpose it is usually a design requirement that multiple individual RPDs fit into a single node enclosure.

Given the above, Remote PHY devices that are built with a single downstream and a single upstream (frequently called 1x1) or a single downstream with dual upstreams (1x2), such that MSOs can place a single one in the node or place additional units when capacity demands require it. Newer silicon designs include more capacity at lower power levels, making it possible to develop RPDs that contain multiple downstream ports and multiple upstream ports, such as 2x2 and 2x4 designs in a single Remote PHY device.

4.3.2. Inside plant

In contrast to the outside plant environmental characteristics, Remote PHY devices developed for inside plant have other constraints that make the design of such devices very different than for RPDs developed for outside plant. The key differences in the design of RPDs for inside plant are as follows:

- Rather than the physical volume of the allocated space, the layout is a primary concern so that dense set-ups in a rack are possible, including cabling distribution in the front and/or back of the rack

- Designs must support an environment where forced air is used for heat dissipation, requiring airflow from front-to-back or back-to-front, sometimes allowing airflow from side-to-side and/or in a vertical direction

- Powering frequently requires DC power supplies, but AC power supplies are used in other cases

In the case of inside plant RPDs, these are frequently implemented in one of two different form factors:

- Modular, where RPDs are individually removable in a chassis-based design, or

- Fixed, where the entire set of RPDs are part of a monolithic device

The modular design is generally used in larger headends where the ability to replace a defective individual unit is a paramount concern. The fixed design is targeted for a smaller facility, or even a cabinet, where space is the primary concern.

4.4. Generational considerations

One final consideration for this paper is the evolution of the Digital Access Architecture to support new generations of equipment.

The initial implementation of the DAA components included support for DOCSIS 3.1. The main component that is specifically developed and implemented for DOCSIS 3.1 is the Remote PHY device, which incorporates ASIC devices which support up to DOCSIS 3.1, but will not support DOCSIS 4.0 functionality.

As newer parts of the DOCSIS 4.0 specifications are implemented, such as FDX, the RPDs will have to be swapped out in order to expand their support for DOCSIS 4.0. This process is akin to what had to be done with CCAP linecards in the past, where either upstream and/or downstream linecards are swapped over time as new versions become available. And, as part of this upgrade, the older equipment is reused in other locations where the newer equipment is not yet needed.

However, for the remaining DAA components, if these are implemented in software on general purpose compute platforms, these should be upgradeable to support newer DOCSIS specifications such as DOCSIS 4.0 and/or other enhancements by simply expanding the functionality implemented in software and downloading it to the platforms in which they run.

In fact, the process for upgrading the DAA components to support changes in the functionality including enhancements to DOCSIS, becomes easier than ever before given the nature of the DAA platform, especially when the DAA implementation includes a minimalistic Remote PHY device at the edge of the HFC network and virtualized components for the remainder of the DAA components.

5. Conclusions

Demand for more narrowcast service capacity has driven many changes for MSOs to reclaim spectrum. Splitting nodes to reduce the number of homes passed in an HFC node's footprint is another method used to provide more narrowcast capacity. Enabling narrowcast services on this reclaimed and newly added spectrum requires more equipment in the headend. Given the trajectory of growth using existing equipment will create demand for more headend space. Deploying DAA allows for this growth without expanding facility footprint.

DOCSIS 3.1 created a demand for higher MER out of the node to take advantage of the highest modulation orders available. Transitioning to a digital forward link is the enabling function to achieve this demand. Moving the entire PHY layer to the node has become the standard method to implement a digital forward link.

Deploying DAA in the outside plant allows for more efficient infrastructure utilization between services. DAA and commercial services can share physical fiber due to common DWDM wavelengths and spacing. There is also the possibility of a converged Ethernet switching network to serve both DAA nodes and commercial customers. Having Ethernet inside of a DAA node allows for efficient PON deployments sourced from the same node.

DAA migration scenarios were outlined showing how RPNs can be initially deployed and scaled to offer more bandwidth in their lifecycle. This includes capacity on both the RF and ethernet sides.

Moving from specialized CCAP hardware to a virtualized system can lead to a more scalable system where scale can be added without the step function of adding new CCAP chassis. Feature velocity can be improved on a virtualized system.

Deploying DAA can be done with a CCAP core that provides all needed services or with discrete components that allow for a multi-vendor platform that includes smaller functional components. Implementing discrete components requires well behaved components that adhere to a well-defined specification. Key DAA functions include a GCP principal core, a DOCSIS core, a narrowcast and broadcast video source, an SCTE 55-1 or 55-2 OOB source, and PTP timing distribution.

Remote PHY devices are available for both inside and outside plant applications. Power level and segmentation capability / capacity are key attributes of outside plant RPDs. Density and serviceability are key attributes of inside plant RPDs.

Even though current generation RPDs lack support for FDX operation the other DAA components may have the ability to be upgraded for FDX operation using the current hardware. This enables an easier transition to FDX on an as needed basis later by swapping hardware on desired nodes while utilizing the existing platform.

Abbreviations

A/D	analog-to-digital
AC	alternating current
AM	analog modulated
ASIC	application specific integrated circuit
BC	Broadcast

CBR	constant bitrate
CCAP	converged cable access platform
cDVR	cloud digital video recorder
D/A	digital-to-analog
DAA	distributed access architecture
dB	Decibel
DC	direct current
DOCSIS	data over cable services interface specification
DTA	digital terminal adapter
DS	downstream
EIA	Electronics Industry Association
Gbps	gigabit per second
GCP	Generic Configuration Protocol
GHz	gigahertz
HD	high definition
HFC	hybrid fiber-coax
HSD	high speed data
Hz	Hertz
IP	Internet protocol
ISDN	integrated services digital network
Kbps	kilobits per second
MER	modulation error ratio
MSO	multiple system operator
NC	narrowcast
OFDM	orthogonal frequency division multiplexing
ONU	optical network unit
OOB	out-of-band
OTN	optical termination node
PHY	physical
PON	passive optical network
QAM	quadrature amplitude modulation
RF	radio frequency
RPD	remote PHY node
RPN	remote PHY nodes
SC-QAM	single carrier QAM
SD	standard definition
SDV	switched digital video
SNR	signal-to-noise ratio
US	upstream
VBR	variable bitrate
VCS	video compression standards
VOD	video on demand
WDM	wave division multiplexing