SCTE CABLE-TEC EXPO 23
CELEBRATING 40 YEARS

2023 Fall Technical Forum
SCTE • CABLELABS • NCTA

DNS Cowboys, On the Edge of a New Frontier

Lowering Customer DNS Latency with vCMTS Edge Compute

A Technical Paper prepared for SCTE by

Charlie Helfinstine
Sr. Principal Engineer
Comcast
1800 Arch St. Philadelphia, PA
267-858-8002
charles_helfinstine@comcast.com

De Fu Li, Distinguished Engineer, Comcast

Eric Stonfer, Sr. Principal Engineer, Comcast

Joe Crowe, Principal Engineer, Comcast

Table of Contents

List of Figures

List of Tables

1. Introduction

With cable wireline users demanding low-latency services for new applications like hosted gaming and virtual reality (VR), Multiple System Operators (MSO) are lowering the latency of critical network services like Domain Name System (DNS) by hosting them on edge compute infrastructure, such as Virtual Cable Modem Termination Systems (vCMTS).

vCMTS platform is a multi-site, Kubernetes edge compute platform deployed at scale to deliver the Xfinity 10G Network experience. It has been focused on delivering best-in-class network access services to residential and business customers but is being expanded to support new network service capabilities such as DNS. No other computing platform has been deployed this close to customers inside the Comcast Network. If edge computing is the new frontier, then vCMTS platforms are literally at the edge.

Our paper and presentation will outline work underway to host DNS residential caching resolvers on the vCMTS and enhance the multi-tenant capabilities of the platform to host third-party applications. In addition, the integration of vCMTS DNS into the residential DNS cache operations, engineering, and future scaling will be described. Moving network functions like DNS caching closer to the network edge will lower overall DNS latency and enhance the DOCSIS 10G Network experience.

2. The Drive for Lower Application Latency

Since the advent of Transmission Control Protocol / Internet Protocol (TCP/IP), latency has been critical for determining user experience with network applications. From the responsiveness of online games and interactive media to the needs of video conferencing, limiting the network delay is a key component for a modern network. The overall latency is an aggregation of component delays and propagation delays between network devices spanning from end-to-end on every packet transit. Limiting the number of components, the delay through network devices, and the distance between them all play a part in reducing the latency experienced by the end user.

Applications attempt to limit the number of round-trips traffic must take across the network; if there are ways to shorten the path, the effect can be multiplied across the frequent traffic needed for continuous interactive applications. Ongoing research in low latency technologies is opening the opportunity for new applications such as edge-cloud hosted gaming.

Customer applications that strive for lower latency may be optimized to have quick responses, but they are reliant on common network services like DNS to not act as a latency bottleneck and a source of unexpected experience. The effects of DNS latency can impact any application, and care must be taken to enhance the network latency of DNS to keep pace with the push for lower latency.

2.1. Emerging Markets for Lower Latency

Basic responsiveness is important for all applications, but there are widespread and growing new needs for low network latency from online gaming and teleconference meeting systems. Recent advances in network latency enable hosted gaming environments that compete with the console and mobile gaming markets. The time between a user's controller input and the response of the game, alongside interactions with other players, are critical for the entertainment value of a game. In online meetings, the added delay between participants in a conversation contributes to communication difficulties and loss of the natural rhythm of speech.

Longstanding low-latency applications like content delivery networks and video-on-demand systems are also driving lower latencies due to the competitive market for the services. Any improvement in network latency for a provider's services can be a differentiator in a crowded space.

2.2. Emerging Technologies for Lower Latency

A number of recent network initiatives are driving lower latency at almost every point in the network. Whether the work is in response to the growing needs of applications or has enabled the viability of new uses, there is a focus on optimizing queuing and congestion management to push the most throughput and lowest latency out of TCP and QUIC transport protocols.

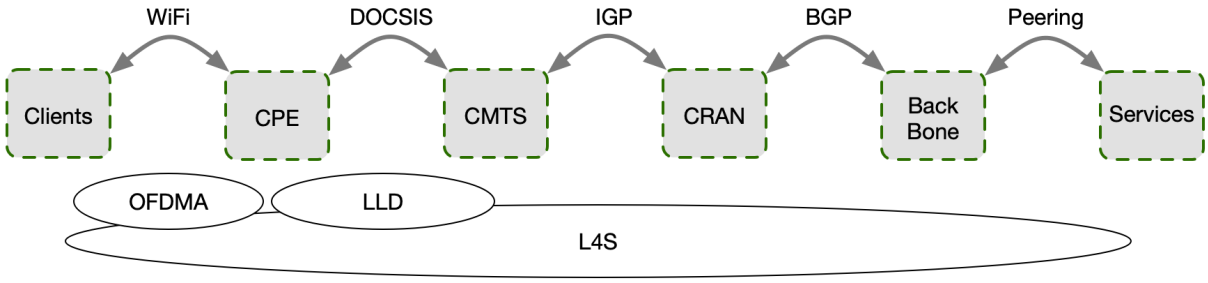

Figure 1 – Low Latency Technology Areas

The Low Latency, Low Loss, Scalable Throughput Internet Service (L4S) architecture described in RFC9330 is one such project that is developing a systemic approach for lowering end-to-end latency in congested networks. It describes the use of Explicit Congestion Notification (ECN) congestion flow-control signaling as a way to hold off packet drops and limit the retransmit and TCP window size scaling effects on overall throughput. However, to feel the benefits of ECN signaling, care must be taken to ensure that appropriate queue management is enforced on the ECN flows end-to-end to ensure optimal delivery of latency-sensitive traffic mixed among all traffic. There is significant engineering work needed to implement L4S services across large networks with standard queuing paths, with the goal of lower latency achieved when complete end-to-end paths are upgraded to support the queue management features required.

Another project driving down latency closer to the network edge is Low Latency Data Over Cable Service Interface Specification (LLD). As part of DOCSIS 3.1, LLD lowers the queueing delay and media acquisition delay for application traffic from sources that are termed Non-Queue-Building applications. Such applications include interactive gaming systems and online meeting tools that attempt to send lower capacity time-sensitive flows across the network and are less likely to overflow queues in the path. LLD will support these apps with a low-latency queue. Measurements show that low-latency queue performance can lower latency from the standard 10ms down to 1ms.

Even the last link to client devices, Wi-Fi, is lowering latency with Orthogonal Frequency-Division Multiple Access (OFDMA) modulation included in Wi-Fi 6 and 6E. OFDMA provides subchannel access between wireless endpoints that can allow better utilization of available spectrum and queuing of traffic. It's possible to lower application traffic latency over OFDMA for lightweight flows that can fit into the subchannel multiple access. The same apps that may be considered Non-Queue-Building applications in LLD are the types of applications that may best utilize the OFDMA subchannels.

Across L4S, LLD, and Wi-Fi 6, there is a shared strategy of low-latency queueing. Each technology focuses on a different segment of the network, but together they can enable better end-to-end latency for applications like hosted gaming or online meetings.

2.3. Lower Latency Network Services

Alongside application data, network services like DNS benefit from lower latency. Most network interactions start with DNS resolution, which needs to complete before further application flows may start, making the DNS resolution time additive to the overall application performance. DNS latency can be improved with caching located close to customer devices and by leveraging the lower latency queueing capabilities of the network in the future.

3. Shrinking DNS Latency in the Network

Providing DNS customer caching services at MSO service provider scale with low latency is a challenge and must be pursued with a goal of zero downtime. The needs of low-latency applications are pushing service providers to embrace new approaches for DNS that can lower latency at the network edge while integrating into existing DNS architectures and maintaining Service Level Agreement (SLA) expectations.

With the emerging LLD and L4S protocols providing a path forward for low-latency applications to traverse the network, DNS service appears to be a good candidate to use the low-latency queues when available. Standard DNS services are low-bandwidth flows using Non-Queue-Building protocols with short packet sizes that will lend well for minimizing queue depth in low-latency queues. As LLD and L4S adoption increases, DNS's use of the low-latency systems should increase in tandem.

In the interim, DNS latency may be improved by managing the location of caching services provided in the network. Moving the customer caching services closer to the end users provides a clear path to lowering the latency observed for DNS services.

3.1. DNS Customer Caching Locations in Provider Networks

DNS caching services for residential and business customers can be hosted at a number of different locations with differing effects on latency. Service providers who operate a network edge can deploy DNS services centrally off their backbone, in main data centers, in regional network sites or datacenters closer to the network edge, or in Customer Premises Equipment (CPE) on site with the end user.

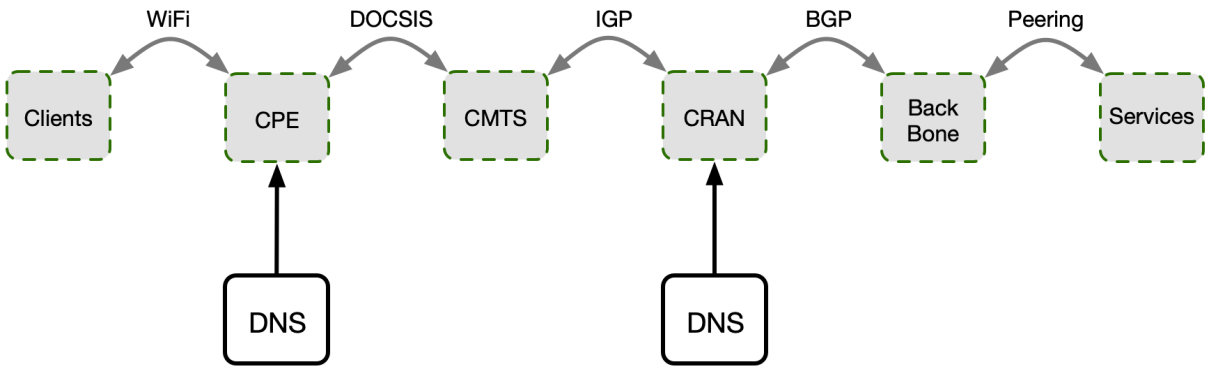

Figure 2 – DNS Caching Locations

By counting hops and considering the speed of light through fiber, the natural assumption is that the closer to the end user, the lower the network latency possible for the DNS service. However, there is a complexity to different deployment strategies. DNS cache hit rates are affected by the number of subscribers sharing a given cache view, and placing a cache too close to a few subscribers will benefit the resolution of domains that are widely used by almost all clients, but there is a long tail of DNS entries supporting important applications that are less frequently queried yet need fast resolution.

Placing the physical infrastructure for DNS services also increases in cost the closer you get to the edge of the network. When a service can be deployed in a limited number of sites, the cost per unit of DNS service is lower than a corresponding deployment of the same capacity that is located in exponentially more sites that are geographically closer to customers - due to the expense of installation and deployment being roughly similar per site independent of capacity.

Due to the automation of changes and upgrades, the number of sites supported operationally is less important than in the past, but the development effort and the room for error still increase with the number of sites. In an environment where little or no operational intervention is needed for DNS services, the number of DNS sites also has less relevance. For instance, hosting DNS caching services on customer premises equipment is a common approach that is very close to customer devices, however, conducting operational DNS functions such as cache flushes or deploying config changes becomes very difficult.

3.2. Current DNS Infra and the Latency Provided to Customers

MSOs commonly provide residential and commercial DNS resolution services across their footprints with a combination of regional sites and centralized sites. As an example, the Comcast network is subdivided into Converged Regional Access Networks (CRAN) that serve differing numbers of subscribers ranging from hundreds of thousands to millions based on geographical location, and each CRAN hosts a corresponding DNS site. The CRAN resolving service is advertised with anycast routing on the well-known addresses of 75.75.75.75 and 2001:558:FEED::1 as the primary DNS address given by DHCP, and a centralized copy is advertised on 75.75.76.76 and 2001:558:FEED::2 as the secondary DNS address.

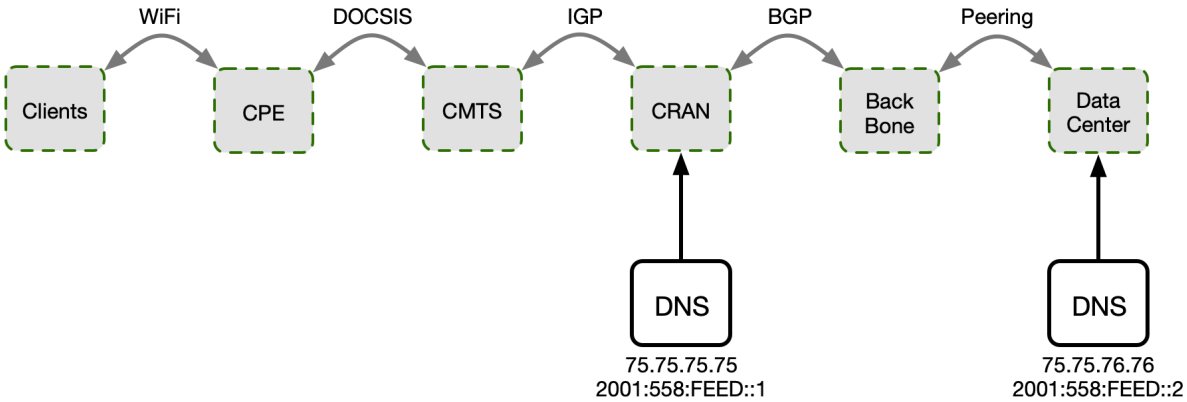

Figure 3 – DNS Service Locations

The regional site deployment model provides close, low-latency DNS service to customers and high cache hit rates from the client scale that maps to a given site. DNS latency data has been shared in more technical detail before, but for context, Comcast customers commonly see DNS latencies ranging as shown.

Table 1 – Comcast DNS Latency

DNS Anycast Address	Typical Latency (ms)
75.75.75.75	18
75.75.76.76	25

Source: RIPE Atlas

It should be noted that Comcast does not cache DNS records on CPE devices due to scaling concerns around flushing stale cache entries. Additionally, Comcast resolvers provide Domain Name System Security Extensions (DNSSEC) validation of signed domains and encrypted DNS services that do not scale well at the CPE level.

4. Leveraging vCMTS Edge Compute for Lower DNS Latency

With the drive to low-latency applications, MSOs are working to lower the DNS latency seen by customers by placing new DNS service components closer to the network edge. The virtual CMTS (vCMTS) platform, deployed as a key component of the Xfinity 10G network, offers a computing environment located as close as possible to the customer prem without being on-site and can host DNS services with significantly lower latency.

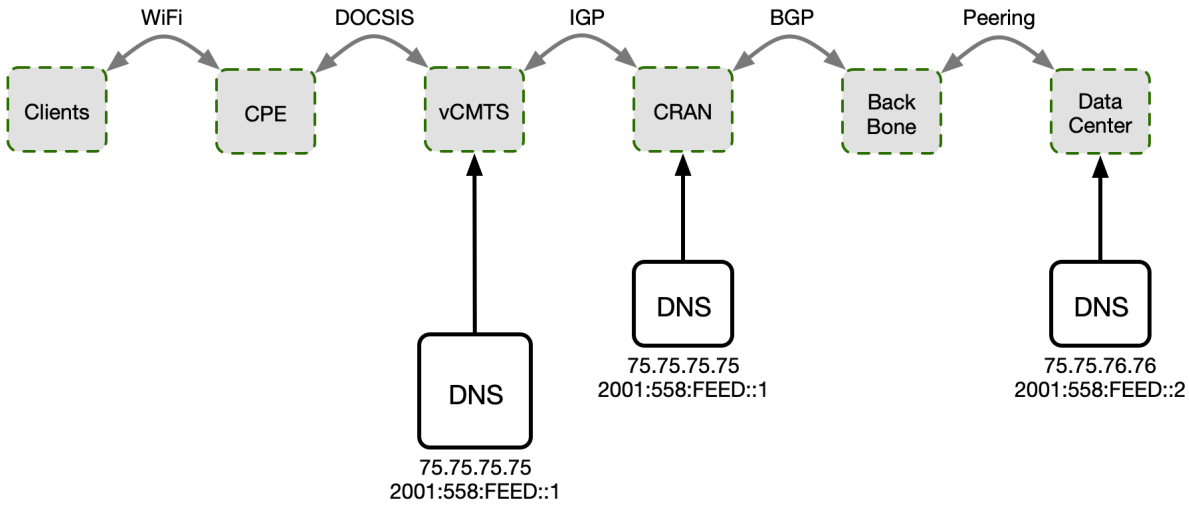

Figure 4 – DNS Service Locations with vCMTS

4.1. vCMTS Edge Cloud Background Information

The broadband access over the cable access medium is defined by the Data Over Cable Service Interface (DOCSIS) specifications. The cable modem is the customer-premises equipment (CPE). The cable modem termination system (CMTS) is the headend equipment. Earlier generations of the CMTS were built on custom hardware platforms. As part of the DOCSIS specification umbrella, the distributed access architecture (DAA) specification splits the CMTS into two distinct components: the physical function and the core function. The Remote Physical Layer (PHY) Device (RPD) provides the PHY function, while the core functions consist of CMTS operating on the MAC and upper layers. With the advance of data center technologies, the CMTS core function is realized in software running on Commercial Off-The-Shelf (COTS) servers and virtualized. Hence, the virtual CMTS (vCMTS) is referred to as a software implementation of the DAA CMTS core function.

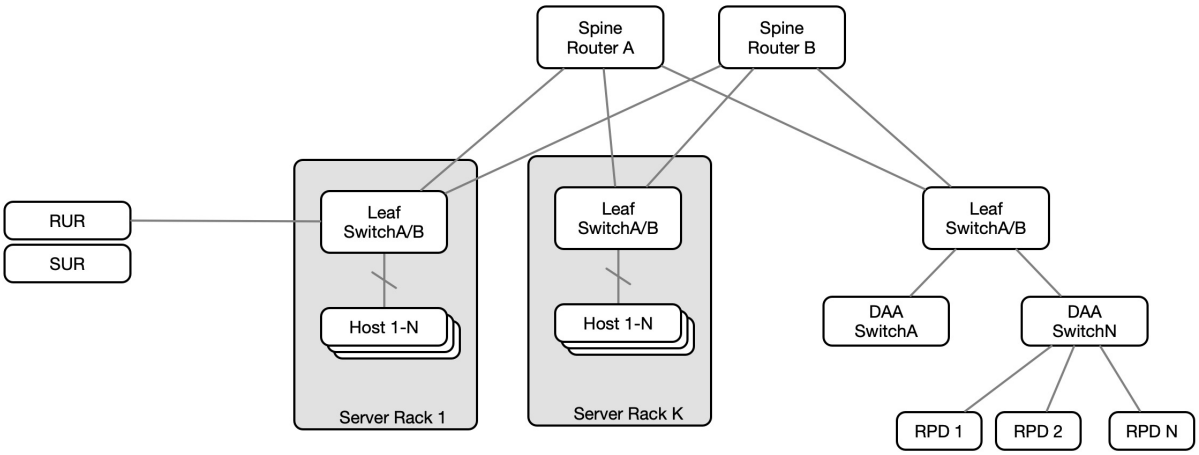

Figure 5 – vCMTS Edge Coud (Simplified View)

To host the vCMTS software applications, private edge clouds physically located in selected headends are required. Figure 5 shows a simplified diagram of the vCMTS edge cloud platform. The RPDs are connected to the DAA aggregation switch. There are multiple racks of servers that host the vCMTS and supporting applications. The vCMTS software is instantiated or orchestrated to run on the host; there are hundreds and thousands of these vCMTS workloads (instances). The network traffic flows between the workloads on the host and RPDs interconnected via the datacenter leaf-spine switch fabric.

4.2. Deploying DNS on vCMTS Edge Cloud Platform (High Level Design)

As mentioned earlier, the virtual CMTS (vCMTS) platform is a multi-site, Kubernetes edge compute platform deployed at scale to deliver the DOCSIS 10G Network experience. It has been focused on delivering best-in-class network access services to residential and business customers but is being expanded to support new network service capabilities such as DNS.

The vCMTS applications are deployed and orchestrated using Kubernetes. The microservice workloads are logically partitioned into Kubernetes namespaces. This is illustrated in Figure 6. The workloads for the vCMTS core function are orchestrated to its own namespace. Similarly, there are many software infrastructure services, such as a REDIS cluster for in-memory databases.

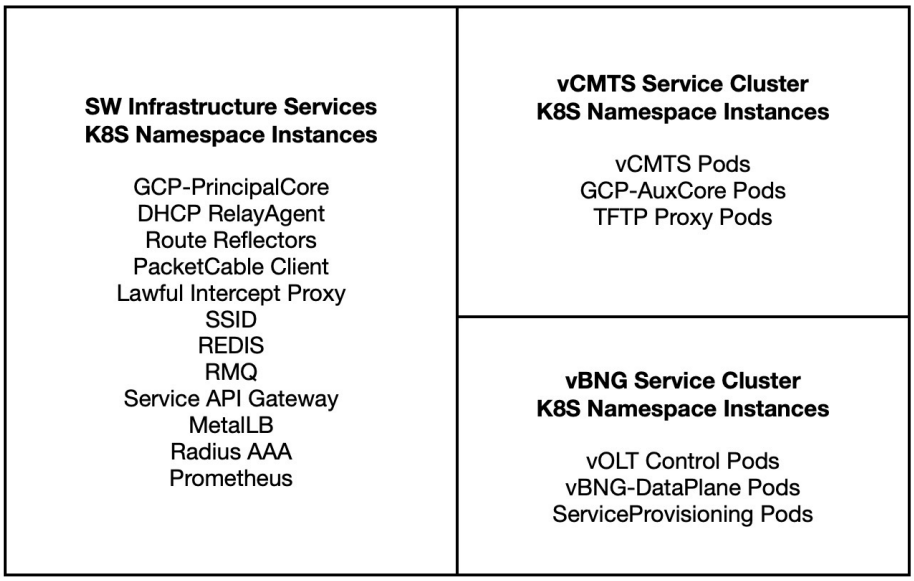

Figure 6 – vCMTS Namespace Partitions

The bulk of vCMTS workloads perform network function virtualization (NFV) tasks for DOCSIS core functions which are typically compute and network bandwidth-intensive Data Plane Development Kit (DPDK) applications. In contrast, the infrastructure microservices like REDIS have very different CPU and network bandwidth requirements, more in line with traditional application services.

Despite being a network service, the DNS cache workload functions more like a traditional application than an NFV component that is better suited for the infrastructure microservice category. Hence, one consideration is deploying it in the infrastructure namespace, where it can be managed as other shared components.

With the vCMTS disaggregated forwarding plane and Kubernetes container network interface, careful integration of the DNS service is required. Existing anycast DNS addresses will be advertised by vCMTS route reflectors locally in the routing table of the vCMTS to forward client DNS traffic to the ingress service for the local DNS caching pods. Customer devices will use the same DNS service addresses of 75.75.75.75 and 2001:558:FEED::1, however, the vCMTS DNS pods will perform the DNS caching and recursive resolution instead of the regional network DNS anycast services.

The local anycast route advertisement is subject to status monitoring performed on the DNS pod, and any interruption in DNS service will cause the local anycast route to be withdrawn. Route filtering in the vCMTS table ensures that the best route to the anycast addresses after a local anycast route is withdrawn will be the regional CRAN resolvers that provide DNS services at scale today. vCMTS instances in the same site will not be used as fail-over resources for DNS services inside the site. Typically, capacity is reserved for failover scenarios and growing demand in anycast systems, but by planning the fail-over paths, the capacity reservations can be absorbed by the scaled systems in the CRAN.

When creating load models for the vCMTS DNS service, an analysis was performed to predict the DNS traffic volume expected from a set number of subscribers. This volume is expected to grow with the numbers of devices in households and use characteristics of popular applications, however, as a baseline calculation, measurements were taken through 2022 and 2023 on observed queries per second counts over Comcast's DNS customer caching service per region and indexed to the subscriber counts per region. The

results were remarkably consistent across the footprint and yielded a capacity model that is scaled for implementing DNS services on the vCMTS.

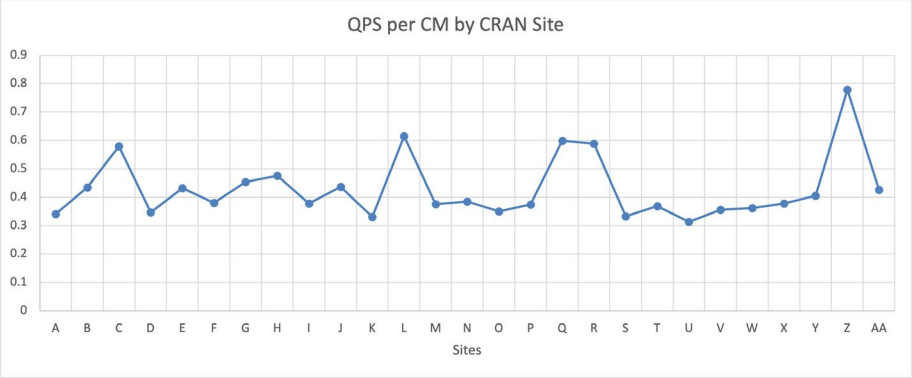

Figure 7 – DNS Utilization per CM

Functionally, the vCMTS DNS service supports the same feature-set as the CRAN customer caching resolvers with DNSSEC validation for signed domains and dual-stacked services among many others. The uniformity of the service requires that operational commands such as DNSSEC negative trust anchors are applied in both the CRAN locations and the vCMTS locations when needed. Operational reliability expectations are the same for DNS services hosted on vCMTS as the services hosted in CRAN locations. DNS system performance, status monitoring, and alerting, capabilities need support and equal vigilance.

5. Conclusion

To meet the push for low-latency services, MSOs can expand DNS services to the vCMTS edge compute platform. Development work is ongoing to expand the DNS service closer to end users and support rising next generation low-latency applications.

Abbreviations

CM	Cable Modem
CMTS	Cable Modem Termination System
COTS	Commercial Off-The-Shelf
CPE	Customer Premises Equipment
CRAN	Converged Regional Access Network
DPDK	Data Plane Development Kit
DOCSIS	Data Over Cable Service Interface Specification
DNS	Domain Name System
DNSSEC	Domain Name System Security Extensions
ECN	Explicit Congestion Notification
IP	Internet Protocol
L4S	Low Latency, Low Loss, Scalable Throughput Internet Service
OFDMA	Orthogonal Frequency-Division Multiple Access
RPD	Remote PHY Device
SLA	Service Level Agreement
TCP	Transmission Control Protocol
vCMTS	Virtualized Cable Modem Termination System

Bibliography & References

[RFC9330] Briscoe, B., Ed., De Schepper, K., Bagnulo, M., and G. White, "Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture", RFC 9330, DOI 10.17487/RFC9330, January 2023, https://www.rfc-editor.org/info/rfc9330.

CableLabs: Low Latency DOCSIS: Technology Overview 2019

