Practical Deployment Lessons of a Centralized Virtualized CMTS

A Technical Paper prepared for SCTE•ISBE by

Asaf Matatyaou
Vice President, Solutions and Product Management, Cable Access Business
Harmonic, Inc.
4300 North First Street, San Jose, CA 95134
1-408-490-6834
asaf.matatyaou@harmonicinc.com

Richard J. Walker
Vice President, Engineering & Info Technology & Services
Shared Services
2700 Oregon Road Northwood, OH 43619
1-419-724-3735
rjwalker@sharedsvcs.com

© 2018 SCTE•ISBE and NCTA. All rights reserved.

Table of Contents

Title **Page Number**

Table of Contents 2
Introduction 3
Architecture Description 3
Deployment Details 5
Real-World Considerations 6
1. Compute and Network Resource Location 7
2. Networking 8
   2.1. Layer 2 or Layer 3 Converged Interconnect Network 8
   2.2. Traffic Prioritization and Capacity Management 9
   2.3. Redundancy 9
3. Timing 9
   3.1. Remote DOCSIS Timing Interface 9
   3.2. DOCSIS Latency 10
4. Operations 11
Conclusion 11
Abbreviations 12
Bibliography & References 12

List of Figures

Title **Page Number**

Figure 1 - I-CCAP, PHY Shelf and Remote PHY Node Deployment Comparison 4
Figure 2 - Traditional I-CMTS Deployments Across Multiple Hub Locations 4
Figure 3 - vCMTS Deployed in a Centralized HFC Architecture 6
Figure 4 - Adding Remote PHY Nodes to a PHY Shelf Deployment 8

Introduction

The promise and potential of virtualizing a cable hub has been discussed over the past few years. There are many opportunities when discussing virtualization. One possible starting point is virtualizing a CMTS in a centralized deployment model in a headend or hub.

While the benefits and general considerations of a virtualized CMTS (vCMTS) have been described in past technical papers, this paper will focus on real-world experiences and lessons learned deploying a vCMTS in a centralized architecture. Areas of consideration when enabling virtualization via shifting from a hardware-based CMTS to a vCMTS will include the use of Remote PHY as a protocol/specification between a vCMTS Core and a PHY Shelf in a centralized facility (e.g. headend or hub), and usage of IEEE-1588. The ease of expandability will be highlighted, including node splits with PHY shelves, as well as Remote PHY Nodes (RPN). Lastly, the usage of streaming telemetry will be explored in comparison to legacy monitoring techniques.

Architecture Description

The architecture described in this paper is made possible by recent advancements in standards and technology, specifically the Remote PHY specifications published by CableLabs and the virtualization of the CMTS into commercial off-the-shelf (COTS) x86-based servers. CableLabs issued the first set of Remote PHY specifications in June 2015 and most recently updated the specifications in May 2018 with the tenth revision of the specifications.

"In a Remote PHY Architecture, the classic integrated CCAP (I-CCAP) is separated into two distinct components. The first component is the CCAP Core and the second component is the Remote PHY Device (RPD)."[1] The RPD consists of the physical layer functionality defined for an I-CCAP, with the remainder of the I-CCAP functionality residing in the CCAP Core. The CCAP Core is logically the combination of a CMTS Core and EQAM Core, and is connected to the RPD via IP transported over digital fiber. In this paper, the CCAP Core is implemented as a virtual CMTS (vCMTS) with a separate legacy EQAM pre-existing in the operator's network, and will be referred to as "vCMTS."

Since 2015, SCTE technical papers, such as "Transforming the HFC Access Network with a Software-Based CCAP" and "Real-World Deployment of a Virtual Cable Hub", have defined virtualization and the benefits of a vCMTS. The benefits and general considerations for vCMTS are beyond the scope of this paper.

"Remote PHY" is an implementation of a Distributed Access Architecture (DAA), but not necessarily restricted to DAA deployments. In fact, the term "remote" doesn't restrict the RPD from physically being co-located with the vCMTS, and the CableLabs specifications describe two examples where the RPD and RF may be located in the network, in the headend/hub or in an optical node.

Figure 1 shows an I-CCAP deployment architecture, as well as Centralized and DAA deployment architectures, both of which use the Remote PHY signaling to communicate between the vCMTS and the RPD (existing in the PHY Shelf and RPN). Remote PHY signaling includes the Downstream External PHY Interface (DEPI), Upstream External PHY Interface (UEPI) and the Generic Control Plane (GCP).

---

[1] Remote PHY Specifications, CM-SP-R-PHY-I10-180509, pg. 10

© 2018 SCTE•ISBE and NCTA. All rights reserved.

The benefits of Remote PHY and a detailed description of the specifications, including the signaling, are beyond the scope of this paper and are well documented in the industry over the past few years.

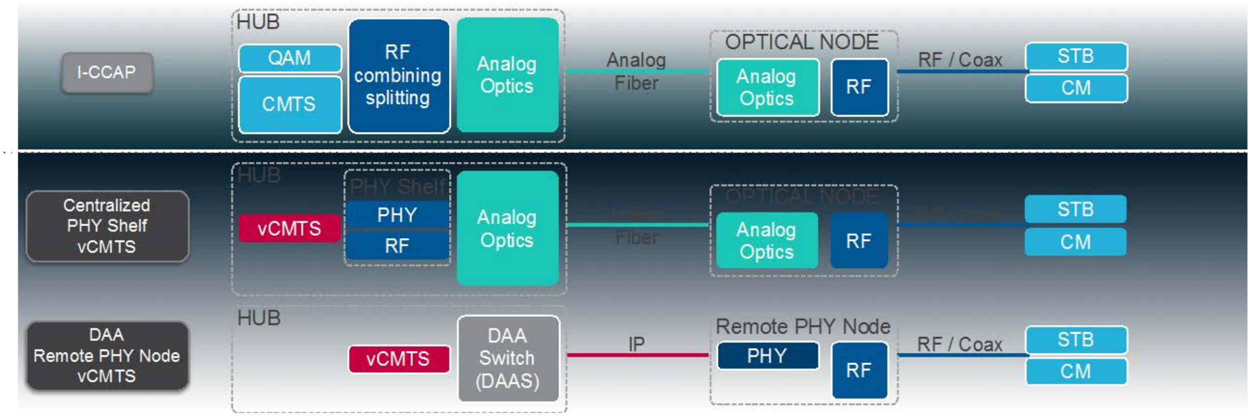

Figure 1 - I-CCAP, PHY Shelf and Remote PHY Node Deployment Comparison

This paper focuses on a Centralized deployment architecture, whereby another variation of the PHY is located in secondary hubs. The term "Centralized" is used to describe this deployment architecture as the existing HFC and analog optics in the field are leveraged. Figure 2 shows a traditional deployment with I-CMTS chassis deployed in each hub location within the operator's footprint.

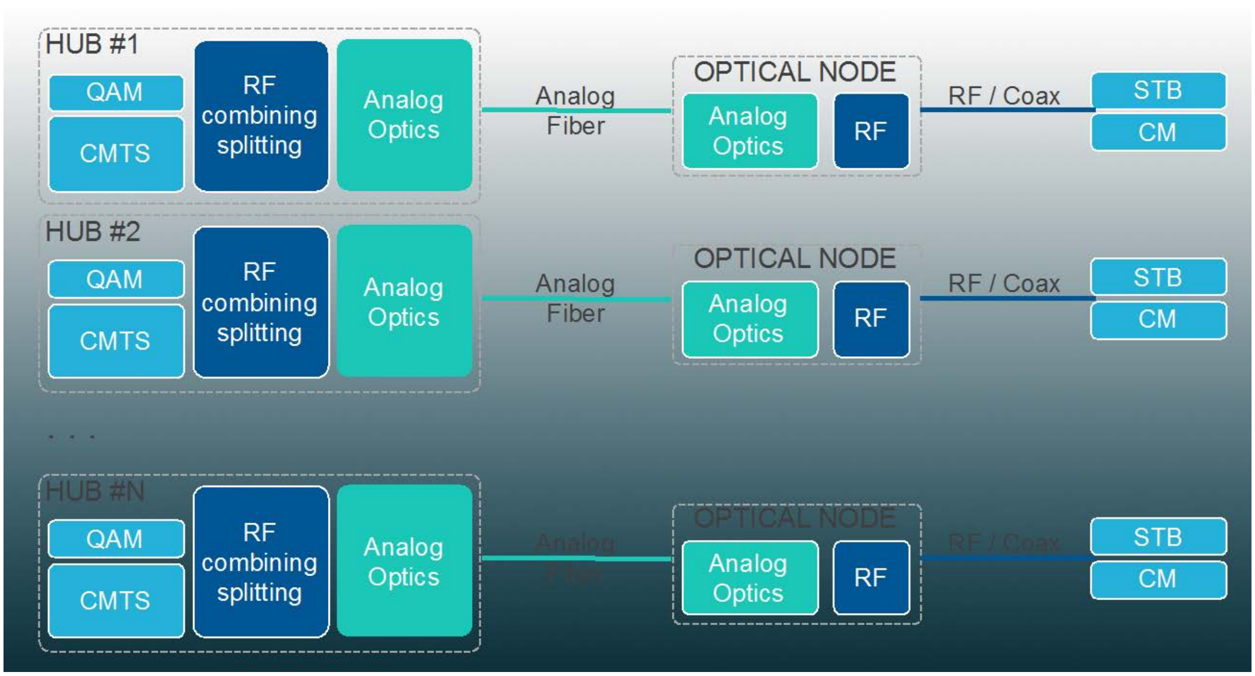

Figure 2 - Traditional I-CMTS Deployments Across Multiple Hub Locations

This traditional I-CMTS deployment was the starting point before the transformation to deploy vCMTS and PHY Shelves. The transformation to a vCMTS deployment was driven by the benefits of virtualization, such as[2]:

- Reduced space, power and cooling
- More frequent and shorter development cycles
- Sustainable capacity growth, elastic scalability and increase flexibility
- Improved Total Cost of Ownership, including reduced operational (OpEx) and capital expenditure (CapEx)

In particular, with vCMTS running on COTS x86-based 1-RU servers, the entirety of the CMTS Core functionality was consolidated to a single centralized location. The PHY Shelves with the RPD functionality were deployed in secondary hubs where the legacy I-CMTS chassis were previously deployed. These PHY Shelves connected to the existing HFC infrastructure. This transition delivered immediate value while leveraging all other legacy infrastructure, such as broadcast and VOD services being processed and delivered by existing EQAMs.

Further future value may be attained as this is a step towards deploying DAA, with initial small scale and eventual ramp up of RPNs connected to the same centralized vCMTS hub location over an IP network. More servers and capacity can be added, as needed, to scale at the chosen pace over time.

Deployment Details

The deployment details covered will describe the ending point after centralizing and consolidating the vCMTSes into a single hub, deploying PHY Shelves in secondary hubs and leveraging the existing HFC infrastructure.

Let’s describe each device in this deployment type and where the device is located.

1. vCMTS: the CMTS Core functionality is implemented on a set of COTS x86-based servers. All vCMTS Cores are located in a single, centralized hub and are connected to the Converged Interconnect Network (CIN).
2. Core Routing Engines (CRE): the switch fabric connecting the vCMTS Core servers to the rest of the CIN. All the CREs are located in a single, centralized hub and are connected to the CIN over a Layer 3 network.
3. Core routers: large-scale core routers connecting the access network with the core backbone network, located in a single, centralized hub.
4. Distributed Access Architecture Switches (DAAS): aggregation switches connecting the RPDs in the secondary hubs with the vCMTS Core servers (via the core routers), located in the secondary hubs.
5. PHY Shelves and RPDs: many RPDs share a single highly-available chassis in a PHY Shelf, which connect to the vCMTS Core servers over the CIN, and output RF over the existing HFC infrastructure. The PHY shelves are located in different secondary hubs.
6. IEEE-1588 PTP Grandmasters: “Remote DTI provides timing synchronization between CCAP Cores and RPDs based on the IEEE 1588v2 standard. The protocol supports the basic synchronization between the CCAP Core and Remote PHY Device for DOCSIS/video/OOB services.”[3]

---

[2] Real-World Deployment of a Virtual Cable Hub, pg. 5
[3] Remote PHY Specifications, CM-SP-R-PHY-I10-180509, pg. 27

© 2018 SCTE•ISBE and NCTA. All rights reserved.

Figure 3 shows the devices and connectivity between the vCMTS Core servers over the CIN, which is "the network between the CCAP Core and the RPD. The CIN encompasses either or both the hub access network and the optical access network. The CIN can contain both Layer 2 switches and Layer 3 routers."[4] In this deployment type, the CIN traverses between the centralized vCMTS hub and the secondary hubs over multiple switches, leveraging a Layer 3 network.

One of the benefits in this deployment type is that different RPDs and PHY shelves may connect to the same or different vCMTS Core servers in the vCMTS hub, which allows flexibility and scalability in growing capacity. For example, multiple secondary hubs (and the RPDs in those hubs) can be connected to the same vCMTS Core server.

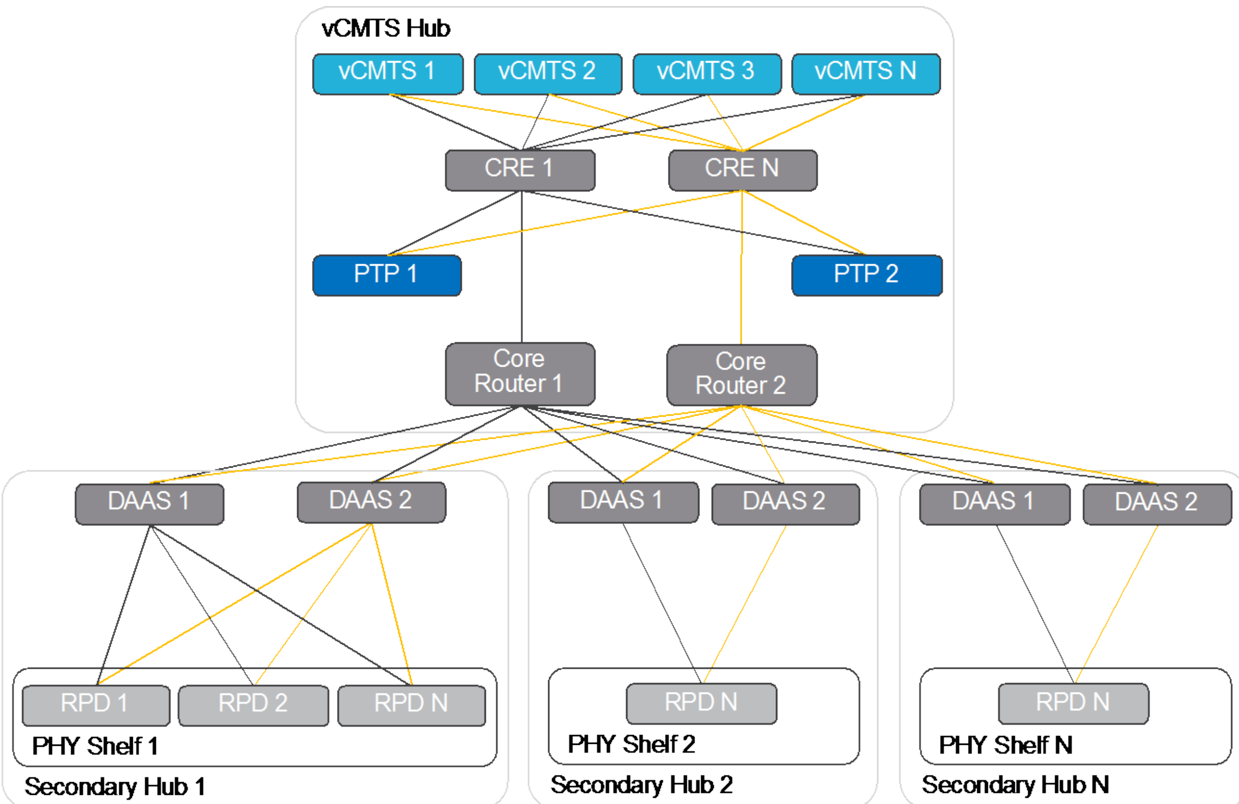

Figure 3 - vCMTS Deployed in a Centralized HFC Architecture

Real-World Considerations

While there are many benefits and opportunities with this new Remote PHY-enabled vCMTS in a centralized HFC deployment architecture, there are real-world considerations which should be considered. In particular, the following topics are described: compute, networking, timing, RPD or PHY type and operations.

[4] Remote PHY Specifications, CM-SP-R-PHY-I10-180509, pg. 24

1. Compute and Network Resource Location

One of the key considerations when evaluating this architecture is where the CMTS Core is located, for two reasons. The first is that with I-CCAP, the CMTS Core and RF had to be co-located. I-CCAPs are chassis-based products and are scaled between 13 and 16-RU per I-CCAP chassis. There was no choice in the past as I-CCAPs would be deployed at each hub location, regardless of the size and scale required to support the nearby footprint of subscribers. The Remote PHY specifications enabled separating the CMTS Core from the RF and provides the operator an opportunity to decide where the CMTS Core should be located.

The second reason for CMTS Core location determination is virtualization. A vCMTS Core solution provides the operator an opportunity to determine where the vCMTS Cores should be located, with options such as installing the vCMTS Cores at each hub location with the RF (similar to I-CCAP) or consolidating the vCMTS Cores at a few hubs or even a single centralized hub location. Having at least two vCMTS Core server locations may also provide geographical redundancy.

In this real-world deployment example, the vCMTS Core servers were consolidated in an existing hub location, which would be the single hub location for vCMTS Core servers now and in the future. With a single vCMTS location, future expandability of distributed RPNs can be accomplished over the same deployment architecture, connecting each RPN to the appropriate DAAS in a nearby secondary hub, which may already be connected to an existing PHY Shelf. This expandability option is very efficient, as PHY Shelves and RPNs can connect across the CIN to existing vCMTS Core servers and DAAS, providing the operator the option to expand the compute resources of vCMTS Cores and network resources of the DAAS as capacity demands.

© 2018 SCTE•ISBE and NCTA. All rights reserved.

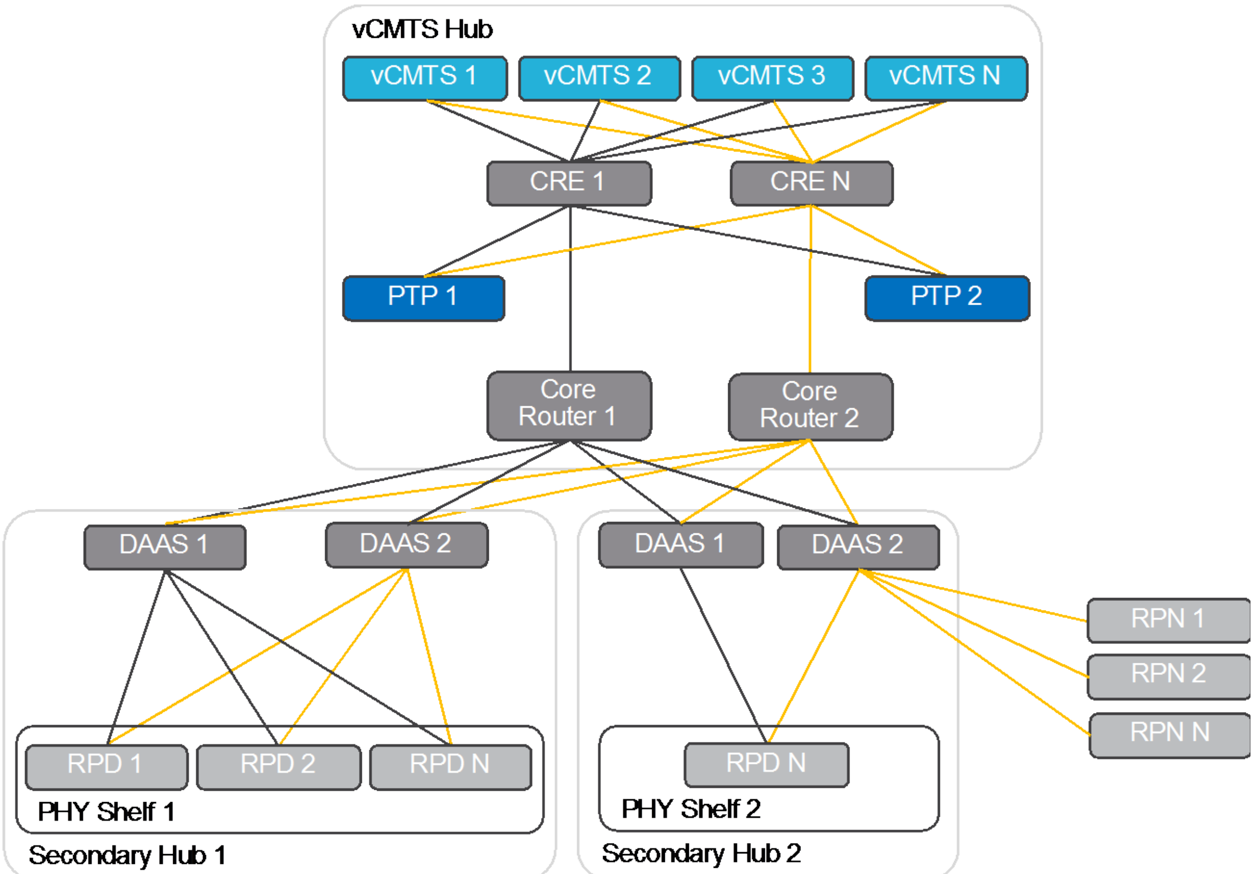

Figure 4 - Adding Remote PHY Nodes to a PHY Shelf Deployment

More flexibility in deployment location and scalability requires increased attention on the scalability of each element in the end-to-end network. While expandability is easier and provides quicker time-to-market to add capacity, operators must pay attention to the scale limits of the compute and networking resources separately, as each type of resource type may require additional devices when those limits are reached. On the other hand, when scale limits are reached for a particular resource, those can be expanded in a focused method and the operator won't have to scale everything at once. For example, when I-CCAP chassis reaches any of its scale limitations, another I-CCAP needs to be installed. In comparison, when a DAAS reaches a scale limitation, such as port count, another DAAS switch can be added without adding additional vCMTS Core servers. Nevertheless, an operator must pay attention to each element's specifications and plan a network for existing and additional subscribers.

2. Networking

2.1. Layer 2 or Layer 3 Converged Interconnect Network

There are many real-world lessons learned in the area of networking, as separating the vCMTS Core from the RPD not only spreads the CIN network elements across a variety of locations, and as described before, but it also spreads across different network resource instances and different network configurations. In other words, different deployments need to consider how many hops exist between the vCMTS Core and the RPD, and if the CIN is layer 2 or layer 3.

In this real-world deployment experience, layer 3 networking was selected for its benefits of greater flexibility and protection, as well as it is simpler to manage and scale between the CRE, core router and DAAS with this deployment architecture. Layer 3 networking has embedded control plane capabilities which provide these benefits to the operator, as well as easier configuration across the different network devices.

2.2. Traffic Prioritization and Capacity Management

Another important deployment consideration is that this architecture may expose shortcomings in the existing operator network and force improvements in IP network robustness and traffic prioritization. It is simpler to deploy when it can be guaranteed that the CIN is congestion free. However, if this can’t be guaranteed, capacity and congestion management of the CIN network devices are necessary.

Capacity management is important to determine when expanding CIN network resources are necessary and installing these network resources before congestion occurs. If congestion does occur, capacity management and traffic prioritization need to be calculated and evaluated. Traffic prioritization effectively allows higher priority control and user packets to survive the congestion. There are multiple ways DOCSIS and IP-based traffic prioritization can be used to maintain the Quality of Service (QoS) when manageable congestion occurs in the CIN. However, if congestion reaches to the extent that is impactful to critical control packets, such as DOCSIS MAP MAC Management Messages (MMMs), no amount of traffic prioritization will resolve the impact to stable network operations.

2.3. Redundancy

Network redundancy is another deployment decision with many options available, including:

1. Link redundancy
2. Chassis redundancy
3. Line card redundancy within a chassis

The decision criteria for each network device type (CRE, core router and DAAS) is based on failure domain size (how many subscribers are impacted), the built-in physical redundancy within each network device and the cost of the incremental networking equipment. In this deployment example, lower-cost CRE and DAAS network elements are 1-RU devices and are deployed with redundant links and with two redundant instances to the vCMTS Cores and PHY Shelves, so there is protection in case a single CRE or DAAS fails. On the other hand, the core routers are carrier-class chassis-based devices with redundant line cards and don’t require link redundancy to each CRE, since a separate network path via another core router is available in the CIN should a core router fail.

3. Timing

3.1. Remote DOCSIS Timing Interface

With all Remote PHY deployments, timing specifications such as R-DTI need to be adhered to, regardless of the location of the RPD. “The MHAv2 version of DTI (i.e., R-DTI) defines how to distribute phase and frequency information from the CCAP Core device to remote PHY devices within the HFC network.”

“For Ethernet based networks, IEEE 1588 allows both phase and frequency information to be transferred between nodes across an existing packet network with switches or routers, thus making it ideal for R-DTI.

© 2018 SCTE•ISBE and NCTA. All rights reserved.

In order to reduce any phase offset introduced by latencies through the network, IEEE 1588 defines a protocol for calculating the latency across sections of the network, and then compensating for those latencies. The latency calculations assume that the link is symmetric, and therefore the protocol works well for traditional full duplex Ethernet networks. IEEE 1588 also defines a protocol for determining the latency through any intervening switches or routers within the network, but the device is to be IEEE 1588 capable [referred to as PTP aware]. If the devices are not IEEE 1588 capable, the phase offsets and convergence times within the network will be greater [referred to as PTP unaware]." [5]

This real-world deployment experience is across a PTP unaware network, as not all existing network devices were IEEE 1588 capable. While the convergence times within the network are greater, as expected, they have not been operationally significant to justify immediate replacement of all network devices to be IEEE 1588 capable. However, it is critical to evaluate and consider the jitter and latency conditions of the CIN regardless of selecting an PTP aware or unaware mode. Both PTP deployment modes demand meeting jitter and latency requirements, which can be impacted by the CIN network device capabilities, and the number of network hops and congestion conditions.

As timing is critical for Remote PHY operation, there are a couple of other options to consider when deploying IEEE 1588 PTP grandmaster(s) in the network, which transmit the synchronization information to the other clocks in the same network. The first option is the reliability of the PTP grandmaster, as there exist a range of products which are small (SFP form factor) and less reliable as compared to full carrier-grade products which have redundant input/output clock (IOC) cards. The second option to consider is whether to use the best master clock algorithm (BMCA), which determines which is the highest quality or "best" clock within the network, in case the grandmaster clock quality is compromised or fails.

3.2. DOCSIS Latency

Latency is an important parameter in any QoS system, including DOCSIS-based solutions, keeping in mind that end-to-end latency measured between a subscriber's CPE and the end-point server extends beyond the DOCSIS portion of the network. DOCSIS has different latency in the downstream and upstream directions, which is a common trait of point-to-multipoint access technologies.

I-CCAP latency measurements are well known and are described in the MULPI specification, with a minimum latency associated with the DOCSIS MAC protocol for best effort traffic. The minimum latency budget consists of the worst-case round-trip propagation delay (variable), queuing delays within the CMTS, processing delays within the CMs and downstream delays caused by the PMD-layer framer and FEC interleaver.[6]

The expected and measured behavior is that ping times increase linearly as distance increases between the CMTS Core and CM, attributed to the one-way propagation delay of 0.8 msec. per 100 miles (0.5 msec. per 100 km.). However, maximum bandwidth is not impacted as this distance is increased.

There is no difference in DOCSIS MAC and PHY processing delays between I-CCAP and Remote PHY deployments. However, this topic is relevant with Remote PHY deployments as the CMTS Core may be centralized further in the network than the RF (located with the RPD) and may extend the distance, as well as the number of network hops between the CMTS Core and the CM. The transit time in both directions is comprised of these two contributing factors: propagation through the fiber and coax, and the transport processing delays. Real-world experience has shown that the dominant contributor to increased

[5] Remote DOCSIS Timing Interface, CM-SP-R-DTI-I07-180509, pg. 6, 16
[6] MAC and Upper Layer Protocols Interface Specification, section 7.2.1.6

latency within the DOCSIS portion of the network is in fact the propagation delay, with the increased number of network devices in the CIN minimally impacting the transit time and associated latency.

As with an I-CCAP, any Remote PHY-based CMTS Core needs to account for the MAP advance time supporting the maximum transit time between the CMTS Core and the CM.

The impact of end-to-end latency and the contribution of DOCSIS latency within the network to best-effort services is beyond the scope of this paper.

Nevertheless, with Remote PHY-based deployments, the round trip and latency performance are more in the operator's control in comparison to I-CCAP. The operator can maintain and, in some cases, improve latency performance of the system. Conversely, the operator may deliberately choose an installation scenario that will reduce latency performance in order to gain in other aspects, such as centralizing vCMTS Cores in a single hub location.

4. Operations

Technology is great, but it must be supported by real-world operational practices which support deployment of any new technology. In the experience of deploying a vCMTS in a centralized HFC architecture, there is an immediate opportunity to start by consolidating monitoring of the vCMTS Cores in a single consolidated hub or single network operations center (NOC). This consolidation allows monitoring of any service group within the set of vCMTS Core servers. Additionally, with a software-based virtualized CMTS Core, there are many points of inspection in the software which can be made visible based on field experience and extended over time, based on new findings. Continuous monitoring improvements will also be augmented over time with configuration, deployment and automation capabilities.

Conclusion

Cable operators have an existing footprint which they continue to grow and improve, but can't be overhauled overnight. Technologies such as Remote PHY and virtualization continue to extend the tool set which cable operators can use to stay competitive with capacity growth demands and challenging market environments. This paper has described a starting point which takes an existing I-CCAP centralized HFC deployment and changes a few architectural elements in the network with immediate benefits while paving a path to DAA. This approach leverages an installed base of vCMTS Core servers and will ultimately support a hybrid set of centralized and distributed dense PHY shelves, smaller and more remote PHY shelves and remote PHY nodes.

© 2018 SCTE•ISBE and NCTA. All rights reserved.

Abbreviations

CapEx	Capital Expenditure
CCAP	Converged Cable Access Platform
CMTS	Cable Modem Termination System
COTS	Commercial Off-The-Shelf
CPE	Customer Premise Equipment
CPU	Central Processing Unit
DAA	Distributed Access Architecture
DEPI	Downstream External-PHY Interface
DOCSIS	Data Over Cable Service Interface Specification
Gbps	Gigabits Per Second
GCP	Generic Control Plane
HFC	Hybrid Fiber-Coaxial
HW	Hardware
I/O	Input/output
MAC	Media Access Control
NFV	Network Function Virtualization
NIC	Network Interface Controller
NOC	Network Operations Center
OOB	Out-of-band
OpEx	Operating Expenditure
OS	Operating System
PHY	Physical
PNM	Proactive Network Maintenance
RF	Radio Frequency
RU	Rack Unit
SCTE	Society of Cable Telecommunications Engineers
SDN	Software Defined Networking
SW	Software
TCO	Total Cost of Ownership
TTM	Time to Market
UEPI	Upstream External-PHY Interface
vCMTS	Virtual CMTS
vCPE	Virtual CPE
VOD	Video on Demand

Bibliography & References

DOCSIS 3.1 MAC and Upper Layer Protocols Interface Specification, CM-SP-MULPIv3.1-115-, May 9, 2018, Cable Television Laboratories, Inc.

Matatyaou, Asaf. Real-World Deployment of a Virtual Cable Hub. Publication. San Jose: Harmonic, 2017. Web.

Matatyaou, Asaf. Transforming the HFC Access Network with a Software-Based CCAP. Publication. San Jose: Harmonic, 2015. Web.

Modular Headend Architecture v2 Technical Report, CM-TR-MHAv2-V01-150615, June 15, 2015, Cable Television Laboratories, Inc.

Remote DOCSIS Timing Interface, CM-SP-R-DTI-I07-180509, May 9, 2018, Cable Television Laboratories, Inc.

Remote PHY Specification, CM-SP-R-PHY-I10-180509, May 9, 2018, Cable Television Laboratories, Inc.

© 2018 SCTE•ISBE and NCTA. All rights reserved.