VOL. 5 NO. 2 **DECEMBER 2019**

# JOURNAL OF NETWORK OPERATIONS

SCTE · ISBE STANDARDS

NETWORK OPERATIONS

scte.org • isbe.org

NETWORK OPERATIONS

SCTE · ISBE

Society of Cable Telecommunications Engineers
International Society of Broadband Experts

# JOURNAL OF NETWORK OPERATIONS

**VOLUME 5, NUMBER 2**
**December 2019**

Society of Cable Telecommunications Engineers, Inc.
International Society of Broadband Experts™
140 Philips Road, Exton, PA 19341-1318

NETWORK OPERATIONS

# Table of Contents



## Technical Papers





**Editorial Correspondence:** If there are errors or omissions to the information provided in this journal, corrections may be sent to our editorial department. Address to: SCTE Journals, SCTE•ISBE, 140 Philips Road, Exton, PA 19341-1318 or email journals@scte.org.

**Accepted Submissions:** If you have ideas or topics for future journal articles, please let us know. Topics must be relevant to our membership and fall under the technologies covered by each respective journal. All submissions will be peer reviewed and published at the discretion of SCTE•ISBE. Electronic submissions are preferred, and should be submitted to SCTE Journals, SCTE•ISBE, 140 Philips Road, Exton, PA 19341-1318 or email journals@scte.org.

**Subscriptions:** Access to technical journals is a benefit of SCTE•ISBE Membership. Nonmembers can join at www.scte.org/join.

**From the Editors**

Welcome to Volume 5 Issue 2 of the *Journal of Network Operations*, a publication of collected papers by the Society of Cable Telecommunications Engineers (SCTE) and its global arm, the International Society of Broadband Experts (ISBE). In this issue we continue to focus on next generation cable networking technologies such as wireless, cloud-native CMTS, more efficient and higher capacity EPON, higher-reliability HFC powering for new technologies such as 5G, and network disaggregation. All of these play an important role in today's and tomorrow's cable networks.

HFC plant powering is a topic that isn't necessarily top-of-mind when it comes to next-gen technologies. But, without reliable and well-engineered network powering, those next-gen technologies won't work as intended. Are you thinking about powering 5G small cells from your HFC plant? Upgrading to 1.8 GHz? Or rolling out a distributed access architecture (DAA) such as remote PHY? If so, it would be a good idea to read Rob Anderson's paper, "Upgrading and Preparing the Access Network for 5G," which takes us through a review of the basics of plant powering and shows just how important it is to consider what happens when newly-installed devices place additional powering demands on the network. As well, in the event of a loss of grid power, ensuring sufficient backup time via better battery monitoring and health will keep those new technologies on-line.

Passive optical networks have been evolving over time, supporting different types of access services such as fiber-to-the-home in the residential marketplace, and fiber-to-the-office or fiber-to-the-tower in the enterprise and business space. Indeed, we now see 10 Gbps Ethernet PON (10G-EPON) as part of the network toolkit. As bandwidth and capacity needs continue to grow, what's next? Marek Hajduczenia shows us in "Nx25G-EPON: The Evolution of PON to Multi-Channel High-Capacity P2MP Access System with Native Channel Bonding." Hajduczenia does a deep dive, starting with the physical layer and wavelength allocation plans, moving next to the physical coding sublayer. He discusses advanced forward error correction, and the multi-channel reconciliation sublayer and how it provides support for channel bonding and skew mitigation, wrapping up with Multi Point Control Protocol (MPCP) and its extensions for improved upstream scheduling efficiency.

If a DAA deployment is in your plans, you likely know that rolling out remote PHY or remote MAC/PHY (aka flexible MAC architecture) means that the physical layer and MAC layer are separated. The MAC layer electronics might remain in a core in the headend or hub while the PHY layer electronics are installed in shelves or nodes. Alternatively, the MAC layer electronics might be placed somewhere else. What if that somewhere else is in, say, the cloud? In other words, move the MAC functionality to a virtualized form of a software-based cable modem termination system (CMTS) on a general-purpose server. Gurpreet Singh's paper, "Cloud-Native CMTS: Challenges and Opportunities," outlines the benefits of a cloud-native CMTS as well as the challenges both vendors and operators can face and the coordination required when going "cloud-native."

If you were to ask typical cable subscribers to describe their cable service, their descriptions would include Wi-Fi, but probably wouldn't differentiate the wired cable part from the wireless part. For many subscribers, their in-home Wi-Fi is their cable service. If you were to ask cable operators about Wi-Fi, you'd almost certainly hear just how important reliable Wi-Fi is for their subscribers, and the extent to which Wi-Fi related service calls are part of the daily routine. Authors Kurt Lumbatis and Navene Kannan, in "Quantifying Wi-Fi Coverage; 'CQ-Score' An Adaptive Quality Metric to Assess

 

Wi-Fi Coverage in a Network," discuss how to quantify the quality of a wireless device's connection to a Wi-Fi access point. This becomes even more important in homes with multiple APs. The authors propose a new tool for characterizing coverage, called a coverage quotient (CQ) score.

A favorite saying among many of us in the technical field is "the only constant is change." Bill Beesley and Francois Moore begin their paper, "Operationalizing Disaggregate Networks," with almost those same words. Not familiar with disaggregation? Beesley and Moore explain the term as "breaking apart what historically had been a large, multiservice, chassis-based optical transport system into its functional components." They discuss its benefits and operational considerations (and there is a lot to consider here). The authors argue that disaggregated networks are the future of networking and will give service providers faster availability of new technologies, as well as lower capital and operational costs.

We are grateful for the individuals who contributed to the *Journal of Network Operations*, including the authors, reviewers, and the SCTE·ISBE publications and marketing staff. We hope you enjoy this issue of the *Journal*, and that the selected papers provide inspiration for new ideas and innovations in cable network operation. If you have feedback on this issue, have a new idea, or would like to share a success story please let us know at journals@scte.org.

SCTE·ISBE *Journal of Network Operations* Senior Editors,


Ron Hranac

Technical Marketing Engineer – Cable Access Business Unit, Cisco Systems, Inc.

SCTE Fellow

Daniel Howard

Principal, Enunciant LLC.

SCTE Senior Member


 

# Upgrading and Preparing the Access Network for 5G

## Considerations for Higher Outside Plant Powering with Increased Reliability

A Technical Paper prepared for SCTE•ISBE by

Rob Anderson, Director of Product Management
Alpha Technologies Services, an EnerSys company, SCTE•ISBE Member
3767 Alpha Way
Bellingham, WA, 98226
randerson@alpha.com
360-392-2293

# Table of Contents



 

# List of Figures



 

# 1. Introduction

Using the access network for powered backhaul has been achieved by many cable operators. Surveillance cameras, Wi-Fi access points and small cell radios benefit today from hybrid fiber/coax's (HFC's) ultra-reliable backup power and DOCSIS®-enabled backhaul data services. However, the next wireless wave, 5G, will be more challenging. 5G is engineered to be smarter, faster and more efficient than 4G. It will place additional demands on the HFC network. Operators can meet new 5G network demand by preparing today.

The network impact of the upcoming 5G wave must be considered for operators to prepare their networks for these new services. The number of connected devices required for full 5G support will be daunting. The 5G "massive machine type communications" (MMTC) service category requires support for up to one million connected devices per $km^2$. The number of radios needed to support MMTC will vary by topology. One thing is certain, for even a portion of new 5G radios to be HFC powered, enhancements to the HFC power grid will be required.

5G also defines an ultra-reliable low-latency communications (URLLC) service category for bandwidth and latency sensitive applications. Today's DOCSIS 3.1 has inherent non-deterministic latencies that can measure from 20 ms to 200 ms under heavily loaded conditions, making it incompatible with URLLC services. Fortunately, advanced work underway at CableLabs® is addressing this issue.

The primary focus of this paper is to review the access network relative to powering new devices and services. We begin with an overview of HFC network power. Next, we look at the CableLabs 10G initiative and review increased demands on HFC network power associated with technologies and equipment needed to achieve the objectives of 10G. Then we dive into an in-depth review of HFC power and how backup power is provided to the coax network.

# 2. Access Network - Powering Overview

## 2.1. HFC Architecture

Power is fundamental to all aspects of the broadband network, from headend facilities through the outside plant to individual premises. Because utility power disruptions are unpredictable, operators have historically utilized uninterruptible power supply (UPS) systems for reliable power. In addition, status monitoring systems provide real-time notifications of critical events, enabling timely, manual intervention to avoid disrupting customer service. Monitoring systems and power supply intelligence support methods of power management, thereby allowing operators to address power-related issues before they become service-affecting events. The coax segment of HFC networks functions to transport both RF signals and electrical power used by network equipment along the coax. UPS systems condition and backup AC utility power prior to being injected into the HFC coax. A typical HFC network is depicted in Figure 1.

 

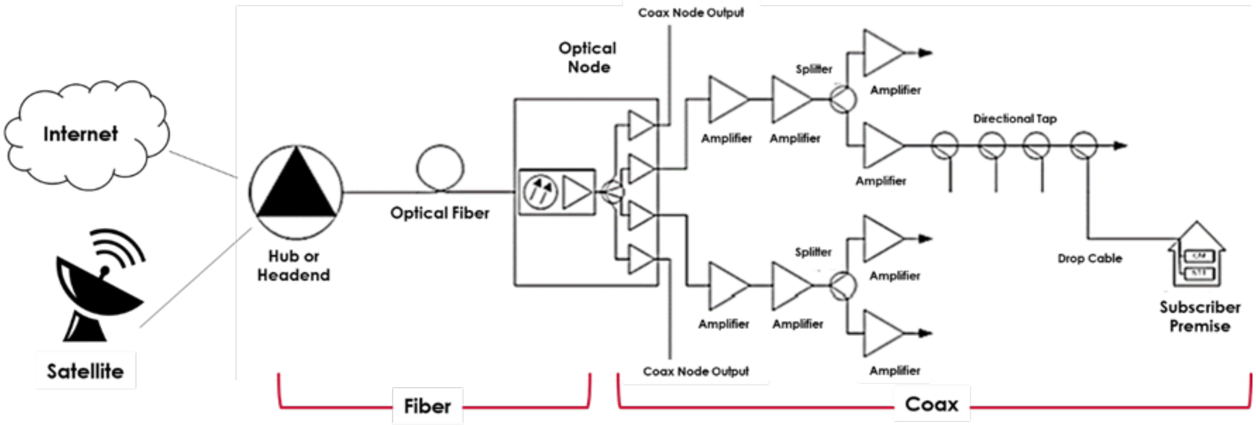


**Figure 1 – HFC Network Block Diagram**

The diagram depicts two primary network segments: a fiber segment and a coax segment. The fiber segment uses fiber optic cable to transport information between the headend or hub and multiple optical nodes (ON). ONs convert between optical and electrical signals and convey those signals to multiple coax network segments. Each coax segment typically consists of a tree-and-branch structure for signal distribution. Passive splitters subdivide the signal into multiple paths (branches). Amplifiers are located throughout the coax network segment to boost signals, providing appropriate signal levels to the end users, which are typically homes and businesses. Directional taps located near the end users "tap off" the signal from the main coax cable into drop cables which bring the signals into the customer premise.

ONs and amplifiers require power, and additional power is consumed by other devices as well as the coax line loss from Joule heating (a.k.a. $I^2R$ losses). Power and RF signals are both multiplexed onto the coax, eliminating the need for separate power cables. Power for HFC components is provided by the previously mentioned UPS systems. UPS systems are physically located throughout the coax portion of the network where required to provide power for each active network element. UPSs are physically installed on outdoor utility poles, in dedicated ground mounted enclosures and in secured utility areas of multiple dwelling units (MDUs). An example of a typical utility pole mounted broadband UPS is shown in Figure 2.

 


**Figure 2 – Typical Utility Pole Mounted Broadband UPS**

The UPS converts utility power (120 VAC or 240 VAC in North America) to 90 VAC for insertion into the coax. Early cable networks used 30 VAC then later 60 VAC for network power. Today, networks almost exclusively use 90 VAC power with a few older 60 VAC networks still in operation. Unlike utility service which provides sinusoidal AC power, the broadband UPS produces a quasi-square wave or trapezoidal shaped power output. The reason for this is discussed in a later section. Energy storage elements, typically batteries, within the UPS system, enable the power supply module in the UPS system to provide continuous, reliable power to the HFC network during utility disruptions. The output of the UPS is connected to a power inserter, which acts like a reverse directional tap, to inject power onto the coax cable. Some operators use the term "shunt" to describe the process of injecting power or shunting power onto the coax.

## 2.2. Broadband UPS Capacity

Within the HFC network each amplifier, splitter and some taps can be configured to pass power through itself and on to the next device in the network or to block power from passing through itself. The decision to pass or block power within a specific network device is determined by the operator and is based on criteria which may include:

- Do downstream devices require power?
- Does the UPS have sufficient capacity to power additional devices?
- Are additional powered devices planned for this network segment such that UPS power capacity needs to be reserved?
- Will powering additional devices cause this UPS to exceed the operator's maximum powering policy guidelines?

In practice, HFC powering architectures are complex. The following example uses a simplified HFC network segment to illustrate some important powering concepts.

 

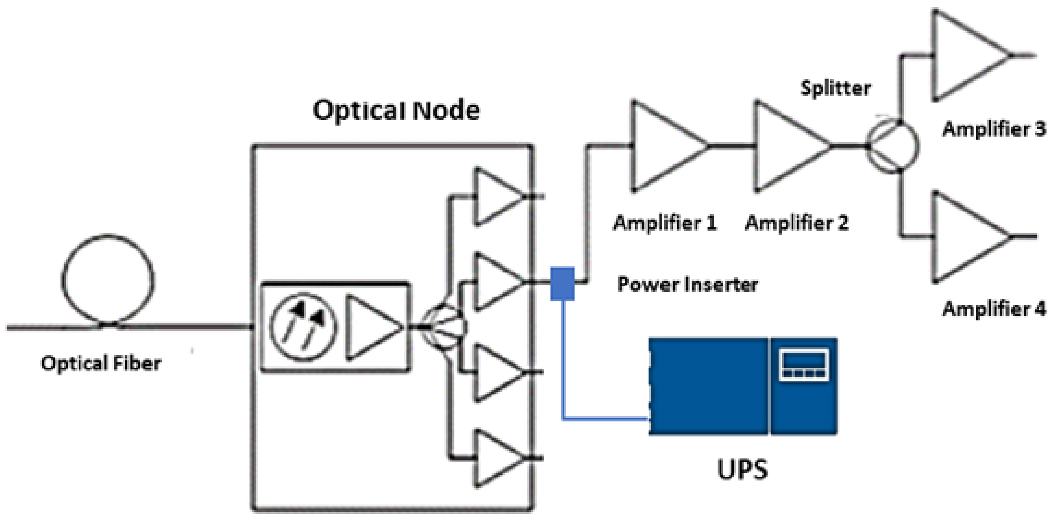


**Figure 3 – Network Powering Example**

Figure 3 – Network Powering Example, represents a simplified portion of an HFC coax network segment. In this example a broadband UPS is connected near one output of the ON. The ON is configured to be powered from this coax segment. Amplifiers 1-4 are powered from the UPS. Amplifier 1 and Amplifier 2 are each configured to pass power through their chassis, enabling downstream devices to be powered from the pictured UPS. The splitter is configured to pass power through both outputs. Amplifiers 3 and 4 are powered from the UPS and are configured to block power from passing to their respective outputs.

Assume that the ON requires 100 W of power and Amplifiers 1-4 each require 70 W of power to operate. Also assume the UPS output is configured to 90 VAC and that it has the capacity to provide up to 1350 W of power. For simplicity, ignore coax line loss. The total power required is calculated as the sum of power required from each network active:

$$P(Active) = P(ON) + P(amp1) + P(amp2) + P(amp3) + P(amp4)$$

$$\text{Total Power (Active)} = 100\text{ W} + (70\text{ W} \times 4) = 380\text{ W}$$

Next, factor in coax line loss: assume a common 0.625 inch diameter coax cable which has a typical resistance of 0.0011 ohm/ft is deployed. Further, assume the network has 1,000 ft coax spans between each element listed here:

- ON to Amplifier 1
- Amplifier 1 to Amplifier 2
- Amplifier 2 to Splitter
- Splitter to Amplifier 3
- Splitter to Amplifier 4

Using these assumptions, the total coax length between the UPS and amplifier 3 or the UPS and amplifier 4 is 4000 ft.

 

Let's calculate the power lost in a single 1000 ft segment of coax. Using Ohm's Law:

$$P(loss) = I^2R$$

Where:

P(loss) = power lost from coax line resistance, measured in watts. Note: this energy is converted to heat, hence the term Joule heating for $I^2R$ losses

I = current through the cable, measured in amperes

R = resistance of the length of cable, measured in ohms

Taking amplifier 1 in isolation, we calculate the power lost in the coax segment between the UPS and amplifier 1 as follows:

$P(loss) = (70\ W/90\ V)^2 \times (0.0011\ ohm/ft \times 1{,}000\ ft) = 0.67\ W$

In this example we used the Ohm's Law relationship: I = P/V for the first term.

This calculation is an ideal approximation. In real-world calculations we must account for cumulative voltage drops across each coax segment, i.e., the 90 VAC at the UPS output is reduced through each coax segment. The voltage drop is proportional to both cable resistance current per the relationship:

$$V(drop) = I(cable) \times R(cable)$$

HFC active elements, including ONs and amplifiers, have typically been constant power devices (early ones were constant current), although newer units may support adaptive power operation via the Adaptive Power System Interface Specification (APSIS™). Either way, as input voltage to a device is reduced due to coax line resistance, the device's current will increase to maintain the required power load (P=VI). As the current increases, power loss through the coax increases. Recall this relationship discussed earlier: $P(loss) = I^2R$.

In our example, the five active components require 380 W to operate. Additional energy consumed (lost) due to voltage drops across the various coax segments can be calculated to be an additional 38 W. While this additional power consumption from coax line loss is only 10% of the load in this case, with greater power demand per device such as to be expected from the generic access platform (GAP), the current goes up on each coax segment and the power lost due to $I^2R$ losses can exceed that due to an additional device. The calculations have been omitted for brevity, but the important take-away from this example is that network powering requires careful planning, especially if upgrades to devices as well as new devices are to be deployed. A UPS with the spare capacity to power a new device to be installed near the UPS may not have the capacity to power that same device when it installed some distance away through a length of coax due to the additional coax loss that will occur when power is consumed farther away from the UPS. The combination of voltage drop through the coax and the resulting increased current draw from the constant power load device could place the network in a state of power overload; in extreme cases this state will cause the power network to temporarily fold-back resulting in loss of service to customers. Avoiding this situation is discussed in upcoming sections.

 

# 3. Powering the 10G Network

In early 2019 CableLabs introduced the 10G platform. The 10G platform is a combination of technologies intended to deliver 10 Gbps symmetrical, low latency Internet. Technologies under the 10G umbrella include data transport using optical, wireline and wireless media. 10G is less about any one technology and more about the user experience. A user enjoying a 10G connection could be connected from their smart phone to a 5G small cell that is backhauled over DOCSIS to a distributed access architecture (DAA) optical node that is co-located with an edge server, which is hosting applications locally to provide low latency interaction. The network infrastructure upgrades that are needed for this 10G experience all require additional HFC power. This section focuses on HFC network powering needs for new technologies used to support 10G.

## 3.1. Distributed Access Architecture (R-PHY and FMA)

A distributed access architecture decentralizes and virtualizes headend and network functions. The physical RF network interface elements of the headend based converged cable access platform (CCAP) or cable modem termination system (CMTS) are moved to the ON housing. A 10 Gbps optical Ethernet connection between the headend and ON enables communications with the new distributed elements. Specific DAA implementations currently take on one of two flavors: remote PHY (R-PHY) or flexible MAC architecture (FMA). A DAA block diagram is shown in Figure 4.

The remote DAA element containing R-PHY or FMA technology is generically referred to as the remote PHY device (RPD). Powering the RPD is typically straightforward. The network fiber and coax physical interconnects remain unchanged as the RPD functionality is added to the ON housing. RPD powering comes via one of the coax connections. Because of the critical function performed by the ON and the RPD, and the larger number and/or value (as in business users) of customers that typically depend on these devices, reliable power is considered very critical and a UPS is often located a short distance away. The ON in the earlier example required 100 W to operate. The new RPD upgraded ON requires approximately 50% more power or about 150 W. Actual power requirements vary by manufacturer and features.

 

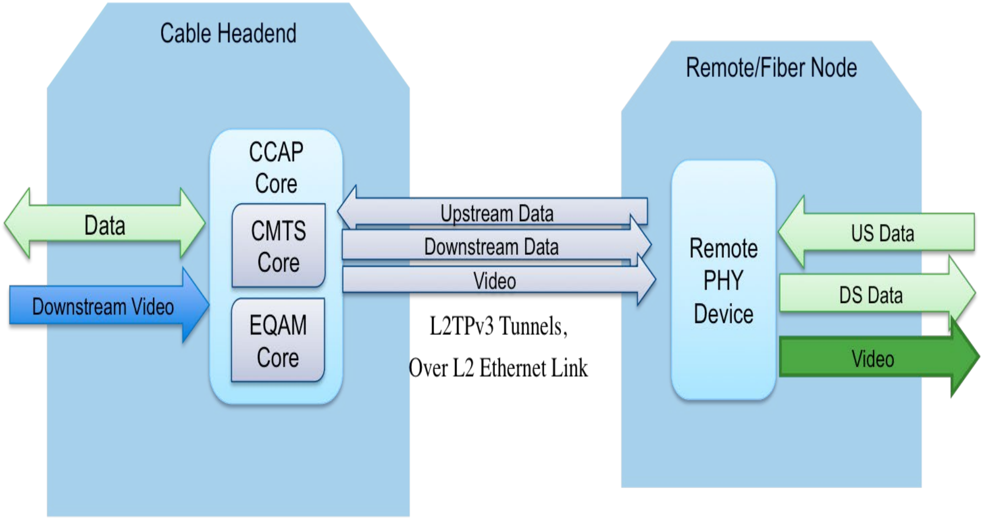


**Figure 4 – DAA Structure**

The transition from analog ONs to digital RPDs has created a new powering challenge. If power to an analog ON is disrupted, the ON output may glitch momentarily but will resume operation once power is restored. RPDs are complex software-driven systems. A power interruption causing the RPD to reset could result in the RPD being offline for several minutes while it reboots and re-establishes communications with its CCAP core counterpart. Since utility power disruptions are inevitable, a primary function of the UPS is to condition the network power, providing seamless backup power irrespective of the utility grid condition. Utility power disruptions come in many forms: outages, brownouts, surges and frequency shifts are a few of the more common powering anomalies. The UPS maintains continuous frequency and phase synchronization with utility power. When the UPS detects utility power disruptions, it switches its output from utility power to internally generated power that is phase and frequency synchronized with the utility power grid. This synchronization assures a clean transition to UPS generated power. An oscilloscope image of a broadband UPS line to backup power transition is shown in Figure 5.

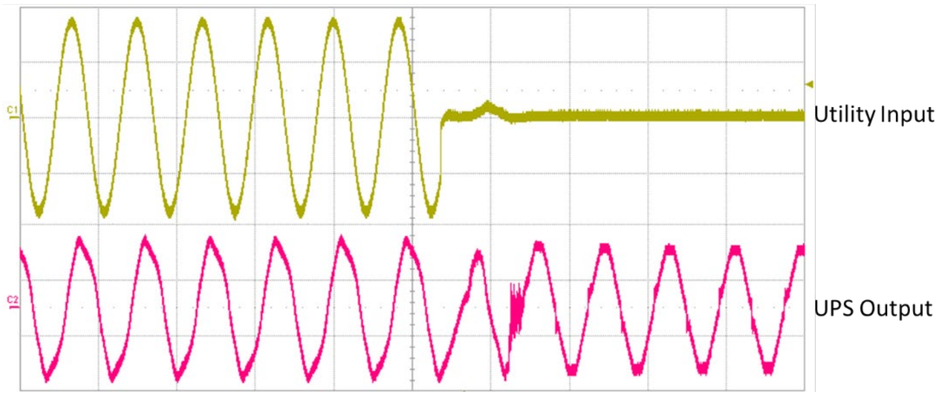


**Figure 5 – UPS Transition**

 

Operators should be aware of two considerations related to power quality for RPD-equipped nodes. First, modern broadband UPS systems can accomplish near seamless transitions between line power and backup power. From the oscilloscope image, some transition noise is visible but is easily filtered by the RPD input power pack. Second, older model UPS units may not be capable of providing line to backup power transfers as precisely as shown in this example. This may be due to older, less robust designs or it may be a result of aging components within the UPS. The takeaway here is that to avoid RPD powering issues, make sure that broadband UPS systems are up to date and that the RPD input power pack has been qualified for use with the various power transition waveforms that can occur with today's typical broadband UPS.

### 3.2. Coherent optics

Coherent optics is a technology enabling one optical wavelength to be used for concurrent transmission and reception of high-speed data. Coherent optics is not a new technology. It has been used for years in data centers and for long-haul data transport. Coherent optics technology currently enables data transfer speeds up to 400 Gbps. Coherent optics technology has not historically been used in access networks because the technology is temperature sensitive, making it unsuitable for outdoor applications. Recent developments have resulted in extending the temperature range of coherent optics, thereby making the technology appealing for outside plant (OSP) applications. Specifically, temperature hardened coherent optic small form-factor pluggable (SFP) style modules will soon be utilized in RPDs and will thus replace the current 10 Gbps Ethernet backhaul line with a 100 Gbps link. 200 Gbps capable modules are targeted for development in the near future. The following diagram depicts an RPD-equipped node utilizing coherent optics for communications back to the headend and CCAP core.

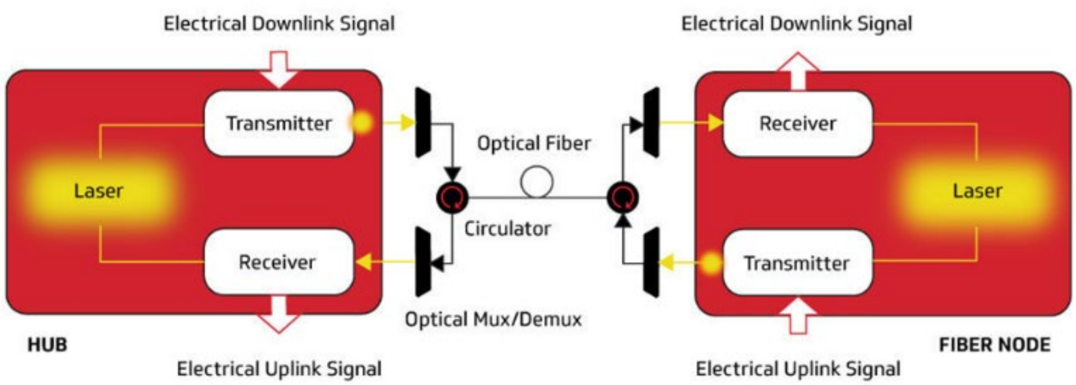


**Figure 6 – RDP with Coherent Optics Backhaul**

The addition of a 100 Gbps coherent optics module to an RPD-equipped node is expected to require about 15 W of additional power per device. This is based on early development module power measurements from optical vendors.

 

## 3.3. Edge Computing

Edge computing is described by CableLabs as follows: Edge computing brings the power of cloud computing closer to the customer premises at the network edge to compute, analyze and make decisions in real time. The goal of moving closer to the network edge—that is, within miles of the customer premises—is to boost the performance of the network, in particular to reduce the network latency, and thereby enhance the reliability of services and reduce the cost of moving data computation to distant servers, which mitigates both bandwidth and latency issues.

The edge computing concept is still under development. However, it represents an important aspect of the 10G experience. Early discussions have proposed that the edge computing server be co-located with the RPD-equipped node, perhaps within the same housing. Under this scenario, the additional power required to operate the edge server falls to the broadband UPS. As edge computing concepts evolve in the coming months the new HFC powering requirements will become clear. Current industry discussions envision an embedded server, co-located with each RPD-equipped node and requiring an additional 20 to25 W of power per ON location.

The SCTE GAP working group is defining a mechanical housing and associated backplane that could be used to support a co-installed RDP and edge computing module. This concept is being promoted and prototypes developed by at least one vendor. The additional network power consumption due to GAP devices being deployed has been estimated to be significantly greater than merely adding RPD functionality to an existing node, however the specification is still in development.

## 3.4. Deep Fiber and Full Duplex DOCSIS

With full duplex (FDX) DOCSIS the upstream and downstream traffic concurrently use the same spectrum, thereby doubling the efficiency of spectrum use. Upload speeds are significantly improved, enabling higher speed symmetrical communications and lower latency. The downside is that today’s RF amplifiers do not support full duplex communications. This requires either removing amplifiers from the full duplex signal path or developing new amplifiers that support full duplex operation. Full duplex compatible amplifiers are under industry discussion and may be an option in the coming years. Removing amplifiers from the FDX path is being done today by at least one operator in North America by re-architecting their HFC plant. New fiber is being installed deeper in the access network and new RPD-equipped nodes are being installed. With this design, the size of a service group is targeted at 50 homes per node, which enables a true N+0 architecture and paves the way for FDX to be implemented between the RPD-equipped node and the end user.

For N+0 network powering it was originally assumed that small capacity UPS systems would be co-located with one or two RPD-equipped nodes. This made sense in light of the fact that there would be no downstream amplifiers requiring power. Unfortunately, this assumption was proven unfeasible during N+0 field trials. The cost and complexity of installing new UPS systems with accompanying permits and utility power connections made this approach impractical. Instead, existing and new coax cables, along with new power feeder cables are used to carry power from existing UPS systems to new RPD-equipped nodes. While the latter approach eliminates cost and complexity of deployment, it also adds more coax loss to the network, thereby increasing the total power requirements of the network.

 

### 3.5. Extended Spectrum DOCSIS

Extended spectrum DOCSIS utilized the RF spectrum from 1.2 GHz to 1.8 GHz for downstream traffic. The upstream-downstream RF split is reassigned from one of several lower spits to an extended split. The current extended split under discussion is 5 MHz to 683 MHz upstream and 804 MHz to 1794 MHz downstream. HFC RF split options are shown here:

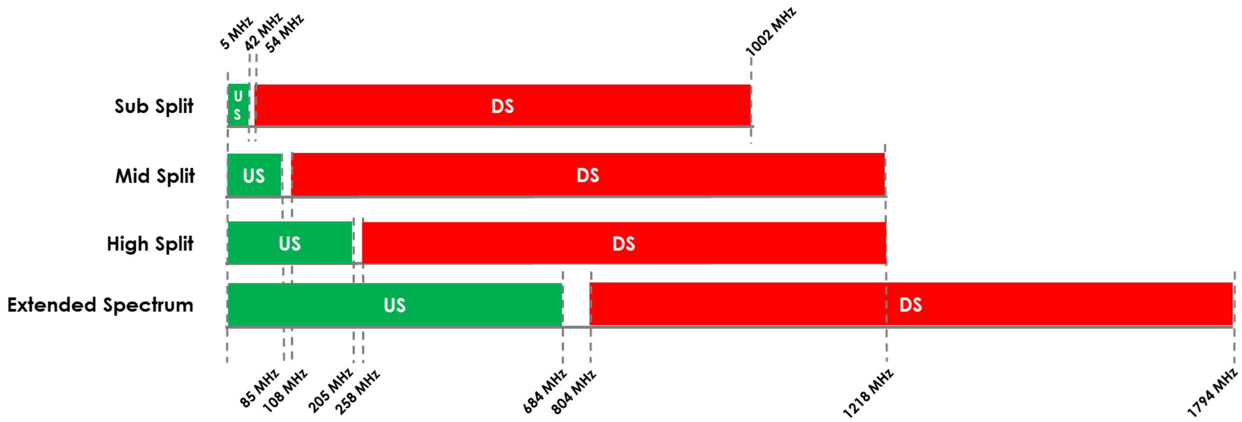


**Figure 7 – Extended Spectrum RF Splits**

HFC powering methods for extended spectrum HFC is expected to be similar to powering a standard spectrum HFC network with one exception. New extended spectrum amplifiers will require additional power to support transmission in the 1.2 GHz to 1.8 GHz range. It's too early to estimate how much additional power will be required for amplifiers to support the extended portion of the spectrum, primarily because some of the proposals include using lower order (and thus lower power consuming) modulations in the highest frequencies. Either way, operators considering extended spectrum operation will need to plan their network powering budget accordingly. As a next step beyond extended spectrum DOCSIS, industry technical papers and proposals are discussing extending the RF spectrum to 3 GHz which would provide up to 25 Gbps data speeds. HFC power requirements for this 3 GHz ultra-spectrum will be determined as supporting technology develops. Again, the take-away is that the powering requirements for the HFC network will only go up as networks are upgraded to support higher capacity, lower latency, and increased reliability. And since adding new power supplies to minimize coax losses is typically not cost-effective, the coax losses will also go up in the network as the power consumption (and thus the current) goes up.

# 4. Powering 5G Small Cells with HFC

The approach to powering 5G small cells from the HFC network is similar to powering other wireless devices. A gateway device interfaces the HFC coax power and DOCSIS backhaul to the radio's power and Ethernet interfaces. Operators will face new challenges as they consider deploying 5G small cells in densities needed to support the 5G MMTC service category. A simple small cell deployment example is explored here to illustrate some of the power complexities to consider.

Figure 8 illustrates three small cell radios powered from the HFC network. For this example the following assumptions are used:

 

- The UPS is configured to 90 VAC output and is rated at 1350 W output power.
- The optical node is RPD-equipped consuming 150 W.
- The amplifier consumes 100 W.
- Each small cell consumes 100 W.
- Each gateway consumes 30 W
- There is a 1500 ft coax run between the ON and the amplifier
- There is a 1500 ft coax run between the amplifier and the first gateway / small cell
- There are 500 ft coax runs between the first and second gateway / small cell and between the second and third gateway / small cell.
- The coax is a common 0.625 in diameter cable with resistance of 0.0011 ohms / foot
- The gateway / small cell minimum input voltage = 45 VAC.
- An MSO policy states that any broadband UPS can be loaded to maximum 80% of rated capacity.

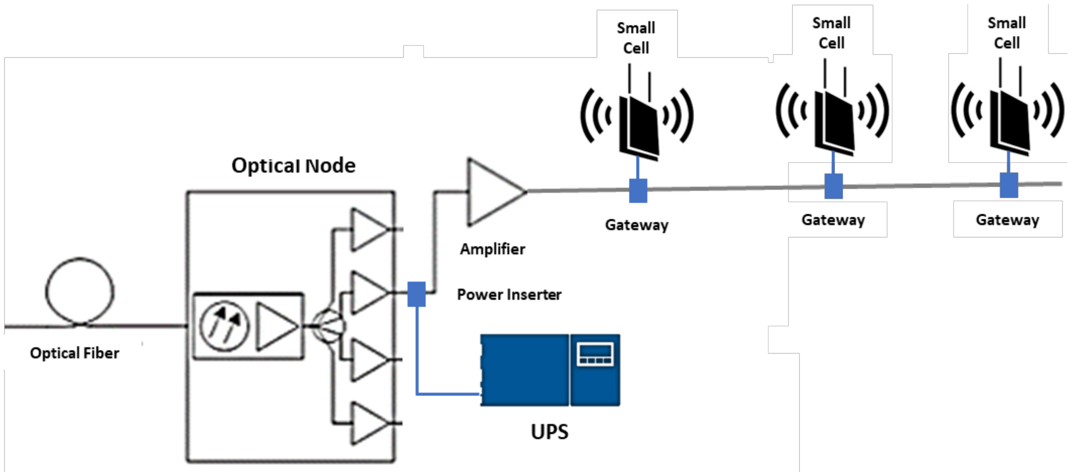


**Figure 8 – HFC Powered Small Cell**

Ignoring cable losses for the moment, the power required from the active devices is the sum:

P(Actives) = P(ON) + P(amp) + 3 x (P(small cell) + P(gateway))

Total Power (Actives) = 150 W + 100 W + 3 x (100 W +30 W) = 640 W

Using the coax cable properties as listed, the network cable loss adds 219 W to the power consumption. Total power consumed is:

P(total) = P(actives) + P(cable loss) = 859 W

This represents 64% of the rated UPS capacity. It's important to note that due to accumulated line resistance in the coax, the input voltage to the third and farthest gateway/small cell has dropped from 90 VAC to 63.6 VAC. And a 30% reduction in voltage means a corresponding increase in the current and the square of that increase in $I^2R$ losses in the coax at that point in the network.

Now add a fourth gateway/small cell near the amplifier and observe the effect on network power:

P(total) = P(actives) + P(cable loss) = 770 W + 314 W = 1084 W

This represents 80% of the UPS rated capacity which is the UPS load limit per operator policy. Input voltage to the farthest gateway/small cell has dropped to 58.7 VAC from the cumulative effects of coax line loss.

Next, if instead of adding a fourth gateway / small cell near the amplifier, what happens if the new equipment is added near the existing third gateway / small cell? The added line loss would result in the end-of-line voltage dropping below 45 VAC. While the equipment attempts to power-up with this reduced line voltage the current increases. The two end-of-line gateways will detect the low input voltage and disable small cell power, thereby leaving both radios disabled. With the radios disabled the coax current decreases and voltage drop reduces (i.e., higher end-of-line voltage). Once network voltage increased above 45 VAC the gateways attempt to power the small cells and the power overload condition is repeated. This situation could cause the UPS to temporarily fold-back, dropping power to the entire network segment. UPS fold-back is reviewed in a following section. To avoid such undesired outcomes, operators are advised to review power network loading parameters before new equipment is provisioned.

# 5. HFC Power – Under the Hood

This section reviews details behind HFC backup power. The operation of ferroresonant transformers is discussed as well as benefits and cautions around using power factor corrected loads in the HFC network. Battery health is inherent to power reliability. The topics of battery health and operation are reviewed.

## 5.1. Ferroresonant Overview

HFC networks have used ferroresonant style UPSs (ferros) for several decades. In addition to providing uninterrupted backup power, ferros provide three features essential to outside plant operation.

First, ferros provide an impressive 1000:1 input to output electrical isolation. This isolation provides critical protection for network equipment used in unpredictable outdoor environments. For example, a 5 kV electrical surge on the utility input of the ferro may produce only a 5 V variation on the output, thereby saving HFC gear from costly damage.

Next, ferros provide output short circuit protection. HFC networks consist of active components separated by spans of fiber optic and coax cable. Damage and wear to power carrying coax network segments can result in electrical short circuits or "faults" in network power. The high impedance of the ferro's LC tank circuit enables it to "fold back" or drop output voltage during fault conditions. When the coax fault is remedied, the ferro output is restored and normal network powering resumes.

The third reason ferros are well suited to HFC network powering relates to the ferro's voltage and current characteristics. Consider two common AC voltage waveforms: a sinusoidal wave typical of a linear UPS and a trapezoidal wave typical of a ferro UPS. Both wave forms are illustrated in Figures 9 and 10. Assuming the HFC equipment's power conversion stage uses a common diode bridge, capacitor and regulator circuit to convert the AC input to an internal DC bus, there will be a minimum instantaneous input voltage during each AC cycle at which the circuit will energize and draw current. For a linear UPS the voltage-current relationship is shown in Figure 9 – Sinusoidal Voltage and Current. Segment (A) represents the portion of each AC cycle where instantaneous voltage is high enough to energize the powered equipment. Current is consumed during this period.

 

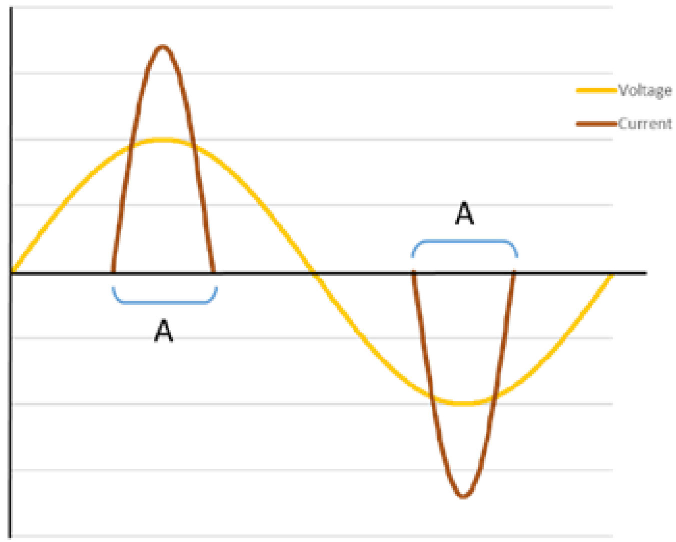


**Figure 9 – Sinusoidal Voltage and Current**

For the ferro UPS, the voltage-current relationship is shown in Figure 10 – Ferro Trapeziodial Voltage and Current. The ferro incorporates a LC tank circuit operating in magnetic saturation, causing voltage to rise quickly to a maximum value then remain constant for a portion of each AC cycle. Segment (B) represents the portion of each AC cycle where instantaneous voltage is high enough to energize the powered equipment. Power is consumed during this period.

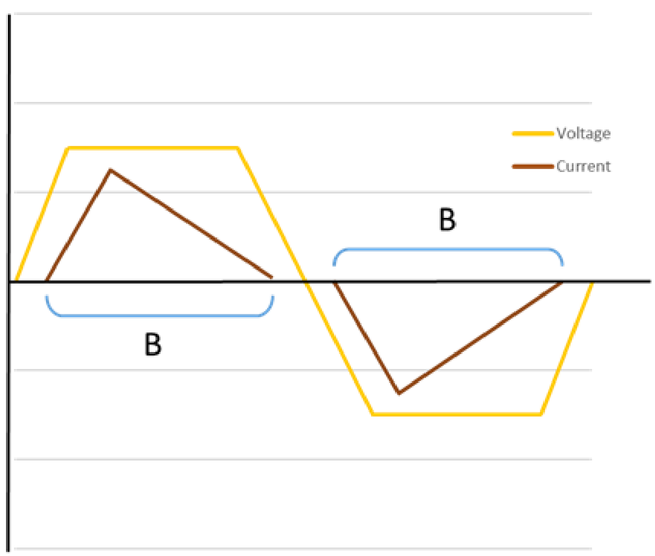


**Figure 10 – Ferro Trapeziodial Voltage and Current**

Because the ferro's output voltage is above our critical threshold for a longer period within each AC cycle than its sinusoidal counterpart, current is consumed and energy provided for a longer period within each AC cycle. This results in lower peak currents for a given constant power load. Also, the ferro's high impedance output naturally limits peak currents. Limiting peak currents results in less voltage drop across coax segments and more usable voltage farther away from the ferro than would be available from an equivalent linear UPS.

### 5.2. Power Factor Correction

Power factor correction (PFC) circuitry is designed to reduce or eliminate any voltage-current phase difference (sinusoidal current) and harmonic currents (non-sinusoidal current) to increase power efficiency. Although PFC is common in equipment connecting to utility power, PFC has not been widely adopted in HFC networks. Implementation would be simple enough: linear power packs within optical nodes and amplifiers would be replaced with power packs implementing PFC. There is added cost associated with the active PFC circuitry, but what are the benefits?

One frequently asked question is whether PFC loads within the HFC network will reduce energy consumption at the ferro and reduce the associated energy bill. Unfortunately, the answer is, not as much as one might think. The large ferro-resonant transformer acts as a passive PFC device insulating the HFC load from the AC line utility and providing an almost constant utility power factor of 0.9, irrespective of the power factor of the various HFC loads. As a result, changing the HFC equipment power pack from linear to PFC style has little effect on utility power cost.

A second potential benefit of PFC enabled HFC loads is the reduction of harmonic currents. This would lower the RMS current and increase usable power. Lower RMS current also means less voltage drop across coax spans, increasing the network's usable voltage reach. A trade-off for these benefits is PFC implementation efficiency. PFC circuits typically utilize an active dual conversion method for managing current. This draws more power than the traditional non-PFC linear power conversion approach and results in some efficiency loss.

To quantify the actual benefit of HFC PFC, the network topology must be considered. Actual efficiency improvements and voltage reach will vary at each load and is determined by multiple factors including load current, input voltage, resistive effects of coax and passives, and the actual power conversion efficiency of the PFC implementation.

A word of caution to those considering implementing PFC into HFC loads: The ferro's trapezoidal voltage and current output wave shapes change dramatically with various load profiles and are implementation specific. PFC implementations must adapt to these differences to be effective. PFC can provide benefits within the HFC network. Due to the complexity of HFC network implementations, the recommended way to qualify a PFC design and to quantify benefits is to use actual network testing.

### 5.3. Backup Power Requires Healthy Batteries

The ability of broadband UPS systems to provide uninterrupted power is directly related to the condition of the UPS system batteries. Aspects of battery health and maintenance relating to OSP UPS operation are discussed here.

Multiple factors impact battery runtime and health. For clarity, runtime as used here is defined as the instantaneous runtime available from a battery or bank of batteries irrespective of environment or history. Battery health is defined as the present maximum capacity of a battery or bank of batteries. Five primary factors affect battery runtime and battery state of health. These are: state of charge, ambient temperature, temperature history, battery age and charge history. This section discusses charge history and its effect on both battery runtime and battery health or capacity. Note that this discussion applies to lead-acid types of batteries, which are commonly deployed in HFC networks due to their lower cost. As new battery technologies become available and cost effectively deployed, the conclusions below will need to be modified based on the charging and lifetime characteristics of the newer battery technology.

### 5.3.1. Effects of Charge on Battery Capacity

Overcharging and undercharging batteries has a significant effect on battery life. Using charging specifications from battery manufacturers, the following discussion illustrates how overcharging or undercharging batteries can cause premature battery failure.

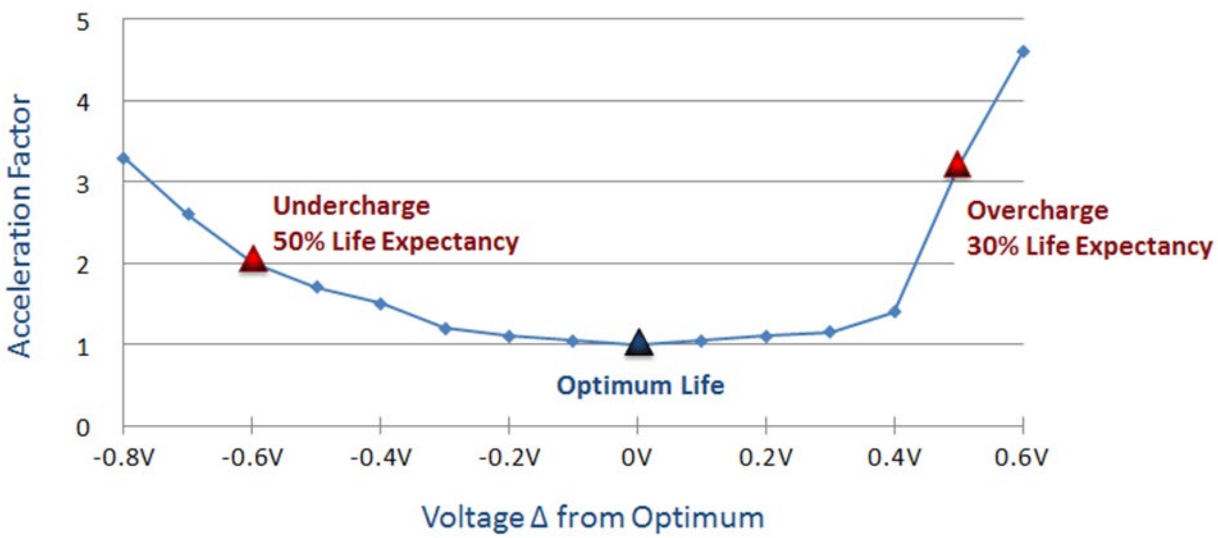


**Figure 11 – Charge Effect on Battery Life**

In this diagram, the X-axis identifies specific voltages under and over the optimum charging voltage. The Y-axis shows the acceleration factor or multiplier on battery life. For example, if a battery were undercharged by 0.6 V, its effective age would accelerate by a factor of 2x resulting in a useful life of 50% of that battery's optimum life. Likewise, if a battery were overcharged by 0.5 V, its useful life would be only 30% of its optimum life.

It should be noted that any charge related degradation effect occurs only while the overcharge or undercharge condition is applied. For example, undercharging a battery by 0.6 V for a period of four months (perhaps the duration until the next preventive maintenance cycle) would result in a loss of two months to the overall life of the battery (using a 2x acceleration factor for 0.6 V undercharge.) Once the charge problem is corrected, no further damage will occur, however, the capacity has nonetheless been permanently diminished.

### 5.3.2. Battery Chemistry and Charge Mismatch

Typical OSP power supply installations include one or multiple battery strings. Each battery string consists of 12 V batteries connected in series. Three series batteries are combined to achieve a 36 V string or four batteries are configured for a 48 V string. The connection between the power supply and the battery string(s) is through a wire harness connected across the entire string (i.e., one connection is at the negative terminal of the first battery (ground), the second connection at the positive terminal of the third battery for 36 V strings or the fourth battery for 48 V strings). In this configuration, power from the batteries and charge to the batteries is routed through the wire harness for the entire string. A three battery, 36 V battery string is shown in Figure 12.

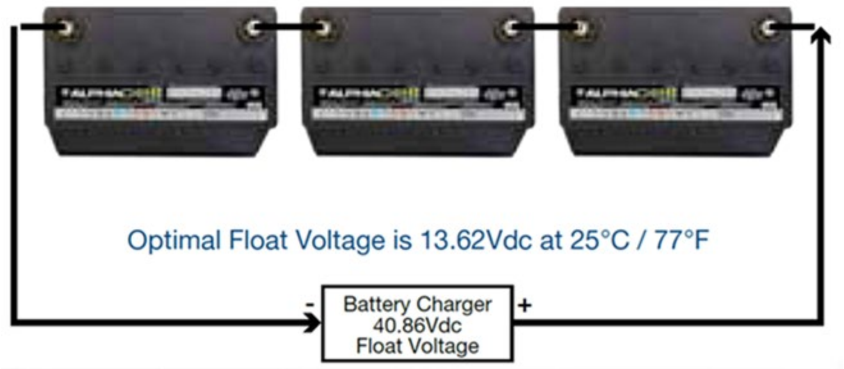


**Figure 12 – Battery Charge Configuration**

Optimal float voltage is also listed for this 36 V string as 13.62 VDC per battery. This value may vary among different battery manufacturers and technologies. The broadband UPS is configured to supply 40.86 VDC across the entire battery string. From the charger's perspective, each battery appears in the circuit as a fixed resistance. Per Ohm's Law, the three batteries, acting as resistors in series, create a voltage divider and the 40.86 VDC is distributed equally across each of the batteries at 13.62 VDC per battery.

Batteries operate via electro-chemical reactions. Time, temperature and charge history can affect this chemistry, thus altering the battery's internal resistance. If each battery in Figure 12 – Battery Charge Configuration has internal resistance values that change at different rates, the 40.86 VDC charge voltage will not be applied equally across each battery. The result of this unequal internal battery resistance is illustrated in Figure 13.

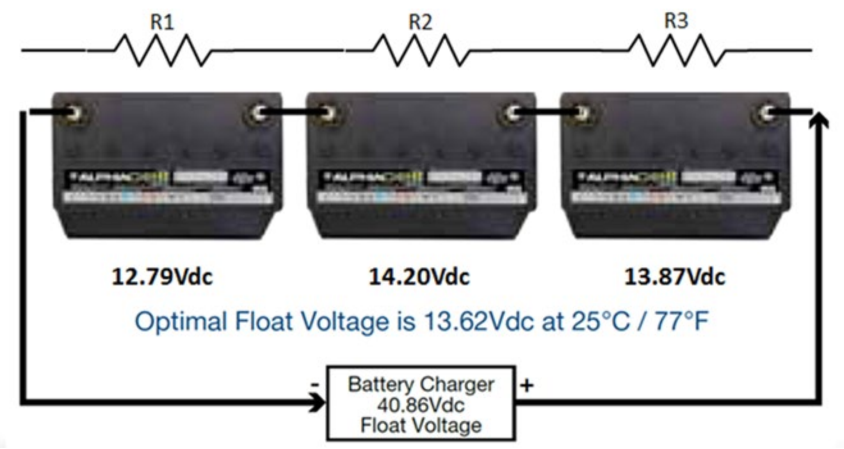


**Figure 13 – Battery Δ Resistance**

 

As the internal resistance of each battery changes at different rates over time, the effective circuit where R1≠R2≠R3 causes the charge voltage to be distributed unequally across the three batteries. The result is that some batteries will be undercharged while others are overcharged. Figure 13 – Battery Δ Resistance, illustrates an example where one battery is undercharged at 12.79 VDC while the other two batteries are overcharged at 14.20 VDC and 13.78 VDC. The total charge voltage of 40.86 VDC is correct, but the battery chemistry has caused an internal string charge variation that will shorten the life of all three batteries over time if not corrected.

OSP monitoring software can be configured to identify disparate voltages within individual battery strings for maintenance before extended battery damage occurs. Operators should consult battery manufacturers for specific voltage threshold parameters to trigger alerts and initiate corrective action. Multiple options exist to mitigate the effects of time on internal battery resistance. Some modern UPS systems are equipped with charge balancing technology that will re-direct charge within a battery string to offset the effect of changes to internal battery resistance. This charge management technology is available from various manufacturers in a variety of configurations. Operators should be aware of the effects of charge imbalance on battery life and determine the best course of action for their situation, especially as new services such as 5G and upgraded devices such as RPD and GAP are deployed.

### 5.3.3. Preventive Maintenance is Essential

In 2016, a US-based cable operator conducted an investigation to determine the root cause of their growing number of OSP broadband UPS alarms. Standby test fail alarms were of particular interest due to the critical nature of this alarm. Across three US cities, this operator identified that 22% of their broadband UPS systems had failed standby tests, due in part, to battery cable corrosion. The following shows representative battery corrosion at one of these installations.

 

**Figure 14 – Battery Corrosion**

The center battery shows excessive corrosion around both the positive and negative battery posts. This type and level of corrosion can occur with some styles of OSP batteries that have not been properly maintained. At the conclusion of this study, the operator identified deficiencies in their local OSP preventive maintenance practices that led to these potentially service-impacting results.

Corrosion will damage battery power cables and battery voltage monitoring sense wires. As battery power cables degrade, electrical resistance increases and the ability of the batteries to provide sufficient current diminishes. Eventually, as this operator experienced, the corrosion will increase cable resistance sufficiently to cause a broadband UPS standby test to fail. Prior to a standby test failure, backup capacity had diminished and the expected runtimes during actual power outages were lower. Had an actual utility outage occurred, standby power would be compromised, and customer loads dropped, potentially before the standby test indicated any problem.

Could this liability have been avoided through more diligent preventive maintenance practices? The answer, of course, is yes. "How frequently should each UPS be visited?" is a frequently debated question. Responses vary between operators and even between systems within the same operator. These answers range from six months to two years when technicians are queried. Conducting PM visits and finding nothing to correct is wasting valuable service resources. Waiting too long between PM visits could result in service impacting situations going unchecked. One often hears reports of PM visits to multiple installations within a geographic area with some locations checking out OK while other sites require extensive maintenance. Clearly, there is no one right answer to the question of PM frequency. Can anything be done to reduce unneeded PM visits while focusing limited resource on locations needing physical intervention?

 

The answer to this question is two-fold:

- First, there are some issues that require on-site inspection to identify and correct. Examples include pest infestations and water intrusion due to physical enclosure damage. Because this category of problems exists, scheduled preventive maintenance visits are required.
- Second, it may be possible to identify a category of future service-affecting problems through analysis of data available from status monitoring systems. This would enable operators to prioritize site visits around locations at high risk of causing future service disruptions. Low risk sites could be visited less frequently and only to inspect and correct issues that are undetectable any other way.

## 6. Conclusion

Several aspects of HFC powering have been discussed. A few key points include:

- Network power design can help ensure that power backup capacity continues to meet evolving network performance and power delivery expectations.
- Supporting new 10G service levels and upgrading existing HFC active devices place additional powering demands on the network. Due to evolving specifications and technology, not all 10G and upgrade powering demands are fully characterized at this time. Current network evolutions have seen trends towards higher power and more attention to backup capacity to support new infrastructure technologies. Operators can revise network powering guidelines to help with future architecture upgrades. Revised guidelines may include larger capacity UPS enclosures to permit additional battery strings to be installed in the future when backup requirements increase. Guidelines could also indicate that new UPS installations should be configured for 240 VAC utility power instead of the more traditional 120 VAC configuration. This assures that the UPS site can accommodate higher capacity UPS systems required to support future increased powering needs.
- Powering new equipment from the HFC plant, such as 5G small cells, must be carefully engineered. UPS capacity and coax line loss changes are critical elements of any network upgrades.
- UPS system must be maintained to assure backup power integrity. UPS batteries require periodic maintenance including physical inspections for environmental damage. Batteries should be remotely tested often enough to discern any negative trends in power capacity.

As operators prepare for and launch devices and services that deliver higher speed and lower latency that are also more robust, early consideration for the powering and backup time requirements for these services will go a long way to minimize unplanned surprises occurring late in the deployment cycle.

# 7. Abbreviations and Definitions

## 7.1. Abbreviations

| | |
|---|---|
| AC | alternating current |
| AP | access point |
| APSIS | adaptive power system interface specification |
| CCAP | converged cable access platform |
| CMTS | cable modem termination system |
| DAA | distributed access architecture |
| FMA | flexible MAC architecture |
| GAP | generic access platform |
| Gbps | gigabits per second |
| HFC | hybrid fiber/coax |
| Hz | hertz |
| MDU | multi dwelling unit |
| MMTC | massive machine type communications |
| ON | optical node |
| PFC | power factor correction |
| PM | preventive maintenance |
| RPD | remote PHY device |
| R-PHY | remote PHY |
| SCTE | Society of Cable Telecommunications Engineers |
| SFP | small form-factor pluggable |
| UPS | uninterruptable power supply |
| URLLC | ultra-reliable low-latency communications |
| VAC | volts of alternating current |

## 7.2. Definitions

| | |
|---|---|
| 10G | A CableLabs® initiative supporting 10 gigabit symmetrical data service. |
| 5G | 5$^{th}$ generation wireless |
| Downstream | Information flowing from the hub to the user |
| Upstream | Information flowing from the user to the hub |

# 8. Bibliography and References


*CableLabs® website. CableLabs.com: 10G Platform; FDX Point-to-Point Coherent Optics; Distributed Access Architecture: Moving Beyond Cloud Computing to Edge Computing*

*Blueprint for 3 GHz, 25 Gbps DOCSIS®; Getting 25 Gbps PON-Like Performance Out of HFC*, A Technical Paper prepared for SCTE•ISBE, Chapman, Jin, 2019

*Access Network Energy Reduction, It Can Be Done,* Broadband Library, Fall 2018, Anderson

*Improving OSP Network Reliability Using Predictive Alarming; A Study of Advanced Battery Analysis Methods.* A Technical Paper prepared for SCTE/ISBE, Anderson & Burgett, 2016

*Generic Access platform (GAP) Requirements*, IPS WG1, Interface Practices Subcommittee, SCTE-ISBE Standards

*Improving Network Reliability Through Effective Power Management*, Society of Cable Telecommunications Engineers, Anderson & Burgett, 2014

 

# Nx25G-EPON

## The Evolution of PON to Multi-Channel High-Capacity P2MP Access System with Native Channel Bonding

A Technical Paper Prepared for SCTE•ISBE by

Marek Hajduczenia, Principal Engineer III, Spectrum Enterprise
14810 Grasslands Dr | Englewood, CO 80112
marek.hajduczenia@charter.com
+1-720-536-1366

 

# Table of Contents

**Title** **Page Number**



## List of Figures

**Title** **Page Number**



 

# List of Tables

## 1. Introduction

The seemingly never-ending growth in bandwidth demand has been clearly visible over the years in different network layers, be that core, metro, or even access. Gigabit access speeds are not as novel as they once were, with a large share of US population having an option to opt in for some variety of fiber-fed high capacity broadband access, be that a direct fiber-to-the-home or a DOCSIS-based cable offering. Two years ago, barely 5% of cable subscribers had gigabit broadband speeds available to them, whereas by the end of 2019 it is expected that more than 90% of cable subscribers will have gigabit broadband speeds available to them. This equates to over 80% of the US population now having access to broadband speeds of 1 Gbps and this is only the beginning.

Various flavors of passive optical network (PON) technology have been used over the years to offer different types of access services, including residential (fiber-to-the-home), enterprise (fiber-to-the-office), or business (fiber-to-the-tower, for example). The support for strict service level agreement (SLA) provides an option to offer differentiated service levels on the same shared PON fiber plant, while mixing different types of customers, and even oversubscribing PON capacity for selected customers, without impacting higher fidelity service types.

As new 10G-EPON rollouts started becoming more numerous and the 10 Gbps PON industry reached the initial stages of technical and economic maturity, the question about the future of PON technology in general resulted in the call for interest (CFI) in July 2015 in Waikoloa, HI USA [NG-EPON CFI]. The CFI and the following work of the NG-EPON study group pointed out the possible high-bandwidth applications that could not be addressed in an economic manner with existing 1G and 10G PON platforms, resulting in the need for a higher data rate system. Operator representatives emphasized the need for preservation of the existing outside fiber plant investment (backward compatibility) and coexistence with already deployed PON systems – something that drove the long debate on the wavelength allocation plans, as outlined in the following sections. It was also crucial to develop a system that allows for continuous capacity growth using a channel bonding mechanism, allowing an operator to add wavelengths (and capacity) to the system without the need to wait for one more generation of a standard to be completed. This particular request drove the development of a channel bonding mechanism in the Nx25G-EPON system.

The resulting system is a more efficient EPON, building on years of experience of operators, silicon developers, optics suppliers, and end users, providing higher serial data rate and supporting the channel bonding functionality. This elegant solution preserves the most attractive aspects of previous EPON generations (Ethernet framing, logical link emulation, multi point control protocol) while adding support for transmission across multiple bonded channels, packet fragmentation and reassembly, and dramatically improving scheduling of multiple logical entities per ONU for improved upstream channel efficiency.

The remainder of this document discusses Nx25G-EPON objectives and the underlying assumptions and then moves to the discussion of the Nx25G-EPON architecture, starting with the physical layer and wavelength allocation plans, through physical coding sublayer (PCS) including more advanced forward error correction (FEC), multi-channel reconciliation sublayer (MCRS) providing support for channel bonding and skew mitigation, ending with Multi Point Control Protocol (MPCP) and its extensions for improved upstream scheduling efficiency, including a new Channel Control Protocol (CCP) for per channel transmitter / receiver control required by operators for operations and administration management (OAM) functions in a multi-channel shared medium system of EPON.

 

## 2. Nx25G-EPON Objectives

The objectives of Nx25G-EPON system are a compromise between the market drivers behind the constantly increasing bandwidth demand, technical feasibility of delivering such high-capacity systems to end-users, and economic feasibility of maintaining the overall system cost at an acceptable level. The original set of objectives to provide a system with the total capacity of up to 100 Gbps in symmetric and asymmetric configurations, while maintaining backwards compatibility with the existing outside fiber plant, proved to be too elusive while attempting to maintain the aforementioned economic feasibility of the resulting solution. While technically feasible, such a system would prove too expensive to deploy in a large scale, even when considering business-class services as the primary driver for such high-capacity access system.

Consequently, the scope of the project was reduced to the development of a system with total capacity of up to 50 Gbps in symmetric and asymmetric configurations, while maintaining backward compatibility with deployed 10 Gbps class PON systems (EPON and GPON alike) and their optical distribution networks (ODNs) designed to 29 dB loss budget. Using a combination of 25 Gbps (downstream or upstream) and/or 10 Gbps (upstream only) per lane, the resulting access system supports symmetric and/or asymmetric MAC data rates of:

- 25 Gbps in the downstream and 10 Gbps or 25 Gbps in the upstream (25G-EPON)
- 50 Gbps in the downstream and 10 Gbps, 25 Gbps, or 50 Gbps in the upstream (50G-EPON)

Collectively, these two systems are referred to as Nx25G-EPON, implying the use of multiple parallel 25G-EPON systems, operating on separate wavelength channels, but supporting channel bonding (mechanism functionally similar to the link aggregation mechanism defined in IEEE Std 802.1AX) for an increased aggregate MAC data rate.

Apart from the effective system data rate, one of the critical objectives of Nx25G-EPON is the coexistence with already deployed PON systems. Historically speaking, 10G-EPON (defined in IEEE Std 802.3av-2009) supported backward compatibility with 1G-EPON (defined in IEEE Std 802.3ah-2004), allowing operators with existing 1G-EPON deployments add 10G-EPON using either wavelength division multiplexing (WDM) for fully independent and parallel operation, or using time division multiplexing (TDM) for time-shared operation in the upstream direction. Initially, Nx25G-EPON objectives required coexistence with 1G-EPON and 10G-EPON on the same ODN, severely restricting the availability of the fiber spectrum in the most desirable transmission region (around 1310 nm, due to minimum fiber dispersion). In the face of increased system cost as well as technical challenges associated with operation outside of the 1310 nm window, the task force relaxed the coexistence requirements, focusing only on one EPON generation, i.e., 10G-EPON. The understanding is that by the time Nx25G-EPON is deployed into the network, 1G-EPON will have been fully upgraded to 10G-EPON (and hence, absent on the ODN), or Nx25G-EPON will be deployed over a separate greenfield ODN intended to operate independently from the previous EPON generations. Given the similarity of the wavelength plans between 10G-EPON and ITU-T-defined PON systems (GPON, XG-PON1, and XGS-PON), coexistence with these systems is also part of the Nx25G-EPON objectives, providing a seamless evolution path for any operator with PON deployments out there, be that IEEE-defined EPON or ITU-T defined xPON. The coexistence-related objectives are therefore defined as follows:

- Wavelength allocation allowing concurrent operation with 10G-EPON, XG-PON1, and XGS-PON PHYs (1575 nm-1580 nm downstream, 1260 nm-1280 nm upstream)

 

- Wavelength allocation allowing concurrent operation of 25G-EPON and GPON reduced wavelength set (1480 nm-1500 nm downstream, 1290 nm-1330 nm upstream) PHYs

Finally, it is also critical for operators to preserve their investment into the ODN that has been already deployed. A substantial share of the overall cost of the access system is associated with the fiber deployment process, including trenching, labor, access rights, etc., leaving the actives (OLT and ONUs, etc.) in a distant second place. Most commonly, ODN today is designed for compatibility with 10G-EPON PR30 power budget, supporting the ODN with 29 dB optical loss. Such a large power budget helps minimize the number of PON systems that have to be deployed, while providing very good coverage of the areas of interest. Clearly, to preserve the ODN investment, Nx25G-EPON must accommodate channel insertion losses equivalent to 10G-EPON-defined PR20 and PR30 power budgets, leaving out the lowest PR10 power budget, which received very little interest from operators.

The final adopted objectives of the Nx25G-EPON project can be found in [NG-EPON Objectives].

## 3. Nx25G-EPON Architecture

The resulting Nx25G-EPON physical layer (PHY) architecture is based on multiple instances of 25G-EPON, where each bonded instance operated at 25 Gbps in the downstream and 10 Gbps or 25 Gbps in the upstream. This results in a number of possible system configurations, i.e.,

- 25/10G-EPON, operating at 25 Gbps in the downstream (single channel) and 10 Gbps in the upstream direction (single channel);
- 25/25G-EPON, operating at 25 Gbps in the downstream (single channel) and upstream (single channel) directions;
- 50/10G-EPON, operating at 50 Gbps in the downstream (two channels) and 10 Gbps in the upstream direction (single channel);
- 50/25G-EPON, operating at 50 Gbps in the downstream (two channels) and 25 Gbps in the upstream direction (single channel); and
- 50/50G-EPON, operating at 50 Gbps in the downstream (two channels) and upstream (two channels) directions.

All dual-channel variants of the system rely on channel bonding, allowing for a single higher speed MAC (for example, 50 Gbps) transmitting over a pair of bonded 25 Gbps channels in a manner that is completely transparent to MAC. The resulting logical system architecture was designed in a very scalable manner, allowing any future revisions of the system to add individual channels to the system, as well as replace channel definitions, data rates, FEC encoding, etc., while maintaining all the benefits of a logically bonded multi-channel system.

It is expected that two classes of Nx25G-EPON OLTs will become commercially available in the future, notably 25G-EPON OLT, supporting 25/25G-EPON and 25/10G-EPON ONUs on the same ODN, as well as 50G-EPON OLT, supporting 50/50G-EPON, 50/25G-EPON, 50/10G-EPON, 25/25G-EPON, and 25/10G-EPON ONUs on the same ODN. The resulting OLT and ONU combinations are shown in Figure 1. Note that despite the number of ONU models, the OLT-side implementation is substantially simplified by keeping all operations on per-channel basis.

 

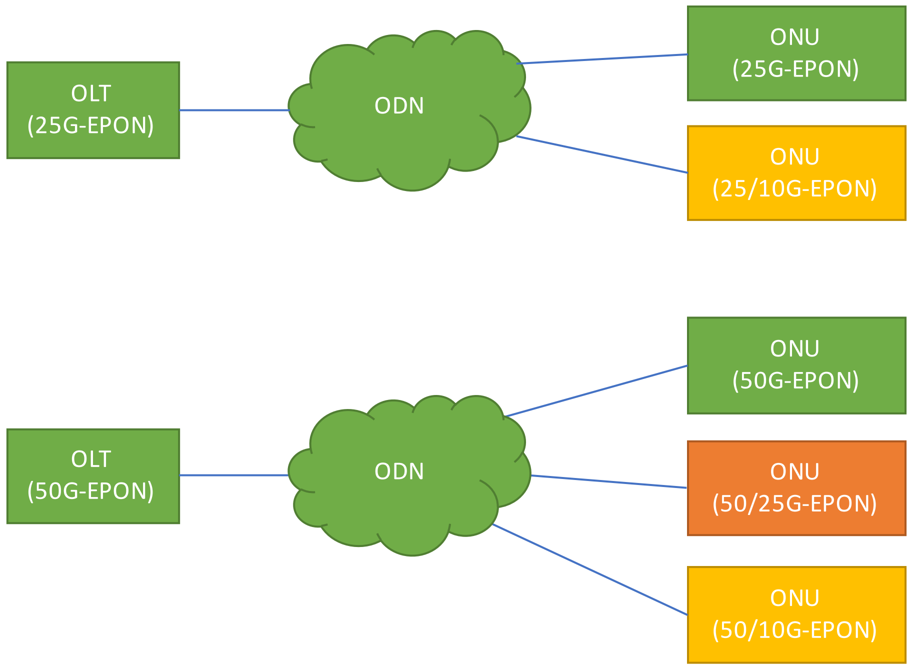


**Figure 1 - 25G-EPON (top) and 50G-EPON (bottom) architecture**

### 3.1. Wavelength Allocation Plan

As indicated before, the wavelength plan design is critical for the backward compatibility and coexistence required for the commercial success of Nx25G-EPON systems. Given the increase in the data rate to 25 Gbps per channel, operation in the low-dispersion O-band region was crucial for the technical feasibility of the resulting system. The use of fixed and if possible – uncooled – directly modulated lasers (DML) in ONUs was also critical for the economic feasibility of the system. Combined with the coexistence requirements in place, these three factors contributed to very long discussions on a wavelength allocation plan for this new system, resulting in three upstream wavelengths and two downstream wavelengths, as shown in Table 1 and in Figure 2.

**Table 1 - Upstream and downstream Nx25G-EPON wavelengths**

| Wavelength Name | Direction | Center (nm) | Range (nm) |
|---|---|---|---|
| DW0 | Downstream | 1358 | ±2 |
| DW1 | Downstream | 1342 | ±2 |
| UW0 | Upstream | 1270 | ±10 |
| UW1 | Upstream | 1300 | ±10 |
| UW2 | Upstream | 1320 | ±2 |

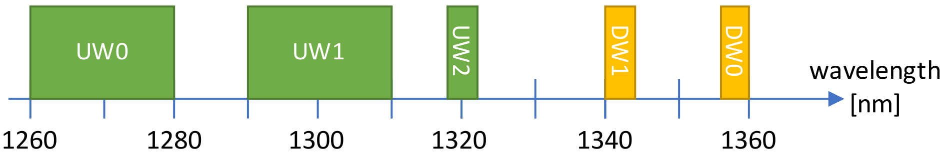


**Figure 2 – Nx25G-EPON wavelength plan**

Two wideband (20 nm wide) upstream wavelength options were selected to keep a brand new 25G-EPON or 50G-EPON as low cost as possible, especially in the case of greenfield deployments, where coexistence with any legacy PON systems is not required. The use of 20 nm and reasonably broadly spaced UW0 and UW1 bands allows for the use of uncooled DML devices at ONUs, keeping the resulting cost of optical modules comparable with 10G-EPON ONU modules, when a full-scale production is achieved. The resulting green-field deployment wavelength plan is shown in Figure 3.

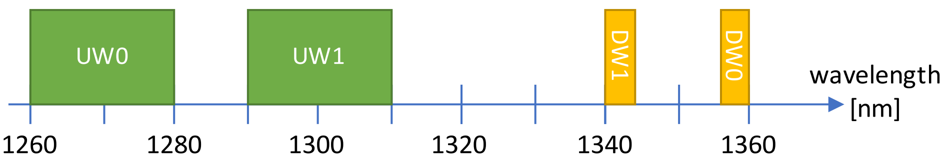


**Figure 3 – Nx25G-EPON greenfield wavelength plan**

For brownfield scenarios where coexistence with first generation PON technologies (1G-EPON and GPON) is required, the upstream wavelength UW0 and downstream wavelength DW0 can be combined to provide a solution that has 10x more capacity. In fact, it is technically feasible for an operator to also use DW1 to achieve an asymmetric 50G/25G-EPON on the same ODN. The resulting wavelength plan is shown in Figure 4.

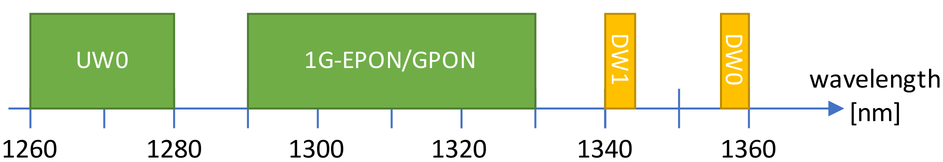


**Figure 4 – Nx25G-EPON brownfield wavelength plan supporting coexistence with GPON/1G-EPON**

For brownfield scenarios where coexistence with second generation PON technologies (10G-EPON, XGS-PON) is required, the use of DW0 and DW1 for up to 50 Gbps in the downstream via channel bonding can be combined with UW1 and more narrow UW2 for up to 50 Gbps in the upstream, as shown in Figure 5.

 

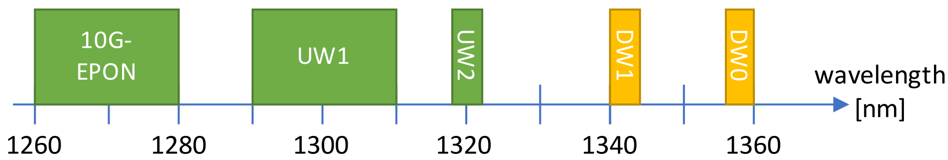


**Figure 5 – Nx25G-EPON brownfield wavelength plan supporting coexistence with XGS-PON/10G-EPON**

Despite the need to support five wavelengths in total, the resulting wavelength allocation plans provide an optimum approach towards coexistence with legacy PON systems (1 Gbps and 10 Gbps systems alike), while also providing the most cost-effective greenfield scenario deployment option. The resulting Nx25G-EPON system provides a clear evolution path for any operator with PON deployments, be that IEEE-specified EPON or ITU-T defined xPON, hopefully achieving the long-debated converged PON platform for all operators around the world.

### 3.2. Physical Coding Sublayer

In the transmit direction, the physical coding sublayer (PCS) in Nx25G-EPON is responsible mainly for converting the data stream received from xMII into codewords, which can then be passed through PMA, PMD, and finally transmitted on the medium. In the receive direction, PCS performs the reverse function, i.e., decodes the received data and recreates the original data stream, hands off to the MCRS and then towards the Ethernet MAC and any associated MAC clients. PCS houses such critical functionalities as line decoder / encoder, FEC encoder / decoder, data detector (ONU transmit path only), scrambler / descrambler as well as gearbox, adjusting the data rates between PCS (bursty transmission) and PMA (continuous data stream). A functional structure of Nx25G-EPON PCS in OLT and ONU, in both the transmit and receive directions, is shown in Figure 6.

A number of PCS functions in Nx25G-EPON are directly derived from their 10G-EPON counterparts, and readers are recommended to examine their details in the respective IEEE Std 802.3av-2009 (covering all 10G-EPON definitions) or [Next-Generation Ethernet Passive Optical Networks: 10G-EPON]. The following sections examine only specific functions that differ substantially from their 10G-EPON counterparts.

 

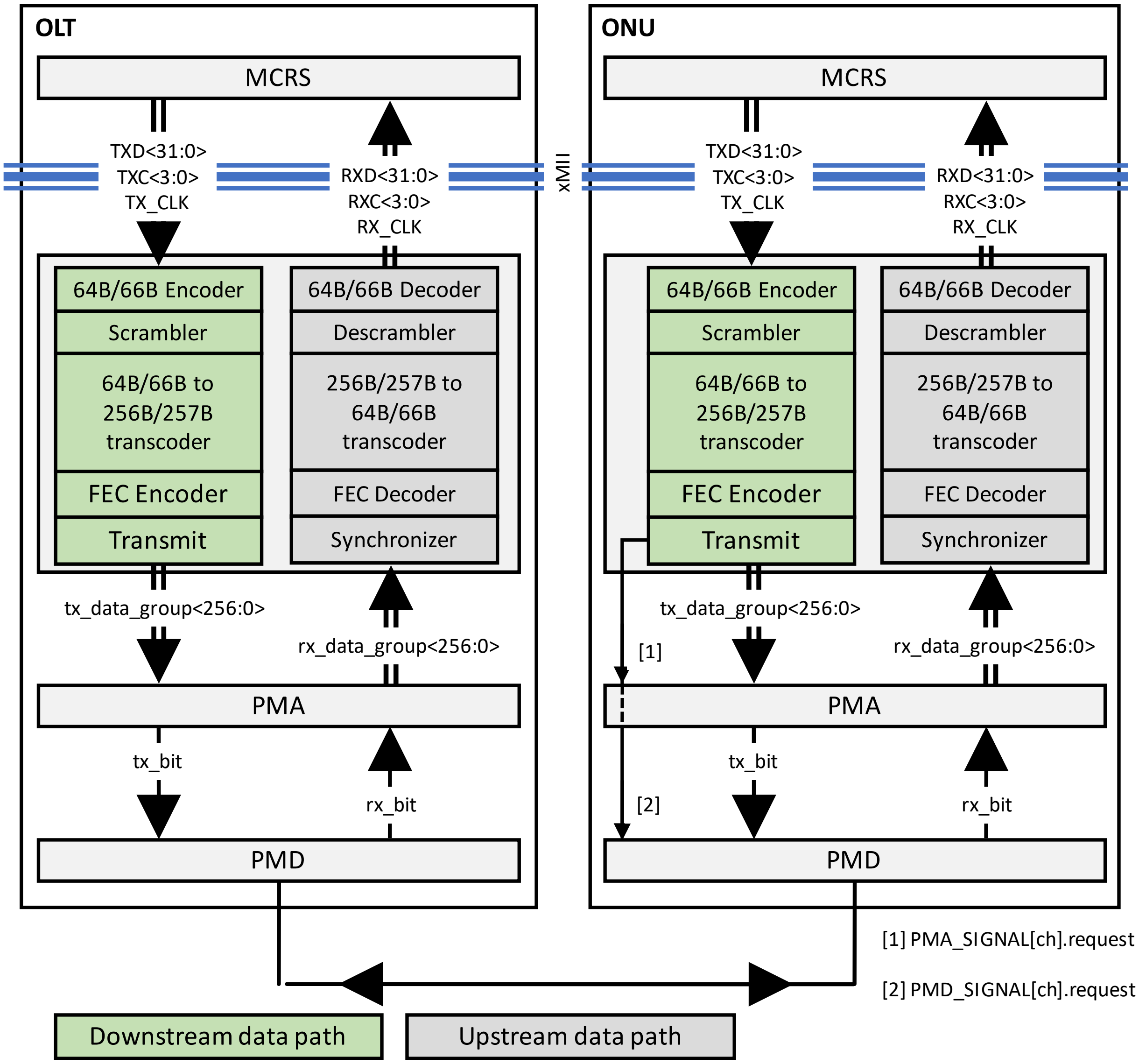


**Figure 6 – Nx25G-EPON PCS functions**

### *3.2.1. 256b/257b Line Coding*

Nx25G-EPON uses a more efficient 256b/257b line coding (note that for historic reasons, IEEE Std 802.3 uses 256B/257B and 64B/66B designations), where four consecutive scrambled 64b/66b blocks are converted into one scrambled 256b/257b block. This line coding results in marginal overhead improvement over 64b/66b encoding: 99.6% efficiency for 256b/257b line code versus 96.9% efficiency for 64b/66b line code. The 256b/257b line coding was originally developed for 100GBASE-R PCS to minimize the line coding overhead and reused in Nx25G-EPON to primarily minimize overhead.

 

### 3.2.2. Forward Error Correction Mechanism

1G-EPON adopted the use of optional, frame-based FEC, while 10G-EPON used a mandatory stream-based FEC. Both mechanisms are poised to provide protection against bit errors occurring during transmission across the optical channel. However, both mechanisms are also quite different in many ways.

A stream-based FEC mechanism processes Ethernet frames and interpacket IDLE symbols as a continuous stream of data symbols, encoding these data symbols into a FEC codeword with a fixed structure and size. Such an implementation is simpler when compared with a frame-based FEC, and it is ideally suited for high speed digital transmissions systems. However, this particular FEC encoding method requires both the transmitter and receiver to use the very same framing structure. A device not supporting the FEC decoding process is simply not able to retrieve data and separate it from parity. This means that all ONUs in an EPON using stream-based FEC must use the very same FEC decoder/encoder and have ability to identify the location of parity data within the incoming data stream. In the stream-based method, the parity symbols generated after each data block are inserted immediately after the FEC parity codeword that they are protecting, resulting in an interleaving pattern of data blocks and parity blocks.

In the frame-based FEC mechanism, the parity symbols generated for each FEC block are grouped together and are appended at the end of an Ethernet frame. Such a parity placement leaves the data frame itself unaltered, representing a major advantage of this particular FEC encoding method. Any device not supporting FEC decoding is still able to receive the data, though will not take advantage of the enhanced FEC bit error protection. In 1G-EPON, adoption of this particular FEC coding method allows for mixing ONUs with enabled and disabled FEC on the same ODN.

Due to the 2.5X serial data rate increase over 10G-EPON, Nx25G-EPON adopted the use of stream-based FEC, though with a few changes relative to 10G-EPON, namely:

- Rather than a classic, fixed rate Reed Solomon code used in 10G-EPON, Nx25G-EPON is equipped with a stronger quasi-cyclic QC-LDPC FEC, using a much longer codeword but also providing a much higher degree of protection against bit errors. The extra coding gain (resulting in perceived improved optical received sensitivity for the given target bit error ratio) provided by this particular FEC code was required to close the more challenging PR30 class power budget without the use of mandatory optical amplifiers.
- Even though the upstream and downstream FEC codeword size was selected to be the same, the upstream channel supports the option of transmitting truncated FEC codewords, improving the upstream channel efficiency and allowing for more granular scheduling. Recall that in 10G-EPON, the upstream transmission burst size had to be granted in multiples of a FEC codeword size to delineate burst boundaries correctly. This is no longer a limitation for Nx25G-EPON, where arbitrary upstream burst sizes may be allocated and the ONU has ability to truncate the last FEC codeword on transmit, and the OLT has the ability to recover data from such a truncated FEC codeword.

Even though several proposals were made to adopt shorter FEC codeword for the upstream channel, providing improved bit error protection and adding to the coding gain, the current system design assumes the same downstream and upstream FEC codeword size to support a low-level loopback mode, facilitating chipset design and debugging. To help with vendor interoperability, the Nx25G-EPON standard does specify a series of golden test vectors, providing a reference point for any FEC decoder and encoder implementation validation.

## 3.3. Logical Layer Operation

Just like 1G-EPON and 10G-EPON before, in the downstream direction Nx25G-EPON operates in a point-to-multipoint (P2MP) mode, delivering logically tagged Ethernet frames to subtended ONUs using a Layer 1 (L1) broadcast across the passive fiber ODN. Each ONU receives a portion of the transmitted downstream optical signal. The number of connected ONUs varies from deployment to deployment, primarily in the function of the provided services (residential versus enterprise), available power budget, as well as the distance between the OLT and the farthest ONU. Typical numbers of connected ONUs range between four and 64, with a typical number of LLIDs per OLT port at 128 / 256, depending on the offered service mix (two to three LLIDs per ONU).Changes in the logical layer tagging in Nx25G-EPON (discussed in the following sections) allow for much more efficient scheduling of individual LLIDs, supporting a much larger number of LLIDs when compared with 1G-EPON and 10G-EPON, while keeping the number of connected ONUs limited by the available power budget. Effectively, in Nx25G-EPON, it is possible to assign, control, and schedule thousands of LLIDs without the negative impact of all upstream transmission overhead associated with polling such a large number of LLIDs.

The downstream channel properties in this PON system make it a shared -medium network: Packets broadcast by the OLT are selectively extracted by the destination ONU, which applies simple packet–filtering rules based on MAC and LLID addresses. The packet filtering rules in Nx25G-EPON build on the 10G-EPON specification but are performed on the LLID value carried in the envelope header /EH/ field instead of the modified Ethernet frame preamble.

Nx25G-EPON introduces a completely new way of looking at the LLID as a logical control entity, with a new set of LLID subtypes defined and associated with a specific function, namely:

- Physical layer ID (PLID): used to control critical Nx25G-EPON operations, such as ONU registrations and arbitration of each ONU's access to the PON medium. All multipoint control protocol data units (MPCPDUs) use the PLID. Each ONU is assigned a single PLID upon the successful completion of the ONU discovery and registration process.
- Management link ID (MLID): used to carry management traffic flows, such as OAMPDUs and CCPDUs. Each ONU is assigned a single MLID upon the successful completion of the ONU discovery and registration process.
- User link ID (ULID): used to carry any customer (client) traffic, with any mapping between client traffic classes and specific service instances left open for custom configuration and at the operator's discretion. ULID values are assigned (provisioned) to an ONU using an appropriate management protocol, typically employing some sort of a network management system (NMS). To further optimize delivery of unicast, multicast, and broadcast services over Nx25G-EPON system, ULID may be unicast (one OLT MAC associated with one ONU MAC), multicast (one OLT MAC associated with a subset of ONU MACs), or broadcast (one OLT MAC associated with all ONU MACs).
- Group link ID (GLID): used to consolidate traffic from a number of LLIDs in a single transmission envelope for improved transmission efficiency. For example, in case of an ONU supporting multiple subscribers in an MTU application, LLIDs supporting a specific traffic class (e.g., best-effort traffic) on a multi-subscriber ONU could be grouped together for improved

 

efficiency. This allows much better per-class traffic scheduling, while still maintaining per-customer polling ability, even in the case of multi-subscriber ONUs.

The downstream channel operation is best depicted in Figure 7, where packets destined to different end subscribers are filtered out by the ONUs from the broadcast data flow.

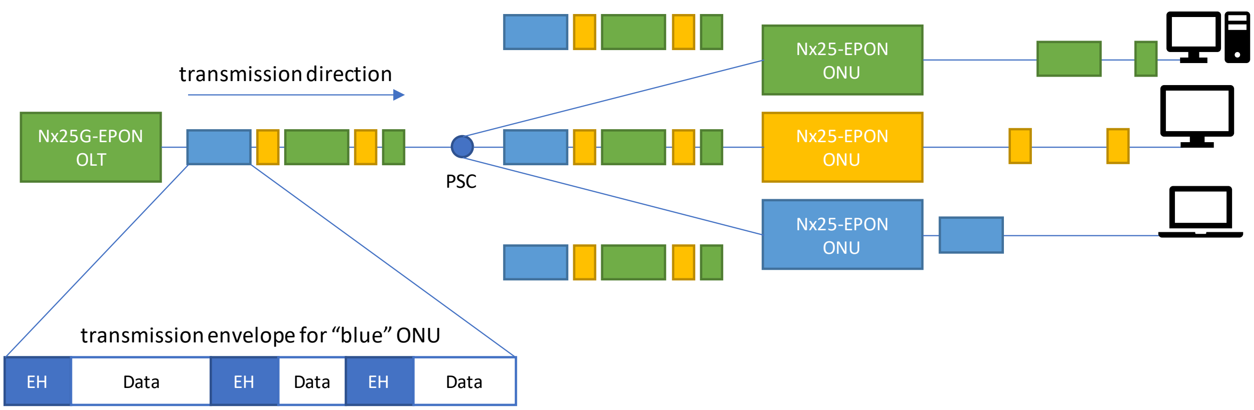


**Figure 7 – Downstream channel transmission in EPON – P2MP operation (broadcast)**

In the upstream direction (see Figure 8 for details), from the ONUs towards the OLT, Nx25G-EPON operates in the multipoint-to-point (MP2P) mode. Any connected and active ONU modules transmit their buffered packets towards the single receiver located in the OLT. Since the specific PON physical constraints do not allow the ONUs to see any data transmissions originating from other subscriber modules, the direct implementation of the carrier sense multiple access / collision detection (CSMA/CD) scheme is not feasible. All ONUs in the given EPON belong to a single collision domain, and thus a centrally managed channel access is required via time division multiple access (TDMA). ONUs in their default state are not allowed to transmit any data unless they are explicitly granted by the OLT.

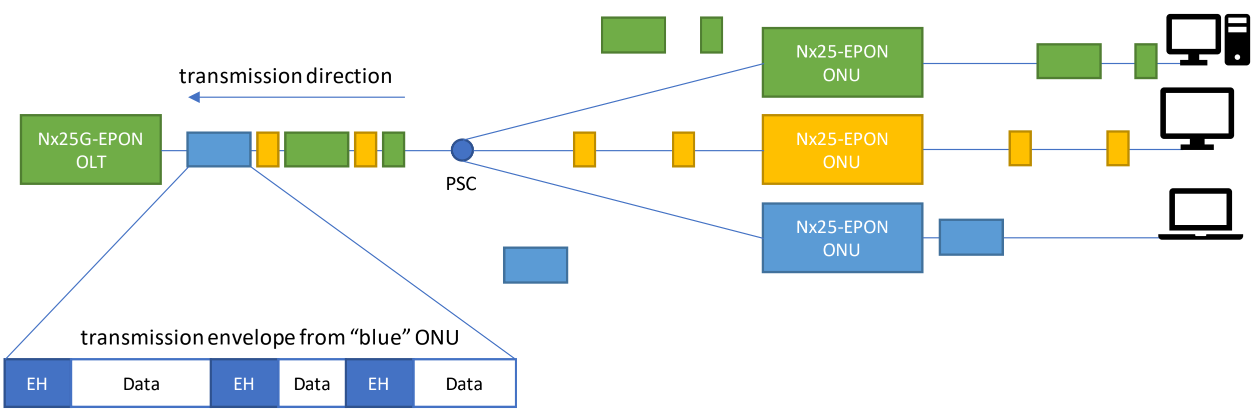


**Figure 8 – Upstream channel transmission in EPON – MP2P operation (TMDA)**

In this way, data collisions are avoided, since the central OLT controller at any moment of time is aware of the scheduled transmissions from individual ONUs. The only exception from this centrally managed upstream channel access scheme is the so–called discovery process, where new and not initialized ONUs are allowed to register in the EPON system. In order to avoid persistent collisions, a random delay mechanism (RDM) is applied to the registration request transmitted by an ONU, offsetting the transmissions originating from ONUs in a random manner and decreasing the collision probability – see Figure 9.

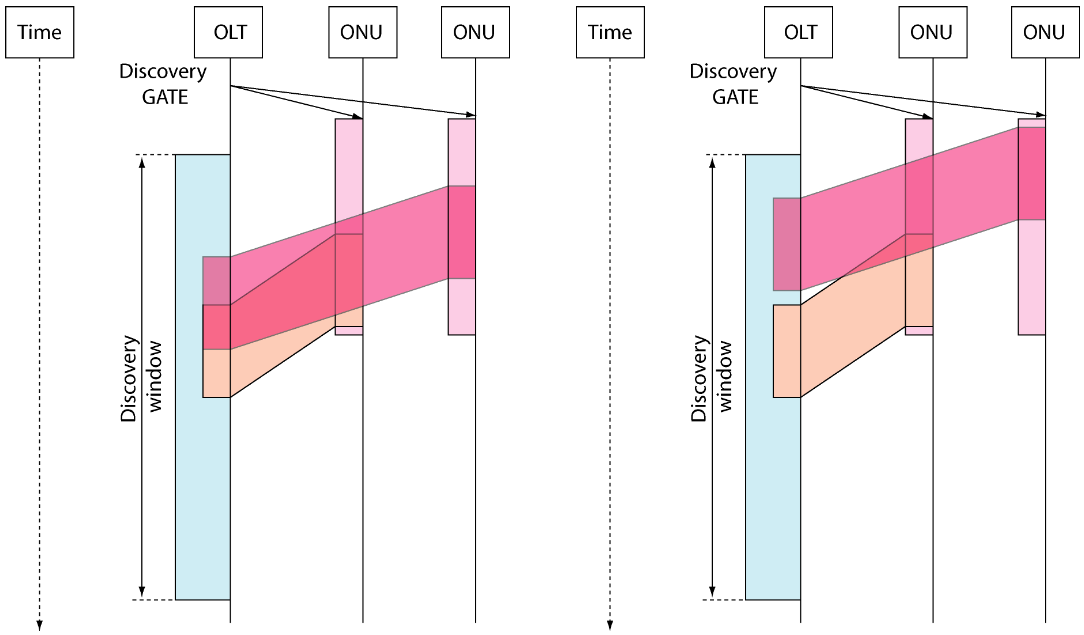


**Figure 9 – Random delay mechanism for EPON discovery**

A multiple access control protocol is required in the upstream direction, since the EPON operates as a multipoint-to-point network and every single ONU talks directly to the OLT. A contention-based media access mechanism is difficult to implement, since ONUs cannot detect a collision at the OLT, and providing the architecture with a feedback loop leading to every single ONU is not economically feasible. Besides, contention-based schemes provide a non-deterministic service, i.e., node throughput and channel utilization may be described only as statistical averages, providing very poor SLA. A summary of changes to MPCP for Nx25-EPON is included in the following section.

## 3.4. Channel Bonding in Nx25G-EPON

Unlike in previous EPON generations, Nx25G-EPON uses logical channel bonding, resulting in data striping across multiple physical channels in both downstream and upstream directions. A new multi-channel reconciliation sublayer (MCRS) was defined for Nx25G-EPON (as well as any future projects needing similar multi-channel bonding capabilities), providing logical channel bonding between any MAC and any number of underlying physical layers (channels), as shown in Figure 10.

 

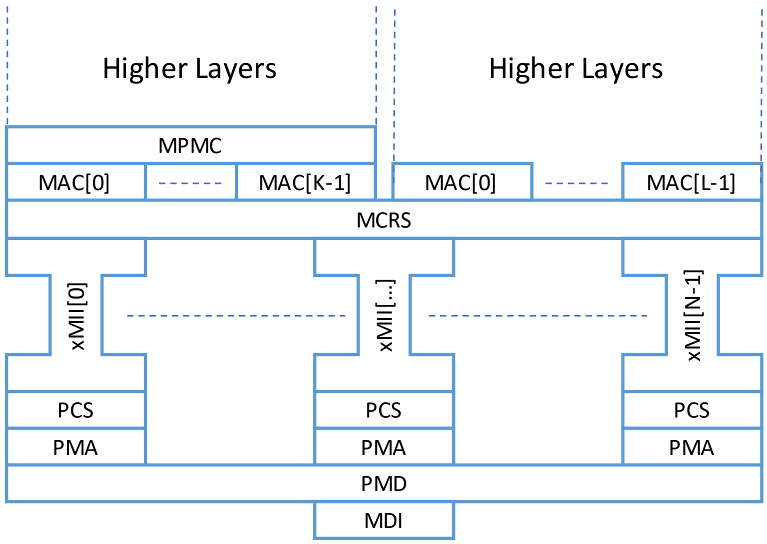


**Figure 10 – Multi-channel reconciliation sublayer in Nx25G-EPON stack**

MCRS adapts the bit-serial data stream received from a standard Ethernet MAC to the parallel format of the PCS service interface, effectively striping any Ethernet frames across multiple physical channels. Each and every PCS instance represents an MCRS channel, carrying data units between OLT and ONU. An MCRS channel operates on envelope quantum (EQ) data units, represented by 72-bit block consisting of eight control bits followed by 64 bits of data. Effectively, a single EQ comprises data from two consecutive medium independent interface (MII) transfers, with re-arranged data and control symbols as shown in Figure 11.

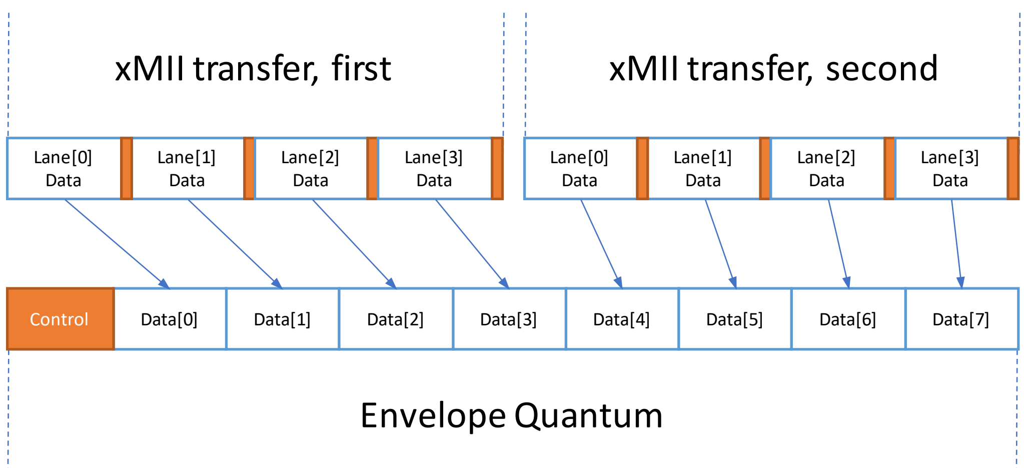


**Figure 11 – Envelope quantum format**

The MCRS encapsulates each Ethernet frame received from the source MAC in a transmission envelope (envelope) structure shown in Figure 12, representing continuous transmission by a specific MAC instance (LLID) on a specific MCRS channel. In a system with a single channel, an envelope includes one or more data frames and may contain at most two partial frames (one at the beginning and one at the end of the envelope) and any number of non-fragmented frames. In systems with multiple channels, envelopes may overlap (see Figure 14) and a frame may be striped over multiple channels with each channel transporting parts of this frame. Note that even though Nx25G-EPON is defined today to operate on two channels, the example shows four channels – the MCRS was defined in a channel independent manner for potential future applications with a larger number of transmission channels.

The envelope header (/EH/) contains a number of additional information about the given envelope, including LLID the given envelope belongs to, the envelope length, and the envelope type (start or continuation) indication. It is worth noting that frames in systems using MCRS may be fragmented, i.e., transmission may be started in one envelope and continued in another envelope as long as the number of fragmented frames is tightly controlled to two at most per envelope.

Each envelope comprises the envelope header /EH/ indicating whether the given envelope is the start of the transmission for the given LLID (envelope start header, ESH for short) or continuation of the transmission for the given LLID (envelope continuation header, or ECH). Every time transmission for a new LLID is started, ESH is transmitted to notify the receiving station of the change of the transmission "context," allowing the receiver to demultiplex transmission to the correct LLID (and associated Ethernet MAC). The envelope header also contains information on the LLID the given envelope is intended to and the length of the given transmission envelope, including all data frames to be transmitted. The data following the envelope header comprises several data octets (/D/) ending with a termination symbol (/T/) followed by an idle symbol (/I/). An example of a transmission envelope with three data frames is shown in Figure 13, where the initial frame to LLID number X is transmitted with /ESH/ headers, while the two following data frames are transmitted with /ECH/ headers.

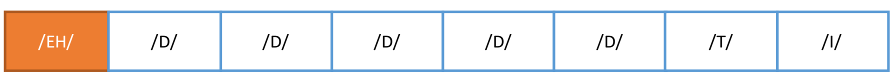


**Figure 12 – Transmission envelope structure**

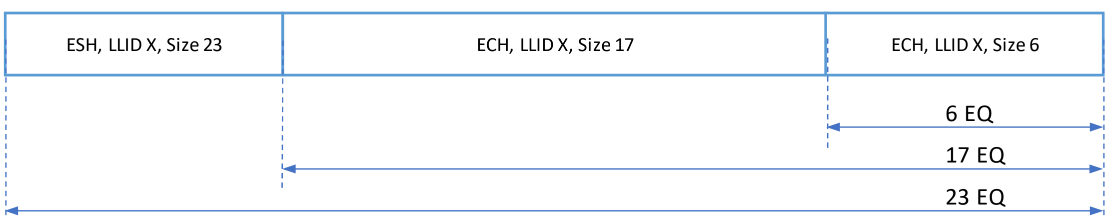


**Figure 13 – Example of a transmission envelope**

The dynamic channel bonding is achieved by interleaving data belonging to a single LLID (i.e., data from a single MAC instance) over multiple envelopes on multiple MCRS channels. Figure 15 shows an example of dynamic channel bonding with partially overlapping envelopes. Each EQ is transmitted on the channel that has the earliest transmission availability.

If there are multiple such channels, the one with the lowest channel index is selected. In other words, the overlapping envelopes are filled with EQs in the increasing order of MCRS channel index, providing a very effective and self-controlling filling mechanism for transmission envelopes.

 

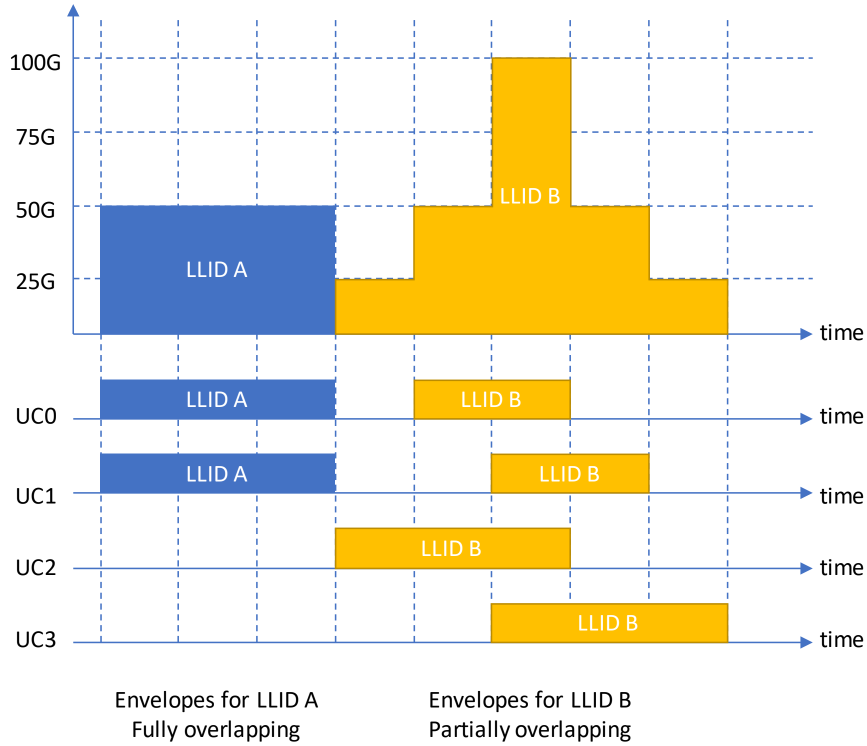


**Figure 14 – Full or partial envelope overlap and the resulting instantaneous data rate**

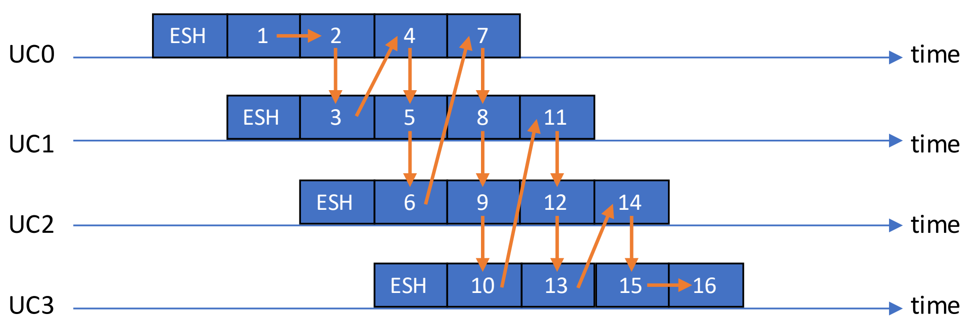


**Figure 15 – Dynamic channel bonding and envelope fill order**

In a multi-channel system using transmission channels with different wavelengths to carry different MCRS channels, each wavelength channel has a different propagation delay in the fiber medium, resulting in a variable propagation delay, causing a certain amount of timing skew between signals received on different MCRS channels. The presence of buffers and state machines in sublayers below the MCRS provides for additional sources of delay that need to be compensated for. To properly restore the order of data transmitted over multiple bonded MCRS channels, the skew between individual MCRS channels needs to be compensated for at the receiving side. The skew remediation mechanism is based on a pair of buffers, controlling and recording the order in which individual EQs are transmitted and recovering the order of these EQs on the receiving end based on the EQ numbering embedded on the transmitting side. In this way, MCRS may compensate for channel skew of arbitrary size, depending only on the depth of transmitting and receiving buffers.

 

The MCRS definitions are data-rate and channel-count independent, allowing bonding of any number of lower-data rate channels to any Ethernet MAC, providing a very flexible and future-proof architecture.

### 3.5. Multi Point Control Protocol in Nx25G-EPON

The MAC control (MACC) sublayer provides real-time control and manipulation of the MAC sublayer, customizing operation and behavior of this PHY agnostic sublayer. Examples of MACC clients include the bridge relay entity, LLC, or other applications characteristic of the particular IEEE 802 network device. In the case of EPON, MACC clients include e.g., discovery client, DBA client, etc. The principles of Multi Point Control Protocol (MPCP) remain largely unchanged from 10G-EPON, controlling the discovery and registration processes as well as scheduling of the upstream bandwidth. Individual MPCP data units (MPCPDUs) have been redefined to support a higher number of scheduling entities (LLIDs) and improve the overall upstream channel efficiency. Detailed definitions of individual messages can be found in Clause 144 in IEEE Std P802.3ca.

Note that there are several subtle changes in the underlying operation of MPCP in Nx25G-EPON as outlined below:

- There are no overlapping grants in Nx25G-EPON. Such grants were used in 10G-EPON to minimize the number of inter-burst gaps and recover some of the associated overhead. In Nx25G-EPON, a single upstream transmission burst may include transmissions from multiple LLIDs, solving the problem in a more elegant manner.
- All granting times in Nx25G-EPON are pre-compensated for the round trip time (RTT), i.e., the round trip delay between the OLT and ONU. This makes each ONU operate as if it were connected back-to-back to the OLT and there were no delay associated with the physical distance between the OLT and ONU. This RTT pre-compensation mechanism simplifies all the time calculations on both the ONU and OLT, where the only added requirement is that the OLT keeps track of any RTT changes and updates its local records accordingly.
- To differentiate Nx25G-EPON MPCPDU structure and operation changes, individual MPCPDUs were allocated new opcode values, allowing for much cleaner delineation from legacy EPON systems.
- A new DISCOVERY MPCPDU was created to provide clearer delineation between the discovery and granting processes. Given the multi-channel nature of Nx25G-EPON, any DISCOVERY MPCPDUs are constrained to the primary channel (first pair of upstream and downstream MCRS channels) established between the OLT and any connected ONU.
- A new SYNC_PATTERN MPCPDU was created to provide a way to configure OLT-optimized synchronization patterns on the ONU. An ONU is still allowed to use the standard-defined default values until the OLT configures new custom values. In certain scenarios, a custom synchronization pattern is believed to be better suited for particular OLT receiver implementations, aiding the burst detection process especially at the more challenging PR30 class power budgets.

### 3.6. Channel Control Protocol in Nx25G-EPON

The Channel Control Protocol (CCP) allows the OLT to query and control the state of individual channels through the exchange of specific CCP data units (CCPDUs): CC_REQUEST and CC_RESPONSE. Individual channels on the ONU may need to be disabled in a controlled manner for a variety of operational reasons, including for example: power saving, scheduled diagnostic / maintenance activities, and optical protection. It is also possible that a channel on the ONU fails and the OLT (and indirectly: NMS) needs to be notified of this fact. Obviously, the reverse is also true, i.e., when a channel is being disabled or fails on the OLT, the ONU should also be reconfigured accordingly to make sure that the disabled / failed channel is no longer in active use.

The ONU channel status discovery is accomplished via an exchange of a CC_REQUEST and CC_RESPONSE CCPDUs. Using the CC_REQUEST CCPDU, the OLT polls the state of all downstream and upstream channels on the given target ONU, without altering the state of any channels.

The ONU may send an unsolicited CC_RESPONSE CCPDU to notify the OLT about any local changes in the channel status, including imminent transceiver element failure, local channel disabling, power failure and resulting channel shutdown. Once the CC_RESPONSE CCPDU from the ONU is received, the OLT saves the channel status information and notifies the NMS of the associated channel status on the given ONU.

The OLT may – at any time – change the state of any of the ONU channels by requesting the transmission of a CC_REQUEST CCPDU, with specific action encoded for each and every downstream and upstream channel for the given ONU. The OLT may enable a channel (state affecting), disable a channel (state affecting), or poll the current channel status (state non-affecting).

The OLT may request the ONU to implement changes for the given data channel in a persistent or non-persistent manner. As the name suggests, any persistent changes are preserved across ONU reset (power cycle) events. Any non-persistent changes are reverted upon ONU reset and re-registration.

# 4. Conclusions

Given the virtually constant increase in bandwidth requirements from end-subscribers as well as changes in commonly utilized networked services, it is expected that by 2025 the existing 10G-EPON deployments will become insufficient to support next generation of multimedia rich digital content at the current split ratios and oversubscription rates. 1 Gbps downstream and upstream speeds to 64 subscribers with the 1:6.4 oversubscription ratio are expected not to be very attractive anymore in the target timeframe, requiring more raw bandwidth capacity while maintaining all the benefit of the shared PON infrastructure. The same holds true for business and enterprise services, though in the case of these services the ability to support multiple instances of services 1+ Gbps without oversubscription will be the critical driver for Nx25G-EPON systems.

When it was first introduced, 1G-EPON represented a substantial step in the evolution of access networking, creating a platform for delivery of bandwidth-intensive applications like IPTV or VoIP. 10G-EPON provided approximately an eight-fold increase in the system capacity (taking the FEC overhead into consideration) at about three times the cost per subscriber, but the observed almost exponential growth in bandwidth demand as well as speed tier wars among service providers result in the need for more access capacity and a scalable solution achieving at least 25 Gbps serial capacity.

Ultra-high definition (4K as well as emerging 8K and 16K standards, including much broader HDR color gamut), video-centric, multimedia-rich services are on the rise, fueled by several years of increased popularity of HD and UHD smart TV sets and their increased penetration in households. It is not uncommon anymore to have at least two concurrent 4K video streams together with background on-demand video content, linear TV broadcast / multicast content, as well as IoT data all sharing a single residential link. In the case of enterprise and business customers, IoT, big data analysis, as well as the trend towards cloud-based services and service mobility all result in a dramatic increase in bandwidth demand, where even small offices require 1 Gbps or more of guaranteed capacity with very strict SLA requirements. Larger companies migrate their network infrastructure to multi-gigabit uplinks, especially in cases of enterprises with centralized databases and cloud-storage for internal use. Similar trends can be even observed in public and private school systems, where multiple campuses are all interconnected to a single data storage location hosting all internal services and resources. It is not, therefore, uncommon anymore for a school district to operate quite advanced and complicated private networks on top of the access infrastructure built on PON or P2P Ethernet access links.

Nx25G-EPON was also designed with other two target applications in mind.

The MDU market is the first of them, focusing on residential and business high density customer locations, where a large fraction of home / offices owners also subscribe to digital services. Such markets exist mainly in certain regions of Europe, as well as in Asia. MDU development is not very popular in the USA, and thus this application does not fit the needs of the American market. The Nx25G-EPON is also naturally suited for deployment in such areas as hospitals, schools and business campuses, as well as governmental and educational institutions, where a large number of wired / wireless users generate a substantial quantity of data traffic which then needs to be delivered to aggregation networks. The improvements in the handling of individual LLIDs and native support for multicast and broadcast services provide more granular SLA enforcement without adding to the upstream channel overhead.

The other target application considered at the time of conception of the Nx25G-EPON effort was mobile backhauling, especially in the light of the emerging 5G technologies requiring substantially more bandwidth than previous generations of mobile data systems, with LTE included. Higher speed Wi-Fi systems (including recently released IEEE Std 802.11ac), resurging interest in small and pico cells, as well as advanced 5G services with expected large volume of smart home and smart office appliances will all drive the need for low cost, high quality SLA systems in the access, providing high capacity at a very competitive cost.

It is expected that Nx25G-EPON equipment will follow the path of system cost increase characteristic of all Ethernet equipment. As the serial data rates exceed 25 Gbps, channel bonding becomes a required system feature due to very large power budgets that cannot be addressed in an economic manner at 50 Gbps and 100 Gbps serial data rates without the use of expensive external optical amplifiers. While the typical expected 10-fold system capacity increase at about three times the port cost may no longer hold for Nx25G-EPON due to the aggregate capacity and system complexity, there are still substantial advantages for operators to migrate to this new system, including the following:

- Much more efficient use of the existing ODN, where higher capacity becomes available without additional investment into the fiber plant.
- Coexistence on the same ODN at least with 10G-EPON systems, where the operator is able to gradually migrate customers from 10G-EPON to Nx25G-EPON, while making no changes to the ODN or any services remaining on 10G-EPON for the time being.
- Ability to serve multiple ONU rates (both symmetric and asymmetric) from a single OLT port in a very bandwidth efficient manner.

The development process of Nx25G-EPONs will keep on driving state of the art engineering in the area of burst-mode receivers, high power laser sources, and ultra-sensitive high data rate photodetectors. Chip integration as well as protocol implementation will also present several challenges yet to be surmounted, mainly in the form of a reliable multi-rate discovery process, OLT receiver design, etc.

It is also anticipated that the rapid stabilization of Nx25G-EPON PMD parameters may potentially bring about a closer cooperation between FSAN/ITU-T and IEEE PON groups, leading to possible convergence of next generation PON systems, at least at the PHY level. This would allow hardware manufacturers to achieve higher production volumes and cut equipment costs, making both PON systems far more economically attractive then when considered separately.

# 5. Abbreviations and Definitions

## 5.1. Abbreviations

| | |
|---|---|
| CCP | Channel Control Protocol |
| CCPDU | channel control protocol data unit |
| CFI | call for interest |
| CSMA/CD | carrier sense multiple access / collision detection |
| DML | directly modulated laser |
| ECH | envelope continuation header |
| EPON | Ethernet passive optical network |
| EQ | envelope quantum |
| ESH | envelope start header |
| FEC | forward error correction |
| Gbps | gigabits per second |
| GLID | group link ID |
| GPON | gigabit passive optical network |
| LDPC | low density parity check |
| LLID | logical link identifier |
| MAC | media access control |
| MCRS | multi-channel reconciliation sublayer |
| MDU | multiple dwelling unit |
| MII | medium independent interface |
| MLID | management link ID |
| MP2P | multi point To point |
| MPCP | Multi Point Control Protocol |
| MPCPDU | multipoint control protocol data unit |
| MPMC | multi point MAC control |
| MTU | multi tenant unit |
| NMS | network management system |
| OAM | operations and administration management |
| ODN | optical distribution network |
| OLT | optical line terminal |
| ONU | optical network unit |
| P2MP | point to multi point |
| PCS | physical coding sublayer |
| PHY | physical layer |
| PLID | physical layer ID |

 

| | |
|---|---|
| PMA | physical medium attachment |
| PMD | physical medium dependent |
| PON | passive optical network |
| PSC | passive splitter combiner |
| QC | quasi cycle |
| RDM | random delay mechanism |
| RTT | round trip time |
| SLA | service level agreement |
| TDM | time division multiplexing |
| TDMA | time division multiple access |
| ULID | user link ID |
| WDM | wavelength division multiplexing |

### 5.2. Definitions

| | |
|---|---|
| downstream | Information flowing from the OLT to an ONU |
| upstream | Information flowing from an ONU to the OLT |

## 6. Bibliography and References

*NG-EPON CFI*, http://www.ieee802.org/3/cfi/0715_1/CFI_01_0715.pdf

*NG-EPON Objectives*, http://www.ieee802.org/3/ca/documents/P802_3ca_objectives.pdf

*Next-Generation Ethernet Passive Optical Networks: 10G-EPON*, Marek Hajduczenia, Henrique J. A. da Silva, editor: Krzysztof Iniewski, 21 September 2010, Wiley

## 7. Acknowledgments

The author would like to thank the following reviewers for their valuable input and comments received during the development of this document: Ryan Tucker, James Graves.

 

# Cloud-Native CMTS

## Challenges and Opportunities

A Technical Paper prepared for SCTE•ISBE by

Gurpreet Singh, Sr. Product Manager, Nokia Corporation, SCTE•ISBE Member
2301 Sugar Bush Rd
Raleigh, NC 27612
gurpreet.singh@nokia.com
(919)-946-1235

 

# Table of Contents

**Title** **Page Number**



## List of Figures

**Title** **Page Number**

# 1. Introduction

Distributed access architecture (DAA) has become a well-known concept in the cable access industry and is shaping the technologies deployed by cable operators to satisfy the ever-growing needs for higher performance, greater reliability, and increased network capacity. DAA can be implemented by splitting the functionality of an integrated-converged cable access platform (I-CCAP) into either remote-PHY, remote-MACPHY or a flexible MAC architecture (FMA).

As part of this network transformation, cable operators can opt to move the DOCSIS® MAC functionality to a virtualized form of a software-based cable modem termination system (CMTS) on a general-purpose server. The virtualization of CMTS functionality was initially driven by the need to reduce power and conserve space within a headend/hub, and to reduce the cost of headend hardware. But now it has become the cornerstone to operational and service agility derived from the benefits offered by software-defined networking (SDN), network functions virtualization (NFV) and cloud-native technologies. In addition, by leveraging cloud-native network functions, cable operators can transform operational and service delivery methods based on similar core principles and a framework successfully used by cloud service and web scale providers.

This paper outlines the key benefits of a cloud-native CMTS for cable operators as well as the many of the challenges both vendors and operators can face when going "cloud-native." Cable operators are urged to take a phased approach to a cloud-based network evolution largely driven by the maturity of the enabling technologies and the processes and frameworks currently in use by many organizations. In doing so, operators will maximize the benefits of DAA and reduce many of the operational challenges that will be encountered in the journey toward a cloud-native cable access network.

# 2. Migration to Cloud-Native CMTS

A CMTS implements multiple network functions such as packet transmit/receive (Tx/Rx), subscriber management, filtering, classification, DOCSIS MAC scheduling, framing, encryption and others.

The non-cloud-native implementations of a CMTS is typically implemented as monolithic applications where each of the individual network functions are stateful and rely on inter-process communications. The transition to cloud-native architecture requires the separation of network functions into individual microservices enabling stateless network functions where possible. This task requires both vendor and operator alike to have strong skills and knowledge of microservices architecture principles. *Figure 1* below shows the key principles and concepts of a cloud-native architecture.

 

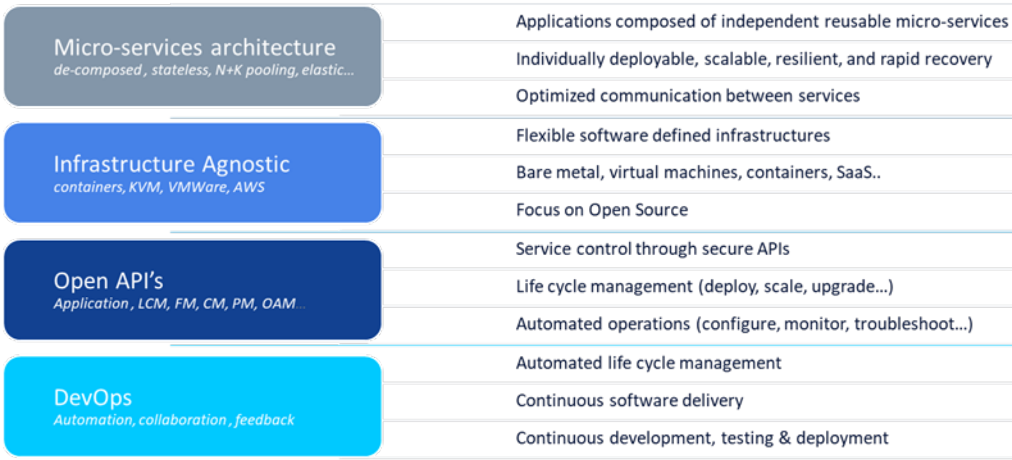


**Figure 1: Key principles of cloud-native architecture**

Tools, skill sets, and a framework supporting each of these aspects are essential for successful migration to cloud-native implementation.

Migration to a cloud-native CMTS architecture, if not properly implemented, can hinder network performance. Moving the state of a network function outside of the microservices container can result in network processing delays. This performance impact can be alleviated by implementing an efficient caching mechanism for frequently accessed data.

This approach also provides for greater service agility so that new applications can be brought to market faster. It should be noted, that if the composition of the network function is not expected to change often, the benefit of a cloud-native architecture may be limited to faster bug fixes and enhancements. Cloud-native implementation of the CMTS may be considered a method for performance and footprint optimization for a virtualized CMTS. Common services and a common control plane can be shared across multiple CMTS software instances thus reducing the compute and storage footprint for the given subscriber and services scale supported by a group of CMTS instances. Cloud-native implementation also enables use cases such as in-service or hitless software upgrades and high availability. These use cases can be cost prohibitive or operationally burdensome to realize when the CMTS functionality is implemented on proprietary hardware as monolithic software. *Figure 2* below shows the workflow for the hitless software upgrade use case enabled by cloud-native implementation of CMTS.

 

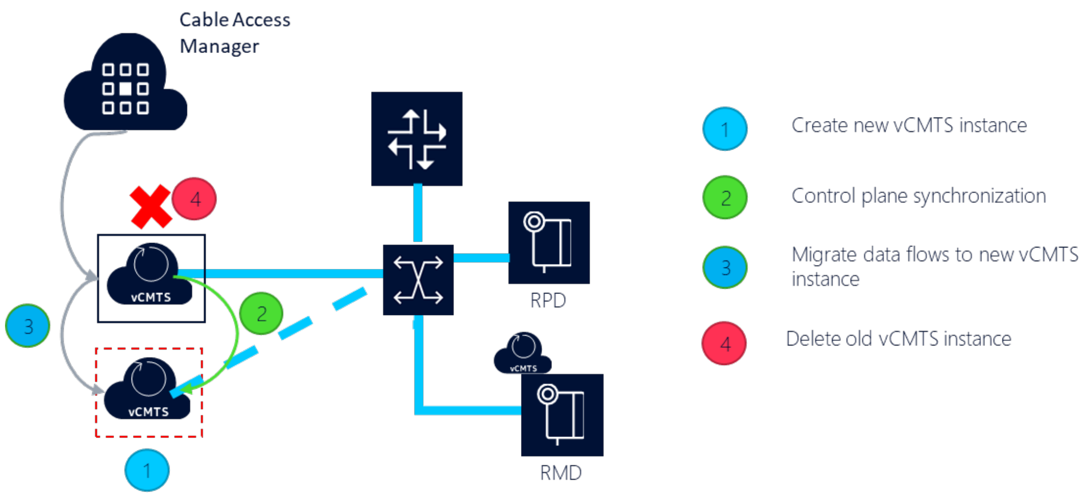


**Figure 2: Hitless software upgrade**

The following are several of the factors to be considered for cloud-native migration:

- Complexity of the monolithic app and communication across the individual network functions
- Degree of difficulty to convert stateful network functions to stateless
- How often the state of the network functions change
- Frequency at which the network functions need to access the state information
- Efficiency of caching mechanisms

Imagine what it would mean to a cable operator's business if customers rarely – if ever – were impacted by maintenance outages or modem resets.

## 3. Impact of Cloud-Native CMTS on Network Slicing

Network slicing in the cable access network is one of the key use cases enabled by cloud-native CMTS.

 

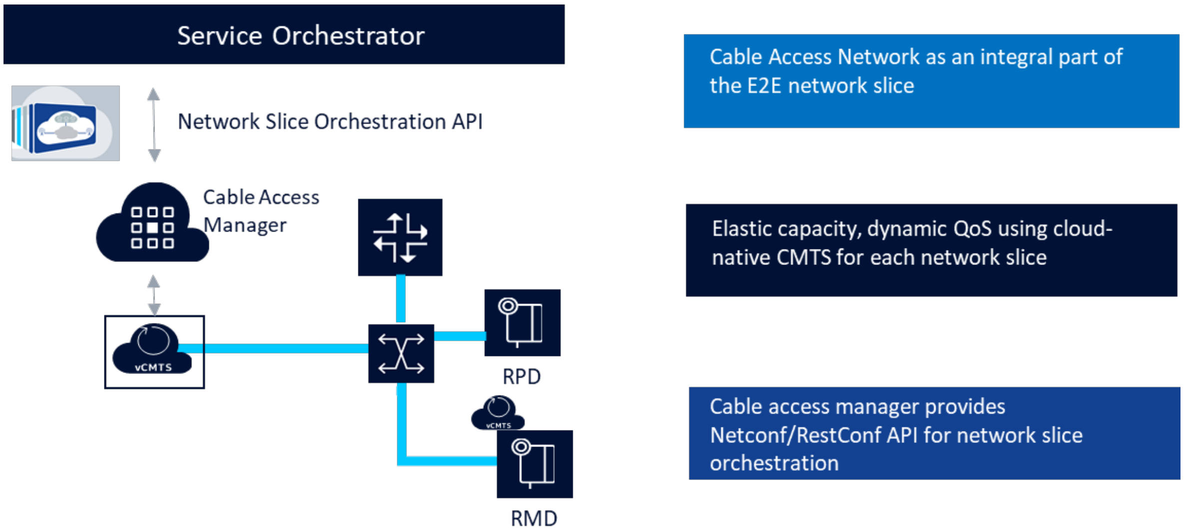


**Figure 3: Cloud-native CMTS and service orchestration API enables E2E network slicing**

Network slicing is an end-to-end (E2E) concept. Network slicing requires that each network segment in the data path — whether radio access network (RAN), cable access, or transport network — has an easy-to-use application programming interface (API) to the orchestrator for policy configuration. Effective network slicing depends on the centralized service orchestrator performing the functions of isolating, orchestrating, and ensuring the service level agreement (SLA) is met accordingly.

In the case of cable access networks, if a CMTS is implemented as an I-CCAP, then the slicing is dependent on the I-CCAP's quality of service (QoS), scheduling, traffic shaping and admission control capabilities. If the CMTS is implemented as cloud-native running on shared resources, then the shared resource allocation for each network slice also plays an important role. Careful allocation of central processing unit (CPU), memory and storage resources are required to make sure that each of the services provisioned on the CMTS instance gets enough resources to guarantee the SLAs. Although advances have been made in the NFV-associated technologies for performance isolation, such as CPU core pinning and smart caching techniques, unaddressed challenges remain nonetheless.

Compared to an I-CCAP implementation, a cloud-native CMTS integrates easily with an E2E orchestration platform and enables network slicing in the cable access part of the data path. The cloud-native CMTS framework provides easy-to-use APIs to the orchestrator for defining network slice policies in the cable access network. If the current CMTS instance cannot satisfy the service request due to lack of infrastructure resources, additional infrastructure resources are allocated by autoscaling the CMTS software instances.

## 4. Infrastructure framework & processes

DevOps and continuous integration / continuous development (CI/CD) are two foundational principles fueling the cloud-native migration of software applications. DevOps frameworks and CI/CD pipelines enable each component or microservice that is part of the cloud-native application to be independently developed, tested, delivered and deployed. Independent lifecycle management of software components enables shorter development and release cycles on the vendor side and shorter acceptance testing, qualification and deployment cycles on the operator side.

 

*Figure 4* illustrates the various stages involved in DevOps. In the networking industry products and solutions are typically developed by the vendors and consumed by the service providers and network operators.

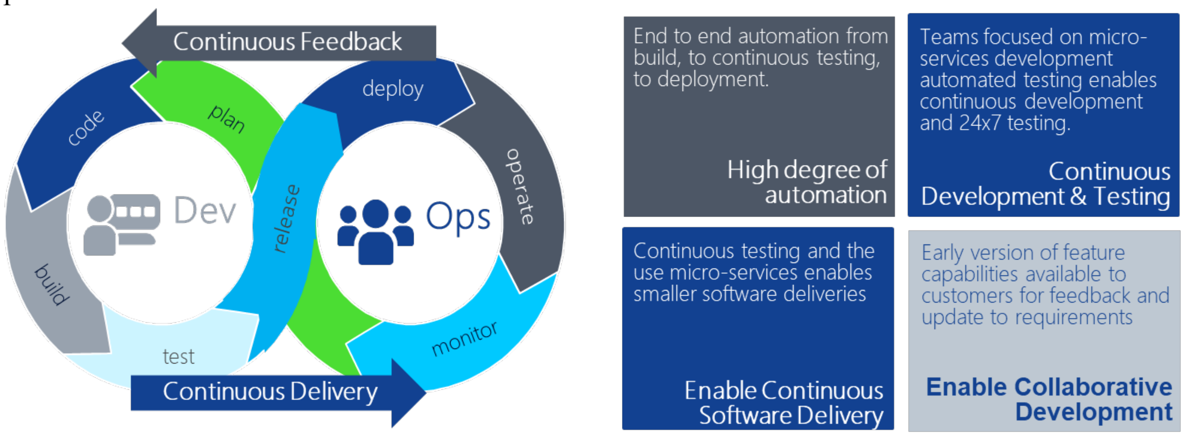


**Figure 4: DevOps stages and core principles**

The left-hand side (plan, code, build, test and release) in Figure 4 is applicable to the vendors and the right-hand side (deploy, operate and monitor) is applicable to service providers and network operators.

The key to shorter cycles in development is the ability to automate the various development and operational stages as much possible without sacrificing quality, reliability and resiliency. This is accomplished with a continuous feedback loop from operations into software development for refining processes and frameworks. Additionally, the speed of execution is enabled by the optimized size of the entity that is being worked on. In the cloud-native world this entity is called a microservice.

If vendors have the processes and frameworks to develop and deliver modular components of the application in a rapid manner (e.g., daily or weekly), but the test cycles and deployment cycles on the cable operator side are still not aligned to the vendor's frameworks, then the benefits of DevOps cannot be realized by the operator. The same is true if the operators are DevOps-ready but vendors are not.

Another challenge faced by the industry is how to bridge the DevOps frameworks for continuous delivery and continuous deployment. The pipeline must cross the boundary from vendor to operator. Technically this can be accomplished by software as a service (SaaS) platforms that provide separate interfaces and front-ends with appropriate capabilities for vendors and operators. Vendors must develop infrastructure and processes that help build the operator's confidence in the weekly builds in terms of quality and reliability. Frameworks and platforms to cross the organization boundaries in an automated fashion open the door for continuous delivery and continuous deployment.

As such, processes must be in place for doing canary builds and deployments before large scale deployments. Automatic rollback and rollout of builds is an essential feature. Similarly, an automated, phased approach of moving only a small portion of the workload first to test the build and then gradually move larger workloads becomes essential as well.

 

To realize the benefits offered by cloud-native applications, associated frameworks that enable DevOps and CI/CD are required by both vendors and network operators.

## 5. Organizational Culture

The importance of cultural and process challenges that will be encountered by both vendor partners and cable operators during this transition to cloud-native applications cannot be underestimated.

Legacy infrastructure and processes are based on consuming significant and infrequent software builds or applications. Testing and validation were done manually and completed over months if not years. Automated testing rarely provided the confidence for the move to an operational environment.

Moreover, the existing operational workforce may not have the skill sets to maintain, troubleshoot or deploy the software applications using the DevOps tools and technologies. Similarly, vendors and operators alike may not be capable enough to deliver or test new software builds on a weekly or daily basis.

Edgar Schein, the author of "Organizational Culture and Leadership" said " Culture is a pattern of shared tacit assumptions that was learned by a group as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems."

Simply, organizational culture is learned through experience. But these assumptions, it could be argued, may not all be valid in an agile organization. The problem set may be different. What has worked in the past may not work anymore because of the external as well as internal factors in play.

DevOps rules and principles within any organization may have to be tweaked or modified to achieve end goals (e.g., superior customer experience) depending on the current structure and culture of the organization. Organizations must embrace fluid learning when on the path to cloud-native applications. Stating the obvious, vendors and service providers must have common goals and objectives and be aligned to this end.

## 6. Phased Approach to Cloud-Native

Cable operators are heavily invested in the I-CCAP infrastructure and historically have relied on node splits, improved DOCSIS modulation iterations, and/or the cultivation of new, unused, or converted spectrum to address the need for more bandwidth. This current strategy, while effective, can be expensive and without constant improvements in video encoder technologies, may even be unsustainable.

Distributed access architectures present an attractive option to future-proof investments and plan for long-term demands. While the CableLabs R-PHY and R-MACPHY specifications are close to maturity, there is a need to achieve multi-vendor interoperability and increased reliability from these newer technologies.

A "one shot" approach of attacking the network layer (DOCSIS MAC and PHY functionality), controller layer (SDN and associated controllers) and management layer (back-office systems) involves high investment and high risk as there are so many different elements of the network being changed at the same time.

*A phased approach, where an operator can move from the network layer to management layer reduces complexities and associated risk. Additionally, moving from I-CCAP to virtualized DOCSIS MAC* *functionality not only requires additional investment for upgrading the access nodes and hub infrastructure, it also involves new operational procedures and skill sets as outlined earlier.*

Figure 5 illustrates the phases of the incremental approach for cloud-native migration of CMTS functionality.

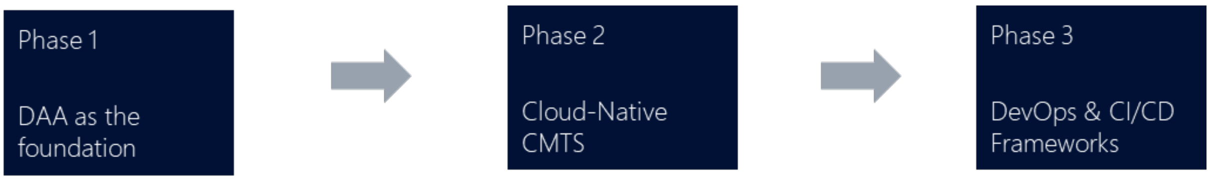


**Figure 5: Phased approach to cloud-native and service agility**

## Phase1: Flexible MAC Architecture for DAA

DAA provides a strong foundation for the transition with minimal change to operational procedures and management layer.  A flexible MAC architecture is one approach to DAA deployments that can mitigate risk and provide the greatest network flexibility. This FMA specification is currently being finalized under the FMA working group umbrella at CableLabs.

FMA presents a solution where the DOCSIS MAC functionality of the CMTS may be virtualized OR placed in the cable access node (R-MACPHY node). DOCSIS MAC functionality may be virtualized and deployed on a commercial-off-the-shelf server or an appliance in the outside plant (OSP), hub, headend or in the data center. Figure 6 shows the different deployment options for DOCSIS MAC functionality of the CMTS.

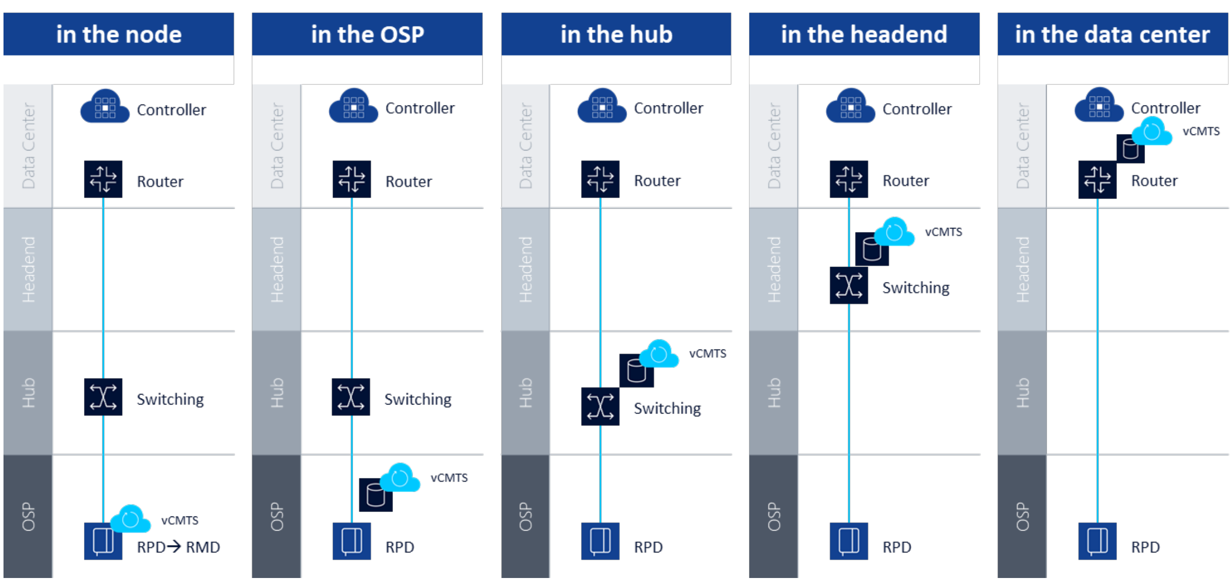


**Figure 6: Flexibility in implementation and deployment of DOCSIS MAC**

 

### Phase 2: Cloud-Native VNF

The second phase involves cloud-native implementation of the DOCSIS MAC based on the maturity of the associated technologies in terms of feature support, performance and scale.

Cloud-native implementations typically involve virtualization of software applications using container technologies where each microservice is implemented as a container that can be instantiated, deployed, upgraded and terminated independently. Similarly, on the vendor side each microservice can be developed, tested and delivered independently thereby accelerating an operator's ability to introduce new features and applications.

The granularity of breaking the CMTS application into microservices depends on the complexity of the monolithic application and the maturity of the framework to manage the lifecycle of each microservice independently. Lifecycle management involves developing, testing, qualifying, deploying and upgrading each microservice component independently.

Figure 7 shows the first level of modularization for a DOCSIS MAC functionality. The high-level categories are DOCSIS control plane, DOCSIS data plane (upper MAC processing and lower MAC processing) and the common services used by the control plane and data plane components. The control plane component does not process data packets and hence the need for compute resources is lower as compared to the data plane components.

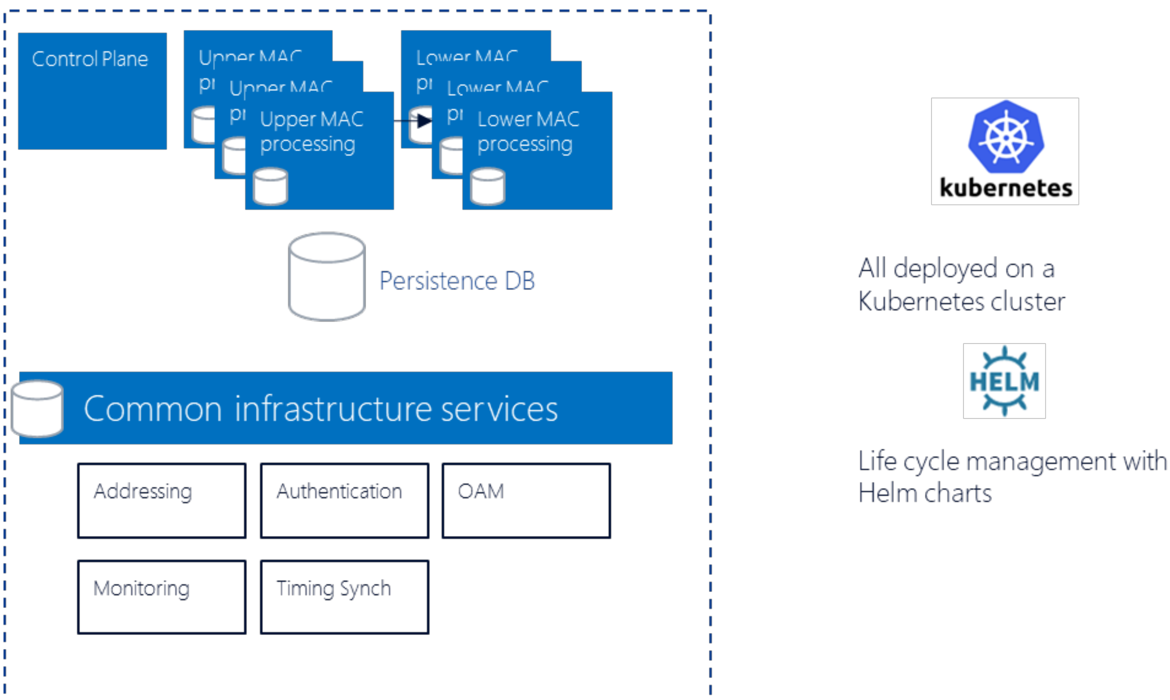


**Figure 7: Cloud-native CMTS decomposition**

This reason lends itself to a possibility of a common control plane component that can be shared across the multiple data plane instances. Similarly, the common service components may be shared across the control and data planes. This first level of modularization presents an incremental and less disruptive approach. Prototyping and learning from the incremental changes provides further insights into how much granularity is enough and what is the corresponding impact on the performance of the overall CMTS functionality in terms of packet processing (throughput), latency and jitter. Upper and lower MAC processing can be further split into more granular components as shown in Figure 8.

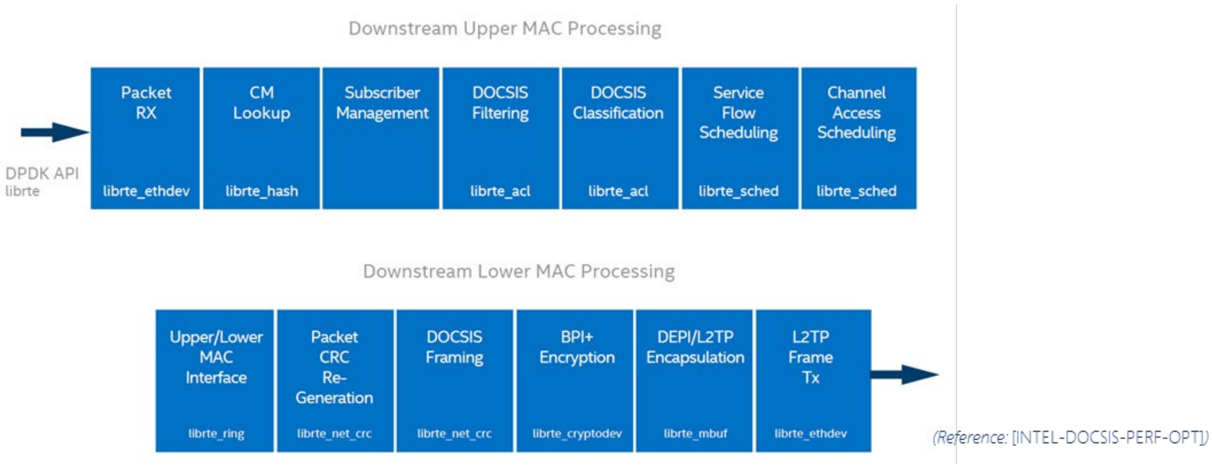


**Figure 8: DOCSIS upper and lower MAC processing components**

Individual components of the VNF – also referred to as VNFC in ETSI NFV terminology – should involve a single function that has a well-defined scope and is independent of the other VNFCs. The functionality for each of the components should be defined based on scalability, redundancy and self-healing capabilities. On-demand capacity scaling of only the component where additional capacity is required is a key characteristic of the cloud-native VNF implementation. Note that there is a balance to be maintained between the size of each VNF component and efficiency of the CI/CD pipelines.

High availability needs to be architected at a more granular level for a cloud-native CMTS as compared to monolithic CMTS implementation. Spreading of VNFC components across servers, racks and data centers can provide enhanced HA and resiliency, but it can have a negative performance impact for latency and jitter. These tradeoffs must be fully understood and vetted prior to embarking on a cloud-native network. [ETSI GS NFV-EVE 011] provides further details on the location independence of the VNF components. It is important to note that location independence should not be interpreted as absolute. Some of the components may need to be co-located to optimize performance, latency and proximity to cache and database infrastructure. At the same time anti-affinity may be required to provide geographic high availability.

[MIGRATING-TO-CLOUD-NATIVE] provides some good techniques and examples for migrating to cloud-native application architectures.

In parallel, it is expected that cable operators will gain experience with the associated DevOps and CI/CD technologies and understand the level of effort and complexity involved in lifecycle management of each microservice component that is part of the cloud-native application.

### Phase 3: DevOps and CI/CD

The third phase involves evaluating, piloting and deploying of DevOps and CI/CD frameworks. Implementing Phase 3 highly depends on collaboration between the vendors and cable operators.

 

Since the dependencies in this phase cross organization boundaries, this will take the longest to mature and stabilize.

Figure 9 illustrates the transformative process from a waterfall to agile, to DevOps + CI/CD process on the right. The industry players are currently at different stages in the process transformation journey with most of them at the agile process stage.

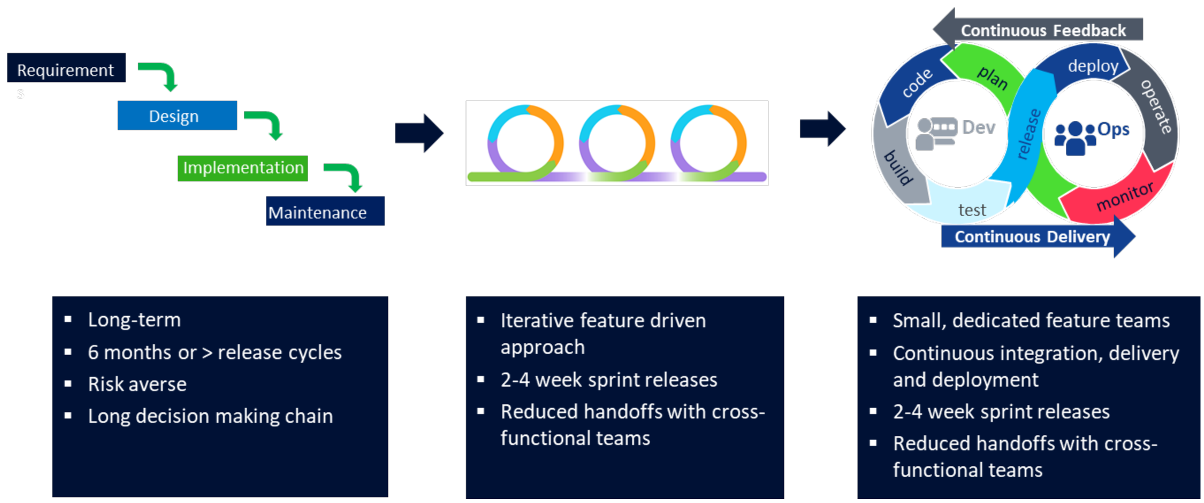


**Figure 9: Process transformation**

***Considering the time and complexities associated with this evolution, it is prudent for cable operators to evaluate solutions that offer incremental benefits at every stage.*** Going for a complete overhaul out of the gate can increase the risk considerably as other market dynamics with competing technologies could impact churn if services and applications don't remain stable and ever-improving.

One of the key enablers to achieve the benefits offered by DevOps + CI/CD platforms is the ability to automate the various processes and activities in the development and deployment pipelines. Automation enables the VNF vendors to deliver incremental modular software builds at a rate that is critical for DevOps and helps the operators to stage, qualify and deploy new builds at a higher pace as well.

Figure 10 shows an example of a CI/CD pipeline with various activities involved.

 

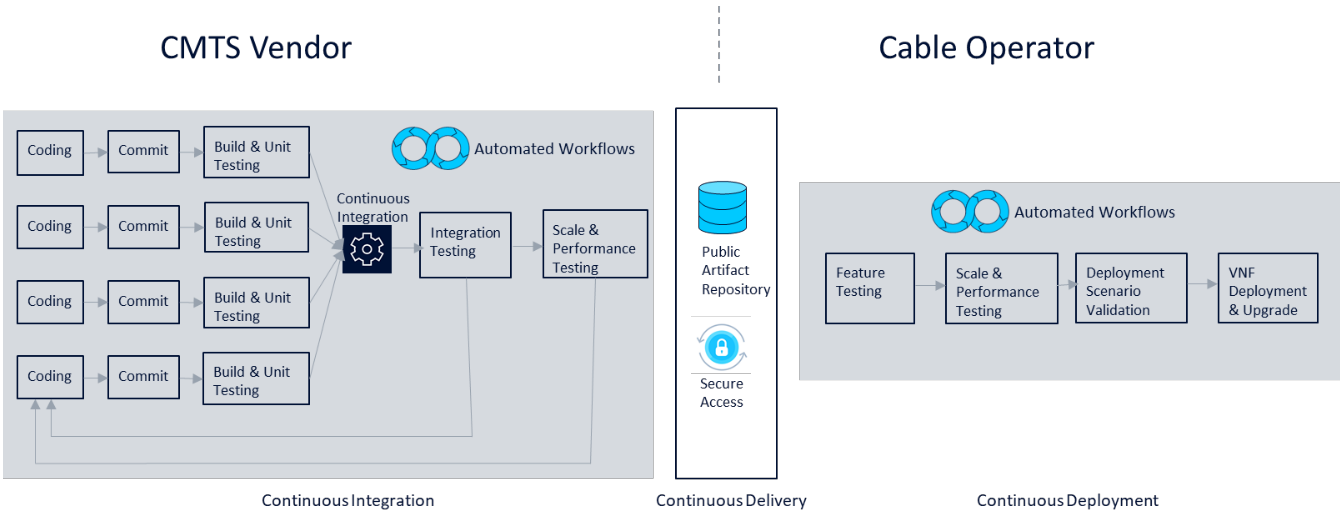


**Figure 10: CI/CD pipeline**

Key factors for an effective CI/CD pipeline implementation include:

- Automated workflows on the VNF vendor side for continuous development and continuous integration
- Automated workflows on cable operator side for staging, build qualification, validating the build for E2E deployment scenarios
- Centralized orchestration for VNF instantiation, VNFC instantiation, VNF upgrades, VNFC upgrades, optimization of NFVI resource allocation, auto-scaling and VNFC migration
- Well-defined interfaces and APIs to vendor and cable operator frameworks for seamless and automated delivery of artifacts
- Automated delivery and digital software supply chain addressing customer subscriptions, entitlements, license key generation
- Secure access to the artifact repository for both vendors and cable operators
- Ability to rollout and rollback builds in a seamless manner

## 7. Conclusion

Implementing a cloud-native CMTS presents a tremendous opportunity for cable operators to re-architect their network. Deploying a cloud-native CMTS using the DevOps and CI/CD platforms discussed in this paper can extend the life of the existing DOCSIS based networks to respond the ever-increasing capacity and performance needs driven by market and consumer demands.

A successful transformation to a cloud-native network relies heavily on the ability of both vendors and cable operators alike to not only build and deploy the cloud-native VNFs but also transform the associated frameworks and processes, including making sure both the operator and the vendor have well-developed, coordinated DevOps cultures in place. It is critical that the industry takes a phased approach to such a significant transformation to avoid both internal and external pitfalls and to capitalize on the real benefits of a cloud-native network architecture.

 

# 8. Abbreviations and Definitions

## 8.1. Abbreviations

| | |
|---|---|
| API | application programming interface |
| CCAP | converged cable access platform |
| CI/CD | continuous integration/continuous development |
| CMTS | cable modem termination system |
| COTS | commercial off the shelf |
| CPU | central processing unit |
| DAA | distributed access architecture |
| DOCSIS | Data-Over-Cable Service Interface Specifications |
| E2E | end-to-end |
| e.g. | for example (*exempli gratia*) |
| ETSI | European Telecommunications Standards Institute |
| FMA | flexible MAC architecture |
| HA | high availability |
| I-CCAP | integrated converged cable access platform |
| MAC | media access control |
| NFV | network functions virtualization |
| OSP | outside plant |
| PHY | physical layer |
| QoS | quality of service |
| RAN | radio access network |
| R-PHY | remote PHY |
| R-MACPHY | remote MAC-PHY |
| SaaS | software as a service |
| SDN | software defined network |
| SLA | service level agreement |
| Tx/Rx | transmit/receive |
| VNF | virtual network function |
| VNFC | VNF components |

## 8.2. Definitions

| | |
|---|---|
| cloud-native VNF | VNF with a full adherence to cloud native principles, or a VNF that is transitioning to cloud native principles.<br>*Reference: ETSI GS NFV-EVE 011* |
| continuous integration / continuous development (CI/CD) | In software engineering, CI/CD or CICD generally refers to the combined practices of continuous integration and continuous delivery. In the context of corporate communication, CI/CD can also refer to the overall process of corporate design and corporate identity *Ref: Reference: Wikipedia* |
| DevOps | A set of practices that combines software development and information-technology operations which aims to shorten the systems |

 

| | |
|---|---|
| | development life cycle and provide continuous delivery with high software quality.<br>*Reference: Wikipedia* |
| distributed access architecture (DAA) | A cable network architecture in which functions that usually reside in the headend or hub are located closer to the user. Moving functions into the network reduces the amount of hardware the headend (hub) needs to house, thus creating efficiencies in speed, reliability, latency and security.<br>*Reference: CableLabs* |
| flexible MAC architecture (FMA) | A subset of DAA in which the MAC electronics can be located in the headend or hub, in the node (or shelf), or elsewhere, and which enables the standardization of the complete disaggregation of the CCAP management, control, and data planes (SDN, NFV, and PNF).<br>*Reference: CableLabs* |
| microservice | Distributed modular software components. |
| network functions virtualization (NFV) | Principle of separating network functions from the hardware they run on by using virtual hardware abstraction<br>*Reference: ETSI GS NFV 003* |
| network function virtualization infrastructure (NFVI) | Totality of all hardware and software components which build up the environment in which VNFs are deployed<br>*Reference: ETSI GS NFV 003* |
| remote MACPHY (r-MACPHY) | A subset of DAA in which the MAC layer electronics and PHY layer electronics are located in a shelf or node. See *flexible MAC architecture*. |
| remote PHY (R-PHY) | A subset of DAA in which the MAC layer electronics remain in the headend or hub, and the physical layer electronics are located in a shelf or node. |
| virtual (or virtualized) network function (VNF) | Implementation of a network function in software that can be deployed on a network function virtualization infrastructure.<br>*Reference: ETSI GS NFV 003* |

## 9. Bibliography and References

ETSI GS NFV-EVE 011: *Virtualized Network Function; Specification of the Classification of Cloud Native VNF implementations*; ETSI NFV

INTEL-DOCSIS-PERF-OPT: *Maximizing the performance of DOCSIS 3.0/3.1 processing on Intel Xeon Processors*; Intel Architecture Study

MIGRATING-TO-CLOUD-NATIVE: *Migrating to Cloud-Native Application Architectures*; O'Reilly; Matt Stine

MONOLITHIC-TRANSFORM: *Monolithic Transformation*; O'Reilly; Michael Cote`

 

# Quantifying Wi-Fi Coverage; “CQ-Score” An Adaptive Quality Metric to Assess Wi-Fi Coverage in a Network

A Technical Paper prepared for SCTE•ISBE by


**Kurt Lumbatis**
Distinguished Software Engineer
Commscope
3871 Lakefield Drive, Suwanee, GA 30024
678 473 2921
Kurt.Lumbatis@commscope.com

**Navene Kannan**
VP, Product Management
Commscope
101 Tournament Drive, Horsham, PA  19044
215 323 2398
Nav.Kannan@commscope.com

# Table of Contents

**Title** **Page Number**



## List of Figures

**Title** **Page Number**



## List of Tables

**Title** **Page Number**



 

# Introduction

Wireless communications (Wi-Fi) has become ubiquitous in homes and in other areas of modern life. Most devices in our home and daily life require Wi-Fi connectivity. The main connection point of these devices to the internet is an access point (AP) which has either a direct or indirect connection to the internet via a wide area network (WAN). In recent years, smart phones, tablets, video set top boxes (STBs), and Internet of Things (IoT) devices have increasingly begun to rely on Wi-Fi as their primary connection point to the internet. With the proliferation of the number and types of these devices, the mobility and distance of these devices from an AP, and the increased bandwidth requirements of devices connecting via Wi-Fi has come the need to quantify the 'quality' of a device's connection to the AP. The 'quality' of a Wi-Fi device's connection becomes important in order to ensure adequate throughput for all devices in a given network, whether it be a home or a business network.

## 1. Problem Statement

With the advent of multiple access points in the home, the need for understanding the performance in the presence of additional access points is crucial. The practical needs are characterized in the following questions, raised from the perspective of a service provider:

1. For the clients being served in a given home, is there adequate coverage? More specifically, assuming capable clients, does the Wi-Fi coverage appear adequate?
2. How do I know if there is a need for an extender?
3. How do I know if the coverage has become better with the extender(s), and to what extent?

The solution to the performance characterization must be meaningful and relate to real-life experience. It should be adaptive accounting for the mobile nature of clients and must be **specific** to that subscriber and that home or network. The solution must not be a cookbook number associated with a device (or even a set of devices).

The purpose of these questions is to arrive at **actionable** data which a service provider may use to provide customer solutions along the lines of:

1. Documentation that the Wi-Fi coverage inside the network is adequate; perhaps highlighting the capability (or the lack thereof) of specific Wi-Fi clients.
2. Documentation that the Wi-Fi coverage in the home may not be adequate for some (or many) of the devices, based on the movement (or location) of such devices and thus an (perhaps additional) extender is suggested.
3. Information to the subscriber that as a result of the addition of the extender there is quantifiable improvement in the coverage, validating the addition.

This paper proposes a novel approach of assessing not only a single device's connection quality to the network, but also allows for assessing the quality of Wi-Fi coverage of an entire network. The assessment being presented here is named the Coverage Quotient (CQ) Score -- *Patent Pending*.

# Quantifying Wi-Fi Coverage

The motivation to assess the quality of Wi-Fi coverage in a given network (or home) is to accurately describe the coverage of not only individual devices, but to determine a total 'coverage score' for a given network. The proposed method overcomes the shortfalls of current methodologies as described in the following sections.

## 2. Current Measurement Methodologies

In current Wi-Fi deployments, very little information is 'known' about a signal being received at a non-AP STA device. The reason is that most devices don't support methods for reporting the signal being received at its location from the AP. Protocol elements are defined in the IEEE 802.11 standard as will be discussed in the following sections. However, most low-cost STA devices don't implement these methods. One additional problem is that, even when available, some of these methods waste valuable Wi-Fi bandwidth in their implementation. Additional methods currently in use have other shortfalls. The following sections discuss these methods.

### 2.1. IEEE 802.11 – Link Measurement

The 802.11k amendment as part of its radio resource measurement specifies a link measurement request/report pair. These requirements are defined in [1] in sections 9.6.7.4 and 9.6.7.5. The link measurement request allows an AP to query a STA device as to the received radio metrics it its location. The link measurement report is the STA's response to the link measurement request and contains two important pieces of information:

- RCPI – indicating the received channel power of the corresponding link measurement request frame at the STA.
- RSNI – indicating the received signal-to-noise ratio at the STA after the link measurement request frame was received.

These two pieces of information quantify the quality of the signal at the STA's location. Sadly, however, very few devices or APs ever implemented support for the link measurement request/response. This method also suffers from the inaccuracy of the values being reported from the STA devices, as most of these devices do not accurately calibrate the receive RF paths. This method does, however, allow for an estimation of MIMO characteristics of the network, if the frame is transmitted as a non-basic-rate transmission. Thus, these measurements can be utilized only as an indication of the signal strength at the STA from an AP.

### 2.2. IEEE 802.11 – Beacon Report

The 802.11k amendment also adds the radio resource measurement capability of a beacon request/report. The Wi-Fi Alliance (WFA) recently codified these measurements as part of their EasyMesh™ protocol[2].

It is outside the scope of this paper to fully document the procedures, however, the purpose of these measurements is to allow an AP to request signal strength measurements from the STA based upon beacon frames (passive scan) or probe response frames (active scan) received by the STA from all APs it 'hears'. The report in part returns:

- BSSID – indicating the MAC address from which the signal was received.

- RCPI – indicating the received channel power at the STA of a beacon or probe response frame received from a particular BSSID (AP).
- RSNI – indicating the received signal-to-noise ratio at the STA when the beacon or probe response frame was received from a particular BSSID (AP).

These pieces of information quantify the quality of a signal at the STA's location for received beacons or probe responses. Again, these measurements suffer from a lack of receiver calibration and understanding of MIMO characteristics of data frames at its particular location since beacons and probe responses are 802.11 management frames and thus are non-MIMO frames, typically transmitted with a very robust modulation type. Additionally, the method must rely upon knowledge of the connection point of the STA (BSSID) in order to determine the link point from the various BSSIDs it reports. Thus, although the report is useful in some respects, it is not useful as a tool to measure network coverage.

### 2.3. Upstream Signal Strength and Noise Floor

The most common method currently in use to determine coverage in a network is to assess the RSSI or RCPI of a STA's transmission to its AP. This method allows the network to assess the coverage in the network based upon a received signal strength from a STA at the AP. The WFA has codified these measurements in the Data Elements Specification. Utilization of upstream signal strength and noise floor measurements present various problems.

1) Mobile devices typically have lower transmission power than the APs to which they are connected.

2) Most APs do not calibrate the receive signal strength, thus the actual received signal strength, (RSSI or RCPI) may be in error by several decibels.

3) The antenna on which the signal is received may have differing path losses. Since most upstream traffic from connected STAs consist of control frames (ACK or block ACK), which are typically transmitted at basic rates utilizing signal replication across multiple antennas (when available), the received signal strength may vary greatly coming from the same connected device at the same location between transmissions.

4) It is an indication of the signal **from** the connected device, not an indication of the signal **to** the connected device **from** the AP. Thus, this method, although viable, doesn't give a true indication of the coverage quality **at** the receiving device **from** the AP.

5) This method also cannot account for differing channel noise characteristics at the STA location. Think of a baby monitor in a bedroom which doesn't affect the AP as much as the STA in the room with the monitor.

For these reasons, utilizing the upstream signal strength and noise floor values is not truly a viable measure of the coverage of a network, only of the power of the STA device.

### 2.4. Downstream Data Link Rate

Downstream link rate is one additional method to assess the actual downstream link rate to individual devices in a network. Individual devices are tracked via downstream link rate (typically in kbps or Mbps) and assessed for how well Wi-Fi is covering the devices in the network. Several problems are apparent in this method:

 

1) The STA MIMO capabilities need to be tracked on a per-connected-device basis and the actual transmission data rate compared with the device's MIMO capabilities, channel bonding utilized for the transmission, and transmission characteristics (preamble type, modulation type, etc.) in order to compare the capabilities of the client device versus what is actually being transmitted to the device.
2) It requires data to be passed and averaged on a consistent basis.
3) It is impossible to know if data has been passed to the device during the measurement interval, or if this was a 'latched' value from a previous transmission.
4) The score must be averaged and stored based on varying transmission rates and possibly changing transmission type (short vs. long preambles, channel widths, MIMO vs STBC mode, PER, etc.).

## 3. Proposed Solution "CQ Score"

The solution being proposed in this paper is a score defined to be representative of the experience of Wi-Fi STAs within a specific network. This solution is referred to as the Coverage Quotient Score. The CQ Score is an adaptive measurement that considers the dynamic behavior of Wi-Fi clients and assesses collective performance across all access points for both the clients themselves as well as the network collectively. The characteristics of the CQ scoring methodologies are as follows:

1) CQ Score methodology assesses the various devices that connect to the network APs on a regular basis.
2) CQ Score methodology learns about the mobility of a device by assessing the dynamic nature of its position, reflected by the collection of downstream MCS transmission rates. This is done by building a 'dynamic range map', predictive of the location profile for the network. The method grooms the data by predicting 'transient extremes' such as is seen when a mobile device is going out of range and subsequently losing association. For instance, someone leaving the home with their mobile phone will continue to have access for some distance until such a time when the Wi-Fi signal is too weak to be recognized. This data point must be ignored, since it is a transient that should not affect the overall measurement.
3) CQ Score methodology applies the concept of a 'data-usage mask' to add weight to the 'dynamic range map' produced above. This improvement allows a data point from a particular device at a particular low MCS rate location associated with low/no data usage (e.g. Dad is mowing the lawn with his device in his pocket), to be set lower if not ignored entirely. Whereas, if Dad is streaming his favorite football game to a video device on the patio at a relatively low MCS rate location, this measurement is given more weight.
4) CQ Score methodology learns about the temporal behavior of devices, in terms of a 24-hour period, to detect patterns associated with specific times of the day. Two sets of such data are relevant, one for weekdays and another for weekends. This data may be retained in a cloud-based service to attain a longer time view.
5) CQ Score is simple to understand and should be amenable for an easy assessment of both individual devices in the network as well as for the total coverage of the network.

### 3.1. Solution Details

This section defines the details of the CQ scoring algorithm including the sample data and calculations involved in deriving the score.

 

### 3.1.1. CQ Score Range Definition

The CQ Score must be intuitive and easily deciphered. The chosen values are 0-9 (although this could be any bounded range). The range of the score must be granular enough to give adequate definition of the coverage, yet not be onerous to decipher. The higher the score, the better the coverage from a Wi-Fi signaling perspective. The values 0-9 were not chosen randomly but correspond to MCS rates as defined in [1] and expanded for legacy rates. Rounding may be employed, or a more accurate representation may be obtained if desired utilizing less stringent rounding. The range may be easily adjusted.

### 3.1.2. CQ Score Data Sampling Frequency

The data sample rate for the CQ Score solution should be configurable and adaptable. However, the sampling frequency should not over-burden the system nor should it cost resolution in the scoring algorithm. In a typical network environment, a sampling frequency of once every second should suffice. This is based on the fact that the average walking speed is 1.4 meters per second. This will average an open-air path loss (or gain) in the sampling interval of approximately 1.4 dB. Moving to less frequent sampling could allow for a higher than desired change in signal strength. For example, at 2 seconds the loss or gain would be 3.9 dB, which is a drastic change in power.

The open-air path loss may be calculated utilizing the formula:

$$\text{Path Loss} = 32.45 + (20 * \text{LOG}_{10}(\text{FREQ}_{MHZ})) + (20 * \text{LOG}_{10}(\text{DISTANCE}_{KM}))$$

Of course, until the next sampling interval there is no way to determine if additional parameters have come into play, such as walking through a doorway which increases attenuation due to additional intervening walls, floors, etc. However, once a second has been determined to be an optimal sampling rate in testing for an active network.

The sampling rate should also be dynamically adjustable within the algorithm. If a network is lightly loaded, exhibiting low data usage, then the sampling rate may be decreased. This is viewed as an optimization to unburden the system with sampled data which statistically remain stable for the network. For example, little or no data is being used in the network during those parts of an average day when people are at work or at school.

### 3.1.3. CQ Score Sample Part 1 – MCS or Data Rate

As inferred above, one of the metrics sampled in order to derive the CQ Score is the downstream MCS rate of data frames going from the AP to the connected device. This metric was specifically selected for various reasons listed below:

1) It is a relative measure of the signal-to-noise ratio of the transmitted signal at the receiving device.
2) It is a true indication of the received signal quality at the device. It accounts for the signal strength, noise floor, and PER of transmissions from an AP to the client device.
3) It is easily obtainable on the AP, and thus does not rely on client capabilities for reporting.
4) It is independent of the number of spatial streams being supported by any given client device.
5) APs have rate shifting algorithms which try to maximize the MCS rate at any given point in time.
6) APs have rate shifting algorithms which will adapt to channel conditions for mobile devices corresponding to greater range, or inhibited channel conditions.
7) An inferred MCS value can be assigned to legacy devices which relieves the stress of understanding client capabilities and translates directly into the scoring method.

8) It produces no additional overhead in the Wi-Fi subsystem for request/report pairs over the air.

Please refer to the tables below for MCS values for various protocols, and the values which should be reported as part of the CQ scoring method.

#### *3.1.3.1. Non-MIMO Rates to MCS Translations*

Since most of the legacy rates (802.11a, 802.11b, 802.11g) may not report actual modulation types or coding schemes, an inference from the data transmission rate to an associated client device is acceptable per the following tables. The values were selected based on the relative 'goodness' of the signal received at the associated client devices.

**Table 1 -- 802.11b Tx Rate to MCS Values[1]**

| 802.11b Rate (Mbps) | Modulation Type | Data Bits Per Symbol | MCS Value |
|---|---|---|---|
| **1** | DBPSK | 1 | 1 |
| **2** | DQPSK | 2 | 3 |
| **5.5** | CCK | 4 | 5 |
| **11** | CCK | 8 | 7 |

**Table 2 -- 802.11 a/g Tx Rate to MCS Values[1]**

| 802.11 a/g Rate (Mbps) | Modulation Type | Coding Rate | MCS Value |
|---|---|---|---|
| **6** | BPSK | 1/2 | 0 |
| **9** | BPSK | 3/4 | 1 |
| **12** | QPSK | 1/2 | 2 |
| **18** | QPSK | 3/4 | 3 |
| **24** | 16-QAM | 1/2 | 4 |
| **36** | 16-QAM | 3/4 | 5 |
| **48** | 64-QAM | 1/2 | 6 |
| **54** | 64-QAM | 3/4 | 7 |

#### *3.1.3.1. MIMO Rates to MCS Translations*

The following table (Table 3) relates the MIMO modulation type to MCS reported values. Note that CQ scoring methods ignore the spatial stream (SS) component, focusing instead on the transmitted modulation type and coding rate for its scoring method. Admittedly, this ignores APs which may utilize STBC methods to replicate a signal across multiple antennas for Tx diversity gain, or those which may transmit on narrower channels in order to gain more spectral power. However, the score will still be valid for coverage, even if the transmission methods are not the most efficient ones. The key is to determine the coverage for the device as evidenced by combination of the modulation type and coding rate.

 

Table 3 - MIMO MCS Values[1]

| HT MCS | VHT MCS | HE MCS | Modulation Type | Coding Rate | Reported MCS |
|---|---|---|---|---|---|
| **0,8,16,24** | **0** | **0** | **BPSK** | **1/2** | **0** |
| **1,9,17,25** | **1** | **1** | **QPSK** | **1/2** | **1** |
| **2,10,18,26** | **2** | **2** | **QPSK** | **3/4** | **2** |
| **3,11,19,27** | **3** | **3** | **16-QAM** | **1/2** | **3** |
| **4,12,20,28** | **4** | **4** | **16-QAM** | **3/4** | **4** |
| **5,13,21,29** | **5** | **5** | **64-QAM** | **2/3** | **5** |
| **6,14,22,30** | 6 | 6 | 64-QAM | 3/4 | 6 |
| **7,15,23,31** | 7 | 7 | 64-QAM | 5/6 | 7 |
| **N/A** | 8 | 8 | 256-QAM | 3/4 | 8 |
| **N/A** | 9 | 9 | 256-QAM | 5/6 | 9 |
| **N/A** | N/A | 10 | 1024-QAM | 3/4 | 9* |
| **N/A** | N/A | 11 | 1024-QAM | 5/6 | 9* |

*Note that for HE, it was decided that since the actual range supported for 1024-QAM is so short, the reported MCS value shall be at the maximum of scoring range.

### 3.1.4. CQ Score Sample Part 2 – Amount of Data Transmitted

The second part of obtaining the CQ Score for a connected device or network is determining a weighting value for the score based on the amount of data transmitted to the device during the scoring interval. The weight accounts for the individual device's data usage on the network and gives the scoring algorithm the ability to 1) assess overall network usage, and 2) more heavily weight client devices which are more active during a given sample interval.

The sample is a simple byte count of the number of data bytes transferred to the connected device during the sample interval. This value along with the sampled or computed MCS value allows for a valid sample to be inserted into the calculations described in subsequent sections.

### 3.1.5. CQ Score Sample Weighting

The CQ Score value will be weighted based on the amount of data being passed to a particular client. This gives "weight" to the individual client's coverage within the realm of the entire network. The weight should be bounded with upper and lower limits. The upper and lower limits should be configurable within the sampling system.

- If the amount of data transmitted to a client is less than the lower limit, then no sample is recorded. This takes into account
    - Disassociated clients
    - Clients which are statistically not utilizing data on the system even if still associated. This could be 100 bytes per second or a value as deemed statistically valid.
    - This value should be non-zero.
    - The lower limit should account for voice traffic, which is low in bandwidth, but high in priority.

 

- If the amount of data transmitted to a client is greater than the upper limit, then the data is capped at the upper limit. This ensures that clients utilizing huge amounts of data don't deplete system resources in the scoring algorithm (i.e., don't blow size constraints for the variables).
  - Clients utilizing huge amounts of bandwidth (> 100 Mbps for example) are capped at 100 Mbps.

The system also utilizes a configurable ***ThroughputWeightScaler***, a simple divisor in order to scale a sample to a reasonable level for storage and later averaging. The typical value for this would be 1000. An example given: this moves from Mbps to kbps in samples stored.

The pseudo code for calculating the weighted CQ Score for an individual sample for a device is shown below

**Bits = (ByteCount * 8) / SampleInterval // Gives a bits per second for the sample period**
**If Bits > SampleUpperLimit then Bits = SampleUpperLimit**
**If Bits <= SampleLowerLimit then Bits = 0**

**SampleScore = Bits / ThroughputWeightScaler**

Some examples for three connected devices are given below.

| STA | REPORTED BYTES | MBPS | SAMPLE SCORE |
|---|---|---|---|
| **1** | 1,250,000 | 10 | 1000 |
| **2** | 625,000 | 5 | 500 |
| **3** | 25,000 | 0.2 | 20 |

Determining the QoS markings of the data being passed to a device could pose an optimization for the weighting value. This optimization allows for voice traffic, which is of lower bandwidth but high priority, to be weighted more heavily within the sampling subsystem. Utilizing this optimization requires the CQ System to become QoS aware, utilizing either deep packet inspection to determine DSCP or CoS frame markings or inspection of Wi-Fi WMM packet buffers which are associated with certain connected devices.

### 3.1.6. CQ Score Data Bins

Once a CQ Score has been calculated and weighted, the next step in the process is to 'bin' the data based on the MCS rate at which the data was transmitted. The data bins relate to MCS values 0 to 9 and correspond to the CQ Sample Part 1. Placing the individual and weighted scores into bins allows for an average CQ Score value to be determined across a "Scoring Interval." The bins are zeroed at the beginning of the score intervals and individual scores are added as they are taken at the sample intervals.

A score table is maintained for each device active and connected to the network. It is outside the scope of this document to delineate determining the status of connected and active devices.

At each sample interval the score determined for each device is added to its respective bin.

**StaBinTable[STA][SampleMCSRate].BinCount += SampleScore**

 

### 3.1.7. CQ Score Interval

A CQ scoring interval is defined corresponding to a time interval at which the CQ Score calculations occur for both individual client devices and the overall network. The scoring interval should be configurable for the system. A typical value would be 60 seconds, or once a minute. At the scoring interval calculations are made to score individual client devices as well as the overall network.

Note, in a mesh deployment scenario mesh devices may be determined by tracking via a network topology table. The connected mesh devices are also STA devices to their root nodes.

At the score interval, summations are made on a per-STA-basis, and a final score value is derived based on the data which is stored in the individual score bins.

### 3.1.8. STA CQ Score Calculation

The actual CQ Score is calculated per the following sequence of calculations:

1) Total the complete score of all MCS bins for the scoring interval.

$$STAbTotal = \sum_{b=0}^{9} STABinTable[STA][b].BinCount$$

2) Calculate per-bin-percentage values for the interval per the following:

***For (m = 0; m <= 9; m++)  {***

***Scorebin.{m} = STABinTable[STA][m].BinCount / STAbTotal * CQScoreScaler.***

***}***

Where:

***m*** = the MCS value (0-9)

***CQScoreScaler*** = A value utilized to scale calculated CQ Scores to within a range of values. This should be a factor of 10. The larger this value, the larger will be the range of the score. Default value is set to 10, and thus the values will be limited to 0 – 10.

***STAbTotal*** is the sum of all bin counts as calculated above.

This calculation gives a scaled value of the percentage of bits passed at individual MCS rates versus the total bits passed during the interval. The CQScoreScaler allows for the scores to be limited in value (0-10 is default).

3) Calculate the final CQ Score.

$$STACQScore = (\sum_{m=0}^{9} Scorebin.\{m\} * BinScaler * (m+1)) - BinScaler$$

Where:

***m*** = the MCS value (0-9)

 

***BinScaler*** = CQScaler / 10

Note: The multiplier of the bin scaler is increased by 1 to alleviate a zero multiplier. The final subtraction removes this offset.

#### *3.1.8.1. STA Score Example Data*

The following table and graph provides an example of the output for the CQ Score calculation.

Note: Data may be easily displayed on a per STA basis to track coverage for individual STA devices.

**Table 4 – Single STA Sample Data**

| Bin | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| Bin Value | 0 | 150 | 350 | 600 | 800 | 1300 | 2700 | 1200 | 400 | 0 |
| Bin Score | 0 | 0.2 | 0.47 | 0.8 | 1.07 | 1.73 | 3.6 | 1.6 | 0.53 | 0 |
| STA CQ Score | | 5.35 | | | | | | | | |

The CQ bin scores can also be easily graphed for visual display

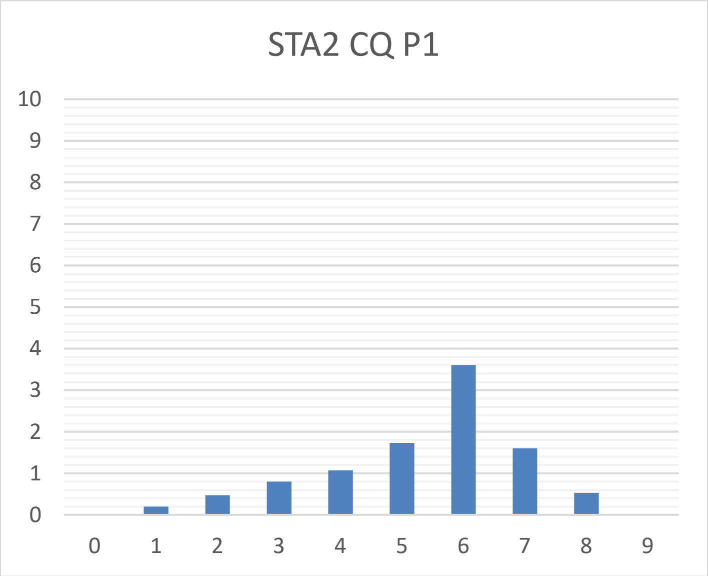


**Figure 1 -- STA CQ Score**

### *3.1.9. Network CQ Score Calculation*

The CQ Score for the entire network is easily obtained from the individual STA CQ Scores, or by following the same series of calculations shown above by summing the individual score bins from all STAs (preferred method). The data can then either be archived in a time-stamped database or passed to a cloud-based system for subsequent processing.

An example exercise for a set of three STAs is given below.

 

**Table 5 -- Network CQ Score**

| | Bin 0 | Bin 1 | Bin 2 | Bin 3 | Bin 4 | Bin 4 | Bin 6 | Bin 7 | Bin 8 | Bin 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| STA1 | 0.00 | 0.00 | 0.00 | 0.00 | 800.00 | 10,900.00 | 23,530.00 | 25,040.00 | 10,060.00 | 20.00 |
| STA2 | 0.00 | 150.00 | 350.00 | 600.00 | 800.00 | 1,300.00 | 2,700.00 | 1,200.00 | 400.00 | 0.00 |
| STA 3 | 0.08 | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 200.32 | 404.64 | 5.32 | 3.00 |
| Total | 0.08 | 150.08 | 350.00 | 600.00 | 1,600.00 | 12,200.00 | 26,430.32 | 26,644.64 | 10,465.32 | 23.00 |
| Network Score | 0 | 0.02 | 0.04 | 0.08 | 0.2 | 1.55 | 3.37 | 3.4 | 1.33 | 0 |

| | |
|---|---|
| Network CQ Score | 6.35 |

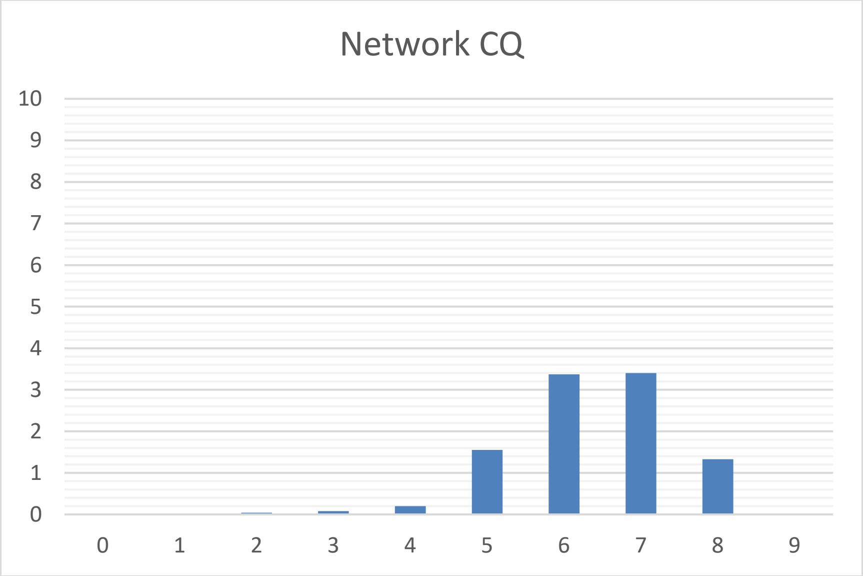


**Figure 2 -- Network CQ Score**

The example demonstrates that STA1 has much more weight on the entire network since it was passing much more data. However, each STA contributed to the overall network score.

Individual STA scores and bTotals are shown below.

| STA | bTotal | STA CQ Score |
|---|---|---|
| **1** | 70350 | 6.46 |

 

| STA | bTotal | STA CQ Score |
|---|---|---|
| **2** | 7500 | 5.35 |
| **3** | 613 | 6.7 |

### 3.2. Advanced Analytics

It is beyond the scope of this paper to detail possible uses of the CQ Score data. However, described below are a few ideas which are immediately apparent.

Advanced analysis of the data can be performed by archiving CQ Score interval data and performing subsequent analysis of the time-stamped data. This can provide hourly scores, daily scores and even allow for flagging low STA CQ scores across timed intervals.

Analytics can be implemented comparing a low CQ score to a bTotal to determine if a STA is demanding large amounts of data at a relatively low MCS rate on a consistent basis. The results of this analysis could lead to a service provider advising a customer that (a) range extender(s) may be necessary in either their home or business network. The same analytic may be used to advise a customer that a Wi-Fi extender has been placed at a bad location from its root device.

Further analytics may be implemented which show the effectiveness of steering mechanisms for STA devices.

# Conclusions

Current metrics utilized to determine the coverage of a Wi-Fi network are lacking in many respects. Many of them don't give the best indication of AP to client coverage. The methods which currently exist can be prone to 1) no STA support for the existing 802.11 methods, 2) difficulty in management based on the various data rates which may be supported on a per transmission basis or 3) the measurement of inadequate metrics to characterize downstream traffic. By utilizing the MCS rate and amount of data being transmitted to an associated STA device, a weighted coverage quotient may be obtained. It is an excellent indication of the network coverage without requiring additional network traffic. The CQ Score can be utilized in many ways to provide the service provider with a detailed view into the network and their customers with guidance in optimizing their network.

 

# Abbreviations

| | |
|---|---|
| AP | access point |
| BSS | basic service set |
| BSSID | basic service set identifier |
| CoS | class of service |
| CQ | coverage quotient |
| dB | decibel |
| DSCP | differentiated services code point |
| Gbps | gigabits per second |
| GHz | gigahertz |
| GW | gateway |
| HE | high efficiency |
| HT | high throughput |
| ISBE | International Society of Broadband Experts |
| ISM | industrial, scientific and medical |
| MAC | medium (or media) access control |
| kbps | kilobits per second |
| Mbps | megabits per second |
| MCS | modulation and coding scheme |
| MIMO | Multiple input and multiple output |
| MU | multiple user |
| mW | milliwatt |
| OFDMA | orthogonal frequency division multiple access |
| RCPI | receive channel power indicator |
| PER | packet error ratio |
| RF | radio frequency |
| RSNI | receive signal-to-noise indicator |
| RSSI | receive signal strength indicator |
| SCTE | Society of Cable Telecommunications Engineers |
| SNR | signal-to-noise ratio |
| SS | spatial stream |
| STA | station [device] |
| TCP | transmission control protocol |
| STBC | space-time block coding |
| Tx | transmit (or transmission) |
| UDP | user datagram protocol |
| U-NII | unlicensed national information infrastructure |
| VHT | very high throughput |
| VI | video |
| VO | voice |
| W | watt |
| WAN | wide area network |
| WFA | Wi-Fi Alliance® |

# Bibliography & References

[1]IEEE 802.11TM-2016 -- *IEEE Standard for Information technology—Telecommunications and information exchange between systems Local and metropolitan area networks— Specific Requirements*; *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications* LAN/MAN Standards Committee of the IEEE Computer Society Copyright 2016.

[2]Wi-FiTM Alliance Multi-AP Specification. © 2018 Wi-Fi Alliance. All Rights Reserved

[3]Wi-Fi MultimediaTM Technical Specification (with WMM-Power Save and WMM Admission Control) © 2012 Wi-Fi Alliance. All Rights Reserved

[4]Wi-FiTM Alliance Data Elements Specification © 2018 Wi-Fi Alliance. All Rights Reserved

[5]IEEE Std 802.1Q™-2011, IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks.

[6]IEEE 802.11ax[TM] *Draft Standard for Information technology— Telecommunications and information exchange between systems Local and metropolitan area networks—Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Enhancements for High Efficiency WLAN*
LAN/MAN Standards Committee of the IEEE Computer Society Copyright 2019

 

# Operationalizing Disaggregate Networks

A Technical Paper prepared for SCTE•ISBE by

Bill Beesley, MSO Strategy and Planning, Fujitsu, SCTE•ISBE Member
2801 Telecom Parkway
Richardson, TX 75082
bill.beesley@us.fujitsu.com
972-479-2098

Francois Moore, Principal Optical Specialist 1FINITY, Fujitsu
2801 Telecom Parkway
Richardson, TX 75082
francois.moore@us.fujitsu.com
972-479-3563

 

# Table of Contents



 

## 1. Introduction

***"Humans are allergic to change. They love to say, "We've always done it this way." I try to fight that. That's why I have a clock on my wall that runs counter-clockwise." Grace Hopper***

One truism in technology is that the only thing that is constant is change. As the telecom industry continues to push more packets with lower latency to their customers all while trying to lower the capital and operating costs to do so, the technologies we utilize have to evolve. It has been my experience that most telecom engineers are supportive of technology changes. Most of them like to get better, faster, or just newer tools to add to their network. Conversely, the operations teams loath change. And new equipment means just that, it means changes to the way they have to deploy and manage their networks. This conflict between the goals of the engineers that select network equipment and the operations teams that have to live with those decisions can delay the adoption of new technologies and in some cases, void the targeted financial benefits. Thus, it is critical to consider the operational aspects of new technologies up front. If you want to achieve the monetary benefits that come with a new technology, whether this is the ability to deliver new services or to deliver existing services more efficiently, you have to consider the operations requirements up front.

One example of a new technology that is driving the need for changes in operations behavior is network disaggregation. Disaggregation is not a new thing. Data center operations have long been separating their control plane from the hardware as well as mixing and matching vendor products. Disaggregation in transport networks, however, has only recently started to gain widespread adoption by service providers. Disaggregating transport equipment brings many benefits to the telecom industry. Telecom service provides gain advantages from increased design flexibility and reduced vendor lock-in. Equipment manufactures gain efficiencies that allow them to bring new technologies to market faster and more economically.

## 2. What is Disaggregation?

For those who are not familiar with network equipment disaggregation, what we are talking about is the breaking apart what historically had been a large, multiservice, chassis-based optical transport system into its functional components. The driver behind this breaking up the chassis is that there are several inefficiencies inherent in the design that constrain innovation and drive costs.

When an equipment manufacturer designs a chassis for a system, they have constrained that system for its entire life. If a newer technology comes along that exceeds the thermal, power or space designs of that chassis, the only option the manufacturer has is to design and build a new chassis. This increases costs that are ultimately passed on to the operator. Even when a new technology can be introduced into an existing chassis, that new design element must be validated for interoperability with every other existing card for that chassis. This also increases development costs and slows the pace at which new technologies can be introduced to market.

## 3. Disaggregation Benefits

There are five primary benefits for the cable operator that are driving the move to disaggregated systems. The first is that disaggregated systems scale more efficiently than monolithic chassis. When deploying chassis, there is a step function for the first deployment and each time you max out the chassis. With disaggregated systems, you can have a low upfront investment because you eliminate the need to purchase a large shelf. Since the line units are also independent of each other, newer technologies can be added to the network without having to replace a legacy shelf system, which means faster adoption and deployments.

The second advantage of disaggregated systems is that, unlike line cards, the blades are not constrained by the shelf. This means that new technologies are not held back if the chassis cannot support a new physical form factor, power, backplane or software control that may be needed in the future. New technologies can be added as they become available without effect on existing deployments.

The third advantage of disaggregated systems is that they more efficiently utilize rack space. Network functions such as routing, switching, transponders and reconfigurable optical add/drop multiplexers (ROADMs) grow at different rates over time. This means a monolithic chassis supporting multiple functions may or may not be optimally populated. This historically has resulted in orphaned slot space that can add up quite quickly. An audit of one large US service provider showed that nearly 60% of their valuable headend space was unutilized. Since disaggregated systems can be deployed in 1 RU increments as services are added, they can be optimally populated to maximize rack space in the headend.

The fourth advantage of disaggregated systems is they allow the operator to build their network on a pay-as-you-grow model. One truism is that network planners rarely get it right long-term. For example, there might be a rural site that is primarily an express path for optical transport between cities that might also be serving a handful of nearby cell towers. The needs of this site on day one are fully served with a simple two-degree ROADM with a few drops. However, what if that small rural location five years from now suddenly blows up with new residential and business development because some large company moved their headquarters and now, we need a few more rings and a lot more drops. Traditionally, if we got it wrong like this it would be resolved by cutting in a whole new system. Since we tend to want to de-risk such costly customer affecting events if we thought there was even a remote chance of this happening, we placed a large system in that location, and it sat mostly idle until it might be needed sometime in the future. With a disaggregated system, it is easy to start small and meet the current needs but still have the flexibility to add additional capacity should the unexpected ever happen in the future. It is a true pay-as-you-grow model because your upfront investment is only what is needed to generate the short-term revenue.

The fifth and final advantage is one that, while quite common in the data center world, has only started popping up on the cable industry's networking radar. This is that disaggregated systems can be open and eliminate vendor lock-in. There really is no good technical reason why systems from multiple vendors cannot be used in our networks. Traditionally, we have sourced large chassis-based systems from single vendors due to the "one throat to choke" mantra. That is to say, "I know the system works because my vendor certified it to do so." In the case of optical transport systems this used to be important because vendor interoperability at the ROADM or line system layer was pretty much non-existent. Packets from vendor A's router to vendor B's router could be passed with a high degree of confidence, but when it came to using one vendor's transponder over another vendor's ROADM, it was generally impractical. Recent industry work to develop standards for Open ROADM and open line systems now make it possible to have true vendor interoperability at the optical layer. Being able to be vendor agnostic in this way allows operators to avoid vendor lock in and to select and deploy best of breed technologies as they come to market.

 

## 4. Operations Considerations

Of course, while the engineers among us are going "Great! Sign me up!", this usually provides the point when we have lost the operations people in the room. By now they are sitting in a corner with their arms crossed muttering things like "Nope! nope!......how am I going to power a bunch of single blades? How will I manage spares? How can I find the right blade in the headend if it is not tied to a chassis? It is too difficult to manage a bunch of independent blades instead of just a single shelf!". The list generally just keeps growing from there.

All of our efforts to build great technological solutions are null and void if we do not seriously consider how it will be operated. So, let us look at that next.

## 5. Powering

Let us first consider the powering of the equipment, because this is usually the biggest concern raised about disaggregation. Most telecommunications facilities utilize a distributed -48V DC power plant. When a new piece of equipment is added to the network a line is run back to what is known as a battery distribution fuse bay acting much like the circuit breaker panel in your home. The work to run this power is done by an electrician, not a network engineer. This means that each time an element is added, there must be coordination between the network engineer and the electrician and usually a charge for the connection. Operations teams are used to powering a single chassis one time and then they can add blades at their leisure. With disaggregated systems we need to rethink how we manage adding power to minimize cost and deployment delays due to the cumulative demands on the electrician's schedule. There are ways to mitigate this and make disaggregation manageable. Some manufacturers have created fixed housings that are simple power aggregation devices into which the 1 RU blades are inserted. The housing can be mounted and powered by the electrician and then technicians can add blades at any time. It should be noted that while this avoids the need to run power for each individual blade, it does eliminate the benefit of space efficiency because the housing is taking up a fixed amount of space regardless of whether it is fully populated. It should also be noted that for some operators, this is seen as a benefit as they use this to claim space in the headend where multiple groups might be placing their equipment. Another approach to the powering dilemma is for the electrician to install connectorized power drops that can be plugged into blades as they are added. Since the connectors eliminate the risk of shorting out the power drop, they can be safely left in place allowing the technician to snap the connector into the blades as they are installed.

## 6. System Management

Another consideration for the operations team is blade management. The concern is that instead of managing a single chassis, operators will now need to manage multiple single blade entities. This is where disaggregation provides a simpler and more flexible operations model, because the operator is no longer locked into the chassis as the only functional management point. Modern network controller software allows operators to take multiple elements and logically integrate them into a single management point. This means that a ROADM and any number of transponders can be logically aggregated together as a single management point. This can be done with one blade, or potentially hundreds without any constraints of the form factor of a chassis. We can also take advantage of disaggregation to add a controller blade to manage multiple blades. This controller behaves in a way a master control card would in a traditional chassis, serving as the single management point, providing database backup and control for the individual blades and managing any provisioning and software upgrade tasks for the system.

 

However, disaggregation allows the operator to decide when they want to pay for and deploy the master controller. It is not required, so where installations are small and a master controller might not be necessary, the operator is not required to pay for one.

## 7. Upgrades

In-service software upgrade is another large recurring activity in network operations. Vendors routinely publish maintenance releases either to provide new major features, incremental enhancements or problem resolution. The disaggregated blade approach provides better operational support by reducing the number of sites requiring an upgrade as well as reducing the amount of time to perform an upgrade.

The first saving lies in the fact that traditional network elements typically include ROADM, transponder and switching functionality. Let us assume a new release provides improvements for the transponder functionality but does not change any of the other network functionality. In order for an operator to maintain a homogeneous network, they are required to upgrade every network node even if these nodes might not be affected by the transponder changes.

Alternatively, a disaggregated blade allow the operator to target the upgrade to only those sites truly affected by the change. In this example, any site where this particular transponder blade is not used does not require an upgrade. This reduction in upgrade requirement can generate a large operational saving for the carrier.

More savings lie in the actual time required to upgrade a site. Traditional network elements have, by nature, large monolithic software loads compared to a specialized blade, which has a much smaller and more specialized software load.

Our network operations center has measured the software download time difference between a traditional network element and a specialized blade. Traditional network elements typically require 25 to 30 minutes for a software download compared to four to 10 minutes for a specialized blade. This represents a reduction by roughly 60% in the time required to upgrade a site, which is both a labor-hour saving as well as a management network infrastructure saving.

Let us translate this advantage into a measurable savings using the following scenario:

- Number of sites: 100
- Percentage of site affected: 60%
- Download time traditional NE: 25 minutes
- Download time specialized blade: 8 minutes
- Number of specialized blade/site: 2 blades

A traditional network element would require upgrading every site, requiring a total of 42 hours of labor. A specialized blade would require 16 minutes per site, based on two blades per site, but only 60 sites would require an upgrade. The total hours for this architecture would be 16 hours, a reduction of 62% compared to a traditional architecture.

 

During our discussions with operations executives, we learned that network elements do not typically age gracefully. This is resulting from the fact that the chassis, its backplane, the processor complex as well as the fans are the first items designed and developed. Once implemented, these items remain fixed in time and all further interface units and switch complexes are designed around these components, which are increasingly out of date and unable to support new technology requirements. However, they persist within the network, because the cost of taking them out and replacing a system carrying customer traffic is very high. As a result, it is common to find systems built around common equipment hardware, which is more than 10 years old and effectively obsolete.

This time anchor has a real consequence for a carrier's operation: retrofit expenses. Over time, these components become increasingly outdated and the availability of new services becomes gated by either backplane speed, power or heat dissipation imposed by the aging common equipment.

At that point, a carrier is facing two unappealing options: cap the existing network and start over or retrofit some of the components such as the CPU or the fans.

Every generation of network elements from virtually every equipment provider has faced this: a CPU retrofit. Every network element will eventually exhaust CPU processing power and/or memory necessitating an upgrade to support continued use. This, on average, typically occurs after six or seven releases.

A CPU retrofit can be extremely expensive to the operator. This requires a physical visit to a site in order to replace two processing card units within a network element. Each visit will typically require the following effort:

- Drive to site from a local support center: 1 hour
- Actual maintenance operation at the site: 2 hours (include remaining on site to monitor recovery)
- Return to support center: 1 hour
- Hardware cost: 2 new CPU units

If we use the example from the previous section, this would mean that the 100 sites would require 400 labor hours as well as 200 new CPU processing card units. Now, technicians might be able to visit multiple sites on a single trip; let us assume an efficiency factor of 30% related to the transportation time. This would lower the total labor-hour to 340 hours.

The actual cost associated to this hardware and labor can vary based on the location and the type of product. Assuming a $150/hour as a loaded labor rate as well as a $2,500 cost for a processor unit, the total cost related to this effort would total around half a million dollars.

A specialized blade completely eliminates these concerns and ensures that these types of expenses will be nonexistent. This is because the specialized blades are independent of each other and each of them is built based on an architecture designed to meet their specific requirements.

## 8. Conclusions

In summary, it is clear that deploying disaggregated networks requires consideration of operational requirements. Although this change is normally received with large concerns from network engineering and operations personnel as demonstrated above, the benefits of disaggregation cannot only be supported but provide significant savings to the operations teams. Disaggregated networks are the future of networking. They will provide service providers with faster availability of new technologies, as well as lower capital and operational costs. These benefits are achievable with minimal impact to operations.

## 9. Abbreviations

| | |
|---|---|
| CPU | central processing unit |
| DC | direct current |
| NE | network element |
| ROADM | reconfigurable optical add/drop multiplexer |
| RU | rack unit |

 

SCTE · ISBE
STANDARDS

NETWORK
OPERATIONS

scte.org • isbe.org