Decoding The Bandwidth Surge During Covid-19 Pandemic

An In-depth Study On DOCSIS Upstream Bandwidth Surge And Their Impact On Video Conferencing

A Technical Paper prepared for SCTE•ISBE by

Ram Ranganathan
Director, Office of CTO
COMMSCOPE
Mississauga, CANADA
905 568 7317
ram.ranganathan@commscope.com

Chris Markovich, COMMSCOPE

Tushar Mathur, COMMSCOPE

Omar Abu-Hijleh, COMMSCOPE

Thomas Cloonan, COMMSCOPE

John Ulm, COMMSCOPE

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

During the COVID-19 pandemic, millions of people around the globe have become far more reliant on their broadband internet connection to stay connected to their family, friends, and co-workers. The broadband internet has emerged as an essential technology that is keeping society together and the economy running in this stressful time.

The growth of broadband internet traffic was bound to happen due to “Work from home” and “Stay at home” instructions around the world during the pandemic. The dependence upon virtual meeting and online collaboration applications has become increasingly important during the pandemic as well. The web apps that are helping us stay connected must work reliably for people to be able to socially connect and stay productive. A few key questions had to be answered quickly during these crises by the broadband operators and vendors – Are the bandwidth surges during the COVID-19 pandemic breaking the broadband Internet connectivity? How are the commonly used applications handling these surges?

There are many popular web applications that drive the conferencing audio and video traffic in any broadband network. Many cable operators have reported a sudden growth of Upstream and Downstream traffic in their networks resulting from these conferencing apps[1]. Due to the tighter constraints that currently exist on the Upstream spectrum, the DOCSIS Upstream is more sensitive to this uptick in bandwidth consumption. In this paper, we will present our research and analysis of the major sources of DOCSIS Upstream bandwidth surge. Specifically, this paper will provide insight into the surge impacts on video conferencing applications and observations on latency, jitter, packet loss, and throughput as DOCSIS Upstream channel utilizations reach an extraordinary tipping point.

2. Impact of bandwidth surge on DOCSIS networks

Key application drivers for the bandwidth surge are quite similar on both DOCSIS upstream and downstream networks, including:

- Work-at-home applications – VPN traffic, video conferencing (Cisco Webex™, Microsoft Teams™, Zoom™ and others) and cloud sync for upstream traffic
- Remote learning applications including video conferencing/Zoom™
- Video streaming (Netflix™, Hulu™, Amazon Prime™, Disney+™ and others) predominantly as downstream traffic
- Gaming applications (more actual gaming traffic on the upstream as well as Twitch broadcasts, while downstream is impacted by Twitch video streaming of games)
- Social networks (Facetime/Skype increased usage between families etc.)
- Increased TCP Acks on the upstream due to increased downstream traffic

Based on our studies, as shown in Figure 1, downstream bandwidth usage increased by 12-25% during the busy hours (typically evening prime-time) and also significantly increased during midday by 26-86% again due to video conferencing related traffic loads. On the upstream, this surge was even more profound as peak periods of upstream bandwidth shifted towards the afternoons and the overall “sleep or quiet” periods of traffic was much lower. Upstream peak usage during typical evening peak periods increased by 20-50%, while midday usage showed a much more substantial increase of 30-150%.

The increased upstream usage is typically an achilleas heel of DOCSIS deployments for Multiple System Operators (MSOs) as today’s deployment are based on low-split with upstream channels in the frequency range of 20-42MHz or 20-65MHz. Future migration to mid or high splits may reduce the impact of such an upstream bandwidth spike.

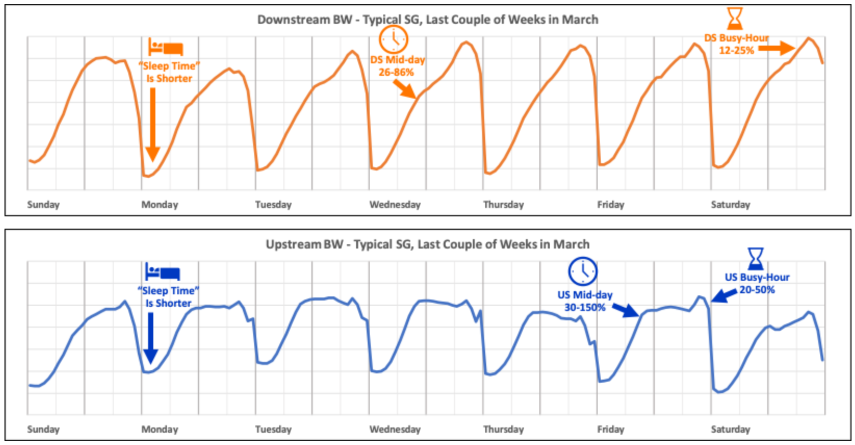

Figure 1 – Downstream and Upstream impact during COVID-19

Based on data analytics collected from multiple MSOs using the CommScope ServAssure program, we were able to observe network utilizations levels before and after the pandemic[2]. DOCSIS channel utilization increased significantly week over week during the pandemic lock-down for both upstream and downstream. While these rapid surges may have pulled back as the lockdown eased, the overall congestion observed, and the significant number of service groups impacted shows the need to be prepared for such bandwidth surges in the future. This paper focuses on impact of upstream DOCSIS channel congestion on applications such as video conferencing. It also studies how key performance metrics (KPIs) such as latency and throughput perform under such constrained channel conditions.

Overall, DOCSIS based cable broadband networks managed to handle the bandwidth surge pretty well. Key factors contributing to this positive result include:

- CMTS scheduling algorithms are adaptive handling congestion with fairness based algorithms
- Most Internet applications, especially high bandwidth ones like streaming video and conferencing apps, are adaptive in nature and adjust well to bandwidth fluctuations
- Mechanisms in modems and gateways like Active Queue Management have worked in tandem with application adaptive behavior to keep latency and jitter to somewhat an acceptable level
- Several MSOs have planned additional capacity headroom in their network planning and deployment (for example, 1.2 times Tmax rate to absorb SLA surges)

However, it should be noted that it was not all positive news when it comes to the sudden increase in bandwidth need. Applications like gaming suffered from increased "lag" (latency/jitter) due to buffer bloat issues. Due to packet delays and drops, video tiling and macroblocking effects were also observed in streaming and video conferencing applications. In addition, for subscribers with data caps for their broadband access, those data caps were hit more frequently during the pandemic.

3. The Experiment

Our study in this paper focuses specifically on video conferencing applications and how they are impacted by increased in-home traffic and by shared traffic within a DOCSIS service group. In a typical cable broadband access network, congestions points are typically encountered in these areas:

- In-home Wi-Fi Congestion: Predominantly caused due to large number of WIFI users or endpoints in the home. Other factors like wireless interference and distance-to-the-WIFI-gateway related issues also has an impact to overall network congestion, latency and packet losses.

- Cable Modem (CM) Congestion: This is often due to increased bandwidth from in-home applications. Algorithms like Active Queue Management on CMs that manage buffer bloat also influence performance KPIs.
- DOCSIS Channel Congestion: This is attributed to increased bandwidth usage in the neighborhood (associated with the specific Service Group) and scheduling algorithms in the Cable Modem Termination Systems (CMTS) that go hand-in-hand in managing this increased bandwidth.

The focal point of the study presented here is on #2 and #3 from the list above, specifically on the DOCSIS upstream channel which is usually more constrained due to limited bandwidth availability.

3.1. Focus on Video Conferencing Applications

As noted earlier, one of the key drivers of Internet traffic during the COVID-19 bandwidth surge was video conferencing. Applications such as Cisco Webex™, Zoom™ conferencing, Google Meet™ and Microsoft Teams™ were used extensively by work-at-home employees, remote learning and other e-learning activities and even as virtual video meetups for social events.

For the sake of this study, we took three of such commonly used applications which allow simultaneous sharing of audio and video along with the ability to be in "presentation mode" with a number of users. The chosen applications will be referred to as App 1, App 2 and App 3 in the remainder of this paper as the intention of our study is not critique the behavior of such applications, but rather to illustrate how diverse the behavior can be and how changing bandwidth conditions can affect the QoE of various conferencing apps.

All three video conferencing apps chosen are capable of running over UDP or TCP as the transport protocol for media traffic, however the default protocol (and the preferred one) is typically UDP. For our experiments, all conferencing apps were using UDP as their transport protocol for media (audio/video) traffic.

It should be noted that outside of audio and traffic, each conferencing application also had its own set of signaling protocols. Some of these leverage existing protocol standards such as Real-time Transport Control Protocol (RTCP) and Session Traversal Utilities for NAT (STUN). For all experiments, we did not make any modifications to the behavior of these signaling protocols but some of the traffic patterns resulting from how these applications reacted to packet loss and latency via signaling were observed.

It was also determined that all three apps have their own flavor of "adaptive streaming" behavior to handle congestion control and ensure that the end user behavior is kept to a minimum due to changing network conditions (and other factors in the streaming path).

3.2. Experiment Goals

Some key questions we attempt to find solutions for include:

1) How do key performance metrics like latency, jitter and application throughput change with increased DOCSIS channel utilization and in-home congestion?
2) What happens to Quality of Experience (QoE) of video conferencing applications when DOCSIS upstream channel is congested (70% or higher capacity used)?
3) How does the QoE change when the congested channel issue is combined with increasing home Internet traffic including multiple conferencing sessions?

3.3. Experiment Approach

Our initial setup in conducting these experiments was to determine how many active cable modems or gateways are “active and interacting” in a typical DOCSIS Service Group during the course of the pandemic. We expected that our previous figures on such active CMs may not be accurate due to the increased activity due to the lockdown effect in the pandemic. Our study of traffic patterns from North American MSOs during the late March and early April timeline showed us that typically 75-80% CMs were active and transmitting during a 15-second window during peak traffic periods. This number (80%) was used in our study to emulate traffic from multiple modems in a service group.

Another key aspect that was considered during these experiments was using emulation versus simulation techniques. Due to the complex nature of traffic patterns and changing server behavior of conferencing applications, it was decided that real-world traffic emulation communicating with real-world conferencing servers was the preferred approach.

As part of the experiment setup, we had an array of traffic generators that was modelled from statistical models derived from real-world, sampled DOCSIS upstream traffic. Specifically, bursty traffic loads were initiated from these traffic generators using statistical models of three key parameters:

- Duration of traffic burst
- Varying bandwidth levels within each traffic burst
- Inter-burst interval – idle time between successive bursts

The DOCSIS upstream was congested to the desired channel congestion levels (70%, 80% etc.) depending on the experiment needs using these traffic generators. Once this was achieved, real-world video conferencing applications were run with multiple users with the target user on a PC behind the CM under test. The test PC was then broadcasting a video stream and the QoE of that media stream was observed. Additional metrics such as latency and throughput were measured during the video conferencing session.

A conceptual representation of the experiment testbed is shown in Figure 2.

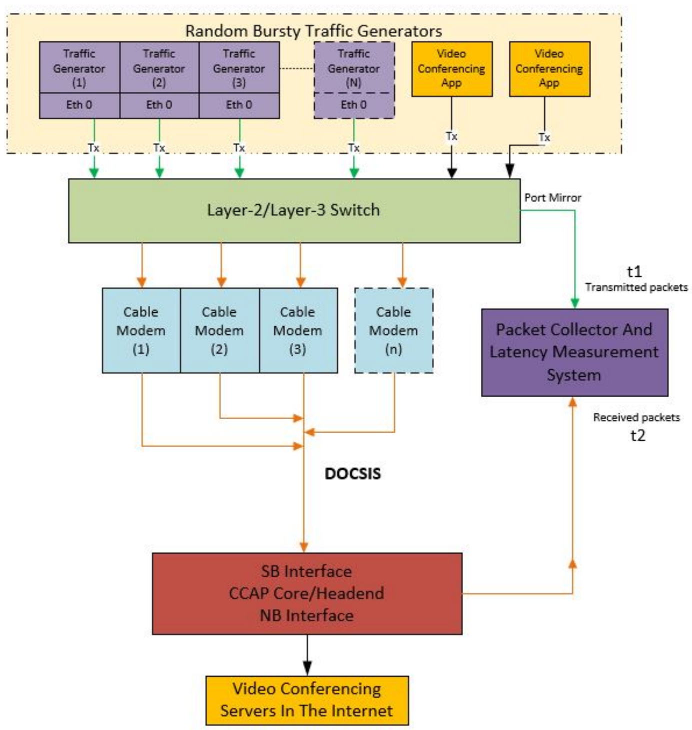

Figure 2 – Testbed Setup

3.4. Inputs

A service group size of 400 cable modems was chosen for the experiment setup. Based on the prior analysis of real-world subscriber activity, 80% of those modems were considered to be active and interacting.

Four service tiers were part of the testbed as shown in Table 1, along with the distribution of CMs across these tiers. 319 of the 320 modems (excluding the CM under test) were then connected to Traffic generators (as Customer Premises Equipment/CPE devices) and real-world modelled bursty traffic were sent from these generators based on the experiment's need to achieve the desired DOCSIS upstream channel utilization (70%, 80% and so on).

The "**CM under test**" was chosen from service Tier 3 which had a 25Mbps downstream and 5Mbps upstream Service Level Agreement (SLA). This was chosen as the "minimal acceptable upstream bandwidth" needed to run at least a couple of video conferencing applications from the home. Depending on the experiment, devices behind the test CM were one or more PCs running video conferencing applications or traffic generators for throughput/latency/packet loss measurements.

For all video conferencing tests, the test video was chosen to be an animated GIF image (repetitive motion) to account for motion based video transmission and the setup was maintained in a controlled lighting environment with minimal human interference during tests. Fluctuations in video quality (lighting, scene changes etc.) can result in significant variation in transmission bandwidth and the purpose of the controlled environment was to keep any drastic bandwidth fluctuations to a minimum. It should be noted that the animated GIF image was streamed via video during the conference rather than as a "screen share". Screen sharing typically uses less bandwidth than video transmission itself in such applications, so the experiment was focused on the more bandwidth-intensive video stream sharing.

Table 1 – Experiment Parameters

Key parameters	Values
Number of active CMs in SG	320
Service tiers (DS/US bandwidth)	Tier 1: 100Mbps DS/40Mbps US Tier 2: 50Mbps DS/20MBps US Tier 3: 25Mbps DS/5Mbps US Tier 4: 12 Mbps DS/1 Mbps US
Distribution of service tiers	Tier 1: 10% Tier 2: 65% Tier 3: 15% Tier 4: 10%
CM configuration details	DOCSIS 3.1 CMs, default values for AQM/buffer controls, ACK suppression enabled
DOCSIS Upstream Channel capacity	~100Mbps (4xATDMA channels: 3x6.4Mhz channels and 1x3.2Mhz channel) DS is not congested in these tests

3.5. Experiment Setup

Two primary setups were used for these experiments. The first set of experiments were focused on DOCSIS upstream channel loading using traffic generators as shown in Figure 3. Along with the traffic generators loading the DOCSIS channel, a single PC client is behind the test CM on a video conferencing session with 4 other participants. A small "test" stream is used to measure latency and packet loss metrics due to the congestion.

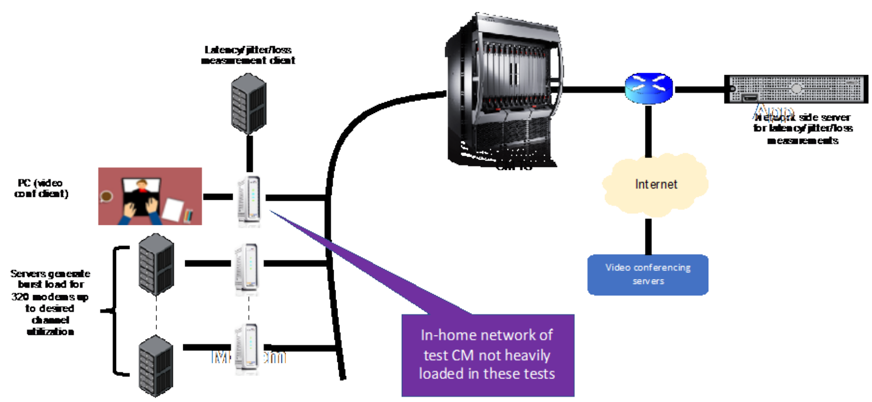

Figure 3 – Setup 1: QoE impact due to DOCSIS upstream

The second setup, as shown in Figure 4, adds multiple conferencing clients behind the test CM and also generates in-home traffic loads to emulate aspects like a speed test or file upload. Overall throughput achieved by the conferencing clients as well as other KPIs were measured using this setup.

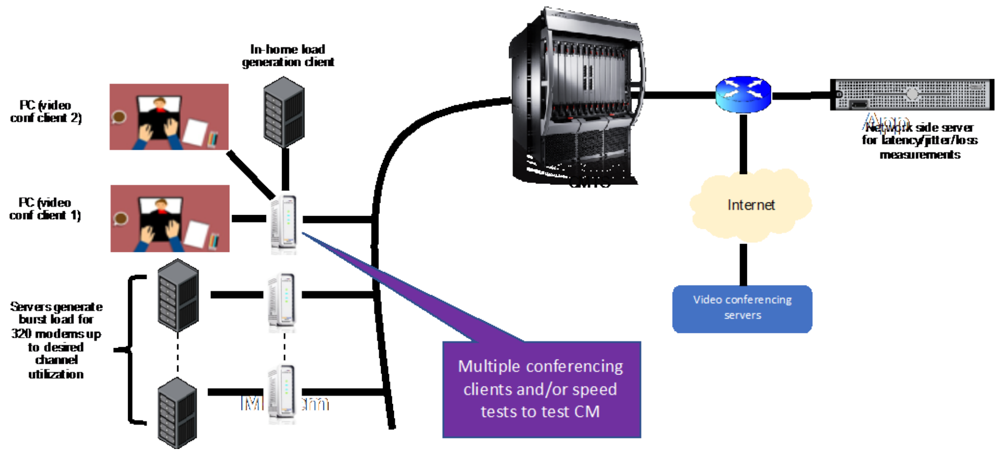

Figure 4 – Setup 2: QoE impact due to in-home and channel congestion

4. Observations and Analysis

4.1. Network Congestion related measurements

To obtain an understanding and a good baseline of how key performance metrics like latency, jitter and packet loss behaved under network load, experiments were conducted loading the DOCSIS upstream with bursty "real-world-like" emulated traffic from the traffic generators. Test streams were injected behind cable modems in every service tier defined in Table 1 to measure (mean) latency, jitter and packet loss at three key channel loading points – 78%, 100% and 122%. It should be noted that this baseline experiment did not have other high bandwidth applications like video conferencing.

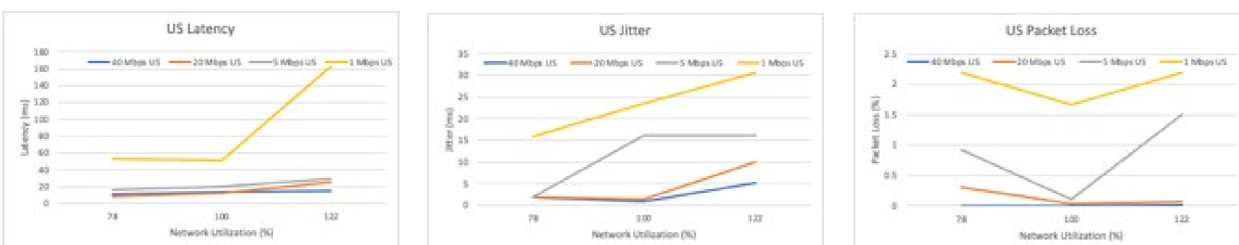

Figure 5 – Latency, Jitter and Packet Loss under DOCSIS upstream load

From Figure 5, it can be observed that (mean) upstream latency was not impacted heavily by channel utilization unless the Tmax was at the lowest performing tier (=1Mbps). The jitter graph shows that even with 5Mbps upstream Tmax, packet jitter starts becoming an issue as channel gets congested. This was further validated during our video conferencing tests which we describe in detail in the sections below. Packet loss in general was pretty low to be significant concern to typical applications.

Now that we have a rough idea on how performance KPIs change under DOCSIS channel conditions, let us focus on specific impact of upstream traffic – both in-home and on the shared DOCSIS channel – to video conferencing applications. All experiments described further were conduction on the "CM under test" which was using the upstream Tmax of 5Mbps – as it indicated the "tipping" point where impact on latencies and jitter can easily be observed.

4.2. Understanding Video Conferencing App Operating Points

All three video conferencing apps studied were adaptive to changing network conditions – available throughput, packet latency and jitter as well as packet drops. Each application had a preferred sweet spot for optimal operation, especially for video streaming bandwidth. Table 2 shows typical operating bandwidth range for the video stream from these apps as well as where they tend to "converge" if they have sufficient bandwidth.

Table 2 – Conferencing App operating "sweet spots"

App #	Operating video throughput range	Optimal Target throughput
App 1	200Kbps – 1.2Mbps	700-800 Kbps
App 2	500Kbps – 2 Mbps	1-1.25 Mbps
App 3	1Mbps – 2.5 Mbps	2.25-2.5 Mbps

Conferencing app behavior for media traffic bandwidth with and without network congestion is shown in Figure 6, Figure 7 and Figure 8. It can be observed that some Apps like App 1 are conservative in bandwidth usage while others like App 3 tend to be aggressive in maintaining good video quality pushing continuously to use bandwidth above 2Mbps against traffic congestion. Others like App 2 follow a more step-wise increase approach not unlike several standard adaptive streaming video protocols.

It was observed that at least 500Kbps was required to have minimal acceptable video quality in terms of frame rate and resolution and ideally that close to 1Mbps was required for good quality video streaming. Below 500Mbps, some apps like App 1 continued to show video streams at very low frame rates or video resolution, while other apps moved to "audio only" mode often against the bursty network load.

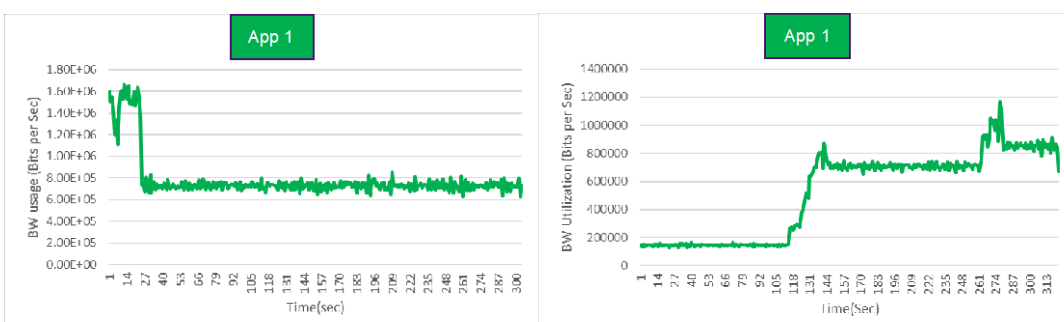

Figure 6 – App 1 without (left) and with (right) heavy network load

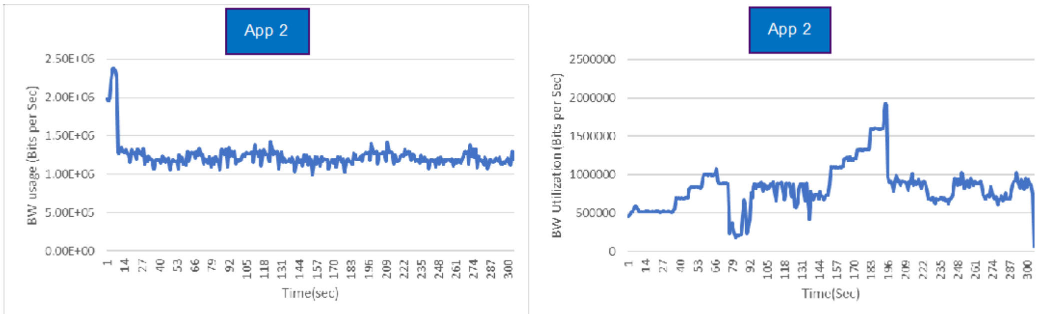

Figure 7 – App 2 without (left) and with (right) heavy network load

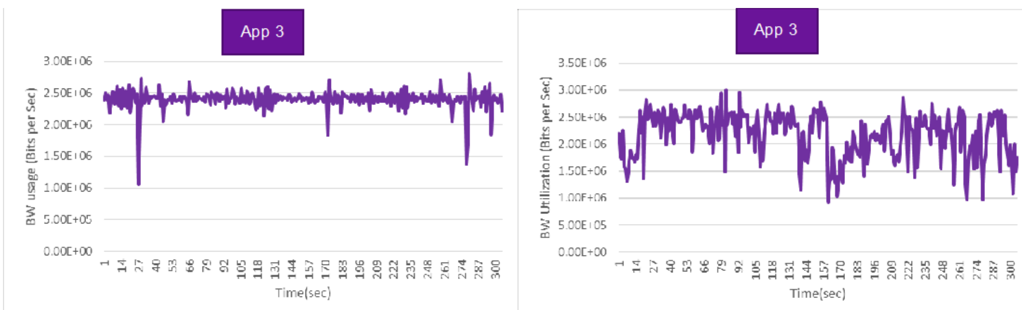

Figure 8 – App 3 without (left) and with (right) heavy network load

4.3. Speed test & Conferencing App throughput behavior

With multiple conferencing clients behind the test CM, a speed test-like behavior was emulated using TCP tools and the achieved throughput of the speed test app was studied. Figure 9 shows the behavior of such speed tests when 2 conferencing client apps (of the same type) are running in conjunction with a speed test behind the test CM. The baseline speed test shows a typical reduction in achievable throughput with increasing DOCSIS channel congestion (the max upstream is 5Mbps for the test CM).

As seen in the figure below, anomalies in app behavior are observed when network congestion exceeds 75-78% of channel capacity. Interestingly, at lower network congestion levels, the adaptive algorithms were trying to maximize bandwidth use and not let the TCP-based speed test achieve adequate throughput. Once the network gets congested, the algorithms tend to back off and this allows for better speed test results.

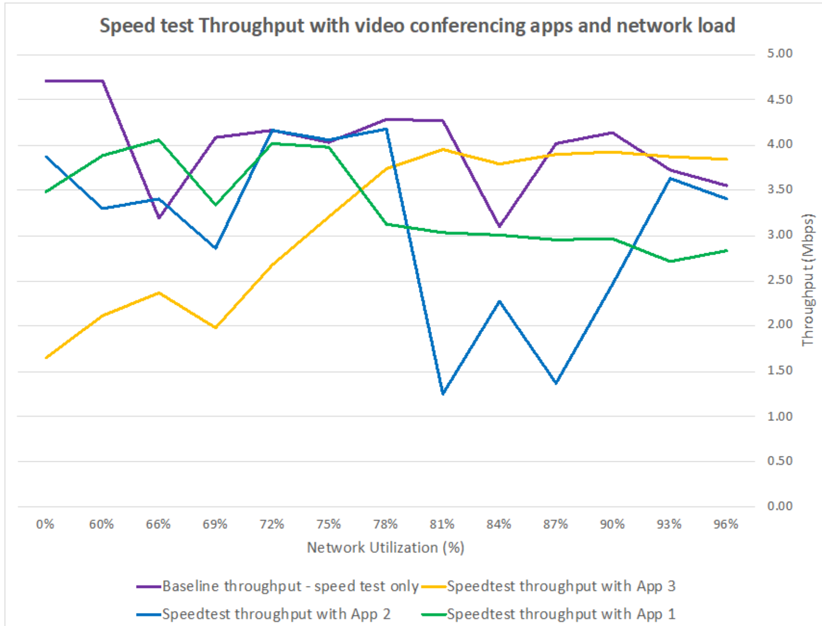

Figure 9 – Speed test along with multiple conference apps

Looking at the throughput dataset from an alternate perspective, Figure 10 shows how much throughput a single video conferencing app can achieve when a speed test type data is running along with other conferencing sessions behind the test CM. When varied with DOCSIS channel load, this gives an indication of how congested the channel can get before there is deterioration in the conferencing app QoE to the user.

Each App shows a distinct behavior against network congestion.

1. App 1 uses low bandwidth (600-800kbps) when the network is not congested. That is the typical baseline throughput for App 1. However, at around the 75% mark, the bandwidth used starts increasing to around 1Mbps – likely due to packet retransmissions needed to maintain the same video resolution and quality. No impact was seen to the actual video quality of the app until the congestion in the network exceeded 90%.
2. App 2 shows a much more classic adaptive video streaming behavior with a sawtooth-like variation with throughput. While it is not visible from the graph itself, looking at the raw statistics of throughput vs latency, it was evident that the two KPIs worked in tandem with each other. When app latency increased, the video dropped in resolution and quality, allowing less bandwidth to be used.
3. App 3 had the highest bandwidth utilization of the three, typical of the nature of the app as described in Section 4.1. However, it was the most yielding of bandwidth usage, once network congestion hit 75% or higher, dropping to as low as 600Kbps.

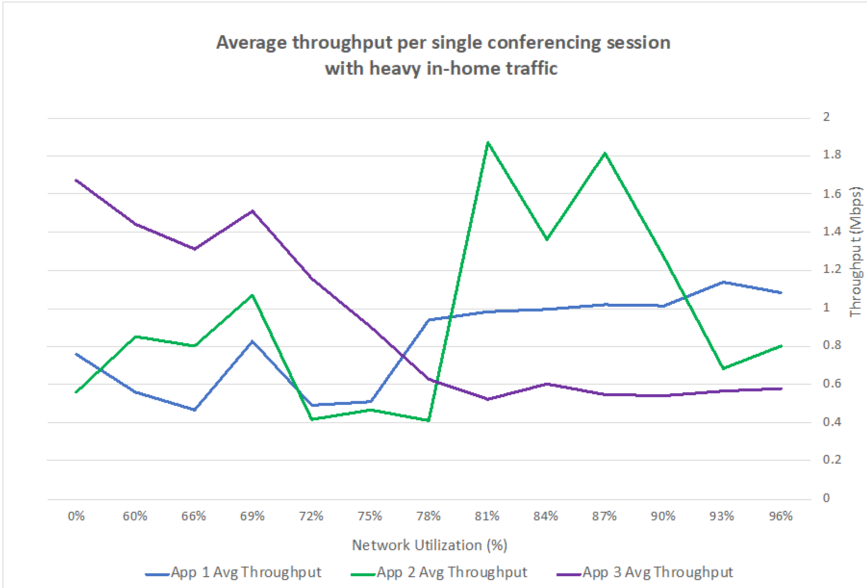

Figure 10 – Speed achieved by a single conference session under CM+channel load

It should be noted that while having a 1Mbps target for app video throughput is optimal, the app is still very functional at lower throughputs around 500Kbps however with reduced frame rates and/or video quality. All apps were very effective in working with congested networks and working for optimal performance. The tipping point of network congestion where apps tended to react rapidly to changing conditions was around 75-90% of channel capacity.

4.4. Video Conferencing Latency measurements

To measure latency of packets through this congested network, a low bit rate UDP stream was injected from a traffic generator behind the test CM and packet transmission times were measured over the DOCSIS upstream (to the Network Side Interface of the CMTS). Latencies of the data sets were computed and compared, both with and without conferencing apps running, as well as in the presence of the speed-test traffic behind the test CM.

Figure 11 and Figure 12 show mean latencies of conferencing apps under varying upstream DOCSIS channel congestion, with and without additional CM loading with speed test traffic. Both set of graphs have two concurrent conferencing sessions behind the test CM.

Looking at Figure 11, where the home network is not congested, App 3 is showing the highest latency impact. This is to be expected as App 3 had a tendency to take as much as 2.5Mbps per video session if the bandwidth is available. Hence 2 concurrent sessions with additional latency measurement traffic can lead to traffic bursts close to the 5Mbps Upstream Tmax of the service tier of the test CM. App 1 is not impacted much by latency and track very close to the baseline latency (when no conferencing apps are running). App 2 starts getting impacted by latency as network congestion hits 78% and the “sawtooth” behavior mimics the throughput graphs of the same application.

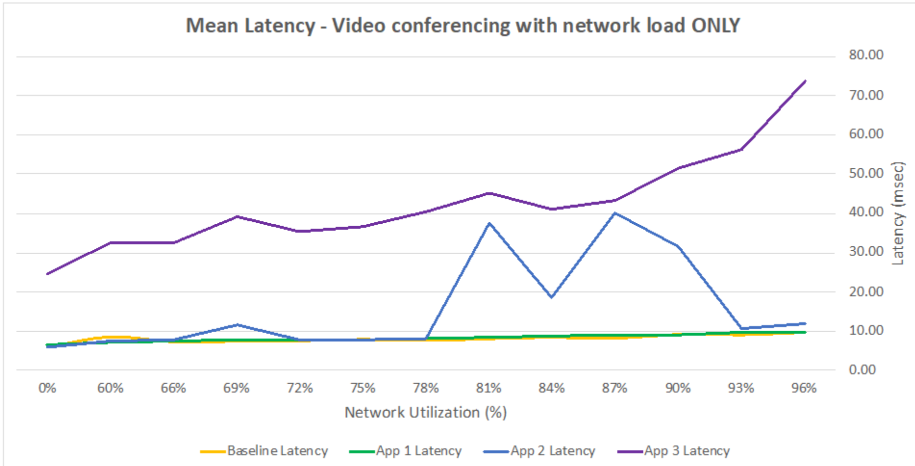

Figure 11 – Mean Latency for conferencing apps under channel loading ONLY

In Figure 12, we can now see that once the home network is loaded, latency impact on conferencing apps are much more prominent. Similar to throughput graphs, the inflexion point where apps tend to adapt their reduce their latency is around 78% of network congestion. App 3 being the most bandwidth hungry tends to be most impacted by heavy traffic loads, while App 2 shows a similar behavior to throughput variations with a sawtooth-like fluctuation in latency under heavy congestion.

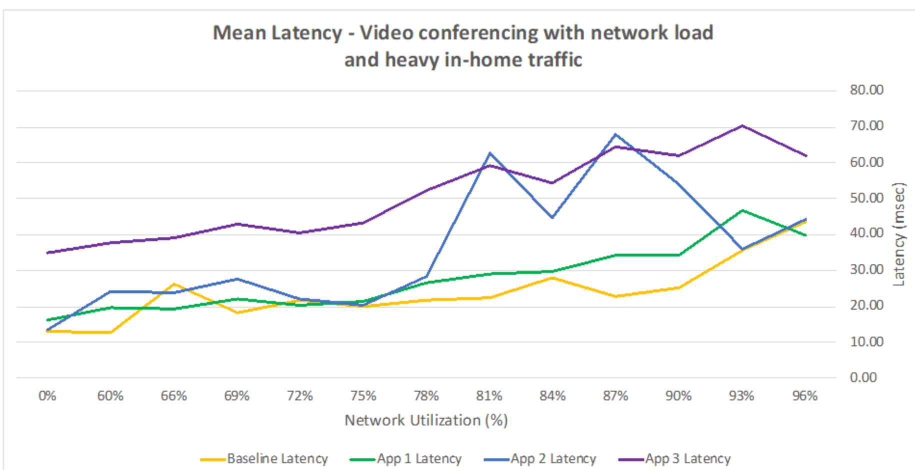

Figure 12 – Mean Latency for conferencing apps under CM+channel load

However, mean latencies only tell part of the story. Jitter or delay variation is key in how these conferencing apps are responding. Most of the packets should be within the latency threshold of 150ms, based on ITU-T G.114 specifications[3] for optimal performance of video conferencing style applications. Let us look at how 95th percentile of these latencies to see how the operating points look – to monitor if most of the video packets are maintained under the recommended 150ms threshold.

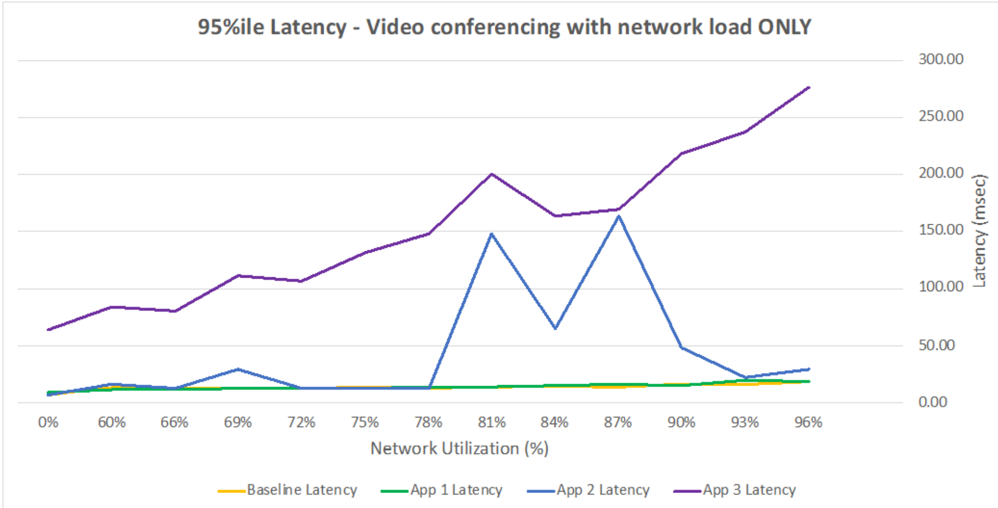

Figure 13 – 95th percentile latency for conferencing apps under channel loading only

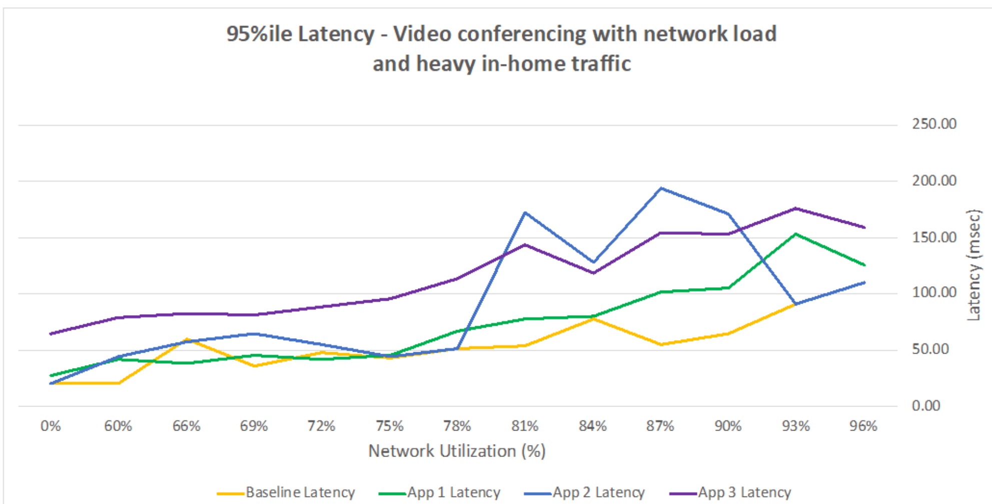

Figure 14 – 95th percentile latency for conferencing apps under CM+channel load

Based on Figure 13, with only network loading, App 1 latency is hardly impacted by network loading only. App 2 starts showing increasing latencies when network load exceeds 78% of channel capacity. However, its adaptive algorithm is able to keep latencies under 150ms by reducing video throughput. App 3 shows increasing latencies beyond acceptable threshold when network congestion is around the 78-80% mark.

Figure 14 shows the 95th percentile latency of conferencing apps when the TCP speed-test is also running behind the test CM. The apps are much more susceptible to latency impact with this increase load at home, but the inflection point seems to be still around the 78-80% of channel capacity. While increasing in-home traffic impacts App 2 more, App 3 is actually showing improved latencies when the congestion at home increased. One possible explanation for this could be the fact these applications are adaptive in nature. The presence of additional traffic at the CM and algorithms like Active Queue Management could result in better latency metrics.

5. Conclusion

Our experiments with various video conferencing applications indicate that these applications adapt well with varying network conditions and traffic patterns. To keep the desired throughput higher than 1Mbps for video quality and less than 150ms of latency for majority of the traffic, our study shows that a DOCSIS upstream channel utilization should not exceed 75-80% of its full operating capacity. Additionally, ensuring optimizations at the cable modem or home gateway level such as Active Queue Management, Buffer Control TLVs and TCP ack suppression will help reduce the impact of a highly congested channel and allow conferencing apps to adapt better to changing network conditions.

Our study focused on a service tier with fairly stringent bandwidth constraints. To mitigate the impacts of home congestion on these applications, it is desirable to move more subscribers to higher tiers. In the medium term, adding any additional available DOCSIS spectrum, leveraging DOCIS 3.1 technologies like OFDM/OFDMA and performing physical node splits for smaller service group sizes can help mitigate pressure from such bandwidth surges. Long term solutions include moving to mid or high splits, moving to fiber-deep and distributed access architectures including Full Duplex DOCSIS and Extended Spectrum DOCSIS solutions.

Abbreviations

AQM	Active Queue Management
ATDMA	Advanced Time Division Multiple Access
CM	Cable Modem
CMTS	Cable Modem Termination System
COVID-19	Coronavirus Disease 2019
CPE	Customer Premises Equipment
DOCSIS	Data Over Cable System Interface Specifications
DS	Downstream (In the context of DOCSIS transmission – to the home)
KPI	Key Performance Indictor
MSO	Multiple System Operators (DOCSIS cable operators in our context)
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PC	Personal Computer
QoE	Quality of Experience
SG	Service Group
SLA	Service Level Agreement
TCP	Transmission Control Protocol
TLV	Type-Length-Value
UDP	User Datagram Protocol
US	Upstream (In the context of DOCSIS transmission – from the home)
VPN	Virtual Private Network

Bibliography & References

1. Jeff Finkelstein, Tom Cloonan, Doug Jones, "MSO Learnings from COVID-19", SCTE, 2020 SCTE Cable Tec Expo, 2020.
2. John Ulm, Tom Cloonan, "Managing the Coronavirus Bandwidth Surge", SCTE, 2020 SCTE Cable Tec Expo, 2020
3. ITU-T Recommendation G.114, One way transmission time, 05/2003