2018

Employing H.264 Coarse and Medium Grain Scalable Video to Optimize Video Playback over Passive Optical Networks

David Moore
Georgia Southern University

Employing H.264 Coarse and Medium Grain Scalable Video to Optimize Video Playback over Passive Optical Networks

An Honors Thesis submitted in partial fulfillment of the requirements for Honors in the Department of Electrical & Computer Engineering

By
David L. Moore

Under the mentorship of Dr. Rami J. Haddad

ABSTRACT

In this work, we propose the use of Coarse Grain Scalable (CGS) and Medium Grain Scalable (MGS) H.264/AVC video to optimize video playback on passive optical networks (PONs) by investigating network performance metrics such as data delay, video delay, and video delay jitter. Video playback is improved by sequentially dropping layers of scalable video. Dropping just a single CGS enhancement layer results in improvements of up to 57% for both data and video delay. However, video delay jitter benefits the most with an improvement ranging from 47% to 87%. Surprisingly, dropping subsequent CGS enhancement layers does not significantly improve the PONs performance. In order to remedy this effect, our focus switched to employing the H.264/AVC MGS video standard. Though video traffic delay is the primary object of optimization in this work, the proposed algorithm’s impacts on other network performance metrics such as data traffic delay and video traffic delay variance (jitter) are analyzed as well. Video playback is improved by employing an adaptive scalable video layer dropping algorithm which drops a progressively larger number of scalable video layers as network utilization increases as measured by the moving average of the video packet delay. The influence of the algorithm's three parameters on its performance is investigated in detail, and the results of the optimized adaptive dropping algorithm are compared to baseline static dropping algorithm.

Thesis Mentor:________________________

Dr. Rami J. Haddad

Honors Director:_______________________

Dr. Steven Engel

April 2018
Electrical & Computer Engineering
University Honors Program
Georgia Southern University

CONTENTS

1. ACKNOWLEDGMENTS

I would like to extend my most sincere gratitude to my mentor, Dr. Rami J. Haddad, who has been my constant support and supervision throughout my college experience. It was Dr. Haddad who approached me during my freshman year and offered me a position in his lab so that I might work towards my honors thesis. I am extremely grateful that Dr. Haddad saw potential in me and gave me opportunities to grow so that I might accel. Dr. Haddad always took the time to hear out my concerns and followed through no matter how busy he was. It has been my distinct pleasure to work under you, Dr. Haddad.

I would also like to thank Dr. Steven Engel, Dr. Francis Desiderio, and Mrs. Erin Martin of the University Honors Program for their unwavering support. The University Honors Program has provided a constant challenge from the annual Experiential Learning Projects to this Honors Thesis. The University Honors Program has provided the push that has been my spark of inspiration.

Finally, I would like to thank my family and friends for their constant love and support and Georgia Southern University for being a nurturing environment in which to test my Eagle wings and soar.

2. INTRODUCTION

Cisco predicts that by the year 2021, 82% of all IP traffic will be video. As this is up substantially from the 73% of all IP traffic which was video in 2016, the proportion of IP traffic which is video is growing rapidly [1]. Compounding the problem, annual global IP traffic will reach 3.3 ZB per year by 2021 which is a sizeable increase from the annual global IP traffic of 1.2 ZB per year in 2016. Taken together, these two statistics clearly show a substantial increase in the sheer volume of video traffic. The size of standard video frames has doubled several times in the past with the most recent example being the transition from FHD (1920x1080) to 4K (3840x2160). The video frame size doubling adds to the pseudo-exponential growth of video-induced bandwidth demand. The existing network infrastructure will soon need to be expanded in order to accommodate this influx of new IP video traffic. In response to the growing prevalence of video traffic, our focus has turned to optimizing video traffic over Passive Optical Networks (PON), a financially feasible high-capacity access network infrastructure and the preferred next generation access networks to replace the copper access networks [2].

The passive optical network structure has an optical fiber that connects the Optical Line Terminal (OLT) at the service provider side with multiple Optical Network Units (ONUs) at the user side through a passive splitter. This star topology mandates the use of a centralized Medium Access Control (MAC) protocol to control the scheduling of the upstream traffic from the ONUs to the OLT. The centralized protocol operates using a cyclical polling process to service all the optical network units. However, the polling process can increase the upstream queueing delay which is critical when streaming delay-sensitive traffic such as video traffic.

To overcome this limitation, exhaustive grant sizing using queue size prediction techniques were proposed [2-5]. When these techniques are applied, the ONU provides predictions for future traffic to be considered in sizing and scheduling grants. These techniques were proven to be effective in reducing video queueing delay, and video queueing delay jitter but caused data queueing delay to increase.

Scalable Video Coding (SVC) is a subset of the H.264/AVC (Advanced Video Coding) standard which provides support for scalable video. The three main types of scalable video are temporal, spatial, and Signal-to-Noise Ratio (SNR) scalability. Temporal scalability allows the bit-rate of the content to be adjusted by changing the frame rate. Halving the frame rate is a simple yet effective way to reduce the bandwidth of a video stream. Spatial scalability allows the resolution of the frames to be changed. SNR scalability allows the quality of the video to be changed without changing either the frame rate or the resolution. Instead, the spectral content of the frame is broken down using the Discrete Cosine Transform (DCT) and the various spectral components are sorted into layers which allow each frame to be decoded with only partial spectral content.

In previous work, a scalable video transmission system has been implemented using H.264 with scalable video coding (SVC) and bit division multiplexing [6]. The resulting scalable video transmission had an improved spectral efficiency. Scalable video has also been used in a peer-to-peer (P2P) communication system with hierarchical network coding [7]. By using these two together, initial traffic delay and wasted bandwidth were both reduced. Baccichet et al. showed that SVC allows for a fairer distribution of video quality with respect to each user's connection speeds [8], and [9]

showed that SVC in conjunction with rate adaptation helped fulfill the delay constraints of P2P interactive group communication applications.

Additionally, SVC has been utilized on opportunistic networks [10, 11]. Video layers were transmitted with redundancy proportional to their importance in the process of decoding resulting in a viewing experience that started off low-quality but was steadily enhanced as transmission continued. SVC was used in conjunction with the path diversity of video distribution networks to adjust the streaming strategy to match bandwidth fluctuations [12]. Ghareeb et al. did similar work though they optimized the traffic using Quality of Experience (QoE) evaluations [13]. Scalable video was used to improve a BitTorrent-based P2P television system [14]. It was shown that SVC coupled with conservative chunk selection effectively prevented frame freeze and stalling.

SVC was used in a P2P video on demand system where it helped the system adapt to the heterogeneity of the connected devices [15]. A wireless IP network streaming system was developed to take advantage of the improved multicast and robustness of SVC [16]. Various applications of SVC encoding were pointed out particularly in the context of wireless networks. The work reported in [17] proposed a comprehensive Video MAC Protocol (VMP) utilizing SVC to accommodate triple-play services over fiber-wireless networks effectively. Even though significant advancements in the utilization of scalable video over networks were reported, the literature still lacks work addressing the utilization of SVC in PONs.

Multiple studies have used MGS video to adapt the necessary bit-rate to the bit-rate of the channel. Kim, Fujii, and Lee proposed using MGS to allow the reduction of transmitted packets in order to better serve a real-time streaming service [19]. They

combined this packet dropping with active modulation and found the optimal combination of these schemes using Lagrange Optimization by evaluating the number of received video packets received without error. Joo and Song also investigated the use of MGS video in a time-varying wireless environment, though their focus was on developing an effective IPTV system capable of achieving an optimal tradeoff between the number of subscribers receiving the MGS enhancement layers and their IPTV service quality [20]. Hannuksela et al. proposed an algorithm in which MGS enhancement layers are appended to the end of the base packet payload until the Maximum Transmission Unit (MTU) is reached [21]. This approach achieved a 0.3 to 0.5 dB gain in average luma Peak Signal-to-Noise Ratio as compared to the typical packetization where both algorithms generated the same number of packets. Sohn et al. developed an MPEG-21-based Adaptation Decision-Taking Engine (ADTE) to exploit the qualities of MGS video and better match the available bit-rate [22]. Mansour, Krishnamurthy, and Nasiopoulos created bit-rate and distortion prediction models for MGS video [23]. Taking into consideration both the mean absolute difference (MAD) of the motion prediction and the quantization parameter (QP), the models accurately predict the size and distortion of both base and enhancement layer MGS packets allowing the channel bit-rate to be matched accurately.

Studies have also focused on evaluating and improving the performance of the MGS standard itself. Gupta et al. performed a large-scale study of the rate-distortion (RD) and variability-distortion (VD) characteristics of CGS and MGS video [24]. One of their many results is that MGS decoded video frequently exceeds the RD performance of single layer SVC video for low to medium quality video, though the relative RD

performance drops significantly for high-quality video. Similarly in [25], a priority-based hierarchical extraction method is proposed which out-performs standard MGS extraction schemes. It is recommended that no more than five MGS fragments are used, and it is noted that the proposed priority-based hierarchical extraction method outperforms the standard algorithms.

In this work, an adaptive dropping algorithm is presented to improve the received quality of video over PONs. Though many of the previous works have used MGS scalability to adapt the size of a stream to fit the channel's capacity, this has never been done in the context of PONs while taking note of the resulting data packet delay, video packet delay, and video packet delay jitter. Therefore, we are proposing the utilization of scalable video traffic over PONs to trade-off quality of video with network performance. To the best of the authors' knowledge, optimizing scalable video traffic over PONs has never been attempted. This is an interesting problem given the centralized nature of the PON's MAC protocol.

3. PROPOSED SOLUTION

In this work, the effect of scalable video traffic on the passive optical network video queueing delay, data queueing delay, and video queueing delay jitter is investigated. To the best of our knowledge, no previous investigation of the effect of scalable video on PON networks was ever conducted except for by us [26-27]. It is claimed that dropping enhancement layers from the scalable video will reduce data and video queueing delays as well as reduce the queueing delay jitter. This will serve as the basis for developing optimal algorithms for utilizing scalable video over PONs.

In the first part of this work, the effects of dropping entire coarse grain scalable (CGS) layers (referred to as static dropping) are investigated. Since the top layer of the CGS video is much larger than any of the enhancement layers below it, dropping the top layer had a much more significant impact on the data traffic delay, video packet delay, and jitter than dropping any of the other enhancement layers. This rendered CGS video unsuitable for the sort of optimized layer dropping that had been planned as evidenced by the CGS frame statistics in Table 1. However, a replacement was found in MGS video. The layers of the utilized MGS video traces have similar average packet sizes (as seen in Table 2), so the impact of dropping the top enhancement layer is comparable to the impact of dropping the bottom enhancement layer.

Layer	Average Frame Size (Bytes)	CoV of Frame Sizes	Mean Frame Bit Rate (bits/sec)
0	153.96	1.136	36951.5
1	428.59	1.299	102862.5
2	1137.31	1.427	272955.3
3	4025.22	1.153	966053.7

Table 1 - Frame Statistics for Each Layer of the 4-Layer CGS Encoding of Gandhi

Layer	Average Frame Size (Bytes)	CoV of Frame Sizes	Mean Frame Bit Rate (bits/sec)
0	478.47	1.812	114833
1	581.92	0.701	139661
2	333.63	1.464	80071
3	228.44	1.438	54825
4	274.84	1.499	65961
5	207.41	1.648	49778
6	131.04	1.451	31489

Table 2 - Frame Statistics for Each Layer of the 7-Layer MGS Encoding of Gandhi

MGS scalable video made optimized enhancement layer dropping possible. In order to better gauge the effectiveness of the optimized dropping, baseline simulations were run by dropping a static number of enhancement layers throughout the simulation. Results were obtained for each video and each set of enhancement layers. Simulations were run with just the base layer (layer one), with all of the layers (layers one through seven), and with all of the layers in between.

In contrast to static dropping which always drops a specified number of enhancement layers, adaptive dropping works by dropping a varying number of

enhancement layers depending on how congested the network is. Network congestion is measured by averaging the five most recent video packet delay values. This measured average value of video packet delay is then compared to a set of predetermined benchmarks. If the video packet delay is below the threshold denoted MinB (Minimum Bound), then no enhancement layers are dropped. Similarly, if the video packet delay is above the threshold denoted MaxB (Maximum Bound), then all of the enhancement layers are dropped. The thresholds for dropping intermediate numbers of enhancement layers (1 layer, 2 layers, etc.) are calculated by linearly interpolating the MaxB and MinB values. Figure 1 illustrates the adaptive dropping algorithm.

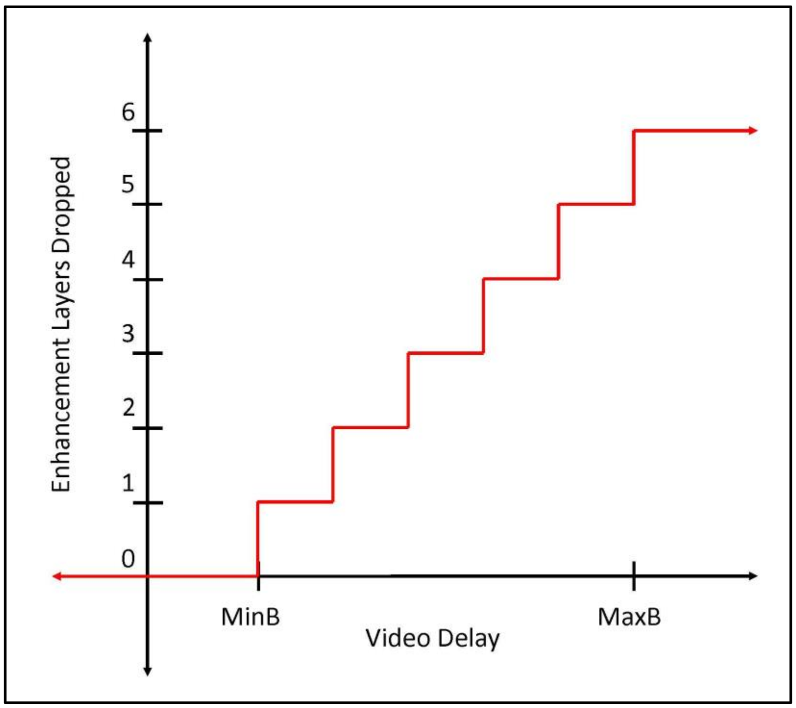

Figure 1 - Algorithm for Adaptive MGS Layer Dropping

4. SIMULATION MODEL

A set of simulation experiments were conducted, using a CSIM-based EPON discrete event simulator, to compare the difference in performance measures (data queueing delay, video queueing delay, and queueing delay jitter) for the level of video scalability. The PON has a transmission rate of 1 Gb/s which services its 32 Optical Network Units (ONUs), and each ONU has its data and video queues. A Limited grant size shared between the data and video queue was used to allocate the video and data grants. Priority Serviced First (PSF) scheduling policy was used to provide differentiated service to video queues over data queues [5].

To ensure rapid convergence of data traffic, a Poisson traffic source was used to generate the packets for the data queues. The following quad-modal data packet size distribution was used: 60% 64 bytes, 4% 300 bytes, 11% 580 bytes, and 25% 1518 bytes. For video traffic, video traces from the Arizona State University video trace library [18] were used to generate the packets for the video queue. Experiments were conducted using the following three H.264/SVC CGS video traces: Gandhi, Star Wars IV, and Terminator. All three videos utilized the G16B15 Group of Pictures (GoP) format which is IBBBBBBBBBBBBBBB and had a delta quantization parameter (DQP) of 10. The videos were formatted using the Common Intermediate Format (CIF) with a resolution of 352x288, and their frame rate was 30 *fps*. The video frames were broken down into SNR layers using H.264 CGS. Table 3 highlights the statistics of the video traces used in this experiment.

Video	Layer	Compression Ratio	Average Frame Size (Bytes)	Mean Frame Size (kbps)	Peak Frame Size (kbps)
Gandhi	0	987.65	153.96	36.95	361.9
Gandhi	1	489.69	428.59	102.86	1345.0
Gandhi	2	200.55	1137.3	272.95	3509.0
Gandhi	3	57.49	4025.2	966.05	8564.2
Star Wars IV	0	1331.47	114.2	27.4	409.0
Star Wars IV	1	489.69	310.53	74.52	828.2
Star Wars IV	2	200.55	758.24	181.97	2087.8
Star Wars IV	3	57.49	2644.97	634.79	6225.8
Terminator	0	860.57	176.7	42.4	298.8
Terminator	1	295.54	516.26	123.90	1416.5
Terminator	2	104.26	1458.56	350.05	4003.0
Terminator	3	27.4	5550.7	1332.16	10787.0

Table 3 - Frame Statistics for Each Layer of Several CGS Encoded Videos

The topology of PONs has a substantial impact on the effectiveness of dropping packets. PONs employ a point-to-multipoint topology with the internet service provider connecting to an Optical Line Terminal (OLT) at one end and the consumers or small groups of consumers connecting to one of *n* Optical Network Units (ONU) at the other end. The OLT is connected to the *n* ONUs by a 1:*n* splitter as seen in Figure 2. While downstream transmissions are sent to all of the ONUs, the ONUs must share upstream link time via a decentralized MAC protocol.

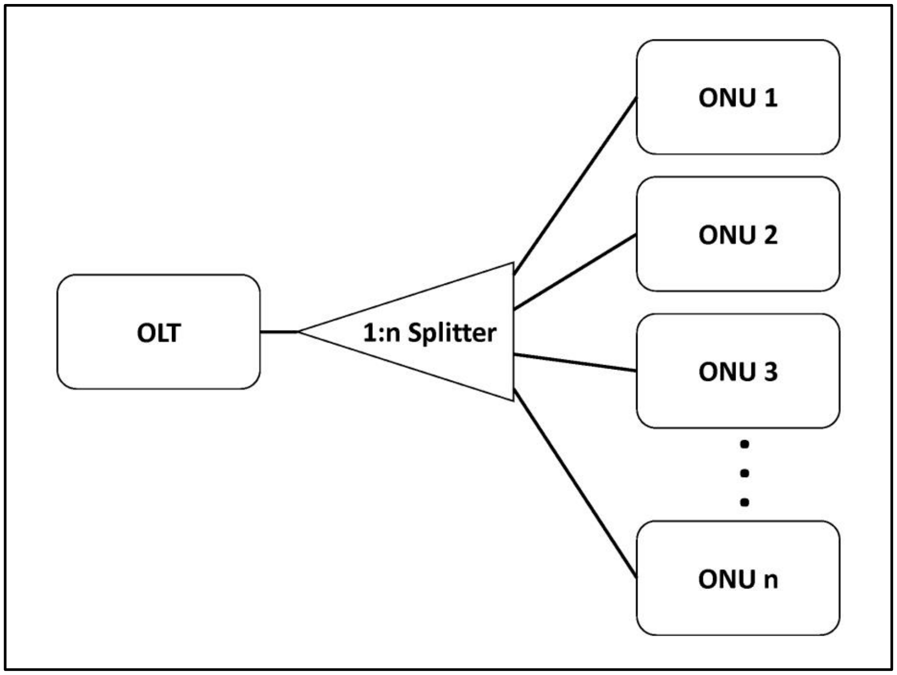

Figure 2 - PON Topology

In order to model the complex interactions of a PONs link layer, a CSIM-based EPON discrete event simulator was used. Video trace files obtained from Arizona State University's video trace library were used to model the effects of sending video packets through a PON without actually encoding the video and sending it through a physical PON [18]. The video traces files that were used were Citizen Kane, Gandhi, Indiana Jones, Silence of the Lambs, Star Wars IV, and The Terminator. Each one of these videos was encoded using MGS with a 1,2,2,3,4,4 split and explicit QP cascading. Additionally, they were all CIF (352x288) format with a base layer QP of 35 and an enhancement layer QP of 25.

5. DISCUSSION OF RESULTS

To validate our initial claim, we conducted a set of experiments to quantify the effect of scaling the video by dropping 1, 2, or 3 enhancement layers compared to unscaled video. Through these experiments, we first quantified the reduction in the average data queueing delay. Figures 3(a,b,c) illustrate the average data packet delay for the 3 videos (Terminator, Gandhi, Star Wars IV) using different video scalability levels. These figures graph the queueing delay of data packets (which excludes video packets) vs the utilization of the network as measured in Gigabits per second. Since the simulator models a 1 Gigabit per second PON, the x-axis values of 0.1 and 0.2 all the way up to 0.9 represent filling 10% and 20% all the way up to 90% of the PON with data traffic. As the network becomes more and more congested with data packets, there is less bandwidth for each packet, and delay increases as reflected in Figures 3(a,b,c).

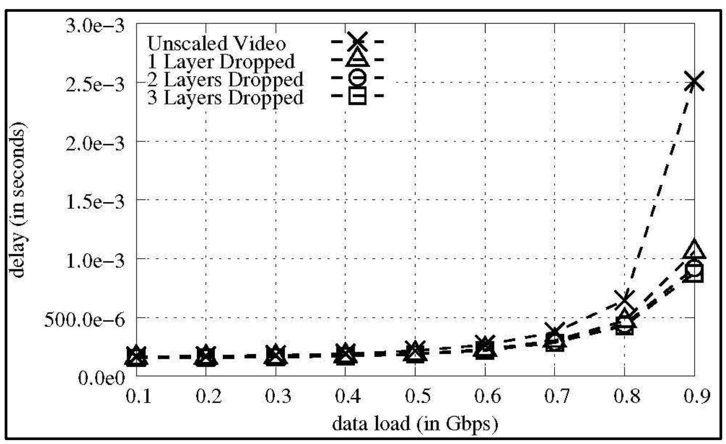

Figure 3a - Average Delay of Data Packets in Terminator

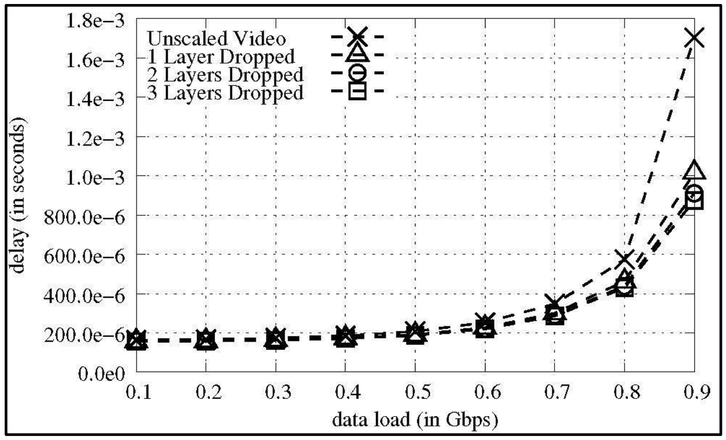

Figure 3b - Average Delay of Data Packets in Gandhi

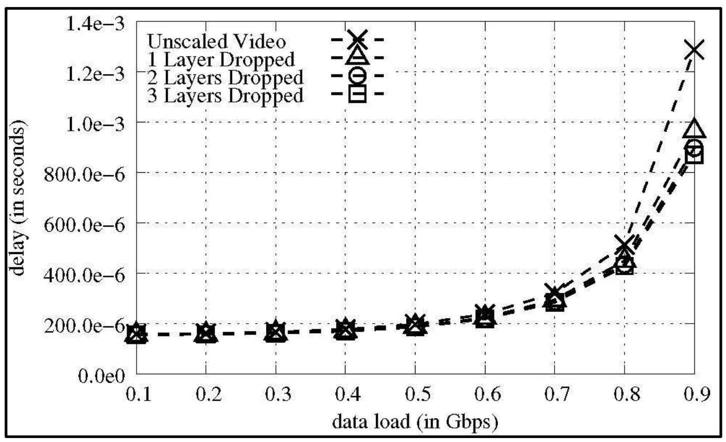

Figure 3c - Average Delay of Data Packets in Star Wars IV

From these figures, it is observed that dropping a single enhancement layer provides a significant reduction in the data packet queueing delay compared to using unscaled video mainly when the PON utilization is relatively high. The reduction in data packet delay is represented visually by the gap between the line with Xs and triangles. However, when more layers are dropped the reduction in the data packet queueing delay is very marginal. This is represented by the space between the triangles, circles, and squares. This will limit the need to drop subsequent enhancement layers after dropping the first one. As an example, for the Terminator video with a data traffic load of 0.9

Gbps, the average data queueing delay, when unscaled video is used, is 2.51 ms. However, when one enhancement layer is dropped, the average data queueing delay drops to 1.05 ms, which is a 57.8% reduction. When two enhancement layers are dropped, the average data packet queueing delay for the same load configuration drops to 0.92 ms, which is a 63.4% reduction. However, this is only 5.6% difference compared to dropping one enhancement layer.

As for video packet queueing delay, Figures 4(a,b,c) illustrate the average video packet delay for the 3 video traces (Terminator, Gandhi, Star Wars IV) using different video scalability levels. The x-axis represents the utilization of the network, and the y-axis represents video delay. As layers are dropped, the lines shift down slightly which indicates a decrease in delay independent of the utilization of the network.

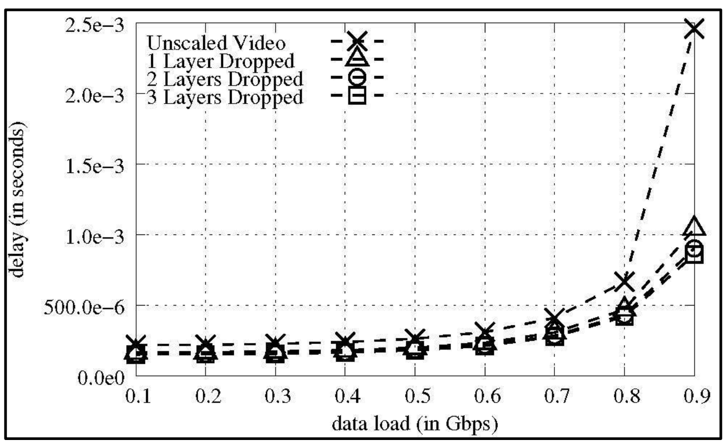

Figure 4a - Average Delay of Video Packets in Terminator

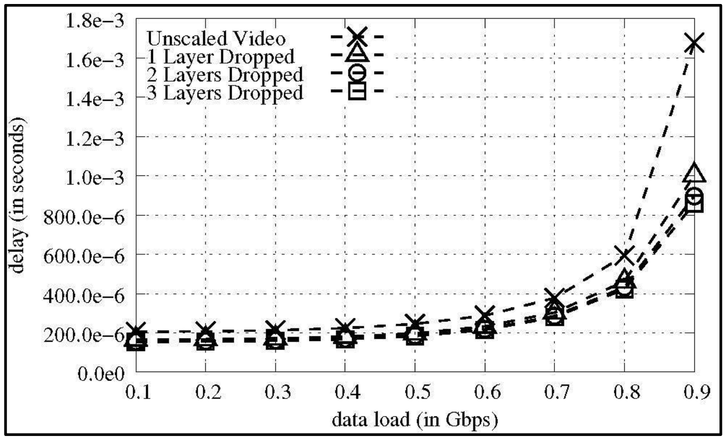

Figure 4b - Average Delay of Video Packets in Gandhi

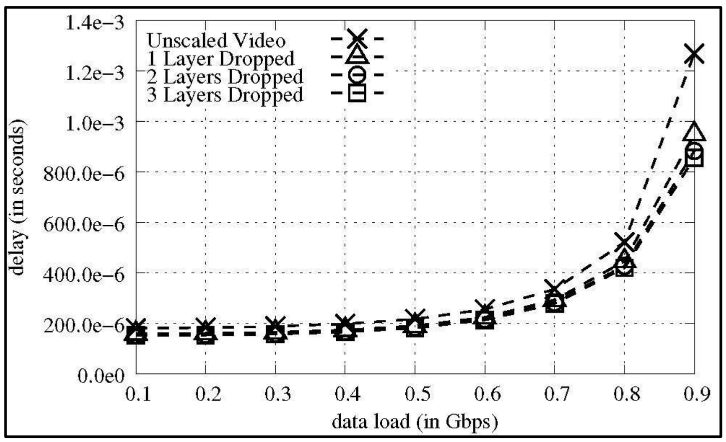

Figure 4c - Average Delay of Video Packets in Star Wars IV

The video packet queueing delay exhibits a similar effect as the data packet queueing delay in which dropping one enhancement layer significantly reduced the average video packet delay compared to using unscaled video. However, when more layers are dropped the reduction in the average video packet queueing delay is minimal. From the video queueing delay prospective, this also limits the need to drop enhancement layers after the first one. As an example, for the Terminator video with a data traffic load of 0.9 Gbps, the average video queueing delay, when unscaled video is used, is 2.46 ms. However, when one enhancement layer is dropped, the average video queueing delay

drops to 1.04 ms, which is a 57.6% reduction. When two enhancement layers are dropped, the average video packet queueing delay for the same load configuration drops to 0.9 ms, which is a 63.3% reduction. However, this is only 5.7% difference compared to dropping one enhancement layer. Even at low-network utilization, dropping a single enhancement layer had a noticeable effect on reducing the video packet queueing delay. This was not the case for data packet queueing delay. As an example, for the Terminator video with a data traffic load of 0.1 Gbps, the average video queueing delay, when unscaled video is used, is 220 μs. When one enhancement layer is dropped, the average video queueing delay drops to 168.8 μs, which is a 23.3% reduction.

Though video packet queueing delay certainly has a negative impact on viewing quality, it does not tell the whole story. Indeed, video packet queueing delay variance (or jitter) also has a significant role in determining viewing quality. Therefore, Figures 5(a,b,c) depict the average video packet delay jitter for the 3 video traces (Terminator, Gandhi, Star Wars IV) using different video scalability levels. The x-axis is still network utilization though the y-axis has changed to video packet jitter which is measured in seconds squared rather than just seconds.

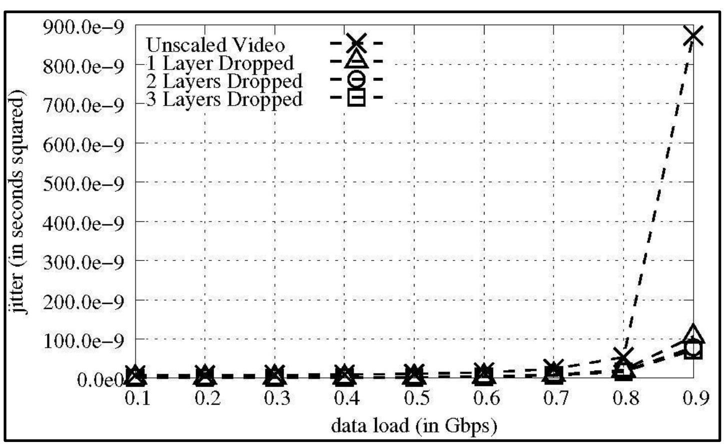

Figure 5a - Jitter of Video Packets in Terminator

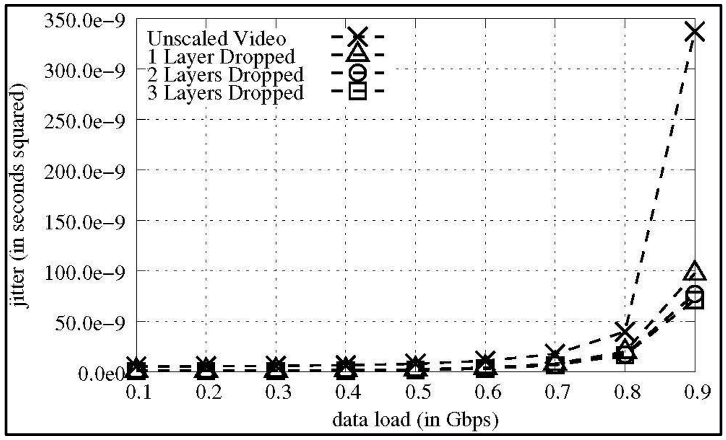

Figure 5b - Jitter of Video Packets in Gandhi

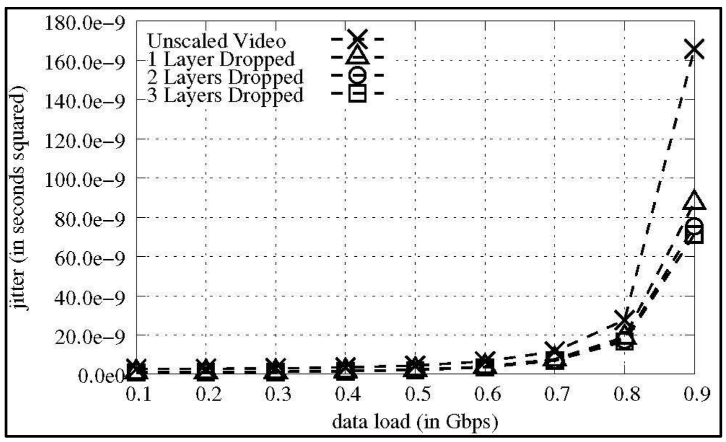

Figure 5c - Jitter of Video Packets in Star Wars IV

From these figures, it is observed that most tangible advantage is the reduction of video delay jitter, which plays a major role in optimizing the video start-up delay, the number of playback freezes (buffering), and the video packet loss rate. Therefore, reducing the video delay jitter can significantly improve the video playback by sacrificing video quality. As depicted in Figures 5(a,b,c), dropping one enhancement layer reduced the average video delay jitter drastically compared to using unscaled video. However, this reduction did not scale linearly as more enhancement layers were dropped.

As an example, for the Terminator video with a data traffic load of 0.9 Gbps, the average video delay jitter, when unscaled video is used, is 873.7 ns^2. However, when one

enhancement layer is dropped, the average data queueing delay drops to 106.2 ns^2 which is an 87.8% reduction. When two enhancement layers are dropped, the average video delay jitter for the same load configuration drops to 78.75 ns^2 which is a 91% reduction. However, this is only around 3% difference compared to dropping one enhancement layer. Moreover, the percentage of reduction in video delay jitter is relatively sizable under any network utilization. For example, the Terminator video with a data traffic load of 0.1 Gbps, the average video delay jitter, when unscaled video is used, is 7.39 ns^2. However, when one enhancement layer is dropped, the average video delay jitter drops to 1.44 ns^2, which is an 80.4% reduction.

Table 4(a,b,c) shows a summary of the reduction in the data delay, video delay, and video delay jitter for all three video trace files at 0.3 Gbps, 0.6 Gbps, and 0.9 Gbps data load traffic.

Data Load (Gbps)	Terminator (%)	Gandhi (%)	Star Wars IV (%)
0.3	7.72	5.25	2.78
0.6	14.67	10.33	6.13
0.9	57.84	40.35	25.04

Table 4a - Percentage Reduction in Data Delay for One Layer Scaling

Data Load (Gbps)	Terminator (%)	Gandhi (%)	Star Wars IV (%)
0.3	23.81	19.38	12.17
0.6	24.2	19.25	12.31
0.9	57.56	40.26	25.13

Table 4b - Percentage Reduction in Video Delay for One Layer Scaling

Data Load (Gbps)	Terminator (%)	Gandhi (%)	Star Wars IV (%)
0.3	78.75	73.75	56.51
0.6	69.09	60.87	40.71
0.9	87.84	71.02	47.28

Table 4c - Percentage Reduction in Video Delay Jitter for One Layer Scaling

As Tables 4(a,b,c) show, the benefits of dropping a single layer are quite substantial. Notably, all of the reduction values are positive. That means that the network performance impact of dropping a single layer was positive for all three of the metrics analyzed though it undoubtedly decreased the received quality of the video. While this dramatic reduction is beneficial for the network, it is not very interesting from an optimization perspective. Network engineers should be able to drop varying number of enhancement layers depending on the congestion of the network. However, CGS video is unsuitable for this sort of optimized dropping since the only layer that has a substantial impact on network performance is the first one. Thus, network engineers would only have two options: send the whole video or drop the first layer. This small number of options us to pursue a more flexible form of scalable video.

For the second part of this work, we performed a second set of static dropping performance evaluations. This section marks the transition to using seven-layer MGS videos rather than three-layer CGS videos. For all subsequent simulations, testing was done with the seven-layer MGS videos Gandhi, The Terminator, and The Silence of the Lambs. As the closely spaced baseline results in Figures 6, 7, and 8 show, the seven-layer MGS video provided a much better testbed for our adaptive dropping algorithm. In order to establish a set of baseline results so that the adaptive dropping has context, a set of

simulations were run each with a different number of enhancement layers. The data packet delay, video packet delay, and video packet delay jitter were calculated for each of these simulations, and the results are displayed in Figures 6, 7, and 8. The data delay values for each of the Gandhi simulations start off very close together and spread apart as the network utilization increases. The same is seen for both video packet delay and video packet delay jitter. The simulation with the base layer (BL) and enhancement layers (EL) one through six (1-6) has the highest data packet delay, video packet delay, and video packet delay jitter. Conversely, the simulation with only the base layer which had dropped all six enhancement layers throughout the simulation has the lowest data packet delay, video packet delay, and video packet delay jitter.

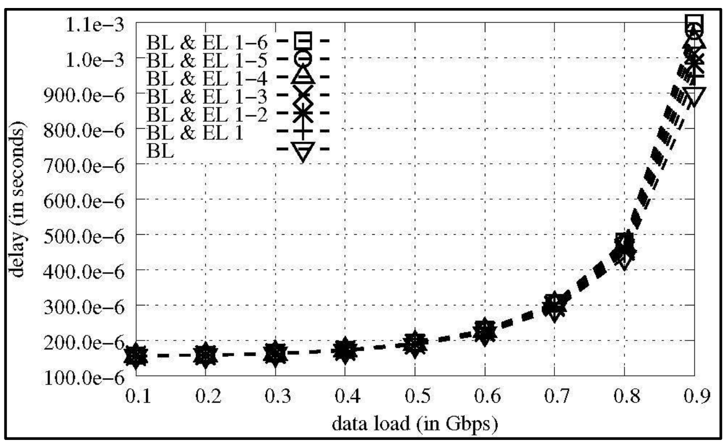

Figure 6 - Data Packet Delay using Static Dropping for Gandhi

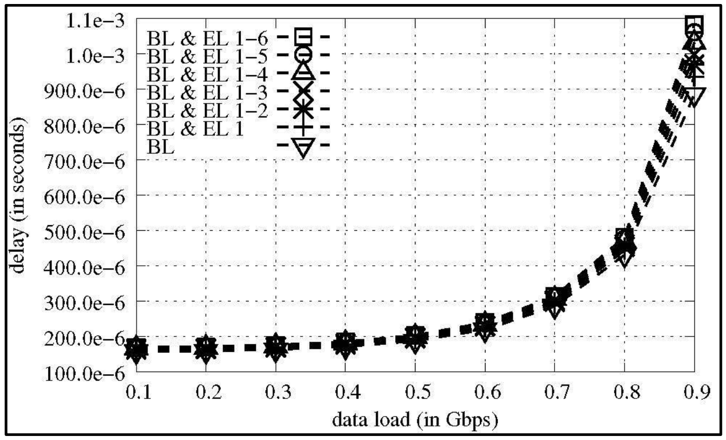

Figure 7 - Video Packet Delay using Static Dropping for Gandhi

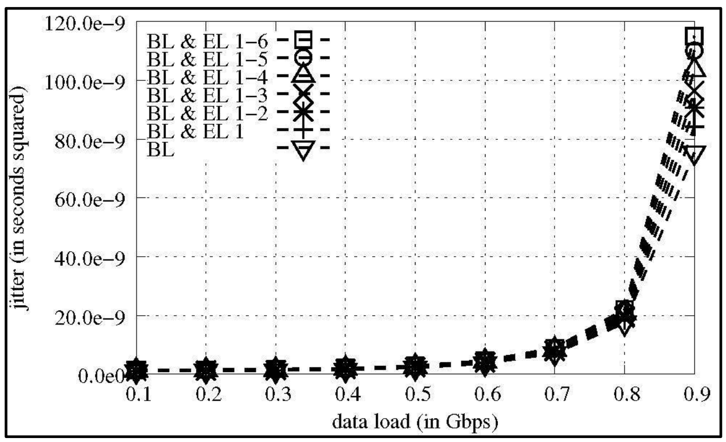

Figure 8 - Video Packet Delay Jitter using Static Dropping for Gandhi

Along with the baseline static dropping simulations, adaptive dropping simulations were also run for Gandhi. With the lower bound fixed at 0.0001 and the number of moving average values set to 5, the upper bound was varied from 0.0008 to 0.0016 in increments of 0.0001 for a total of nine simulations. Increasing the algorithm's upper bound while leaving the lower bound constant increased the total length of the adaptive dropping window and increased the step size which is the amount of delay required to transition from dropping *n* to *n+1* enhancement layers. Figures 9, 10, and 11 show the data packet delay, video packet delay, and video packet delay jitter of each

simulation respectively.

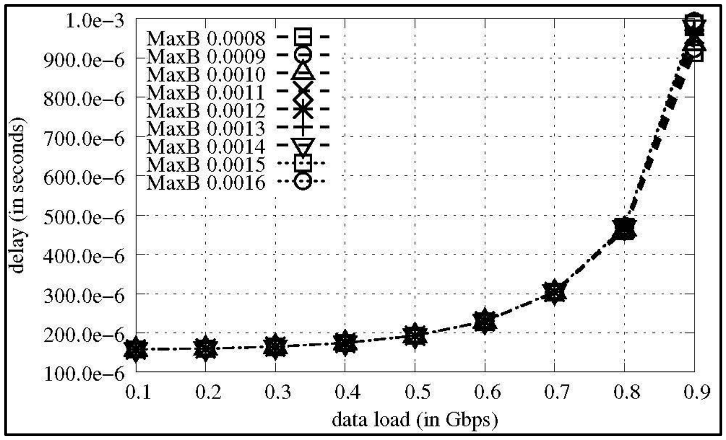

Figure 9 - Data Packet Delay using Adaptive Dropping for Gandhi

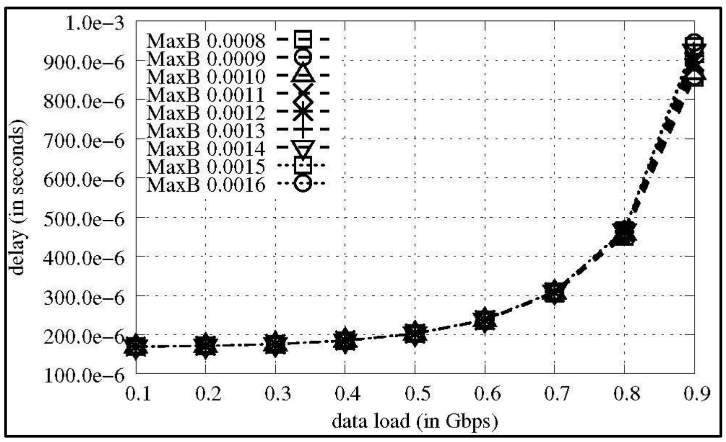

Figure 10 - Video Packet Delay using Adaptive Dropping for Gandhi

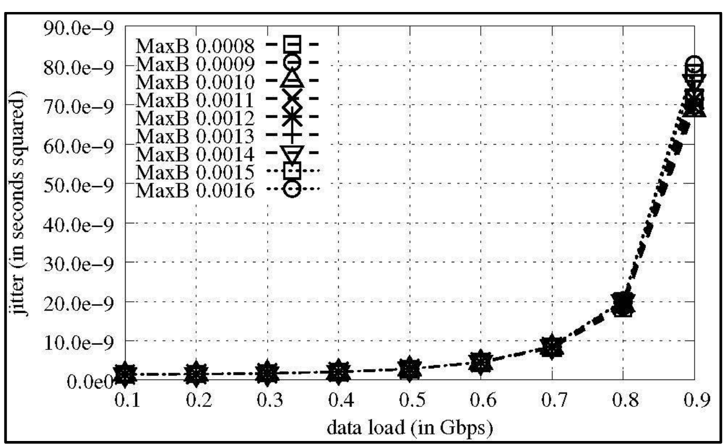

Figure 11 - Video Packet Delay Jitter using Adaptive Dropping for Gandhi

As the simulation with upper or maximum bound (MaxB) 0.0016 has the lowest data packet delay, video packet delay, and video packet delay jitter. The simulation with maximum bound 0.0008 has the highest data packet delay, video packet delay, and video packet delay jitter. Simulations with these two parameters are further investigated as a representative set of the simulations. In Figures 12(a,b,c), 13(a,b,c), and 14(a,b,c), the adaptive dropping simulations with maximum bounds 0.0008 and 0.0016 are compared to the static dropping simulations where all of the enhancement layers are transmitted and all of the enhancement layers are dropped. It is interesting to note that the adaptive dropping simulation with maximum bound 0.0008 actually outperforms dropping all of the enhancement layers with respect to both video packet delay and video packet delay jitter and comes close with respect to data packet delay. However, this reduction in data packet delay, video packet delay, and video packet delay jitter is achieved by dropping packets, so these benefits must be weighed against the number of packets dropped.

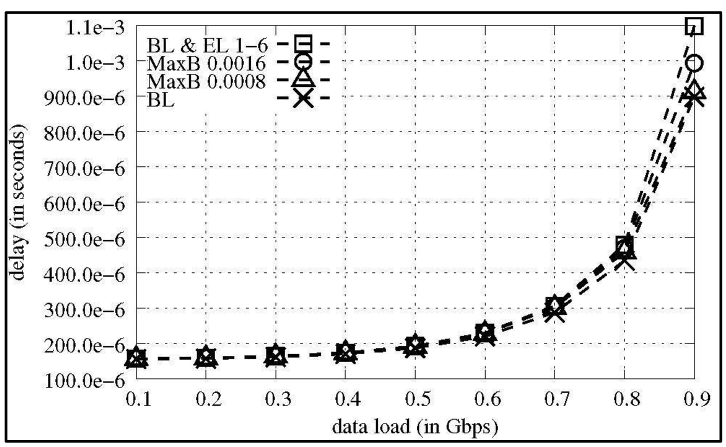

Figure 12a - Data Packet Delay Dropping Comparison for Gandhi

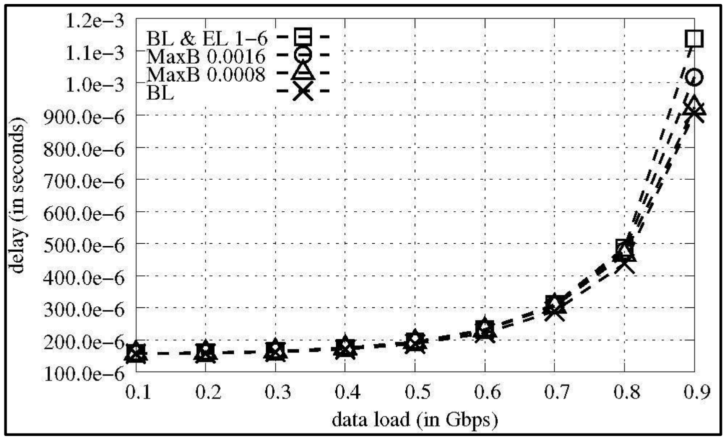

Figure 12b - Data Packet Delay Dropping Comparison for Indiana Jones

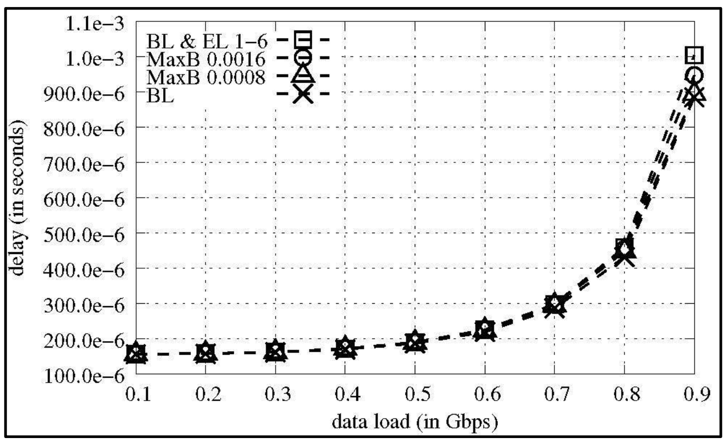

Figure 12c - Data Packet Delay Dropping Comparison for Silence of the Lambs

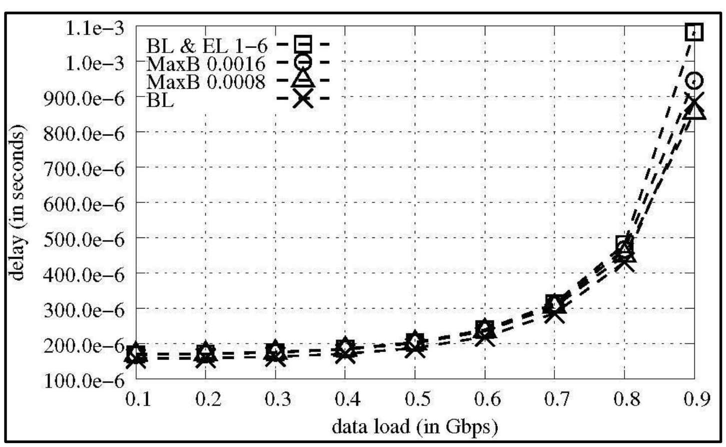

Figure 13a - Video Packet Delay Dropping Comparison for Gandhi

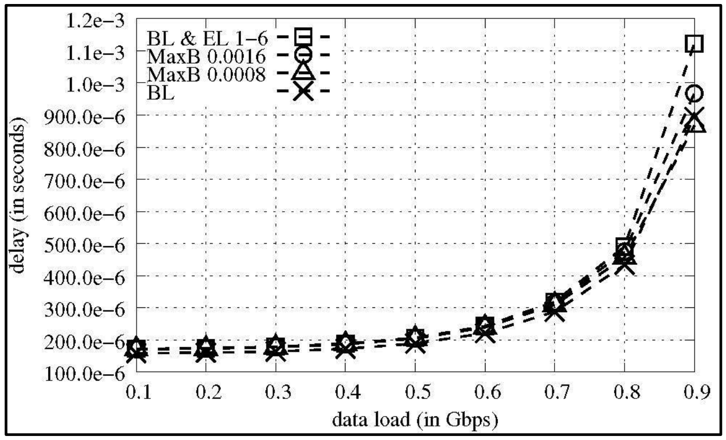

Figure 13b - Video Packet Delay Dropping Comparison for Indiana Jones

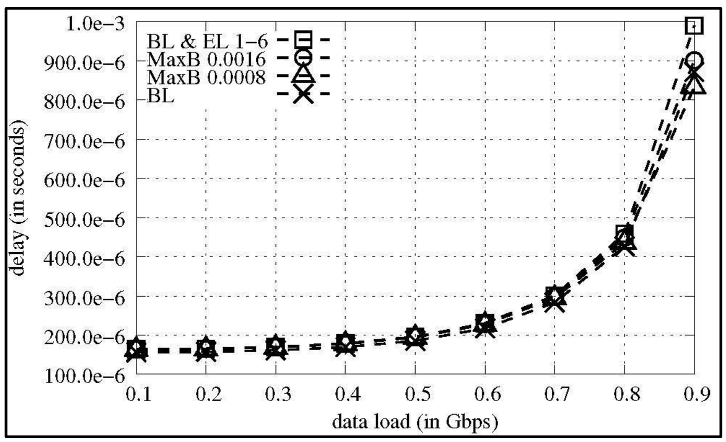

Figure 13c - Video Packet Delay Dropping Comparison for Silence of the Lambs

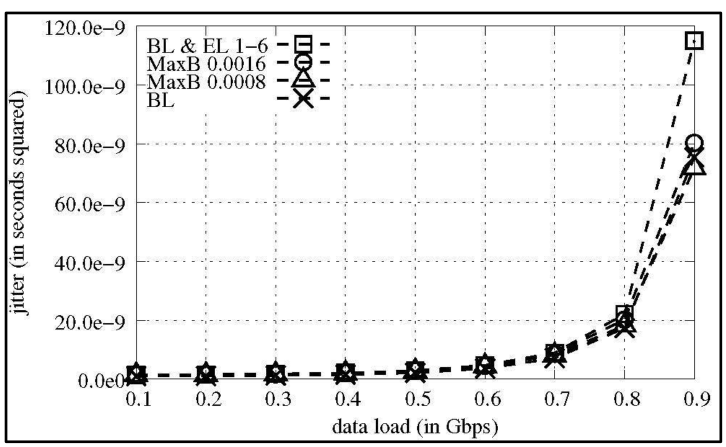

Figure 14a - Video Packet Delay Jitter Dropping Comparison for Gandhi

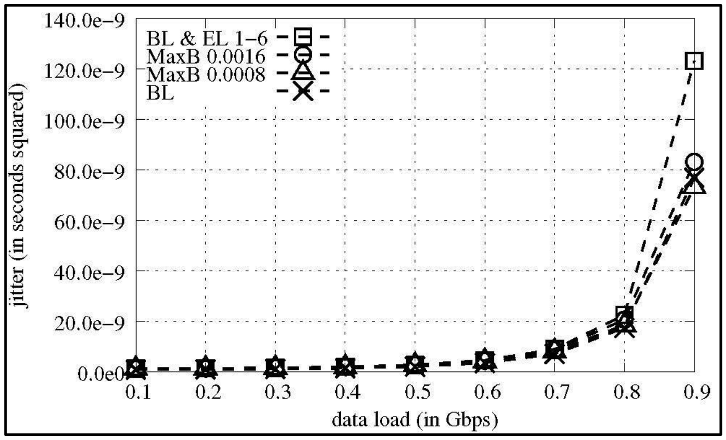

Figure 14b - Video Packet Delay Jitter Dropping Comparison for Indiana Jones

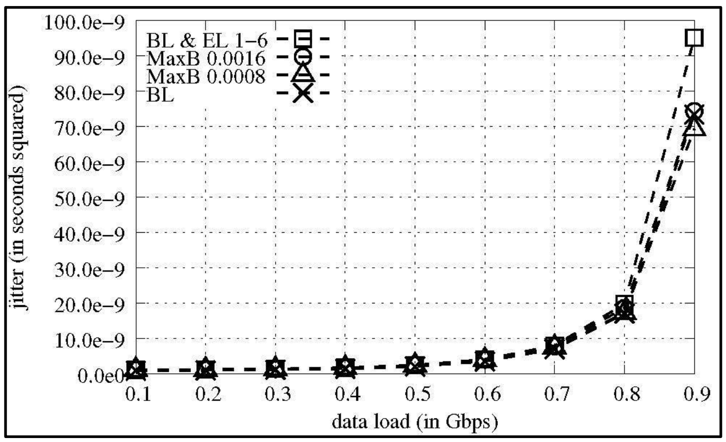

Figure 14c - Video Packet Delay Jitter Dropping Comparison for Silence of the Lambs

Figures 15(a,b,c) show the total amount of data dropped from the whole video which includes the base layer and all of the enhancement layers. Transmitting only the base layer causes the most data to be dropped because all of the enhancement layers are dropped. As the enhancement layers are added back in one at a time, the amount of data dropped decreases until all of them are added back in at which point no packets are dropped. Since the adaptive dropping algorithm drops packets based on the moving average of the video delay, very few packets are dropped when the network utilization is low. Progressively more packets are dropped as network utilization increases. The

adaptive dropping algorithm drops the most when network utilization is at 0.9 Gbps where the dropping with upper bound 0.0008 drops almost all of the enhancement layers and the dropping with upper bound 0.0016 drops the equivalent of between three and four enhancement layers. This inconsistent dropping allows the adaptive dropping to achieve high-quality video when the PON is only lightly utilized and low data packet delay and video packet delay when the PON is heavily utilized. Interestingly, using the adaptive dropping with 0.0008 has lower data packet delay and video packet delay than transmitting just the base layer despite dropping less data.

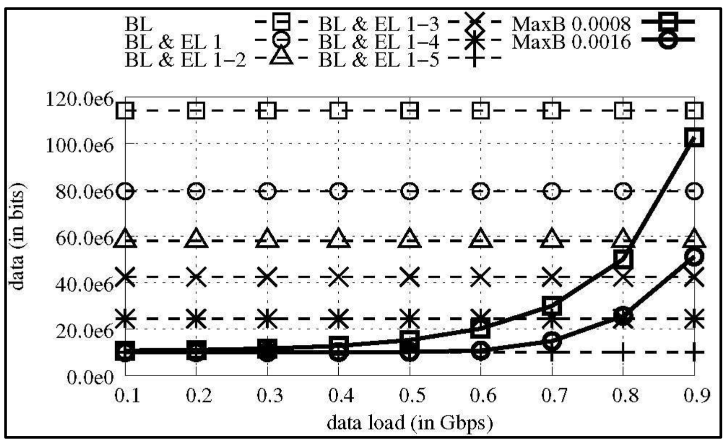

Figure 15a - Data Dropped per Load Level for Gandhi

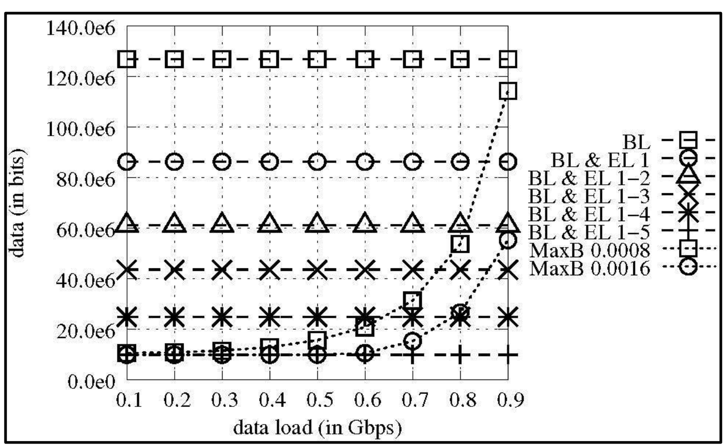

Figure 15b - Data Dropped per Load Level for Indiana Jones

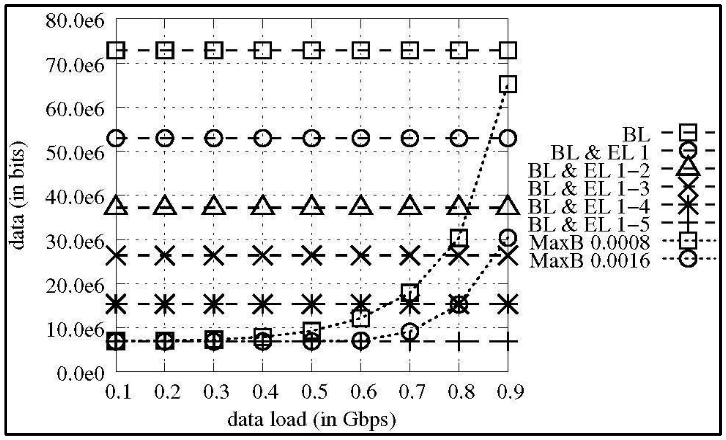

Figure 15c - Data Dropped per Load Level for Silence of the Lambs

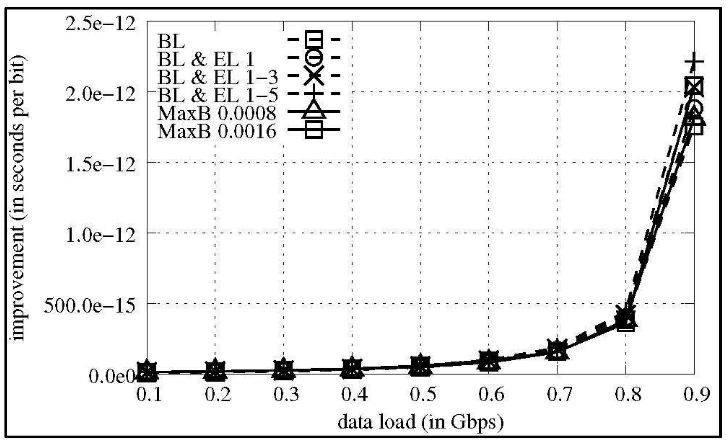

Figure 16a - Data Packet Delay Dropping Benefit for Gandhi

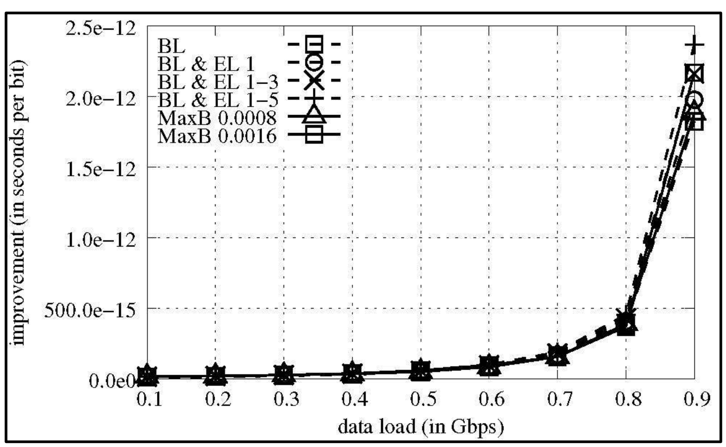

Figure 16b - Data Packet Delay Dropping Benefit for Indiana Jones

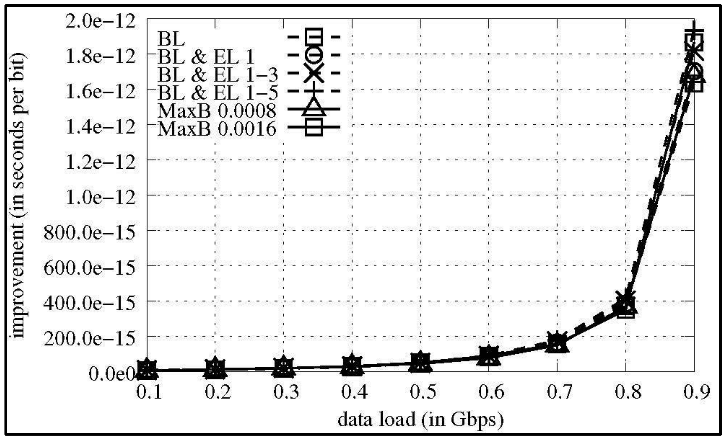

Figure 16c - Data Packet Delay Dropping Benefit for Silence of the Lambs

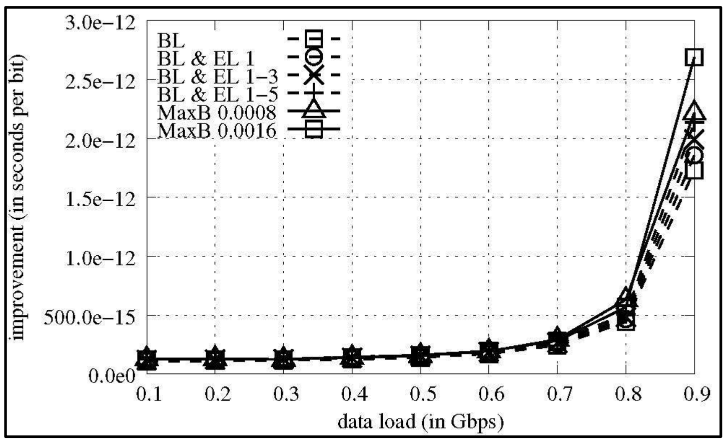

Figure 17a - Video Packet Delay Dropping Benefit for Gandhi

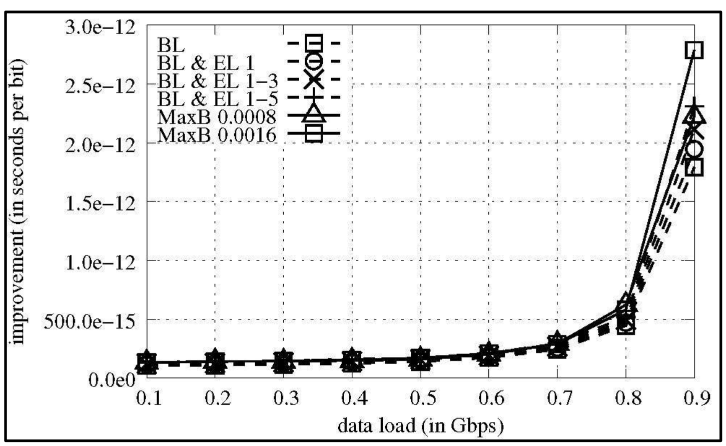

Figure 17b - Video Packet Delay Dropping Benefit for Indiana Jones

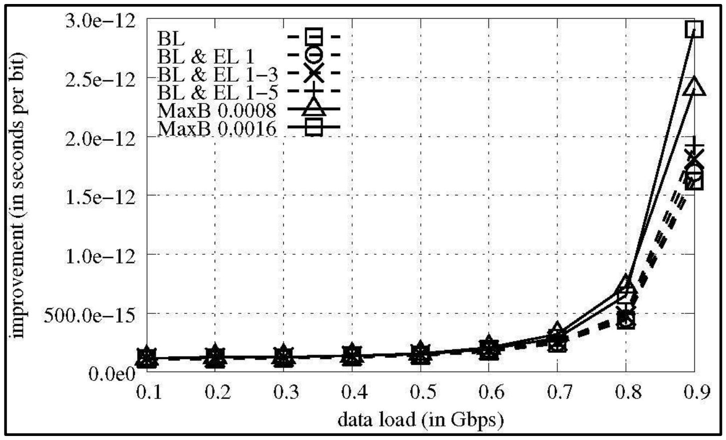

Figure 17c - Video Packet Delay Dropping Benefit for Silence of the Lambs

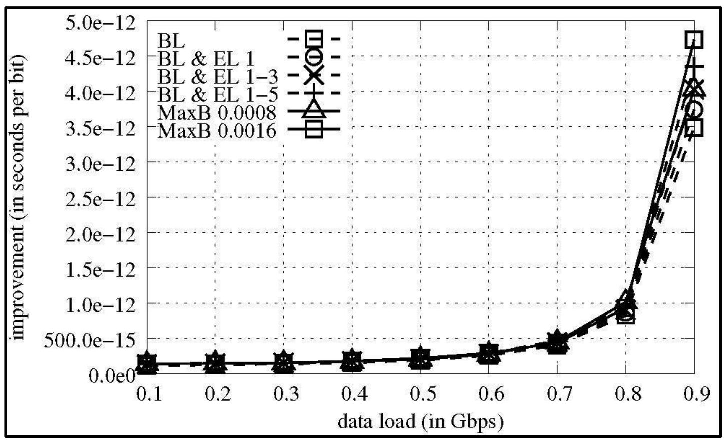

Figure 18a - Combined Data and Video Packet Delay Dropping Benefit for Gandhi

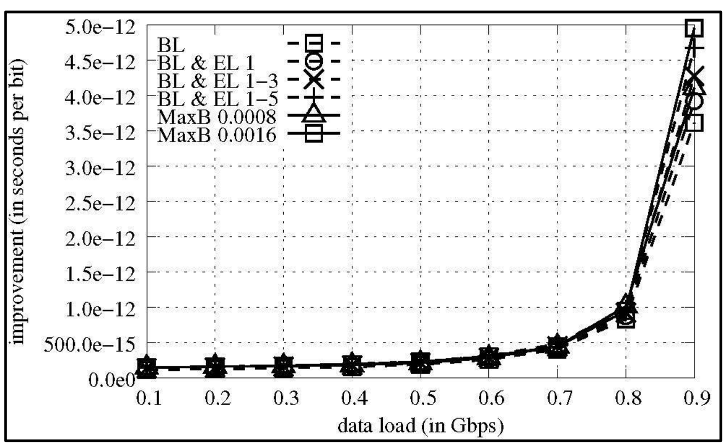

Figure 18b - Combined Data and Video Packet Delay Dropping Benefit for Indiana Jones

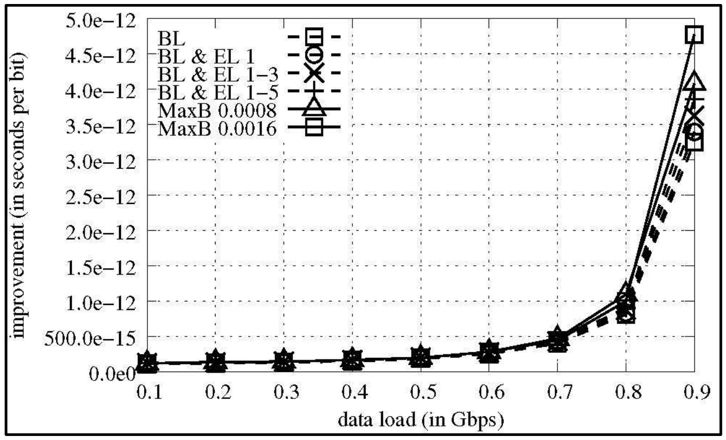

Figure 18c - Combined Data and Video Packet Delay Dropping Benefit for Silence of the Lambs

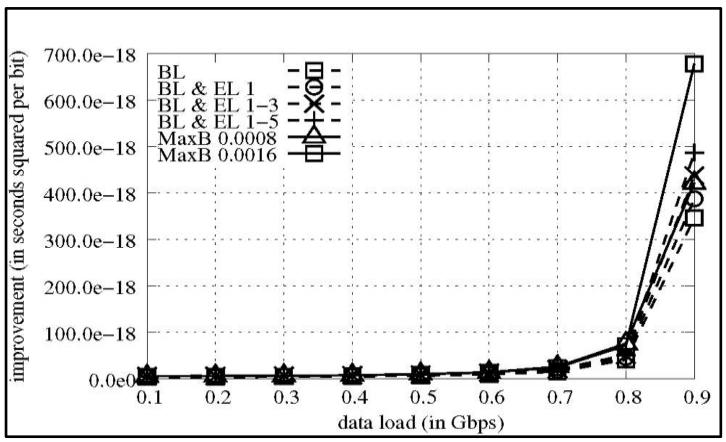

Figure 19a - Video Packet Delay Jitter Dropping Benefit for Gandhi

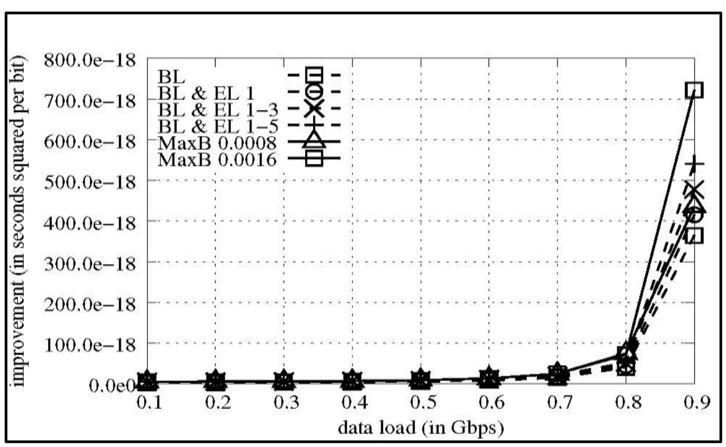

Figure 19b - Video Packet Delay Jitter Dropping Benefit for Indiana Jones

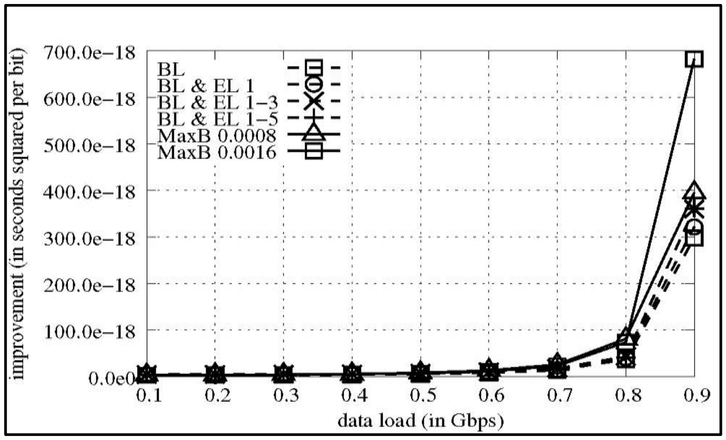

Figure 19c - Video Packet Delay Jitter Dropping Benefit for Silence of the Lambs

To better visualize the benefits of dropping video packets, a cost analysis was performed. In Figures 16(a,b,c), the amount of data packet delay improvement with respect to transmitting the entire video was divided by the number of bits dropped in order to achieve that delay improvement. In Figures 17(a,b,c) video packet delay improvement was divided by the number of bits sacrificed, and in Figures 19(a,b,c) the video packet delay jitter improvement was divided by the number of bits sacrificed. It is interesting to note that dropping less layers (and therefore less bits) is the most efficient type of dropping. Dropping only the sixth enhancement layer leads to the most delay improvement per bit for data, video, and jitter for the static dropping methods. Similarly, MaxB 0.0016, which dropped the least packets of all the adaptive dropping runs, is roughly as efficient as statically dropping the sixth, fifth, and fourth enhancement layers when considering data delay. It is also substantially more efficient than only dropping the sixth enhancement layer when considering video delay and video delay jitter. When the efficiencies of data packet delay and video packet delay are added, as seen in Figures 18(a,b,c), the adaptive dropping with MaxB 0.0016 is still substantially better than the second-best option of dropping only the sixth enhancement layer. The relatively small

amount of data dropped by the adaptive dropping with MaxB 0.0016 is depicted graphically in Figures 15(a,b,c) where even at its peak MaxB 0.0016 only drops somewhere between three and four enhancement layers.

6. CONCLUSION

We investigated the use of scalable video to optimize video playback over PONs and proposed sequential dropping of enhancement layers from the video traces in order to reduce data queueing delay, video queueing delay, and video delay jitter in the network. Dropping the first CGS enhancement layer is much more beneficial to the network than dropping the second or third enhancement layers. This benefit is maximized when network utilization is at its peak. This applies to data delay and video delay, but the most improved network metric is the video delay jitter. This is important as video delay jitter is a critical network performance metric for video playback. It is also noted that, while the delay reductions due to dropping an enhancement layer during low utilization to data delay are marginal, they are noticeable for the video delay.

We also investigated the effects of statically and adaptively dropping MGS scalable video layers in a PON on data packet delay, video packet delay, and video packet delay jitter. Sequentially dropping whole enhancement layers decreased data packet delay, video packet delay, and video packet delay jitter with the gaps between the layers growing larger as the network utilization increases. The maximum bound of the adaptive dropping algorithm was shown to simultaneously decrease the amount of data dropped and increase the network gains per bit dropped as the maximum bound increased. This trend was consistent across all maximum bound values tested with the highest maximum bound (0.0016) dropping the least data and providing the most significant benefits per bit dropped. Though the efficiency of adaptive dropping is lower than static dropping for data packet delay, the gains seen in video packet delay more than compensate, and video packet delay jitter improvement per bit is substantially increased.

7. WORKS CITED

1. Cisco, “The Zettabyte Era: Trends and Analysis,” Cisco Visual Networking Index, 2016.
2. M. P. McGarry, M. Reisslein, and M. Maier, “Ethernet Passive Optical Network Architectures and Dynamic Bandwidth Allocation Algorithms,” IEEE Communications Surveys Tutorials, vol. 10, no. 3, pp. 46–60, Third 2008.
3. R. J. Haddad, M. P. McGarry, and P. Seeling, “Video Bandwidth Forecasting,” IEEE Communications Surveys Tutorials, vol. 15, no. 4, pp. 1803–1818, Fourth 2013.
4. R. J. Haddad and M. P. McGarry, “Feed Forward Bandwidth Indication (FFBI): Cooperation for an Accurate Bandwidth Forecast,” Elsevier Computer Communications, vol. 35, no. 6, pp. 748–758, Mar 2012.
5. R. J. Haddad, M. P. McGarry, and Y. Luo, “Forecast Grant Sizing for Video Delivery over Passive Optical Networks,” in 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), July 2013, pp. 1–6.
6. H. Jin, K. Peng, J. Huang, and C. Pan, “Scalable Video Transmission System: Design and Field Trial,” in 2015 International Wireless Communications and Mobile Computing Conference (IWCMC), Aug 2015, pp. 1481–1485.
7. S. Huang, M. Sanna, E. Izquierdo, and P. Hao, “Optimized Scalable Video Transmission over P2P Network with Hierarchical Network Coding,” in 2014 IEEE International Conference on Image Processing (ICIP), Oct 2014, pp. 3993–3997.
8. P. Baccichet, T. Schierl, T. Wiegand, and B. Girod, “Low-delay Peer-to-Peer Streaming using Scalable Video Coding,” in Packet Video 2007, Nov 2007, pp. 173–181.
9. Y. Snchez, T. Schierl, C. Hellge, and T. Wiegand, “P2P Group Communication using Scalable Video Coding,” in 2010 IEEE International Conference on Image Processing, Sept 2010, pp. 4445–4448.
10. M. Klaghstan, D. Coquil, N. Bennani, H. Kosch, and L. Brunie, “Enhancing Video Viewing-Experience in Opportunistic Networks Based on SVC, An Experimental Study,” in 2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Sept 2013, pp. 3563–3567.
11. M. Klaghstan, N. Bennani, D. Coquil, H. Kosch, and L. Brunie, “Contact-based Adaptive Granularity for Scalable Video Transmission in Opportunistic Networks,” in 2014 International Wireless Communications and Mobile Computing Conference (IWCMC), Aug 2014, pp. 773–778.
12. M. Ghareeb, A. Ksentini, and C. Viho, “A Multipath Video Streaming Approach for SNR Scalable Video Coding (SVC) in Overlay Networks,” in 2011 IEEE Consumer Communications and Networking Conference (CCNC), Jan 2011, pp. 605–610.
13. M. Ghareeb, A. Ksentini, and C. Viho, “An Adaptive QoE-based Multipath Video Streaming Algorithm for Scalable Video Coding (SVC),” in 2011 IEEE Symposium on Computers and Communications (ISCC), June 2011, pp. 824–829.
14. P. Wydrych and P. Choda, “Locality- and Quality-Awareness for P2PTV Systems based on Scalable Video Coding (SVC),” in 2015 IEEE International Conference on Communications (ICC), June 2015, pp. 5319–5324.
15. Y. Lahbabi, E. H. Ibn Elhaj, and A. Hammouch, “Quality Adaptation using Scalable Video Coding (SVC) in Peer-to-peer (P2P) Video-on-Demand (VoD) Streaming,” in 2014 International Conference on Multimedia Computing and Systems (ICMCS), April 2014, pp. 1140–1146.
16. T. Schierl, C. Hellge, S. Mirta, K. Gruneberg, and T. Wiegand, “Using H.264/AVC-based Scalable Video Coding (SVC) for Real Time Streaming in Wireless IP Networks,” in 2007 IEEE International Symposium on Circuits and Systems, May 2007, pp. 3455–3458.
17. N. Ghazisaidi, M. Maier, and M. Reisslein, “VMP: A MAC Protocol for EPON-based

Video-dominated FiWi Access Networks," IEEE Transactions on Broadcasting, vol. 58, no. 3, pp. 440–453, Sept 2012.
18. P. Seeling and M. Reisslein, "Video Transport Evaluation with H.264 Video Traces," IEEE Communications Surveys and Tutorials, in print, vol. 14, no. 4, pp. 1142–1165, 2012, Traces available at trace.eas.asu.edu.
19. D. Kim, T. Fujii, and K. Lee, "Adaptive Modulation Method for MGS Video Streaming," in 3rd European Workshop on Visual Information Processing, July 2011, pp. 122–127.
20. H. Joo and H. Song, "H.264 MGS Coding-based IPTV System Architecture over WiMax Network," in 2010 IEEE International Conference on Multimedia and Expo, July 2010, pp. 1114–1119.
21. M. M. Hannuksela, H. Zhu, H. Li, and M. Gabbouj, "Congestion-Aware Transmission Rate Control using Medium Grain Scalability of Scalable Video Coding," in 2010 IEEE International Conference on Image Processing, Sept 2010, pp. 2929–2932.
22. Hosik Sohn, Hana Yoo, Young Bok Lee, Cheon Seog Kim, Wesley De Neve, and Yong Man Ro, "MPEG-21-based Scalable Bitstream Adaptation using Medium Grain Scalability," in TENCON 2008 - 2008 IEEE Region 10 Conference, Nov 2008, pp. 1–5.
23. H. Mansour, V. Krishnamurthy, and P. Nasiopoulos, "Rate and Distortion Modeling of Medium Grain Scalable Video Coding," in 2008 15th IEEE International Conference on Image Processing, Oct 2008, pp. 2564–2567.
24. R. Gupta, A. Pulipaka, P. Seeling, L. J. Karam, and M. Reisslein, "H.264 Coarse Grain Scalable (CGS) and Medium Grain Scalable (MGS) Encoded Video: A Trace Based Traffic and Quality Evaluation," IEEE Transactions on Broadcasting, vol. 58, no. 3, pp. 428–439, Sept 2012.
25. B. Grkemli, Y. adi, and A. M. Tekalp, "Effects of MGS Fragmentation, Slice Mode and Extraction Strategies on the Performance of SVC with Medium Grained Scalability," in 2010 IEEE International Conference on Image Processing, Sept 2010, pp. 4201–4204.
26. Moore, D., Haddad, R. "Utilizing Scalable Video to Optimize Video Playback over Passive Optical Networks", 2018 International Conference on Computing, Networking and Communications (ICNC), Maui, HI, 5-8, Mar, 2018.
27. Moore, D., Haddad R. "Employing H.264 MGS Video to Optimize Video Playback over Passive Optical Networks", SoutheastCon 2018, St. Petersburg, FL, Apr, 2018.